In the multicultural and multilingual South African context, differences in socio-economic and educational background and development opportunities complicate psychological assessment (Claassen, 1997; Foxcroft, 1997; 2004). In this complex context, the measurement of learning potential provides additional information in the cognitive domain and has shown positive results in terms of psychometric properties and practical utility (De Beer, 2006; 2010b; Lidz, 1987; Murphy & Maree, 2006). Measurement of learning potential implies that assessment includes a learning experience or assistance, and typically adopts a test-train-retest approach. Measurement is therefore expanded to include two sets of measures, as well as a learning opportunity relating to the test task. Such assessments are also referred to as dynamic assessment (DA). DA allows for learning experiences to take place during assessment with a view to measuring learning potential, thus measuring not only the present level of performance of individuals but also the projected or potential future levels of performance these individuals may be able to attain if relevant learning opportunities can be provided.

This approach to assessment is generally associated with Vygotsky’s (1978) theory of the ‘zone of proximal development’ (ZPD). This theory distinguishes between current performance (without help) – also referred to as the ‘zone of actual development’ (ZAD) – and performance that can be attained when relevant learning opportunities are provided, the ZPD. In DA, this same distinction is made in terms of a focus on measures obtained in a pre-test (unassisted performance) and a post-test (performance after learning or with assistance). Of importance is the fact that in interpreting the results of learning potential assessment of persons with varying educational qualifications, and across a wide range of ability/performance levels, learning potential is defined as the combination of current and projected future (potential) performance, and not only in terms of the improvement score (De Beer, 2000a; 2010a). The focus is not only on whether the individual will generally be able to profit from learning, but more specifically on what level of learning/training he or she will be able to cope with – or alternatively, to what degree he or she seems able to cope with a particular level of training offered. At a practical level, current and projected future (potential) levels of performance can be compared to the opportunity to
evaluate whether the individual is currently already at or close to the required level (that is, the target level of training), or shows the potential to perform at or close to the required level.

Traditionally, intelligence quotient (IQ) scores have been seen as immutable and static, a view which has contributed to the strong emotional reactions often associated with this domain. These scores are, however, subject to changes and improvement – generally referred to as the Flynn effect (Flynn, 1987) – where scores on tests have shown increases over time (Wicherts, Dolan, Carlson & Van der Maas, 2010). These changes – which occur without any purposeful intervention – are normally ascribed to various factors, such as increased test-wiseness (Rushton & Jensen, 2010) and other environmental factors (Flynn, 1987) – for example, improvement of educational opportunities and socio-economic standing. Furthermore, IQ gains over time have shown the largest gains to occur in culturally reduced tests and tests of fluid intelligence (Flynn, 1987).

At the heart of DA or the measurement of learning potential is the provision of a learning experience within the assessment. Dynamic testing or learning potential testing focuses on providing learning experiences that might improve performance, and scores have been reported to increase by 0.5 to 0.7 standard deviations (Te Nijenhuis, Van Vianen & Van der Flier, 2007). In DA, the aim of providing a learning experience to allow for improvement in the level of performance by focusing on measurement of fluid ability could therefore allow for optimal improvement in an environment where further hints, guidelines and strategies focused on improving performance are provided.

Although the concept of dynamic testing is generally well supported, its practical use has been hampered by problems concerning, inter alia, lengthy testing times, high costs, a lack of standardised procedures, problems with measurement accuracy, a limiting focus on underachieving populations and a sparseness of validity information available (Grigorenko & Sternberg, 1998; Sternberg & Grigorenko, 2002).

This chapter begins with a brief history of DA, and then provides specific information on the Learning Potential Computerised Adaptive Test (LPCAT), as an example of a South African learning potential test that uses modern psychometric and assessment techniques to overcome some of the limitations and problems generally associated with DA (Kim-Kang & Weiss, 2008). The development of the LPCAT is described, and typical strategies for use of the scores obtained from it are explained. Furthermore, empirical psychometric results for the LPCAT in the South African context are presented. Lastly, the features and challenges of the LPCAT are discussed.

A short history of DA and the development of the LPCAT

The history of learning potential (or dynamic) assessment goes back quite far (Wolf, 1973). The well-known Binet-Simon test developed around the turn of the 20th century can be regarded as the very first learning potential test
Its aim was to identify individuals who could improve their performance when they were afforded a relevant learning opportunity. This kind of assessment again became prominent in the 1970s and 1980s. Since then, numerous researchers have been involved in DA and various approaches to DA have evolved. Lidz (1987), Murphy and Maree (2006) and Haywood (2008) provide more detail of the history of, and different approaches to, DA. Two broad approaches can be identified – on the one hand, the more clinically and diagnostically oriented approach, with remediation as the main aim, and on the other hand, the more measurement- or psychometrically oriented approach, with accurate assessment and obtaining good psychometric properties as the main aim (De Beer, 2010a). The LPCAT falls within the latter of these two broad categories.

Assessment research in the South African context has shown that individuals from disadvantaged educational and socio-economic backgrounds often under-perform on standard cognitive tests (Claassen, 1997; Owen, 1998). These standard cognitive tests often include a large proportion of language-based questions, as well as education-related content such as numerical reasoning. Research has furthermore shown that nonverbal figural item content is fairer to disadvantaged individuals (Hugo & Claassen, 1991). Learning potential assessment – using nonverbal figural content only – provides an alternative measurement approach that can provide additional information to that which can be obtained from standard static tests.

