AN INVESTIGATION INTO THE INTERNAL STRUCTURE OF THE LEARNING POTENTIAL CONSTRUCT AS MEASURED BY THE APIL TEST BATTERY
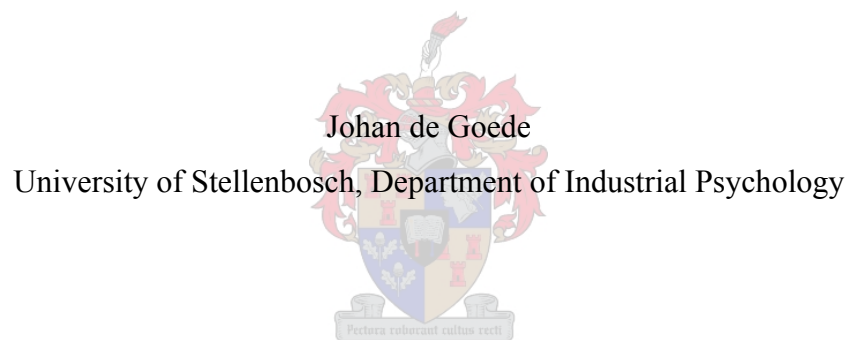
Johan de Goede

University of Stellenbosch, Department of Industrial Psychology

Thesis presented in partial fulfilment of the requirements for the degree of Master of Commerce at the University of Stellenbosch

Supervisor

Prof CC Theron

December 2007

Declaration

I, the undersigned, hereby declare that this thesis is my own original work and that all sources have been accurately reported and acknowledged, and that this document has not previously, in its entirety nor in part, been submitted at any university in order to obtain an academic qualification.

Signature:

Date: 31/07/2007

ABSTRACT

This thesis presents an investigation into the internal structure of the learning potential construct as measured by the APIL Test Battery developed by Taylor (1989, 1992, 1994, 1997). The measurement of learning potential, a core or fundamental ability, as opposed to abilities heavily influenced by exposure to previous opportunities is important in the South African environment. The importance of the assessment of learning potential can be explained partly in terms of the necessity of equalling the proverbial 'playing field' and ensuring that previously disadvantaged individuals are not becoming more disadvantaged by further being denied development opportunities and partly in terms of attempts to compensate and correct for a system that clearly oppressed the development of important job related skills, knowledge and abilities in certain groups. Such attempts at accelerated affirmative development will, however, only be effective to the extent to which there exists a comprehensive understanding of the factors underlying training performance success and the manner in which they combine to determine learning performance in addition to clarity on the fundamental nature of the key performance areas comprising the learning task. In this study the internal structure of the learning potential construct as measured by the APIL Test Battery was investigated through structural equation modelling and regression analysis. Overall, it was found that both the measurement and the structural model fitted the data reasonably well. The study, however, was unable to corroborate a number of the central hypotheses in Taylor's (1989, 1992, 1994, 1997) stance on learning potential. Moreover, the analysis of the standardised residuals for the structural model, suggested that the addition of one or more paths to the existing structural model would probably improve the fit of the model. Modification indices calculated as part of the structural equation modelling could, however, not point out any specific additions to the existing model. Regression analysis resulted in the conclusion that the inclusion of the two learning competency potential measures together with the two learning competencies measures in a learning potential selection battery is not really warranted. The use of information processing capacity as a predictor on its own seems to be indicated by the results of this study. Recommendations for future research are made.

OPSOMMING

Die hoofdoel van hierdie studie was om die interne struktuur van die leerpotensiaalkonstruk soos gemeet met die APIL Toets Battery, ontwikkel deur Taylor (1989, 1992, 1994, 1997), te ondersoek. Die meting van leerpotensiaal, 'n inherente/fundamentele vermoë, eerder as 'n fokus op die meting van vermoëns afhanklik van blootstelling aan vorige geleenthede, is uiters belangrik, in Suid Afrika. Die belang van die meting van leerpotensiaal kan ten eerste verklaar word in terme van die noodsaaklikheid om die spreekwoordelike speelveld gelyk te maak en om te verseker dat voorheen benadeelde individue nie verder benadeel word omdat hulle steeds ontwikkelingsgeleenthede geweier word nie en tweedens, in terme van pogings om die effek van 'n sisteem wat die ontwikkeling van belangrike vaardighede, kennis en vermoëns in sekere groepe in Suid Afrika onderdruk het, teen te werk en te korrigeer. Sodanige pogings tot versnelde regstellende ontwikkeling sal egter slegs slaag in die mate waartoe daar 'n omvattende begrip bestaan van die onderliggende redes vir sukses in opleiding en die wyse waarop die onderliggende redes vir sukses kombineer om leerprestasie te bepaal, asook die sleutelprestasieareas wat die leertaak uitmaak. In hierdie studie is die interne struktuur van die leerpotensiaalkonstruk, soos gemeet deur die APIL Toets Battery, deur middel van strukturele vergelykingsmodellering (structural equation modelling) en 'n regressieontleding ondersoek. Oorkoepelend is daar gevind dat beide die metings- en strukturele model die data relatief goed pas. Die studie kon egter nie 'n aantal van die sentrale hipoteses in Taylor (1989, 1992, 1994, 1997) se standpunt oor leerpotensiaal bevestig nie. Daarbenewens het 'n ondersoek van die gestandaardiseerde residue aangetoon dat die toevoeg van een of meer addisionele bane tot die bestaande strukturele model waarskynlik die passing van die model sal verbeter. Modifikasie-indekse bereken as deel van die strukturele vergelykingsmodellering kon egter geen spesifieke toevoegings tot die bestaande model uitwys nie. Regressieontleding het tot die slotsom gelei dat die insluiting van die twee leerbevoegdheidspotensiaalmetings saam met die twee leerbevoegdheidsmetings nie werklik geregverdig is nie. Die resultate van hierdie studie skyn daarop te dui dat informasieverwerkingskapasiteit op sy eie as voorspeller gebruik behoort te word. Aanbevelings vir verdere navorsing word gemaak.

ACKNOWLEDGEMENTS

It is both ignorant and unrealistic to expect that a study of this nature can be completed without the unwavering support and encouragement of others. The overwhelming magnitude of completing this study that often stared me in the face could have easily derailed my efforts, was it not for my incredible support network.

Firstly, I have to acknowledge the effort and humbly thank Prof. Callie Theron from the University of Stellenbosch. Prof. Theron calmly and patiently guided me throughout the process, continuously encouraged me and gave me the answers when I had none. He is a great man, a true academic, an exceptional mentor and a phenomenally insightful human being. Without Prof. Theron's intricate understanding of the subject matter and his subtle and supportive manner, I am highly doubtful of whether I would ever have gotten to the point of writing this acknowledgement.

Secondly, I have to thank my close family; my wife Robyn, for her exceptional ability to support, drive and motivate me towards my goals and dreams and never doubting my ability, my father Steph, and especially my mother Marie. Not only did they provide me with the encouragement to push through to the end, but they also spent countless hours proof reading, listening to me expressing my frustrations and verbalising my discontent.

