A COMPARISON OF LEARNING POTENTIAL RESULTS AT VARIOUS EDUCATIONAL LEVELS

Marié de Beer
Department of Industrial Psychology
UNISA

Paper presented at the 6th Annual Society for Industrial and Organisational Psychology of South Africa (SIOPSA) conference, 25-27 June 2003.

ABSTRACT

In the measurement of reasoning ability, persons from poor educational and/or socio-economic backgrounds may be at a disadvantage when standard cognitive tests are used. Standard cognitive tests often rely quite heavily on crystallized abilities which are influenced by prior learning experiences. Internationally and locally there have been calls to also consider the learning potential of individuals, since this allows to some extent for disparities in prior learning opportunities. Criticisms of learning potential measures have included the long administration time as well as the limited empirical research results. This presentation provides learning potential results of groups from various educational levels, with specific reference to construct and criterion validity results.

Debate about fairness in assessment has been around for a long time. Cognitive tests in particular have borne the brunt of the criticism of psychological tests (Nell, 1999). Nevertheless, due to their proven predictive utility in terms of overall job performance, they are generally perceived positively for inclusion in selection batteries (De Jong & Visser, 2000). Psychological tests that can be used for all cultural groups without discriminating against subgroups are needed in education and industry. Although in South Africa psychological tests in general have been viewed with considerable skepticism, in particular by those from previously disadvantaged groups, the last few years have seen a renewed appreciation for the contribution that these instruments can make in fair and equitable decision making - provided off course that the tests comply with legislative requirements (Employment Equity Act, 1998; Van der Merwe, 2002).

In an attempt to provide more equitable cognitive assessment, the measurement of learning potential has received increasing attention. Measurement of learning potential has as its aim to provide a learning opportunity within the test administration, in order to provide fairer assessment of in particular the disadvantaged groups. By providing a learning opportunity within the assessment, the focus is not only on the level of performance that the individual can reach at present (with possible limitations of his/her background), but also on providing an indication of the potential future levels of performance that can be reached if relevant training can be provided. Measurement of learning potential typically involves a test-train-retest strategy with some form of help or

training provided as part of the assessment process and thus, in particular, provides useful information for training and development as well as for affirmative action. While the measurement of learning potential has been receiving increased attention locally and internationally for providing alternative and supplemental information in the cognitive reasoning domain, researchers have lamented the fact that insufficient empirical research is hampering its progress (Grigorenko & Sternberg, 1998; Gupta & Coxhead, 1988; Guthke, 1992, 1993a, 1993b). The criticism against learning potential assessment – also known as dynamic assessment or dynamic testing because it involves some interaction or learning during testing – has mostly focused on the fact that they usually take much longer than standard tests to administer and that there is as yet limited empirical evidence in support of their psychometric properties.

South African researchers have contributed both in the development of instruments for the measurement of learning potential (De Beer, 2000a, 2000b; Taylor, 1994b) and also in research contributing to the available information on the validity of dynamic assessment measures (Boeyens, 1989; Lopes , Roodt & Mauer, 2001; Shochet, 1992, 1994; Taylor, 1992, 1994a, 1994b; Van Eeden, De Beer & Coetzee, 2001). Murphy (2002) provides an extensive overview of South African research in dynamic assessment.

Although some South African researchers have shown positive results with the use of dynamic assessment methods (Boeyens, 1989; De Beer, 2002; Shochet, 1992, 1994; Lopes, Roodt & Mauer, 2001), ongoing research is required to provide solid evidence in support of the use of these measures. In particular, Grigorenko and Sternberg (1998) stated that despite the obvious advantages and usefulness offered by dynamic assessment techniques, convincing empirical data are needed to ensure its further and ongoing general use. In the present study, the construct and criterion-related validity of the Learning Potential Computerised Adaptive Test (LPCAT) for groups at various educational levels was investigated.

The LPCAT is a dynamic, computerised adaptive test for the measurement of learning potential that was developed in South Africa. It uses non-verbal figural reasoning items in a test-train-retest format in an attempt to measure learning potential in the fluid reasoning ability domain so that language proficiency or formal academic qualifications should not impact significantly on performance (De Beer, 2000a, 2000b).

METHOD

Participants

Six groups with differing educational backgrounds were used to assess the utility of LPCAT learning potential results at different levels. A short description of each of the groups follow.