A focus on learning potential is important in the South African context, where the vast majority (72 per cent) of the population aged 20 years and older have completed less than secondary education (Statistics South Africa, 2008). This means that a large number of individuals who may need to be assessed are at some disadvantage when measures rely on language proficiency and educational material – as is often the case in standard cognitive assessments. Learning potential assessment is not intended to replace any other assessments, but can provide additional information not available in standard tests to improve decision-making relating to the training and development of individuals, screening for selection and vocational appointments or for training opportunities, and career-related assessment and guidance. For specific aptitudes or choices of particular fields of study, other measures can provide the relevant information (such as aptitude, intelligence, personality and interest-related assessments, amongst others). Learning potential assessment results indicate the level of reasoning (albeit nonverbal figural reasoning) that the individual is currently capable of, as well as the potential future levels of such reasoning that the individual is likely to attain if he or she is afforded relevant learning opportunities. Hence, if the focus is on the development of individuals and improvement of their educational levels, or identification of the appropriate levels of training to provide for future development, learning potential assessment results provide useful additional information.

The concept of learning potential is in line with legislation (the Employment Equity Act No. 55 of 1998) regarding psychological assessment in South Africa. It makes allowance for the fact that not everyone has had the same educational and socio-economic opportunities, and acknowledges the research that has shown that these factors are related to performance in standard cognitive tests.
In the case of the LPCAT, instructions to administer the test have been translated and are available in the *User's Manual* in all 11 official South African languages (De Beer, 2000a), which allows for administration to individuals who may have limited English proficiency, provided that test administrators are fluent in the particular language of the individual being tested.

**An overview of the LPCAT**

The LPCAT is a dynamic and adaptive learning potential test focused on the measurement of learning potential within the general fluid reasoning ability or ‘*g*’ domain. It is ‘intended to serve as a screening instrument that can be used mainly to counter inadvertent discrimination against disadvantaged groups’ (De Beer, 2000b, p.1). It uses nonverbal figural material (Figure Series, Figure Analogies and Pattern Completion) in the test items to exclude language and scholastic content, since these item types show less bias in multicultural assessment, whereas verbal scales in particular often underestimate the cognitive ability of African-language examinees (Claassen, De Beer, Hugo & Meyer, 1991; Hugo & Claassen, 1991; Owen, 1998). Responding to such nonverbal figural pattern items requires common reasoning skills such as identification, comparison and recognition of relations (see Figure 10.1 for an example item).

**Figure 10.1** LPCAT example item

Because of the practical need in South Africa for instruments that can be group-administered, and used to identify (often disadvantaged) individuals over a broad spectrum of ability who show the potential to benefit from further training and development, a link was made between DA and Computerised Adaptive Testing (CAT) based on item response theory (IRT). These modern psychometric methods (IRT for item analysis and CAT in test administration) were employed in the development of the LPCAT (Embretson, 1996; 2004). At the core of IRT methods are three features: item difficulty and individual ability are measured on the same scale; item characteristics are sample-independent; and individual abilities are item-independent (Embretson, 1996; Weiss, 1983). This makes possible a form of CAT in which a unique set of items is selected for each individual during test administration, so that items presented to each individual are continually and interactively selected from the bank of available items to match the estimated...
ability of the individual at that point in time (Weiss, 1983). IRT furthermore allows for accurate measurement of difference scores, and CAT shortens the testing time (Kim-Kang & Weiss, 2008; Sijtsma, 1993a; 1993b; Van der Linden, 2008a; 2008b; Weiss, 1983).

IRT-based analysis was used to perform bias analysis of items (in terms of gender, culture, language and level of education) with a large (N = 2 450) representative sample (De Beer, 2000b). Classical test theory as well as IRT item analysis was performed, and items that did not meet the criteria in terms of measurement properties or differential item functioning (DIF) were discarded in the compilation of the final test (De Beer, 2000b; 2004). The item characteristic curves (ICCs) of different subgroups were compared to determine the extent of DIF (see Figure 10.2). The base scale in Figure 10.2 (indicated by theta (θ)) depicts both difficulty levels of items and ability levels of individuals, while on the Y-axis the probability of an individual at a specific level of ability answering this item correctly is shown as P(θ). The difficulty level of the item (indicated by the letter b) is determined by the theta-level (ability level) where the probability of a correct response is 0.5 (Weiss, 1983).

Figure 10.2 DIF analysis – culture group comparison

Two separate but linked adaptive tests were used for the pre-test and post-test respectively, and total testing time is approximately one hour (for details on the development of the LPCAT see De Beer, 2000b; 2005; 2010a). Advantages of using computerised adaptive methods are that testing time is shortened, and the results are available immediately after completion of the test. However, although the test is administered on a computer, candidates need to use only the space bar and enter key – hence computer literacy is not a requirement for its administration. The results are presented in graph form (see Figure 10.3) from which a report can be (manually) prepared. The levels of performance in both the pre-test and the post-test should be noted, as well as the pattern and gradients of the graphs. The training that is provided between the pre-test and the post-test is aimed at elucidating the applicable reasoning strategies, by providing more example questions in which the basic principles, building blocks and general
strategies for answering the particular types of questions are provided. While practice with further questions might have some effect, use of IRT methods of measurement results in more accurate measurement of the latent trait concerned – in this case, level of fluid general reasoning ability. Furthermore, no questions are repeated in the pre-test and post-test, which precludes the undue effect of memory on performance in the post-test.