Thirdly, I would like to thank the South African Police Service Training College in Philippi. Without their willingness to open their doors and allow me to gather the data that serves as the foundation of this study, it would almost certainly have been impossible to complete the research in its current form. I would also like to express my gratitude to Deon Meiring, from the South African Police Service, for negotiating with the South African Police Service decision-makers to consent to me using their new recruits to gather my data.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION AND OBJECTIVES OF THE STUDY

1.1   INTRODUCTION

To succeed in the global environment it is required of countries to show consistently high economic growth. By maintaining such growth a country gains a competitive advantage and prevents economic stagnation and poverty. Consistently high economic growth can only be achieved if the production of goods and delivery of services in and by a country is done effectively, efficiently and productively.

Productivity can best be achieved if people and other resources are grouped together in organisations. Organisations are formed so that society may accomplish goals, which would be impossible, if everyone acted individually (Gibson, Ivancevich & Donnelly, 1997; Jones, 2001). Organisations are entities that allow people to co-ordinate their actions in order to accomplish specific goals through the identification and realisation of opportunities to satisfy needs (Gibson *et al*., 2001). Thus, the main reason why organisations exist is to produce goods and deliver services in a productive manner, so that real economic value is added to the benefit of shareholders, the government and the broader community. Ultimately, organisations have to accept co-responsibility for a country's economic situation and contribute to a country's global competitiveness.

In relation to our global counterparts, South Africa does not compare well in terms of competitiveness and is currently in the 46th position on the international competitiveness list (Sidirpoulus, Jeffrey, Mackay, Forgey, Chipps & Corrifan, in Cross, Marais, Steel & Theron, 2002). It seems, especially, the ineffective and inefficient production of goods and delivery of services that impact negatively on both the country's economic growth and the country's global competitiveness.

The way in which organisations create real economic value is through a three-cycle input-, conversion- and output process (Jones, 2001). The input obtained and used by organisations include, amongst others, human resources, information and knowledge, raw materials, and capital. The value that an organisation creates at the input stage is largely dependent on the way in which the organisation chooses and acquires its inputs. At the conversion stage, the extent of value created is largely dependent on the way in which the organisation uses, applies and manages its obtained human resources and technology. The created value at this stage consists of the quality of skills within the organisation, including the ability of the organisation to learn from and respond to the environment in which it functions. Finally, at the output stage, an organisation, depending on the effectiveness and efficiency of the prior two stages, delivers a product or service, which is sold at a profit. Profit in turn is distributed to the stakeholders, to the government through taxes, to the community through social corporate investment and re-invested back into the organisation to ensure future profits (Jones, 2001).

The extent of success with which an organisation creates value through the three-cycle value creation process is largely dependent on humans who are the carriers of the production factor labour. It is human actions that are grouped together and co-ordinated to form an organisation. Combining other production factors on their own, without human effort would not constitute an organisation. Organisations striving towards consistently high economic growth have to realise that it is people, in the final analysis that makes the competitive difference. For this reason successful organisations are desperately seeking the best employees and investing in the training and development of its people. From the above argument, it is clear that the quality of the South African workforce will, to a large extent, determine our country's future economic growth and global competitiveness.

South African organisations need to realise that only if the people with the appropriate knowledge, skills, abilities and attitudes are in the right place at the right time, thus adding maximum value to the three cycle process, will the country be able to compete globally. The question, however, arises as to how South African organisations can ensure this? The

answer to this question can be found in effective and efficient human resource management (Carrell, Elbert, Hatfield, Grobler, Marx & van der Schyf, 1998).

Effective and efficient human resource management consists of policies, practices, and systems that influence employees' behaviour, attitudes and performance in such a manner that they are aligned with and support the business goals and objectives (Noe, Hollenbeck, Gerhart & Wright, 2000).

The foundation of effective and efficient human resource management consists of the following two main categories of human resource interventions as identified by Milkovich and Boudreau (1994). The first category refers to the regulation of the flow of workers into, through and out of the organisation. This category includes interventions such as recruitment and selection, placement, promotion and the downsizing of the organisation's human resources. The second category refers to the maintenance and development of the current human resource supply. This category would include interventions such as training, motivation, compensation and labour relations (Cross *et al.*, 2002). If both these human resource categories are managed in an effective and efficient manner then they will contribute to an improvement in productivity and in gaining a competitive advantage.
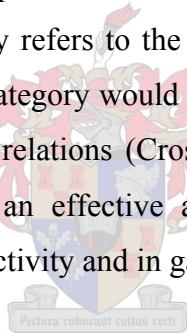
One important human resource intervention, relating to the flow of workers, is personnel selection (Cross *et al.*, 2002). Selection normally implies a situation where there are more applicants for openings than there are positions available for jobs or even training and developmental opportunities. Hence, the primary objective of selection is to fill the available number of vacancies with those applicants who will be most successful in the job or training intervention and, therefore, the subgroup of applicants has to be chosen in a manner that ensures that the average performance on the ultimate or final criterion is maximised (Austin & Villanova, 1992). The ultimate criterion is the criterion construct or latent variable (η) which personnel selection aspires to affect, i.e. job- or training performance. The ultimate criterion should, thus, always be the focus of interest in selection decisions (Ghiselli, Campbell & Zedeck, 1981). This seemingly innocent and too

often forgotten fact, moreover, has significant implications for the interpretation and evaluation of information entering the selection decision.

The fact that interest in selection centres on the criterion, creates somewhat of a dilemma for human resource managers or others who are responsible for selection decision-making. The dilemma is that measurements Y of the multidimensional final criterion (η) are not readily available at the time when the selection decision needs to be made. The only viable solution to the above dilemma would be to obtain a substitute for the criterion (Ghiselli, Campbell & Zedeck, 1981). In other words, selection decisions have to be based on substitute information X, which is available at the time the selection decision needs to be made and which is also relevant to the decision being taken. The relevance of such substitute information is determined by the extent to which an accurate estimate measure of the multidimensional final criterion is achieved.

The only information available at the time when the selection decision is being made, that could serve as such a substitute, would be psychological, physical, demographic or behavioural information on the applicants. Formally X, and therefore by implication $E[Y|X]$, could be considered a substitute for Y if and to the extent that $|\rho[X,Y]| > 0$ [p < 0,05] and if measures of X can be obtained at the time of or prior to the selection decision. The existence of a relationship, preferably one that could be articulated in statistical terms, between the criterion considered relevant by the decision maker and the information actually used by the decision maker as a substitute for the criterion, constitutes a fundamental and necessary, but not sufficient, prerequisite for effective and equitable selection decisions (Guion, 1991; Theron, 2001, 2002). An accurate understanding of this predictor-criterion relationship enables the selection decision-maker to predict expected criterion performance actuarially (or clinically) from relevant, though limited, information available at the time of the selection decision.