Group 1:

Group 1 was an adult learner group (N=194) which consisted of a mixed group of low-literate and literate African adults. Most of this group was male. The group was involved in assessment for vocational training as part of a retrenchment package agreement. The

mean age for this group was 29,7 years. Their level of formal education ranged from grade 1 to grade 12 with a mean level of education of grade 8.

Group 2:
The second group (N=56) were part of a group of grade 7 primary school pupils from previously disadvantaged groups who were involved in a selection process for a special educational project. There were 27 boys (48,2%) and 29 girls (51,8%) in the group. These pupils were specifically identified by their school principals as pupils with high potential. (The latter is important with regard to the level of their performance – as seen in the results reported later.) The results of the total group was available for the construct validity investigation (see Table 2), but it should be noted that since only 19 of the initial group were selected to the training program, the sample available for predictive criterion-related results was very small (see Table 3).

Group 3:
Group 3 consisted of 144 grade 8 pupils from an urban high school. The mean age for this group was 13,2 years. The gender distribution was 56,6% female (N=72) and 43,3% male (N=55) (22 missing values). Furthermore, 41% (N=52) had an African first language while 59% (N=75) had English or Afrikaans as first language. [Note that due to incomplete biographical data the overall sample size differs from the totals used above.]

Group 4:
Group 4 consisted of 166 students that were selected to a bridging training program. In this group 65,5% (N=107) were male and 34,5% (N=56) were female. The mean age for the group was 19,27 years. All these participants were from previously disadvantaged groups and these students have all completed their matric.

Group 5:
Group 5 consisted of 92 first-year Technikon students in the Science and Engineering faculties. Although some effort was made to involve all eligible students, testing was nevertheless voluntary, with the result that the sample cannot be regarded as statistically representative of all Technikon first-year science and engineering students. The mean age of the group was 19,8 years. The language distribution of the group was 50% African home language and 50% English/Afrikaans home language. The gender distribution was 11% female (N=10) and 89% male (N=82).

Group 6:
The last group consisted of 133 fourth-year Accounting students – although the LPCAT results of only 75 students were available. The gender distribution for the total group was 47,4% male (N=63) and 52,6% female (N=70). In terms of cultural distribution, 22,6% (N=30) were Black, 0,8% Coloured (N=1), 12% Indian (N=16) and 64,7% White (N=86). The students were studying at seven different universities – an aspect that could have affected the utility of the academic marks used for criterion data.

Measuring instruments

- Cognitive tests

The LPCAT (Learning Potential Computerised Adaptive Test) was used as learning potential measure for all six groups. The LPCAT definition of learning potential is that it is a combination of the present level of performance together with some credit for improvement shown from that level when relevant training is provided. With the LPCAT four scores are provided, namely a pretest score (present performance level), post-test score (potential future level of performance), difference score (undeveloped potential) and a composite score (reasoned combination of the previous three scores) (De Beer, 2000b). For the present study, the post-test score, which represents the projected future (optimal potential) level of performance, was used as the reported learning potential score together with the other results. The descriptive results for the LPCAT are reported in T-score format (mean of 50 and standard deviation of 10). The LPCAT coefficient alpha internal consistency reliability values range between 0,925 and 0,981 for groups in the standardisation sample (De Beer, 2000b). In terms of other cognitive measures (to investigate construct validity), the Paper-and-Pencil Games (PPG) (Claassen, 1996) was used for group 1. The PPG is a paper-and-pencil test which measures figural, quantitative and verbal skills that are closely related to scholastic achievement and is suitable for the third and fourth school years. It is a group test that serves a screening function. Raw scores out of a total of 50 are provided for both the verbal and nonverbal sections respectively. The total score is the sum of the two scores. The reliability of the test is reported as ranging between 0,78 and 0,95 and correlations of between 0,31 and 0,73 are reported in predicting academic achievement (Claassen, 1996). In the case of groups 3 and 5, the General Scholastic Aptitude Test (GSAT) (Claassen, De Beer, Hugo & Meyer, 1991) was used. In the case of group 6, results from the CPP (Cognitive Processing Profiler) were available. For Group 2 scores on subtests of the Junior Aptitude Test (JAT) were available. For Group 4 the general mental ability scores available were obtained by means of an in-house test for which no psychometric information was available.