Figure 10.3 Example of LPCAT graphic output

Administration of the LPCAT
As a dynamic CAT, the LPCAT has to be computer-administered to allow for the interactive selection of appropriate items for each individual, depending on the specific response pattern and the estimated performance level at the time. For ease of administration – in particular to persons with lower levels of formal qualification – only the space bar and the enter key are used to answer the multiple-choice questions. The use of the interactive CAT is possible, because in IRT the item difficulty levels and individual ability levels are measured on the same scale (Van der Linden, 2008a; 2008b; Weiss, 1983). In CAT, a bank of pre-calibrated items is available for presentation during the testing process. Unlike standard tests, in which all individuals who take the test complete exactly the same items in the same sequence, CAT presents a selection of items unique to each individual, continuously selecting items to be presented on the basis of their difficulty level, matching the individual’s estimated ability level at that point in time. Not only can different items be presented to each individual, but candidates can also receive different numbers of items. A minimum and maximum number
of items are pre-set to be used during test administration. No individual will receive fewer than the minimum number of items, and no individual will receive more than the maximum number of items. Test termination is only partially linked to the number of items; it is also linked to the accuracy of measurement, which in turn depends on the psychometric or measurement quality of items presented. Entry level to the pre-test is set, and thereafter the following steps are repeated until the testing is terminated:

- The first item presented is the item that measures best at the predetermined entry level (that is, the best psychometric quality item available in the bank that has a difficulty level closest to that particular level of ability).
- When the respondent answers the question, three things happen:
  - If the item is answered correctly, the respondent’s estimated ability level is readjusted upwards – assuming that since the question aimed at the entry level of ability was answered correctly, the respondent has a higher level of ability. If the item is answered incorrectly, the respondent’s estimated ability level is adjusted downwards – assuming that since the question aimed at the entry level of ability was answered incorrectly, the respondent has a lower level of ability.
  - The item characteristics of the item presented are also used to calculate an accuracy index, reflecting the accuracy of the ability estimation at that time. A check is done to determine whether the termination criteria are met – if they are, the test is terminated.
  - If the test is not terminated, the next question selected will be the one in the bank that measures most accurately and provides the best information at the current newly estimated ability level.
- When the next item is presented, the process starts repeating – with a check for the number of items presented each time and a check for whether the required accuracy level has been achieved. All respondents will receive the minimum number of items. Thereafter, the test will terminate as soon as the required accuracy level (of the ability estimation) is attained, or as soon as the maximum number of items set have been administered – whichever comes first.

CAT has several positive features, including improving motivation by presenting items of appropriate difficulty level throughout testing, thereby not overwhelming or boring participants with items of an inappropriate difficulty level.

There is no fixed test administration time, due to the adaptive test process described above, but testing generally takes approximately one hour to complete. This includes the introduction, pre-test, training phase and post-test, and on completion the results are available immediately. When testing for various groups is arranged, it usually suffices if test sessions are arranged for one-and-a-half hours apart – since this should generally allow sufficient time for all examinees to complete the test.

**Language versions of the LPCAT**
There are two versions of the LPCAT – a version with either English or Afrikaans text on screen, and a version with no language on the screen – for which
instructions to be read have been translated into all 11 official South African languages (De Beer, 2000a). In order to use the text-on-screen version, a reading proficiency level of at least Grade 6 or 7 in the language of administration is required. All software is installed during the installation process, and the selection of the language for testing is chosen per individual during the test administration process when entering the respondent's details (the options being 'English' or 'Afrikaans' or 'None'). When the 'None' (or 'no language') option is chosen, it implies that the instructions for administration have to be read from the User's Manual in the language chosen. For practical purposes in the case of the latter version, the group should be homogeneous in terms of the language to be used for test administration. In terms of age, it can be administered to respondents aged 11 years and older. In terms of educational level for adults, there is no minimum level required and it can be administered to illiterate adults too.

When the LPCAT is administered to groups, it is essential for all members of the group to complete the same version of the test – either all receiving full instructions and feedback on the example questions with text on the screen in the chosen language (English or Afrikaans) or, for the 'no language' or 'None' option, all seeing only the nonverbal figural patterns on their screens and having the instructions read aloud to them in the chosen appropriate language. Test administration sessions cannot allow for a mixture of the two versions, because those attempting to read instructions or feedback from the screen will be disturbed by the instructions being read aloud for the version in which no text appears on the screen.

The two versions of the LPCAT have different entry levels (in terms of the initial estimated ability level of the individual to start the adaptive testing process). For the version in which the instructions are provided on the screen, the entry level is at the mean level – that is, at a T-score of 50 – which is equivalent to a mid-secondary level (see Table 10.1). Once the first item has been answered, the adaptive process described earlier will commence. For the version of the LPCAT in which no instructions appear on the screen, and in which the instructions are read aloud to the respondents/candidates, the test commences at one standard deviation below the mean – that is, at a T-score level of 40 – which is equivalent to a senior primary level (see Table 10.1). It should be kept in mind that the entry level will not determine or influence the final levels attained in either the pre- or post-test, since the adaptive test administration will ensure that appropriately difficult (or easy) items are administered to match the individual's estimated ability level throughout the test session. Exactly the same introductory practice examples, and example items for the training between the pre- and the post-test, are used for the two versions. In the version in which the instructions and feedback are presented with the text appearing on the screen, the respondents can work independently through the introduction and initial practice examples, pre-test, training and post-test at their own pace. For the ‘no language’ version, instructions and feedback are read to the candidates, who should view the specific (and same) screens while the instructions for that screen are being read from the User's Manual (De Beer, 2000a).
Table 10.1 LPCAT score ranges in relation to NQF levels and educational levels