There exist only two options or approaches to obtain such relevant substitute information. The more traditional construct-orientated approach consists of the following elements. The first element relates to the setting of organisational goals, under which the organisation's

general hiring policy falls. The second element, job design, involves the breaking up of the job into its different tasks, duties and responsibilities that constitute successful job performance. The third element would be the identification and operationalisation of the person-centred constructs (ξ), i.e. knowledge and abilities that determine successful job performance. The necessary knowledge and abilities can be inferred from a job description compiled through job analysis. The presumed interrelationship between these hypothesised determinants and the way they collectively combine in the criterion is postulated in a nomological network or latent structure (Campbell, 1991; Kerlinger, 1986), as a complex hypothesis that explains performance on the job in question (the criterion). The predictive hypothesis should always be justifiable through clear, valid arguments that ξ_i is indeed relevant to η (Arvey & Faley, 1988; Gatewood & Feild, 1990; Guion, 1991; Society for Industrial Psychology, 1998). In its operational form the predictive hypothesis expresses the criterion variable to be predicted, i.e. job success, as a function of the nomological network of variables that serve as substitute predictors at the time of decision-making in the basic form $Y=f(X_i)$. The final element of the more traditional construct-orientated approach entails the application of selection devices or measuring instruments to measure whether a job candidate does indeed possess the required knowledge and abilities. Here it is the presence of the person-centred constructs (or lack thereof) that explains why one person performs better in a specific job than another (Carrell *et al*., 1998, Theron, 2002). The way these hypothesised determinants of performance should be combined is suggested by the way these determinants are linked in the postulated nomological network.

Regarding the second, content orientated approach, the job in question would again be systematically analysed via one or more of the available job analysis techniques (Gatewood & Feild, 1994). This is done to identify and define the behaviours that collectively denote job success if exhibited on the job. Substitute information would then be obtained through low or high fidelity simulations of the job content. These simulations in a selection context necessarily occur off the job and prior to the selection decision. Such simulations would elicit behaviour that, if it would in future be exhibited on the job, would denote a specific level of job performance. Here, it is the ability of the person to cope with the demands of the job (as simulated), that gives an indication of future job performance. If the person is

not able to cope with the simulated demands of the job, then it is more than likely that he or she will also not be able to perform successfully in the job.

Clearly, substantial differences exist between the logic underlying the two approaches in terms of which substitute criterion measures are generated. Although both options obtain substitute criterion information through observable behaviour elicited by a stimulus set (Theron, 2001; Thorndike, 1982), the stimuli in the construct-orientated approach are designed in such a way that a person's response to them is mainly a function of the specific, defined and originally hypothesised construct (ξ) being measured. In the content-orientated approach the stimuli are designed in such a way to elicit the same responses as would have been displayed in the real work situation. However, unlike the construct-orientated approach, the nature of the set of constructs shaping the responses are not known or specified in the content-orientated approach (Theron, 2002).

Despite these differences, both arguments, however, maintain that effective, though not necessarily efficient, selection is contingent on the identification of a substitute (in the form of a differentially weighted combination of measures of the person characteristics that drive job or training success, or alternatively the behavioural competencies which would constitute job or training success) for the ultimate criterion which shows a statistically describable relationship with an operational measure of the ultimate criterion. Both arguments, furthermore, contend that the same condition constitutes a necessary, though not sufficient, condition to achieve fair or equitable employee selection.

Irrespective of the approach used to obtain substitute measures for the final criterion, the following objectives should ideally be satisfied simultaneously by a criterion referenced personnel selection procedure (Guion, 1991; Theron, 2001):

> ➤ The inferences made from predictor scores should be permissible (i.e., the inferences should be valid);
> ➤ The inferences on which selection decisions are based should be fair;
> ➤ The selection procedure should add maximum value (i.e., the selection procedure should optimize utility); and

➢ The selection procedure should minimise adverse impact.

1.2   VALIDITY

The permissibleness of criterion related inferences made from predictor measures in personnel selection are evaluated through empirical validation studies. The main purpose of empirical validation investigations is to determine the extent to which relevant substitute criterion measures are obtained through the use of one of the two selection approaches. Validation is defined as follows in the *Standards for Educational and Psychological Testing*, as published by the *American Educational Research Association, American Psychological Association* and the *National Council on Measurement in Education* (Society for Industrial Psychology, 1998).

> Validity is the most important consideration in test evaluation. The concept refers to the appropriateness, meaningfulness, and usefulness of the specific inferences made from test scores. Test validation is the process of accumulating evidence to support such inferences. A variety of inferences may be made from scores produced by a given test, and there are many ways of accumulating evidence to support any particular inference. Validity, however, is a unitary concept. Although evidence may be accumulated in many ways, validity always refers to the degree to which that evidence supports the inferences that are made from scores. The inferences regarding specific uses of a test are validated, not the test itself. (p. 6)

In its simplest form, validity refers to the extent to which a test or measuring instrument measures that which it intends to measure. In the case of criterion referenced personnel selection the intention is to measure either person centred constructs, which determine job- or training success, or to measure behavioural-/performance constructs which constitute job- or training success. Only if selection instruments succeed in this intention do they supply information relevant to the selection decision. It is crucial that selection instruments do supply information that is relevant to the decision being made in the aforementioned sense since that would determine the extent to which such predictor measures correlate with the criterion and, thus, would determine the extent to which accurate criterion estimates can be derived from them. Validity, thus, refers to the extent to which the available proof supports the performance inferences made from the information obtained from the selection

measures (Anastasi & Urbina, 1997; Arvey & Faley, 1988; Gatewood and Feild, 1990; Theron, 2002).

The nature of the evidence required in justifying the use of the substitute X differs across the construct-orientated approach and the content-orientated approach. With the construct-orientated approach, proof that X provides a construct valid measure of ξ and that Y provides a construct valid measure of η is required. Further, proof is needed that X explains significant variance in Y, thus by implication in η. With the content-orientated approach, proof that X represents a representative sample of the demands that collectively constitute the job and that Y provides a construct valid measure of η, is required. Proof that X significantly explains variance in Y and by implication η is also needed (Theron, 2001).

1.3   FAIRNESS

Fairness is a topic that has been widely debated, discussed and written about in the field of Human Resource Management and Industrial Psychology. The issue of fairness is a complex one. Firstly, because it is difficult to define fairness in psychometric terms, and secondly, because there is a number of fairness models that interpret the issue of fairness differently (SIP, 1998; Theron, 2002).
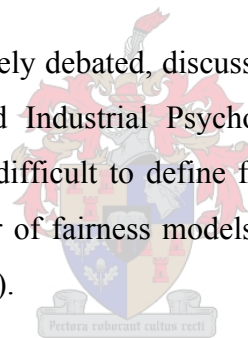
The objective of personnel selection is to add value to the organisation by improving the job performance of employees by regulating the flow of employees in, through and out of the organisation. In other words, to get the highest performing people on the job irrespective of gender, race or culture. In order to achieve this, predictive validity is a prerequisite. It is, however, not a given that when a selection procedure is proven valid that it will also be fair. A valid procedure can still lead to a different interpretation of the probability of job success between different subgroups with equal probability of success or vice versa. (Cross *et al*., 2002). This is because selection decisions are ultimately based on the criterion references' interpretation of predictors (i.e., $E(Y|X_i)$) and not the predictor information *per se*.