- Academic and other criterion results

Where possible, academic results were obtained as criterion measures. For Group 1, results on Adult Basic Education and Training (ABET) literacy and numeracy results were available and obtained at approximately the same time (concurrent validity). For Group 2, subtest results on the Junior Aptitude Test (JAT) were obtained at the same time as the LPCAT results. The academic results for only 19 pupils in this group were available. For the junior secondary school group (group 3), average academic end-of-year results were used. In addition results in an English proficiency and Mathematics proficiency test were also obtained, using the Proficiency Test English Second Language (Intermediate Level (Chamberlain & Reinecke, 1992) and the Test of Basic Numerical Literacy (Venter, 1997). For Group 3 matriculation results were used. For both group 4, 5 and 6 academic results were obtained and used as criterion measures.

Procedure
The groups were tested at different times. Where possible, results from other cognitive tests were also obtained for the purpose of investigating the construct validity of the LPCAT and training or academic results gathered for investigation of criterion validity.

Statistical analysis
Initially, descriptive results were obtained for all the relevant variables and the means of the groups compared. For further visual inspection, distribution of scores of the different groups were also obtained. Correlation and regression analysis were performed to investigate the relationships between the variables.

RESULTS

- Descriptive results

<See Table 1 >

The descriptive results indicate the levels of LPCAT post-test performance for the groups at different educational levels. In considering the magnitude of the differences found between the means of different groups, in addition to investigating the statistical significance of such differences, the effect size of the differences found were also considered (Cohen, 1969) where the difference can be expressed in standard deviation units.

ANOVA results indicate significant differences between the groups with regard to their LPCAT performance levels (see Table 2). The post-hoc Scheffe test indicated that the only groups between which there were *no* significant differences in the LPCAT post-test results were groups 2 and 3 and groups 4 and 5. This can be expected, since group 2 participants were tested at the end of their grade 7 year, while those in group 3 were at the start of their grade 8 year - thus very close together in terms of their educational level. Participants in both group 4 and group 5 had completed their grade 12 qualification and were thus also at the same educational level. In considering the effect sizes when group means were compared, three comparisons were made. Firstly, in comparing the LPCAT results of the adult lower literacy group (group 1) with the senior primary group (group 2), a large effect size (d=1,60) was found. For comparison of group 3 and group 4 and for group 5 and group 6 large effect sizes (d=1,59 and d=0,91 respectively) were found. This indicates that the differences found between the groups at different educational levels are also practically significant.

- Distribution of scores

<See Figures 1 – 2>

The distribution of the LPCAT post-test scores indicate the different levels of performance for the groups at the different levels of education - although within each

group, a fair distribution of scores can also be seen. These results show that, while the LPCAT can distinguish between groups at different levels, there is also adequate variance of scores at each of the different levels.

- Correlation and Regression results

<See Tables 3 -6>

Correlation results show the specific relationship of the LPCAT post-test results with various other variables. The regression results indicate the relationship of a group of independent variables taken together with a particular criterion variable. The correlation results indicate statistically significant relationships between the LPCAT post-test and other cognitive measures – providing support for the construct validity of the LPCAT (see Table 3). With regard to correlations with other (mostly academic) criterion results (see Table 4), statistically highly significant results were found for groups 1 to 4. The magnitude of correlations interpreted as effect size for these groups are between medium and large. For Group 5, statistically highly significant correlations were found between the LPCAT post-test scores and first semester average as well as English proficiency, while a statistically significant correlation was found between the LPCAT post-test and the first year academic average mark. For this group the correlation effect size can be considered medium. For group 6, no statistically significant correlations with academic criteria were found.

In the regression results the LPCAT post-test came out as a significant contributor to the prediction of numeric results for group 1 and average academic performance for group 4 when entered together with the standard cognitive test results that were available.

DISCUSSION

The results indicate that the LPCAT provides useful information in terms of indicating the level of general reasoning ability shown by the individual. It can therefore indicate at what academic level an individual is likely to be able to perform. There is furthermore adequate variance within the different levels to show that it can indicate different levels of performance for persons at approximately the same academic level. In terms of construct validity, results indicate that the LPCAT does measure the typical reasoning ability measured by other cognitive tests. With regard to prediction of (mostly academic) criterion results, useful results are shown at levels up to post-matric, but correlations and accuracy of prediction seems noticeably lower at tertiary level – in particular for the higher university levels. Huysamen (1999) also indicated that the reliability of criterion scores and restriction of range of predictor or criterion scores partly explained some of the problems typically encountered at these levels. Formulae are available for correcting the obtained correlations for the effect of restriction of range, but this does not really offer a satisfactory solution. It should be kept in mind that the particular university level group whose results were used consisted of participants from seven different universities, and the incomparability of academic marks across universities may also have affected the

results. Further research at university level with larger samples from a single institution could provide useful further information.