<table>
<thead>
<tr>
<th>LPCAT T-score range</th>
<th>LPCAT Stanine score</th>
<th>ABET / NQF level</th>
<th>Educational level</th>
</tr>
</thead>
<tbody>
<tr>
<td>20–32</td>
<td>1</td>
<td>ABET level 1</td>
<td>Grades 0–3 (Junior Primary)</td>
</tr>
<tr>
<td>33–37</td>
<td>2</td>
<td>ABET level 2</td>
<td>Grades 4–5 (Middle Primary)</td>
</tr>
<tr>
<td>38–42</td>
<td>3</td>
<td>ABET level 3</td>
<td>Grades 6–7 (Senior Primary)</td>
</tr>
<tr>
<td>43–47</td>
<td>4</td>
<td>ABET level 4/NQF 1</td>
<td>Grades 8–9 (Junior Secondary)</td>
</tr>
<tr>
<td>48–52</td>
<td>5</td>
<td>NQF levels 1–3</td>
<td>Grades 10–12 (Mid- to Senior Secondary)</td>
</tr>
<tr>
<td>53–54</td>
<td>6</td>
<td>NQF levels 4–5</td>
<td>Grade 12+ (Higher Certificate) (Junior Tertiary)</td>
</tr>
<tr>
<td>55–57</td>
<td>6</td>
<td>NQF level 6</td>
<td>Diploma/Advanced Certificate (Tertiary Diploma)</td>
</tr>
<tr>
<td>58–62</td>
<td>7</td>
<td>NQF level 7</td>
<td>3-year Degree/Adv. Diploma (First Degree)</td>
</tr>
<tr>
<td>63–68</td>
<td>8</td>
<td>NQF level 8</td>
<td>Honours/4-year Degree/Postgraduate Diploma</td>
</tr>
<tr>
<td>69–80 (65+)</td>
<td>9</td>
<td>NQF level 9</td>
<td>Advanced Degree (Master’s Degree)</td>
</tr>
<tr>
<td>69–80 (65+)</td>
<td>9</td>
<td>NQF level 10</td>
<td>Advanced Degree (Doctoral Degree)</td>
</tr>
</tbody>
</table>

Results graph and scores of the LPCAT
The LPCAT pre- and post-test results are presented in graph form (see Figure 10.3). The estimated ability/performance levels after answering each question are plotted, and these levels, as well as the number of questions answered, can be seen in both the pre- and post-test plots. In the pre-test, between 8 and 12 questions are adaptively administered from an item bank of 63 questions, while in the post-test, between 10 and 18 questions are administered adaptively from a separate post-test item bank containing 125 questions. The performance level at the end of the pre-test is used as the entry level in the post-test, thereby further improving the accuracy of estimation in the post-test.

The following four scores are presented in the results graph in a T-test form:
- the pre-test score (performance level at the end of the pre-test);
- the post-test score (performance level at the end of the post-test);
- the difference score (numerical difference between pre- and post-tests);
- the composite score (a reasoned combination of the pre- and post-test scores).

Scores are also presented in stanine and percentile format, but these are less useful than the T-test scores. The latter are also used for the interpretation of the level of reasoning shown in the pre- and post-tests in relation to National Qualifications Framework (NQF) or academic levels (see Table 10.1).
Psychometric properties and fairness of the LPCAT

This section provides a summary of some empirical results on the psychometric properties of the LPCAT during its development and validation, as well as in the time since its release in 2000. Preference has been given to studies with larger sample sizes. Information on specific concerns referred to in the Employment Equity Act – that is, reliability, validity and fairness – is provided.

Reliability of the LPCAT

Reliability of CATs is not measured in the same way as that of standard static tests, since individuals completing the test can be given different items as well as different numbers of items to answer, although the scores obtained are on the same scale that measures the latent trait of the particular domain. McBride (1997) indicates that adaptive tests can achieve higher reliability compared with conventional tests in the upper and lower extremes of the ability scale, and at the same time reach a given level of precision, using substantially fewer items than standard tests. This is a result of the items being selected purposefully to match the estimated ability level of the respondent throughout the test. The IRT equivalent to test score reliability and standard error of measurement (SEM) of classical test theory is the test information function. This reflects the level of information available at a particular ability level, as a result of the number and quality of items available at that level in the item bank. The standard error is a function, which means that it is not a single measure over the entire ability range but is calculated at various ability levels, based on the amount of information at different ability levels (De Beer, 2000b).

LPCAT coefficient alpha reliability values range between 0.926 and 0.981 for subgroups based on gender, culture, language and level of education for the standardisation sample of 2 450 Grade 9 and Grade 11 learners, and are reported fully in the LPCAT Technical Manual (De Beer, 2000b). The detail of the test information function is also reported there.

Validity of the LPCAT

Determination of validity of a test generally entails ongoing gathering of information to evaluate the usefulness of test results for various groups in different contexts. It usually requires evidence of the relationships between performance on the test and other independently obtained scores which also reflect the behaviour of concern. Although DA was often criticised in the past for its lack of empirical psychometric evidence, this has changed in recent years (Caffrey, Fuchs & Fuchs, 2008).