Arvey and Faley (1988) define unfair discrimination or predictive bias as follows:

Unfair discrimination or bias is said to exist when members of a minority group (or previously disadvantaged individuals) have lower probabilities of being selected for a job when, in fact, if they had been selected, their probabilities of performing successfully in the job would have been equal to those of non-minority group members. (p. 7)

The regression model of test bias developed by Cleary (1968), has become the standard model of selection decision fairness and it is often recommended that fairness models based on the regression model should be used in studies investigating the fairness of assessment procedures (SIP, 1998). Cleary (1968) defines predictive test bias as follows:

A test is biased for members of a subgroup of the population if, in the prediction of a criterion for which the test was designed, consistent non-zero errors of prediction are made for members of the subgroup. In other words, the test is biased if the criterion score predicted from the common regression line is consistently too high or too low for members of the subgroup. With this definition of bias, there may be a connotation of "unfair", particularly if the use of the test produces a prediction that is too low. (p. 115)

It follows from the regression based interpretation of selection fairness that although it is not a given that when a selection procedure is proven valid that it will also be fair, validity will deteriorate to the extent to which predictive bias exists as defined by Cleary (1968).

1.4   UTILITY

The objective of personnel selection is to add value to the organisation by improving the job performance of employees by regulating the flow of employees in, through and out of the organisation. This implies that the selection procedure must show high utility. According to Dunnette (1966) utility refers to the overall usefulness of a selection procedure and, therefore, contains both the accuracy and importance of decisions about employees. Dunnette (1966) explains further:

Utility does not imply the necessity or even the desirability of reducing all outcomes to a monetary scale, or even necessarily to a common scale, but does imply a careful identification and listing of all possible outcomes- accompanied by a judgmental weighing of the values (both money and human) associated with each. (p. 182)

In addition Boudreau (1991) defines utility analysis as follows:

> Utility analysis refers to the process that describes, predicts, and/or explains what
> determines the usefulness or desirability of decision options and examines how that
> information affects decisions. (p. 622)

Utility analysis is important, because it provides evidence to stakeholders about the effectiveness of the selection procedure (Hough, 2000). The reason for determining selection utility is to show the degree to which the use of a selection procedure improves the quality of individuals selected *vis-à-vis* if the procedure where not used (Gatewood & Feild, 1990). Determining the utility of a selection procedure provides a practical approach where the gain in performance obtained through the use of a specific selection procedure can be expressed in monetary terms (Gatewood & Feild, 1994; Theron, 2002). Most importantly, the purpose of a utility analysis is to provide substantive evidence that the initial investment in a selection procedure does yield substantive returns, significantly larger than the initial investment.

1.5   ADVERSE IMPACT

Adverse impact occurs when members of a group have a reduced likelihood to be selected for a job. Adverse impact is therefore present when there is a substantial difference in the rate of selection between groups that work to the disadvantage of members belonging to a certain group (Guion, 1991).

Normally a selection rate for any group, which is less than four-fifths (4/5) or 80 percent of the rate for the group with the highest selection rate is regarded as evidence of adverse impact (United States. Equal Employment Opportunity Commission, Civil Service Commission, Department of Labor & Department of Justice, 1978). It is important to understand that the comparison group here is the group with the highest proportion of applicants being selected, not the numerically larger group (Guion, 1991). It is moreover important to understand that the selection rates for the various groups are ultimately determined by their expected criterion performance conditional on their test performance

(derived fairly, without systematic prediction bias) and not the selection rates that would have resulted if selection would have occurred top-down on the predictor.

Adverse impact on its own does not constitute discrimination. In employment litigation adverse impact is used to make a *prima facie* case for discrimination, which then transfers the burden of proof to the defendant (Arvey & Faley, 1988; Guion, 1991). If adverse impact is found, the burden of proof is on the employer to demonstrate the job-relatedness of the selection procedure and that the inferences derived from the predictor scores are fair. Alternatively, the employer could show that no equally valid alternative, with less adverse impact, exists. Even though the use of this line of defence is quite widely advocated (Arvey & Faley, 1988; Cook, 1998; Gatewood and Feild, 1990; Guion, 1991), it nonetheless seems questionable. In the final analysis, the cause of adverse impact in personnel selection resides in systematic differences in criterion distributions. Adverse impact in criterion referenced personnel selection can therefore not be avoided by the judicious choice of selection instruments (Schmidt & Hunter, 1981). If adverse impact occurs because of differences in predictor performance across groups, which cannot be justified in terms of differences in criterion performance, it would imply that the criterion inferences derived from such test scores are biased (i.e., the selection decision-making is unfair in the Cleary sense of the term).
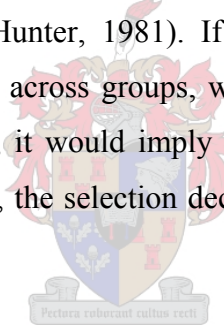
Personnel selection procedures should nonetheless strive to minimise adverse impact, not only in order to avoid litigation, but to ensure that access to job opportunities are distributed across groups in the labour market in proportion to the size of the various groupings and to optimally utilize the human resources available in the labour market.

1.6   SELECTION SCENARIOS

When personnel selection occurs from a diverse applicant group the ideal of simultaneously satisfying the foregoing four objectives is not always attainable. To explore the difficulties involved when selecting from a diverse applicant group, comprising of a previously disadvantaged (majority) group (π_1) and a previously advantaged (minority) group (π_2), it is

useful to graphically create specific selection scenarios, which differs in terms of the nature of predictor and criterion distribution differences across the two groups (Cascio, 1998).