The overall results provide support for the use of the LPCAT as a sifting instrument to assist in decision making for training and development decisions, when combined with other information such as language proficiency, specific aptitude, interest or personality. In particular it provides useful information for the appropriate level of training for individuals. The advantage of the LPCAT is that performance is not reliant on language proficiency or formal academic qualification, making it a culture-fair measure to include in assessment batteries in the South African context.

REFERENCES

Boeyens, J. (1989b). *Learning potential: an empirical investigation. Report PERS-435*, Human Sciences Researches Council. Pretoria: Human Sciences Research Council.

Chamberlain, J.C. & Reinecke, S. (1992). *Manual - Proficiency Test English Second Language Intermediate Level*. Pretoria: Human Sciences Research Council.

Claassen, N.C.W. (1996). *Paper and pencil games (PPG): Manual*. Pretoria: Human Sciences Research Council.

Claassen, N.C.W. (1997). Cultural differences, politics and test bias in South Africa. *European Review of Applied Psychology, 47(4)*, 297-307.

Claassen, N.C.W., De Beer, M., Hugo, H.L.E. & Meyer, H.M. (1991). *Manual for the General Scholastic Aptitude Test (GSAT) Senior Series*. Pretoria: Human Sciences Research Council.

Cohen, J. (1969). Statistical power analysis for the behavioural sciences. New York: Academic Press.

De Beer, M. (2000a). *Learning Potential Computerised Adaptive Test (LPCAT): User's Manual*. Pretoria: Production printers (UNISA)

De Beer, M. (2000b). *Learning Potential Computerised Adaptive Test (LPCAT): Technical Manual*. Pretoria: Production printers (UNISA)

De Beer, M. (2002). Utility of Learning Potential Computerised Adaptive Test (LPCAT) scores in predicting academic performance of bridging students: A comparison with other predictors. Poster presented at the 5th Annual Society for Industrial and Organisational Psychology Congress, Pretoria, 14-17 June 2002.

De Jong, A. & Visser, D. (2000). Black and White employees' fairness perceptions of personnel selection techniques. *South African Journal of Psychology, 30(4), 17-*23.

Employment Equity Act. (1998). No 55. Republic of South Africa.

Grigorenko, E.L, & Sternberg, R.J. (1998). Dynamic testing. *Psychological Bulletin, 124(1)*, 75-111.

Gupta, T.M. & Coxhead, P. (1988). Why assess learning potential? In R.M. Gupta & P. Coxhead (Eds.), *Cultural diversity and learning efficiency: Recent developments in assessment* (pp. 1 - 21). Hong Kong: Macmillan Press.

Guthke, J. (1992). Learning tests - the concept, main research findings, problems and trends. *Learning and Individual Differences, 4(2)*, 137-151.

Guthke, J. (1993a). Current trends in theories and assessment of intelligence. In J.H.M. Hamers, K. Sijtsma & A.J.J.M. Ruijssenaars (Eds.), *Learning potential assessment: Theoretical, methodological and practical issues* (pp. 13-20). Amsterdam: Swets & Zeitlinger.

Guthke, J. (1993b). Developments in learning potential assessment. In J.H.M. Hamers, K. Sijtsma & A.J.J.M. Ruijssenaars (Eds.), *Learning potential assessment: Theoretical, methodological and practical issues* (pp. 43-68). Amsterdam: Swets & Zeitlinger.

Huysamen, G.K. (1999). Psychometric explanations for the poor predictability of the tertiary-academic performance of seducationally disadvantaged students. *South African Journal of Higher Education, 13(1)*, 132-138.

Loopes, A., Roodt, G. & Mauer, R. (2001). The predictive validity of the Apil-B in a financial institution. *Journal of Industrial Psychology, 27(1)*, 61-57.

Murphy, R. (2002) A review of South African research in the field of dynamic assessment. Unpublished MA thesis. Pretoria: University of Pretoria.