The construct and predictive validity for the LPCAT are presented by reporting on results of samples at different educational levels, from low-literate adults to tertiary university levels. A short description of the sample groups is provided below.

i) Group 1: Low-literate adult group (Adult Basic Education and Training (ABET))

A group of low-literate adults was assessed for the purpose of career guidance after their positions were made redundant. The sample (N = 194) was mostly
male and all black. Together with the LPCAT, the Paper-and-Pencil-Games (PPG) (Claassen, 1996) was also administered; this test provides a verbal, nonverbal and total score. For the criterion measure, ABET numeracy and literacy results (Level 1 and Level 3) were used (De Beer, 2000b).

ii) Group 2: Senior primary (Grade 6 and Grade 7 levels)
The longitudinal predictive validity results for two separate groups were investigated (De Beer, 2010b). The first sample group (N = 72) was all female (Grade 6) with a mean age of 11.18 years. The second sample (N = 79) was all male and in Grade 7, with a mean age of 12.44 years. An English proficiency test was also administered (Chamberlain & Reinecke, 1992) to the male sample, while two subtests of the Differential Aptitude Test (DAT) (Claassen, Van Heerden, Vosloo & Wheeler, 2000) were administered to the female sample. For both groups an aggregate score for school academic results was used as the criterion measure (De Beer, 2010b).

iii) Group 3: Junior secondary (Grade 8 level)
A sample group (N = 151) of junior secondary learners with a mean age of 13.2 years was assessed with the LPCAT as well as with the General Scholastic Aptitude Test (adaptive version) (GSAT-CAT) (Van Tonder & Claassen, 1992). An English proficiency measure (Chamberlain & Reinecke, 1992), as well as a test of basic numerical literacy (Venter, 1997), was also administered. School academic results were used as the criterion (De Beer, 2000b).

iv) Group 4: Junior secondary (Grade 9 level)
A group of 253 learners at Grade 9 level was assessed as part of a vocational guidance project. Of this sample group, 96 (37.9 per cent) were male and 157 (62.1 per cent) were female. Three subtests of the DAT Form R (Claassen et al., 2000) were also administered (Verbal Reasoning, Comparisons and 2-dimensional Spatial Reasoning). Academic results in English, Mathematics and Life Orientation were used as criterion measures.

v) Group 5: Senior secondary (Grade 11 level)
A group of 174 learners at a Grade 11 level was assessed as part of a vocational guidance project. For this sample, 63 were male (36.2 per cent) and 111 were female (63.8 per cent). Three subtests of the DAT Form K (Coetzee & Vosloo, 2000) were also administered (Verbal Reasoning, Comparisons and 3-dimensional Spatial Reasoning). Academic results in English, Mathematics and Life Orientation were used as criterion measures.

vi) Group 6: Junior tertiary (Further Education and Training (FET) college first-year level)
A sample group of 75 students was assessed for career guidance purposes. The DAT Form R (Claassen et al., 2000) was also administered. Academic results were used as criterion measures (De Beer, 2008).
vii) Group 7: Tertiary (first-year diploma level)
A first-year sample of engineering and technology students (N = 223) with a mean age of 19.9 years was tested with the LPCAT, as well as with the GSAT-CAT (Van Tonder & Claassen, 1992). Subtests of the Senior Aptitude Test (SAT) were also administered (Owen & Taljaard, 1989). Grade 12 academic results and first-year academic results were obtained, to be used for comparative predictive validity analyses respectively (De Beer, 2000b). (See also Van der Merwe and De Beer (2006) and Van Eeden, De Beer and Coetzee (2001), for other results at this level.)

viii) Group 8: Tertiary (first-year degree level)
A group of applicants for engineering studies at university (N = 382) was tested for screening and selection purposes. Their mean academic results were used as criterion data (De Beer & Mphokane, 2010).

ix) Group 9: Mixed (group from industry)
A sample group from industry (N = 150) was assessed with both the LPCAT and the Raven’s Standard Progressive Matrices (Mann, 2007).

The mean LPCAT scores for the above groups are reported in Table 10.2.

### Table 10.2 Mean LPCAT scores for groups at different educational levels

<table>
<thead>
<tr>
<th>Group</th>
<th>Educational level</th>
<th>N</th>
<th>LPCAT pre-test</th>
<th>LPCAT post-test</th>
<th>LPCAT composite</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>Adult low-literate</td>
<td>194</td>
<td>36.19</td>
<td>37.76</td>
<td>–</td>
</tr>
<tr>
<td>Group 2</td>
<td>Grade 6*</td>
<td>72</td>
<td>50.01</td>
<td>50.87</td>
<td>50.13</td>
</tr>
<tr>
<td>Group 2</td>
<td>Grade 7*</td>
<td>79</td>
<td>54.52</td>
<td>56.10</td>
<td>54.78</td>
</tr>
<tr>
<td>Group 3</td>
<td>Grade 8</td>
<td>128</td>
<td>45.67</td>
<td>47.83</td>
<td>–</td>
</tr>
<tr>
<td>Group 4</td>
<td>Grade 9</td>
<td>233</td>
<td>51.09</td>
<td>52.37</td>
<td>51.30</td>
</tr>
<tr>
<td>Group 5</td>
<td>Grade 11</td>
<td>119</td>
<td>52.50</td>
<td>53.34</td>
<td>52.60</td>
</tr>
<tr>
<td>Group 6</td>
<td>FET first year</td>
<td>74</td>
<td>48.82</td>
<td>49.43</td>
<td>49.00</td>
</tr>
<tr>
<td>Group 7</td>
<td>Diploma first year</td>
<td>159</td>
<td>55.21</td>
<td>56.47</td>
<td>–</td>
</tr>
<tr>
<td>Group 8</td>
<td>Degree first year</td>
<td>382</td>
<td>–</td>
<td>63.96</td>
<td>62.54</td>
</tr>
<tr>
<td>Group 9</td>
<td>Mixed (industry)</td>
<td>150</td>
<td>57.75</td>
<td>58.80</td>
<td>–</td>
</tr>
</tbody>
</table>

Note: * Private school. Sample sizes differ due to missing data.

Data obtained from the above sample groups are reported for construct and predictive validity in the next two subsections.