1.6.1. Scenario 1

The first scenario (see Figure 1.1) depicts a situation where the predictor and criterion distributions[1] of the two groups coincide (i.e. $\mu[Y|\pi_1] = \mu[Y|\pi_2]$ and $\sigma^2[Y|\pi_1] = \sigma^2[Y|\pi_2]$). In such a scenario, if positive validity would be assumed to exist, people with high or low predictor scores also tend to have high or low criterion scores. If, in the investigation of differential validity, the joint distribution of minority and majority predictor and criterion performance scores are similar throughout the scatterplot, no problem exists and the use of the predictor should be continued, because it is possible to satisfy all four of the objectives (Cascio, 1998; Holborn, 1991).
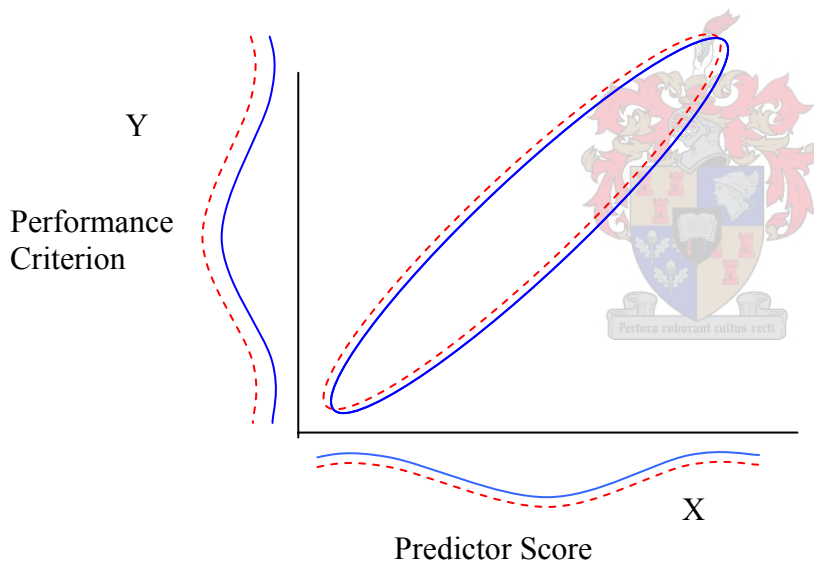
Figure 1.1

Positive Validity, Fair Selection Decisions, No Adverse Impact, And Utility.

If selection is done top-down, based on E[Y|X], then, in terms of the Cleary-interpretation of fairness, the use of a common regression line to base selection decisions on, will not

---

[1] The assumption is that the criterion construct (η) is multi-dimensional and that Y thus is a weighted linear composite representing η. Although it is true that specific dimensions would be more susceptible to ethnic or gender differences and that the dimension weights thus play an important role in determining adverse impact and validity, this aspect is not considered here.

result in systematic non-zero errors of prediction and fair selection decisions will result if the applicants with the highest predicted criterion scores are selected. Also, no adverse impact should be found. The selection procedure will further optimize utility at a fixed selection ratio, validity coefficient and selection cost. The utility will be positive if the monetary value of the performance gain affected by the selection procedure over random selection exceeds the investment required to affect the improvement (Cascio, 1998; Holborn, 1991; Petersen & Novick, 1976; Theron, 2001).

1.6.2. Scenario 2

The second scenario is when there are differences in the predictor distributions between the two groups, but the criterion distributions still coincide (i.e. $\mu[Y \mid \pi_1] = \mu[Y \mid \pi_2]$ and $\sigma^2[Y \mid \pi_1] = \sigma^2[Y \mid \pi_2]$). This scenario could imply the existence of one or more additional determinants of criterion performance on which minority group members on average tend to outperform the majority group. Alternatively the scenario could imply scale bias in the measurement of the underlying predictor construct. This scenario is depicted in Figure 1.2 below (Cascio, 1998).
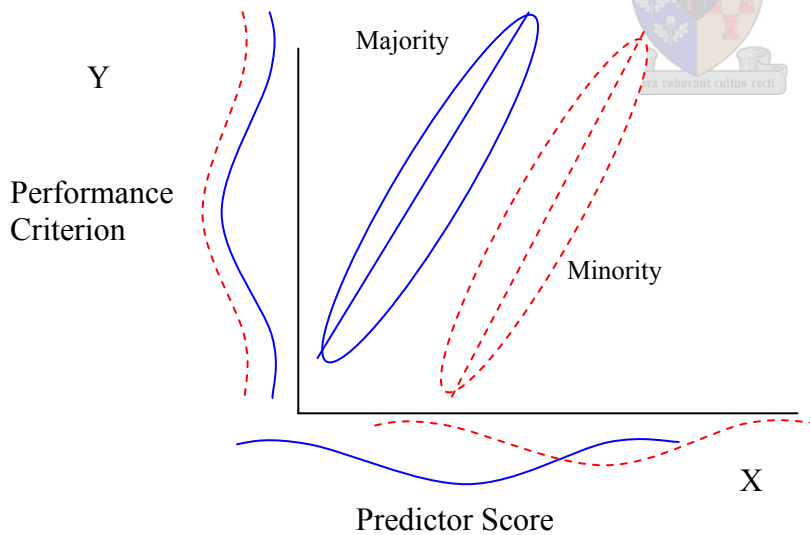
Figure 1.2

Equal Validity, Unequal Predictor Means.

13

In this scenario, if a single predictor cutting score would be (inappropriately) set for both the minority and majority group, the majority group would be less likely to be selected, although the probability of success on the job for both groups are the same. If a top-down selection approach is followed and a single regression line is used to derive E[Y|X] on which decisions are based, it would result in consistent non-zero errors of prediction within each group and selection decisions based on this single regression line would be unfair. Moreover, unfair adverse impact will occur although utility can still be achieved, although not optimized. The procedure could be justified through a criterion related validation study. In this scenario it is, however, also possible to satisfy the fairness and adverse impact objectives. By using separate cutting scores or separate regression lines for the two groups, or more sophisticatedly, by using an appropriate multiple regression equation, which makes provision for the differences in intercept, the fairness objective could still be satisfied. This should, in addition, eliminate adverse impact, while improving utility to its optimum value at a given selection ratio and selection cost. The utility of the selection procedure would be enhanced in that $r(E[Y|X; \pi_i],Y)$ will exceed $r(E[Y|X],Y)$ to the extent to which the combined regression equation resulted in systematic prediction errors. The primary focus should therefore always be on job performance, rather than on predictor performance (Cascio, 1998; Holborn, 1991; Petersen & Novick, 1976; Theron, 2001).
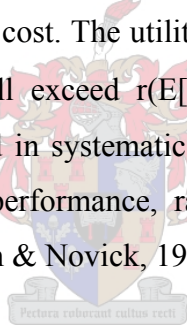
1.6.3. Scenario 3

Scenario three, depicted in Figure 1.3, occurs when there is no significant difference in the predictor distributions, but the members of the minority group tend to perform better on the job than the members of the previously disadvantaged majority group[2] (i.e. $\mu[Y|\pi_1] < \mu[Y|\pi_2]$ although $\sigma^2[Y|\pi_1] = \sigma^2[Y|\pi_2]$). This scenario could imply the existence of one or more additional determinants of criterion performance on which minority group members on average tend to surpass the majority group. If predictions were based on a combined regression equation derived from the combined sample group, systematic under- and over prediction would occur. Through the use of a single simple regression equation, the criterion scores of the minority group would be systematically under predicted, while those

---

[2] The assumption is that the difference in the mean of the criterion distributions of the minority and majority groups is not due to scale bias.