Nell, V. (1999). Standardising the WAIS-III for South Africa: legislative, psychometric, and policy issues. *South African Journal of Psychology, 29(3)*, 128-137.

Shochet, I.M. (1992). A dynamic testing for undergraduate admission: The inverse relationship between modifiability and predictability. In H.C. Haywood & D. Tzuriel (Eds.), *Interactive testing* (pp. 332-355). New York: Springer-Verlag.

Shochet, I.M. (1994). The moderator effect of cognitive modifiability on a traditional undergraduate admissions test for disadvantaged black students in South Africa. *South African Journal of Psychology, 24(4)*, 208-215.

Taylor, T.R. (1992). *Beyond competence: Measuring potential in a cross-cultural situation fairly: Potential in Psychometrics (part two).* Paper presented at the 1992 South African Psychometrics Congress, June 1992.

Taylor, T.R. (1994a). A review of three approaches to cognitive assessment, and a proposed integrated approach based on a unifying theoretical framework. *South African Journal of Psychology, 24(4)*, 184-193.

Taylor, T.R. (1994b). *Learning potential: Theory and practical assessment*. Paper presented at the Annual Department of Industrial Psychology Conference on Fairness in Testing and Assessment (30 August 1994). Pretoria: Holiday Inn.

Van der Merwe, R.P. (2002). Psychometric testing and human resources management. *South African Journal of Industrial Psychology, 28(2)*, 77-82.

Van Eeden, R., De Beer, M. & Coetzee, C.H. (2001). Cognitive ability, learning potential, and personality traits as predictors of academic achievement by engineeringand other science and technology students. *South African Journal of Higher Education, 15(1),* 171-179.

Venter, E.J. (1997). Manual: *Test of Basic Numerical Literacy*. Pretoria: Human Sciences Research Council.

ATTACHMENTS

TABLE 1
DESCRIPTIVE RESULTS FOR LPCAT POST-TEST SCORES *

Group	N	Mean	SD	Min	Max
Group 1	194	37,76	8,998	23	60
Group 2	56	51,11	5,349	33	59
Group 3	128	47,83	7,496	27	61
Group 4	158	57,61	4,825	44	68
Group 5	92	59,15	5,218	45	71
Group 6	75	63,36	3,851	56	74

* Note that due to missing values for certain variables as a result of for instance absences on some of the days of testing, the sample sizes for available LPCAT information may differ from sample sizes for other information

TABLE 2
ANOVA RESULTS FOR LPCAT POST-TEST SCORE DIFFERENCES BETWEEN GROUPS AT DIFFERENT LEVELS

	Sum of squares	df	F	Significance
Between groups	59678,65	5	263,539	0,000
Within groups	31567,268	697		
Total	91245,917	702		

TABLE 3
CORRELATIONS OF LPCAT POST-TEST SCORES WITH OTHER COGNITIVE TESTS

Group	Description	N	R	p-value
Group 1 (Adults - low literacy)	PPG Verbal	110	0,408**	0,000
	PPG Nonverbal	110	0,645**	0,000
	PPG Total	110	0,610**	0,000
Group 2 (Grade 7 - high ability group)	JAT verbal reasoning	56	0,353**	0,008
	JAT nonverbal reasoning	56	0,474**	0,000
	JAT spatial 3D	56	0,448**	0,000
	JAT Mechanical Insight	56	0,246	0,068
Group 3 (Grade 8)	GSAT Verbal	120	0,613**	0,000
	GSAT Nonverbal	120	0,665**	0,000
	GSAT Total	120	0,691**	0,000
Group 4 (Grade 12+)	Non-standard cognitive test (general mental ability)	146	0,661**	0,000
Group 5 (Technikon 1st year)	Blox	75	0,383**	0,001
	GSAT Verbal	76	0,653**	0,000
	GSAT Nonverbal	76	0,693**	0,000
	GSAT Total	76	0,713**	0,000
Group 6 (University 4th year)	CPP verbal	69	0,238*	0,049
	CPP analytical	69	0,369**	0,002
	CPP integrative	69	0,338**	0,005
	CPP learning	69	0.376**	0,001
	CPP metacognitive	69	0,100	0,414
	CPP environment	69	0,283*	0,019