**Construct validity of the LPCAT**
To determine the construct validity of the LPCAT, its correlations with a variety of other cognitive measures for groups at various educational levels are summarised in Table 10.3.
Table 10.3 Construct validity of the LPCAT

<table>
<thead>
<tr>
<th>Group</th>
<th>Educational level</th>
<th>Other measures</th>
<th>N</th>
<th>LPCAT post-test r</th>
<th>p</th>
<th>LPCAT composite r</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>Adult low-literate</td>
<td>PPG Verbal</td>
<td>110</td>
<td>.408**</td>
<td>.000</td>
<td>.411**</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PPG NV</td>
<td>110</td>
<td>.543**</td>
<td>.000</td>
<td>.565**</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PPG Total</td>
<td>110</td>
<td>.610**</td>
<td>.000</td>
<td>.552**</td>
<td>.000</td>
</tr>
<tr>
<td>Group 2</td>
<td>Grade 6</td>
<td>DAT English</td>
<td>72</td>
<td>.263*</td>
<td>.025</td>
<td>.213</td>
<td>.072</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DAT Calc</td>
<td>72</td>
<td>.278*</td>
<td>.018</td>
<td>.280*</td>
<td>.017</td>
</tr>
<tr>
<td></td>
<td></td>
<td>English 1st prof.</td>
<td>79</td>
<td>.405*</td>
<td>.018</td>
<td>.328*</td>
<td>.003</td>
</tr>
<tr>
<td>Group 3</td>
<td>Grade 8</td>
<td>GSAT-CAT VB</td>
<td>120</td>
<td>.613**</td>
<td>.000</td>
<td>.574**</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GSAT-CAT NV</td>
<td>120</td>
<td>.665**</td>
<td>.000</td>
<td>.653**</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GSAT-CAT Total</td>
<td>120</td>
<td>.691**</td>
<td>.000</td>
<td>.664**</td>
<td>.000</td>
</tr>
<tr>
<td>Group 4</td>
<td>Grade 9</td>
<td>DAT Verbal</td>
<td>228</td>
<td>.544**</td>
<td>.000</td>
<td>.500**</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DAT Comparisons</td>
<td>202</td>
<td>.335**</td>
<td>.000</td>
<td>.356**</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DAT 2D</td>
<td>227</td>
<td>.500**</td>
<td>.000</td>
<td>.435**</td>
<td>.000</td>
</tr>
<tr>
<td>Group 5</td>
<td>Grade 11</td>
<td>DAT Verbal</td>
<td>114</td>
<td>.209*</td>
<td>.025</td>
<td>.114</td>
<td>.126</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DAT Comparisons</td>
<td>88</td>
<td>.111</td>
<td>.307</td>
<td>.051</td>
<td>.637</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DAT 3D</td>
<td>108</td>
<td>.524**</td>
<td>.000</td>
<td>.452</td>
<td>.000</td>
</tr>
<tr>
<td>Group 6</td>
<td>FET 1st year</td>
<td>DAT Language</td>
<td>74</td>
<td>.200</td>
<td>.088</td>
<td>.296**</td>
<td>.010</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DAT NV</td>
<td>74</td>
<td>.403**</td>
<td>.000</td>
<td>.363**</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DAT Verbal</td>
<td>74</td>
<td>.389**</td>
<td>.001</td>
<td>.386**</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DAT Calculations</td>
<td>74</td>
<td>.274*</td>
<td>.018</td>
<td>.298*</td>
<td>.010</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DAT Comparisons</td>
<td>74</td>
<td>.193</td>
<td>.100</td>
<td>.112</td>
<td>.334</td>
</tr>
<tr>
<td>Group 7</td>
<td>Diploma 1st year</td>
<td>GSAT-CAT VB</td>
<td>158</td>
<td>.571**</td>
<td>.000</td>
<td>.555**</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GSAT-CAT NV</td>
<td>158</td>
<td>.645**</td>
<td>.000</td>
<td>.626**</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GSAT-CAT Total</td>
<td>158</td>
<td>.668**</td>
<td>.000</td>
<td>.648**</td>
<td>.000</td>
</tr>
<tr>
<td>Group 8</td>
<td>Degree 1st year</td>
<td>ELSA Literacy</td>
<td>309</td>
<td>.481**</td>
<td>.000</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ELSA Numeracy</td>
<td>309</td>
<td>.418**</td>
<td>.000</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maths test</td>
<td>309</td>
<td>.527**</td>
<td>.000</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Group 9</td>
<td>Mixed (industry)</td>
<td>Raven’s SPM (RS)</td>
<td>150</td>
<td>.585**</td>
<td>.000</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Raven’s SPM (TS)</td>
<td>150</td>
<td>.618**</td>
<td>.000</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Notes: * p < .05 ** p < .01. Sample sizes differ due to missing data.

The results indicate that the LPCAT, with its general measurement of ‘g’ fluid ability performance and potential, overlaps with abilities and domains measured by other (cognitive) tests.