of the majority group would be systematically over predicted. Therefore, the use of a single simple regression equation would here result in consistent non-zero errors of prediction within each group and would result in unfair selection decisions while also lowering utility (Cascio, 1998; Holborn, 1991; Petersen & Novick, 1976; Theron, 2001). No adverse impact would, however, occur. This, furthermore, creates the ironical situation that although members of π_1 are systematically disadvantaged by the selection procedure, no *a priori* evidence exists in terms of which a *prima facie* case for indirect discrimination could be made and therefore seemingly no possibility exists of remedying the situation through employment litigation. Moreover, this illustrates the potential danger of trying to ameliorate adverse impact (Hough, Oswald & Ployhart, 2001) by focusing on strategies for reducing subgroup mean differences in the predictor.
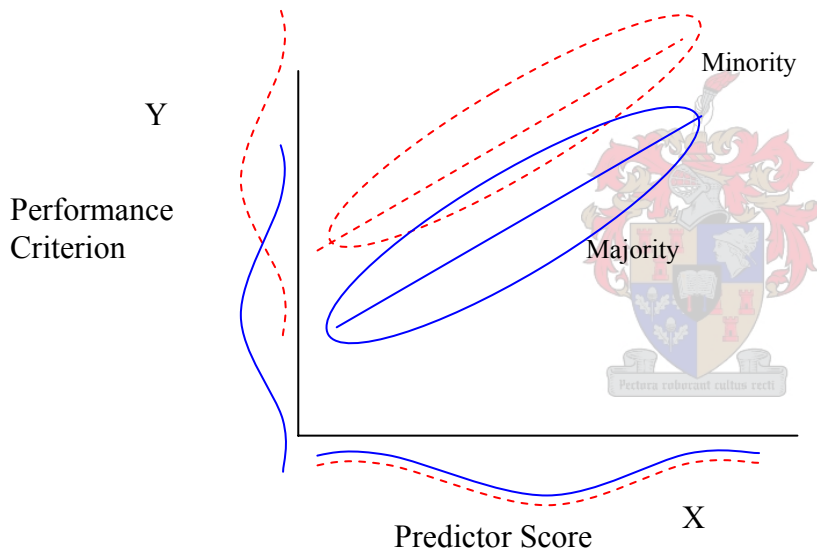
Figure 1.3

Equal Validity, Unequal Criterion Means, With Adverse Impact

Using separate regression lines, or using an appropriate multiple regression equation, to base selection decisions on and selecting those with the highest predicted criterion scores, would result in fair decisions. Adverse impact would, however, now be present if selection is done top-down, but the adverse impact would be fair. Furthermore no equally valid alternative selection instrument would be able to reduce the adverse impact. This is because there is indeed a real difference in the criterion performance between the two groups. Using

an appropriate multiple regression equation, which makes provision for differences in intercept through the inclusion of a group main effect, will also satisfy the utility objective. In other words, in scenario 3, three of the four objectives can be satisfied but not all four. When a valid predictor is used fairly in scenario 3, in a manner which optimises utility, the objective of minimising adverse impact would have to be sacrificed. The important point here is that the adverse impact would not be unfair, although the ideal would have been to avoid it without sacrificing any of the other objectives (Cascio, 1998; Petersen & Novick, 1976; Theron, 2001).

1.6.4. Scenario 4

The last scenario depicts a situation where validity for the minority and majority groups is the same, but the majority group scores lower on the predictor and performs poorer on the job. This scenario, depicted in Figure 1.4, alongside scenario 3, seems to be the most likely scenarios to be encountered in actual personnel selection in South Africa.
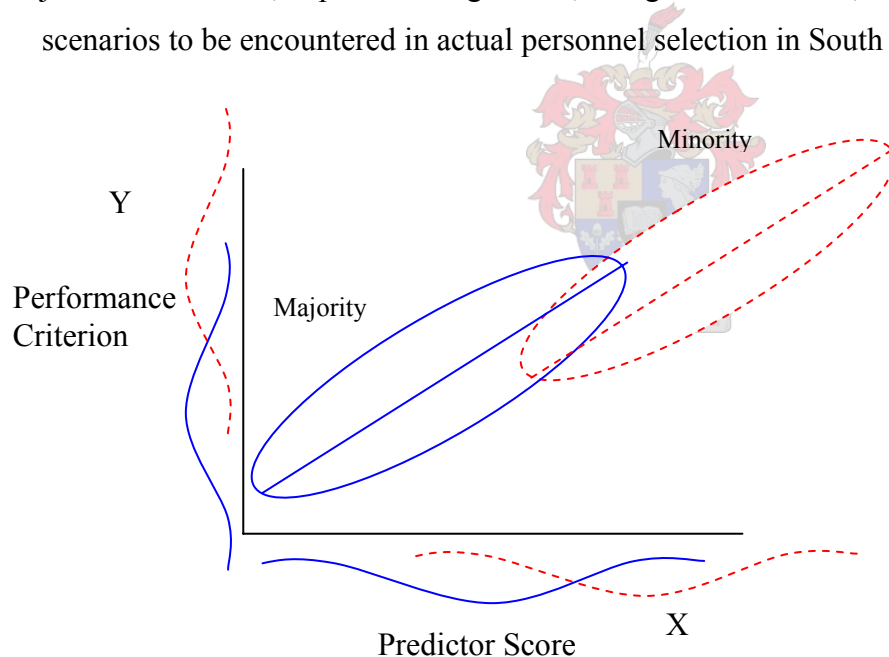
Figure 1.4
Valid Predictor With Adverse Impact

In scenario 4 the use of a single simple regression equation would result in systematic over- and under prediction, as explained in scenario three, and selection decisions would be unfair. Using a single simple regression equation would also cause utility to suffer. If an

16

appropriate multiple regression equation would be used, a top-down approach would still result in adverse impact, even though the selection decisions would be fair. A top-down approach, based on $E[Y|X, \pi_i]$ derived from the appropriate multiple regression equation would optimise utility even though adverse impact would not be minimised. Once again it is important to emphasise that the adverse impact would be fair and defensible as well as unavoidable as long as the utility and fairness objectives have priority over the adverse impact objective (Cascio, 1998; Petersen & Novick, 1976; Theron, 2001). A variation on scenario 4[3] would be to assume that the difference in criterion performance is equal to the difference in predictor performance times the validity of the predictor, so that a single regression line would result in no systematic group-related prediction errors. The utility and adverse impact outcomes would, however, remain the same.

1.7   A NEED FOR THE ASSESSMENT OF LEARNING POTENTIAL

In all four scenarios the assumption was that the selection procedure is equally valid for both groups and that the selection procedure thus could be justified in terms of the relevance of the information provided by the predictor. Available empirical evidence generally supports the assumption that differential validity is not a pervasive phenomenon (Arvey & Faley, 1988; Schmidt & Hunter, 1981).   If the selection decisions are fair in scenario one and two in terms of the Cleary-interpretation of fairness, and if strict top-down selection is followed based on expected criterion performance, then the objectives of minimising adverse impact and maximising utility can subsequently also be satisfied. If no differences in criterion performance would exist, no need for a developmental interpretation of affirmative action would exist.