* $p < 0,05$
** $p < 0,01$

TABLE 4
CORRELATIONS OF LPCAT POST-TEST SCORES WITH CRITERION RESULTS

Group	Description of criterion	N	r	p-value
Group 1 (Adults - low literacy)	Literacy Level 1	182	0,437**	0,000
	Numeracy Level 1	182	0,491**	0,000
	Literacy Level 3	111	0,461**	0,000
	Numeracy Level 3	26	0,610**	0,001
Group 2 (Grade 7 - high ability group)	English	19	0,415	0,077
	Mathematics	19	0,480*	0,038
	Science	19	0,418	0,075
	Year Average	19	0,334	0,162
Group 3 (Grade 8)	English proficiency	128	0,600**	0,000
	Mathematics proficiency	128	0,577**	0,000
	Academic average results	116	0,530**	0,000
Group 4 (Grade 12+)	Numeracy score (ns)	146	0,392**	0,000
	Language score (ns)	146	0,406**	0,000
	2nd Matriculation Mathematics	158	0,525**	0,000
	2nd Matriculation Science	158	0,454**	0,000
	2nd Matriculation Biology	58	0,313*	0,017
	2nd Matriculation Average	158	0,510**	0,000
Group 5 (Technikon 1st year)	Mathematics Semester 1	77	0,138	0,230
	Academic average semester 1	89	0,319**	0,002
	First year average	89	0,230*	0,030
	English proficiency	76	0,460**	0,000
Group 6 (University 4th year)	Accounting	46	0,008	0,956
	Management Accounting	47	0,071	0,635
	Tax	47	0,166	0,264
	Auditing	47	0,008	0,957
	Academic average	46	0,050	0,740

* $p < 0,05$
** $p < 0,01$

TABLE 5 SUMMARY OF REGRESSION RESULTS FOR GROUP 1(ENTER METHOD USED – DEPENDENT VARIABLE NUMERACY LEVEL 1)

ANOVA	SS	df	F (Sign)	Adj R-squared
Regr	878,357	3	13,636 (0,000)	0,258
Residual	2275,907	106		
Total	3154,264	109		
Coefficients	**B**	**Beta**	**T**	**Sig.**
LPCAT post-test	0,244	0,421	3,887	0,000
Cognitive verbal	0,106	0,141	1,425	0,157
Cognitive non-verbal	0,0019	0,044	0,375	0,708

TABLE 6 SUMMARY OF REGRESSION RESULTS FOR GROUP 4 (ENTER METHOD USED – DEPENDENT AVERAGE FINAL MATRICULATION RESULTS)

ANOVA	SS	df	F (Sign)	Adj R-squared
Regr	19864,674	6	33,981 (0,000)	0,579
Residual	13445,285	138		
Total	33309,959	144		
Coefficients	**B**	**Beta**	**T**	**Sig.**
LPCAT post-test	0,520	0,162	2,124	0,035
Numeracy	0,589	0,606	9,664	0,000
General mental ability	2,764	0,288	1,522	0,130
Language	2,994	0,353	1,579	0,117
Science	-11,085	-0,949	-1,956	0,052
Technical	6,031	0,522	1,217	0,226