Predictive validity of the LPCAT
The groups and different measures obtained are summarised in Table 10.4. Academic performance is generally used as it is easier to generalise from these measures.
### Table 10.4 Predictive validity results for the LPCAT at different educational levels

<table>
<thead>
<tr>
<th>Group</th>
<th>Educational level</th>
<th>Criterion measures</th>
<th>LPCAT score</th>
<th>N</th>
<th>r</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Highest corr.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 1</td>
<td>Adult low-literate</td>
<td>ABET Literacy L1</td>
<td>Post-test</td>
<td>110</td>
<td>.437**</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ABET Literacy L3</td>
<td>Composite</td>
<td>111</td>
<td>.461**</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ABET Numeracy L1</td>
<td>Post-test</td>
<td>182</td>
<td>.491**</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ABET Numeracy L3</td>
<td>Composite</td>
<td>26</td>
<td>.610**</td>
<td>.000</td>
</tr>
<tr>
<td>Group 2</td>
<td>Grade 6</td>
<td>Aggregate Academic</td>
<td>Composite</td>
<td>72</td>
<td>.499**</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Grade 7</td>
<td>Aggregate Academic</td>
<td>Post-test</td>
<td>79</td>
<td>.482**</td>
<td>.000</td>
</tr>
<tr>
<td>Group 3</td>
<td>Grade 8</td>
<td>Academic (Sem. 2)</td>
<td>Post-test</td>
<td>118</td>
<td>.524**</td>
<td>.000</td>
</tr>
<tr>
<td>Group 4</td>
<td>Grade 9</td>
<td>English (Academic)</td>
<td>Post-test</td>
<td>233</td>
<td>.340**</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maths (Academic)</td>
<td>Post-test</td>
<td>233</td>
<td>.434**</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Life Orient. (Academic)</td>
<td>Composite</td>
<td>233</td>
<td>.215**</td>
<td>.001</td>
</tr>
<tr>
<td>Group 5</td>
<td>Grade 11</td>
<td>English (Academic)</td>
<td>Post-test</td>
<td>119</td>
<td>.025</td>
<td>.789#</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maths (Academic)</td>
<td>Post-test</td>
<td>119</td>
<td>.005</td>
<td>.957#</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Life Orient. (Academic)</td>
<td>Post-test</td>
<td>119</td>
<td>.063</td>
<td>.494#</td>
</tr>
<tr>
<td>Group 6</td>
<td>FET 1st year</td>
<td>Academic average</td>
<td>Composite</td>
<td>69</td>
<td>.350**</td>
<td>.004</td>
</tr>
<tr>
<td>Group 7</td>
<td>Diploma 1st year</td>
<td>Academic average</td>
<td>Composite</td>
<td>120</td>
<td>.218*</td>
<td>.017</td>
</tr>
<tr>
<td>Group 8</td>
<td>Degree 1st year</td>
<td>Academic average</td>
<td>Post-test</td>
<td>125</td>
<td>.333**</td>
<td>.000</td>
</tr>
<tr>
<td>Group 9</td>
<td>Mixed (industry)</td>
<td>None available</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** For most of the above results, more detailed information and full results can be found in the sources referred to in the sample descriptions above.

* p < .05 ** p < .01. Sample sizes differ due to missing data.

# Although the predictive validity results for the LPCAT have generally shown positive correlations of moderate to large practical effect sizes, the predictive validity correlation results for Group 5 show non-significant correlations. The Verbal Reasoning and Comparisons subtests of the DAT showed similar non-significant correlations with academic performance for this group, and only 3-dimensional Spatial Reasoning results showed statistically significant correlations with the academic results.

With the exception of one group (Group 5), the results show acceptable levels of predictive validity for academic results over a wide spectrum of academic levels.

**Fairness of the LPCAT**

The LPCAT is registered with the Health Professions Council of South Africa as a culture-fair test. The following features of the LPCAT can be deemed to contribute to its fairness in the multicultural and multilingual South African context:

- It focuses on the measurement of learning potential, addressing not only current level of performance but also the projected or potential future level that can be achieved if relevant learning opportunities are provided.
• The content of the test questions contains only nonverbal figural patterns, thereby not requiring language proficiency and not relying on mastery of scholastic content to measure reasoning ability.
• The test instructions have been translated into all official South African languages for the version in which no text appears on the screen and the instructions are read to the respondents, thereby not requiring them to read anything themselves (De Beer, 2000a; 2005). The text-on-screen version is available in English and Afrikaans.
• Computer literacy is not a requirement, since the easy use of only the space bar and the enter key to answer the multiple-choice questions presented simplifies the answering procedure. It allows for the measurement of learning potential of illiterate adults to ensure that appropriate learning and development opportunities are provided.
• During test development, a large and representative sample (N = 2 450) was used for the item analysis and standardisation. This sample was used to perform IRT-based DIF analysis on all new questions compiled with regard to subgroups based on level of education, language, culture and gender. Items not complying with the cut-off in terms of DIF for any one or more of the subgroups were discarded, and not used in the final test.
• CAT allows for items of suitable difficulty level in comparison with the performance (estimated ability) level of the respondent throughout the pre- and post-tests.
• The LPCAT is a power as opposed to a timed test, allowing sufficient time for each question that is presented to be answered, and with no overall set test time.

Practical use of the LPCAT
The LPCAT can be used in contexts in which decision-making involves obtaining information related to required future performance or development and training levels, in terms of the NQF level framework (see Table 10.1). It has shown statistically and practically significant predictive validity for academic results at different levels (basic education, primary, secondary and tertiary level academic results – see Table 10.4).

Practically, the process starts from the end, in the sense that the reason for assessment should be carefully considered first to determine what the required level of performance or the level of training to be completed is. Once this level has been identified (in terms of relevant NQF or educational levels), the LPCAT pre- and post-test results can be compared to this level to determine whether the individual currently (as reflected in the pre-test results) performs close to or at the required level or, if not, whether the individual shows that after a learning opportunity has been presented, he or she is able to function close to or at the required level (as reflected in the post-test results). Smaller improvement scores are an indication that the individual is, in future, likely to perform at similar levels to those currently shown. On the other hand, larger improvement
scores indicate that the individual can be expected to perform at higher levels in the future than those currently shown, provided that relevant learning and development opportunities are provided. Table 10.1 indicates the interpretation of the LPCAT performance levels shown in terms of NQF and/or academic levels.