However, in scenarios three and four all four objectives can no longer be satisfied simultaneously. If selection decisions are fair, in terms of the Cleary-interpretation of fairness, and selection occurs strictly top-down, based on $E[Y|X; \pi_i]$, then the objectives of fairness and utility can be satisfied, but the objective of minimising adverse impact can not

---

[3] The four scenarios clearly represent only a limited sample from an almost infinite number of possible situations that could occur.

be satisfied. In these two cases the objective of minimising adverse impact could be satisfied through quotas or race norming, but only if the utility objective is sacrificed (Theron, 2001). The sacrifice required by top-down hiring within each group (race norming) would depend on the magnitude of the difference in the criterion distributions. According to Schmidt and Hunter (1981):

> … selection systems based on top-down hiring within each group completely eliminates "adverse impact" at a much smaller price in lowered productivity. Such systems typically yield 85% to 95% of the productivity gains attainable with optimal nonpreferential use of selection tests. (p. 1130)

Meta-analytic summaries of criterion differences in the United States indicate a 0,30 standard deviation difference in mean minority and majority group criterion performance (Sackett & Roth, 1996). To the extent that similar conditions would exist in South Africa, race norming presents itself as a viable strategy to combat adverse impact. Ironically this is no longer permissible in the United States in terms of revisions to the Civil Rights Act of 1991 (Sackett, Schmitt, Ellingson & Kabin, 2001). Two considerations, however, argue against a blind reliance on within-group top-down selection. A drop in utility of 5% to 15% can be substantial when projected over number of selectees, time and successive cohorts. More importantly, however, to solely rely on within-group top-down selection would leave the root causes of the performance imbalance, which fundamentally underlies adverse impact, untreated.

Increasing the weights of the work performance dimensions less susceptible to ethnic or gender differences and decreasing the weights associated with dimensions on which larger differences exist would also reduce adverse impact on the composite criterion (De Corte, 1999; Hattrup, Rock & Scalia, 1997). The weighing of performance dimensions should, however, only reflect the relative importance of the various competencies in achieving the objective for which the job exists. The manipulation of criterion composite weights, therefore, does not offer a meaningful solution to the problem of adverse impact (Sackett, *et al.*, 2001).

Although it would not be intellectually honest to attribute the problem of adverse impact on biased selection instruments and/or unfair selection decision-making (Schmidt & Hunter, 1981) and although performance can be maximized fairly despite adverse impact, the problem of adverse impact can nonetheless not simply be ignored. How the human resource function should respond to the problem of adverse impact in selection would depend on why the systematic differences in criterion distributions exist. This is a question that is not raised often enough by human resource managers when contemplating the appropriate response to the dilemma outlined above. This question is, however, very important since remedial actions will only succeed if they deal with the root cause of the problem.

In the South African context it does not seem unreasonable to ascribe the systematic differences in criterion distributions to an environment where past injustices have had a negative impact on the development and acquisition of the skills, knowledge and abilities of certain groups required to succeed. In the past, and even now in the new democratic South Africa, specific groups had and still have easier and more access to opportunities that allow them to develop an array of coping strategies, knowledge, skills and abilities. Access to such opportunities often has the resultant effect that such individuals perform better in conventional assessment situations, in the workplace and in training programmes or educational institutions (Boeyens, 1989; Guthke, 1993; Hamers & Resing, 1993; Taylor, 1989; Taylor, 1992). In contrast there are underprivileged and socially disadvantaged groups which have been denied access to developmental opportunities at home, in school and because of social systems (Boeyens, 1989; Guthke, 1993; Hamers & Resing, 1993; Taylor, 1989). The denial of such opportunities has put these groups at a disadvantage, which only aggravates the adverse impact problem. Advantaged groups will consequently be even more advantaged, being selected for and gaining access to more opportunities, while disadvantaged groups will be more disadvantaged and denied opportunities to develop the necessary coping strategies, knowledge, skills and abilities (Boeyens, 1989). Tests that report standardized mean score differences between ethnic groups on especially measures of cognitive abilities should therefore not be characterized as villains responsible for the problem, but rather as unbiased messengers relatively accurately conveying the consequences of a tragic social system. The solution therefore is not to be found in

strategies to convince the messenger to alter its message as is seemingly suggested by Hough *et al*. (2001) and Sackett *et al*. (2001). The difference in criterion distributions observed between minority and majority groups reflect *bona fide* differences on numerous critical dispositions and attainments required to succeed in the world of work, which have resulted from the systemic denial of access to developmental opportunities. To deny the predictor differences and its impact is to deny the history that caused it.

If the differences in criterion performance between groups can indeed be attributed to past injustices, i.e. the lack of opportunities, then the question should be asked how human resource managers could correct the problem. The answer to this question lies in Milkovich and Boudreau's (1994) second human resource management category that is, maintaining and developing the current human resource supply. Therefore, organisations have to provide individuals who have been denied opportunities in the past with the opportunities to develop the still lacking knowledge, skills, abilities and coping strategies. The need for a developmental interpretation of affirmative action fundamentally lies in the existing differences in criterion distributions where no differences should exist. This argument, however, implies that past social injustices impacted directly on attributes required to perform successfully and not (so much) on psychological processes and structures that play a role in the development of the attributes required to succeed on the job. If past social injustices had the latter, more far reaching impact, rehabilitation of the psychological processes and structures through which critical attributes and competencies develop, would also be required.

Developmental affirmative action opportunities depend on various resources, but these are limited and not everyone can have access to costly developmental opportunities. Hence the need to identify those individuals who show the greatest potential to acquire the deficient attainments and dispositions (Saville & Holdsworth, 2000; 2001), and who would therefore subsequently gain maximum benefit from such development opportunities (Learning Potential Assessment, 2003). Especially in South Africa where organisations really have to affirm through action, there where past social injustices has seriously impacted directly on the attributes required to perform successfully, while still maintaining global

competitiveness, is it essential to identify those that show potential and to provide them with developmental opportunities. Taylor (1992) explains:

> Affirmative action, when implemented correctly, should not involve simply overlooking such skill and knowledge lacunae and advancing people anyway, just because of the colour of their skin. Real affirmative action must include a large development component.

This argument implies a two-stage selection procedure. It actually implies two distinct but linked selection procedures aimed at two qualitatively different criteria. The first selection stage aims to maximize the performance of a selected cohort on a learning performance criterion, whereas the second selection stage is aimed at maximizing performance of the selected cohort on a job performance criterion. Previously disadvantaged individuals who should gain maximum benefit from developmental opportunities would be selected during stage one, and during stage two their learning performance, possibly along with other job related predictors, would be used to assess their, and their more privileged counterparts', suitability for the job in question. Given the less than perfect predictive validity of any selection procedure, this seems a more prudent option than the alternative of selecting previously disadvantaged individuals directly into shadowing positions. This option also seems to have the added advantages that prediction occurs over a shorter distance and more relevant information is available during the job selection phase.