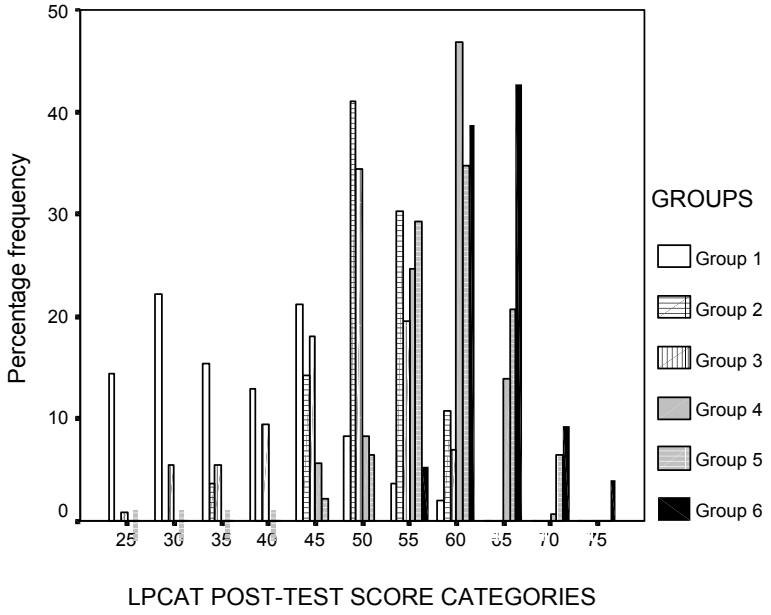

Figure 1: Distribution of post-test scores – all groups together

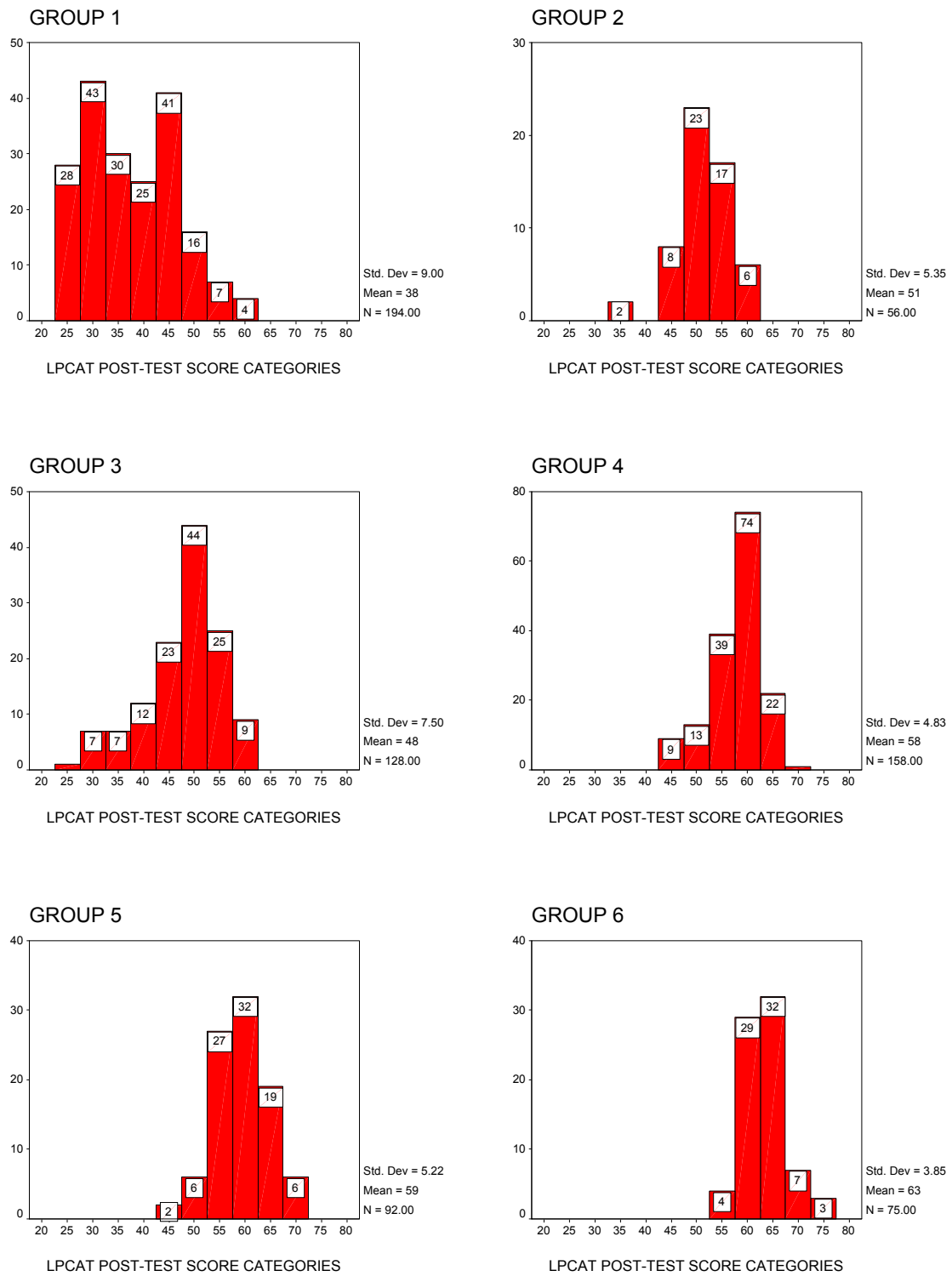

Figure 2: Frequency distribution of LPCAT post-test scores per group