Other measures can be used to identify specific aptitudes, proficiencies or interests, but learning potential assessment can identify the appropriate level at which training and development should be currently targeted and aimed for over time. Due consideration should be given to actual academic attainment, to ensure that appropriate building blocks are put in place over time to assist with optimal development of the individual potential shown.

Measurement of learning potential on the LPCAT is not restricted to individuals of low formal education; it can also be administered to, and its results advantageously used for, individuals up to a postgraduate tertiary educational level. An important prerequisite is to ascertain whether the individual has obtained the specific formal level of academic qualification or training that is required to commence with the level of training offered. Once this has been verified, the difference between the LPCAT levels of performance and the level required for the training offered can be interpreted as the extent of effort that the individual will need to exert in order to achieve success at the required level. If the LPCAT level of performance is lower than the required level, it is interpreted as an indication that more effort will be needed from the individual to achieve success at the required level. The larger the difference, the greater the effort that will be required, or the longer the time it could take to achieve success. In such cases, it is recommended that a step-by-step approach be taken, with training broken down into smaller parts to allow for optimal levels of performance by not overloading the individual in terms of the magnitude (number of subjects) as well as the level of the challenge. If the individual’s test performance is at a higher level than the required level, indications are that he or she should be able to achieve success with moderate effort and within the prescribed number of hours of study indicated.

Features and challenges of the LPCAT

Certain advantages and disadvantages are associated with psychological assessment in general. Some of the advantages include the information that it provides to promote better decision-making, the objective sample of behaviour it represents, and enhancement of a scientific approach to decision-making. The disadvantages include measurement error, the possible effect of poor testing conditions, and the fact that individuals may either try to manipulate results or not be appropriately motivated to ensure optimal results. When using any psychological assessment instrument, it is imperative to be aware of its particular features which may result in specific advantages and disadvantages being present.

Positive features of the LPCAT

As the preceding discussion of the LPCAT has indicated, it has a number of positive features:
• It is considered a culture-fair test. It was developed in South Africa, and has been used internationally, in Africa (South Africa, Mozambique, Namibia, Botswana, Zambia, Ethiopia, Uganda and Gambia), in a number of countries in the East (Sri-Lanka, Cambodia, Nepal and Vietnam) and in Europe (Finland and The Netherlands).

• It has shown satisfactory reliability, and its predictive validity results have compared favourably with those of standard tests.

• Standard training is provided to focus attention of the respondents on the relevant aspects of the task, so that intra- and inter-individual comparisons can be made.

• It allows for equal assessment opportunities, irrespective of the current level of formal qualification (from illiterate to postgraduate levels, with the test adapting to the performance level shown by the individual respondent).

• It can be administered individually or in groups. Instructions are available in all 11 official languages of South Africa (and in French).

• It is quick and easy to administer, and the results are immediately available on completion of the test. Results are presented graphically, which also allows for some qualitative analysis of performance during the pre- and post-tests (see Figure 10.4).

• It is in line with the Employment Equity Act, affording opportunities to those considered disadvantaged educationally and socio-economically and who may therefore not have had opportunities to reach their optimal level of development/qualification.

**Figure 10.4 Example of LPCAT results**
• It adds information that would not be available from static tests. It can therefore assist with the identification of individuals who may otherwise be overlooked (owing to current low levels of attained qualifications), but who have high potential for further development.

**Challenges and problematic issues relating to the LPCAT**

Notwithstanding all its positive features, this form of assessment does pose some challenges and raises problems that need to be addressed:

• The LPCAT could have limited face validity for individuals at higher educational levels, since its content is not related to job performance or training at higher levels. For such groups, it is therefore important to explain how the test works and to justify its inclusion in a particular assessment battery, to ensure that respondents remain motivated and perform to the best of their ability throughout.

• Reliance on computers, and thus on electricity, for administration and to ensure that results are saved could be a problem if power failures disrupt assessment.

• Information provided in the results is only linked to the current and projected levels of fluid reasoning ability shown, and does not provide a direct link to a particular career or job level. Other assessment information would be needed to provide information of the latter kind for career-related guidance and decisions.

• Newer test operating systems pose challenges for the current version of the test administration program. Ongoing updates of software required to maintain compatibility with new operating systems are required. This is discussed below.

**Future developments**

Demands for software programs (and computer-based tests and testing systems) to maintain compatibility with new technology and updated operating systems are ongoing. Interim challenges are addressed by means of bridging or patch programs, but major revisions and updates are also required from time to time. The current test administration program of the LPCAT will be updated in the near future to ensure improved compatibility with new operating systems.

An internet-based Results Analysis Program for processing LPCAT results has been developed. This allows different users from the same organisation to access the same database or specific subfolders of results remotely via the internet. The sharing of runs within a particular user group is also much easier with this program.

Internet-based adaptive test administration is the next development target; this will allow for the use of more automated processes. Other developments in the planning stages are the expansion of the item bank and recalculation of item parameters to ensure current relevance in terms of the interpretation of levels of performance.
Conclusion

As this chapter has shown, the LPCAT can provide a different kind of information than that obtained from standard cognitive tests. However, its results cannot answer all questions relating to the cognitive domain, as it only indicates current and potential levels of cognitive performance in the nonverbal figural domain. Nonetheless, it adds information that enriches the interpretation of individual results when used in an integrated manner with other tests and measures.

References