A need thus exists in South Africa for a method to identify individuals who will gain maximum benefit from affirmative developmental opportunities, especially cognitively demanding development opportunities[4]. Such a method should be one that not only focuses on the level of job performance that the individual can reach at present, but also one that adequately reveals hidden, latent reserve capacities and potential future levels of job performance (de Beer, 2003; Guthke, 1993; Learning Potential Assessment, 2003; Taylor, 1994). Both Taylor (1989) and de Beer (2000) agree that especially in South Africa, with its unique society and the continuous integration into schools, training institutions, industry

---

[4] The assessment of learning potential not only has relevance for selection into affirmative development opportunities though, but could play a role in the admission of employees into any training or development intervention. Learning potential, moreover, should be a valid predictor of performance in any position requiring a substantial amount of action learning

and society, such a fresh, more sensitive, diagnostic technique to assess the capabilities or potential of people from disadvantaged backgrounds, is needed. Ideally, such measures would assess an individual's core or fundamental cognitive abilities and potentialities and not specific job skills that are strongly influenced by past opportunities (Taylor, 1997). This line of reasoning should, however, never lose sight of the fact that only existing attainments and dispositions (Saville & Holdsworth, 2000; 2001) can be assessed. In addition inferences regarding future learning performance and future job performance can only be made from measures of existing attainments and dispositions under a construct orientated approach to selection. Phrased differently, neither learning performance, nor job performance, are random events. Intricate nomological networks of latent variables complexly determine both criteria. People will currently achieve a specific level of learning performance or job performance only if they *currently* satisfy the preconditions set by the nomological network.

Vygotsky (1978) proposed the measurement of learning potential as a method of assessing an individual's core or fundamental cognitive abilities and potentialities. Taylor (1992) defines learning potential as the underlying, (currently existing) fundamental aptitude or capacity to acquire and master novel intellectual or cognitively demanding skills, which is demonstrated through the improvements in performance in response to cognitive mediation, teaching, feedback, or repeated exposure to the stimulus material. Drawing on ideas developed, amongst others, by Vygotsky (1978), Sternberg (1984), Snow, Kyllonen and Marshalek (1984) and Ackerman (1988), Taylor (1989, 1994, 1997) developed a learning potential model, which explicates the latent variables collectively constituting learning potential. In essence, it represents a competency model in that it clarifies the behaviours or learning competencies that constitute learning performance as well as the dispositions or competency potential that determine such performance (Saville & Holdsworth, 2000; 2001).

Based on this learning potential model, a learning potential measure, specifically assessing an individual's hidden latent and reserve potential, reducing the influence of verbal abilities, cultural meanings and educational qualifications has been proposed and developed

by Taylor (1989, 1992, 1994, 1997) in the form of the APIL test battery. Taylor (1997) claims that this learning potential measure is especially suited for application in the following two practical settings. Firstly, it can serve as a useful tool in making fair decisions when job applicants are selected. This claim should, however, be questioned when interpreting fairness as defined above. Allied to this, is the fact that it can also help identify candidates who are likely to cope or master more demanding work roles. Secondly, it can be applied in the educational arena and will help identify candidates who are likely to master new cognitively demanding material in a formal educational or training context.

Earlier it was, however, argued that effective (although not necessarily efficient) selection would be possible if, and only if (a) (substitute) information is available at the time of the selection decision that is systematically related to the ultimate/final criterion of work success (i.e. relevant information); and (b) the nature of the relationship is at least subjectively/clinically, but preferably statistically/actuarially, understood. This would imply that effective selection of previously disadvantaged individuals into formal educational or training is possible to the extent to which there exists a comprehensive understanding of the reasons underlying training performance and the manner in which they combine to determine learning performance in addition to clarity on the fundamental nature of the key performance areas comprising the learning task. The APIL test battery will thus result in effective selection to the extent to which the explanatory model on which it is based successfully explains variance in learning performance.

The primary objectives of this research consequently are to (a) explicate the structural model underlying the APIL test battery and (b) evaluate the fit of the model on empirical data.

The APIL test battery provides dynamic measures of two latent learning competencies and static measures of two latent dispositions, which determine the learning competencies (Taylor, 1989, 1994, 1997). In estimating expected learning performance, these measures would typically be combined in a linear multiple regression model. Given the nature of the structural model underlying the APIL test battery, the question, however, arises whether the

static measures do not become redundant in a model that already includes the dynamic measures.

The secondary objectives of this research consequently is to determine whether the static measures of the two latent learning dispositions would significantly explain variance in learning performance when added to a model already containing dynamic measures of the two latent learning competencies. If the reservation about combining the various measures of the APIL test battery in a linear multiple regression model would turn out to be unfounded, it would imply an alternative conceptualisation of the structural model underlying the battery than that which has been suggested above. Inspection of the modification indices and expected change associated with the fixed paths in the initial model, in conjunction with the significance of the estimated path coefficients in the initial model and the results of the regression analysis will then be used to adapt the model and re-evaluate its fit.

If the structural model is indeed valid, and if the APIL test battery does succeed in selecting those who show a greater probability of succeeding in cognitively demanding developmental opportunities aimed at enhancing the required knowledge, skills, and abilities needed to succeed on the job, and the development programmes do succeed in reducing the differences in the criterion distributions, then adverse impact in job selection should be reduced. Previously disadvantaged individuals should now be significantly less disadvantaged in terms of the required knowledge, skills and abilities. Theoretically, over time, this approach should work towards levelling the playing field so that success or failure in personnel selection can be attributed to previous opportunities or lack thereof to a lesser degree than is currently typically the case in South Africa, without even temporarily relinquishing on the utility objective.

1.8   RESEARCH OBJECTIVES

Given the introductory argument unfolded above, the specific objectives of this research consequently are:

- ➤ To explicate the underlying structural model upon which the APIL test battery was developed, explaining learning performance;
- ➤ To test the model's absolute fit;
- ➤ To evaluate the significance of the hypothesised paths in the model;
- ➤ To investigate the predictive ability of an observed variable linear multiple regression model, regressing learning performance on a weighted linear combination of the two learning dispositions and the two learning competencies;
- ➤ To determine whether the static measures of the two latent learning dispositions would significantly explain variance in learning performance when added to a linear regression model already containing dynamic measures of the two latent learning competencies;
- ➤ To compare the predictive power of the structural model to that of the observed variable multiple regression model;
- ➤ To modify the structural model if necessary; and
- ➤ To compare the fit of the revised structural model to that of the original model.

CHAPTER 2
LITERATURE STUDY

2.1   INTRODUCTION

In this section of the thesis an attempt is made to replicate the comprehensive, systematic and reasoned argument that Taylor (1989, 1994, 1997) put forward in formulating his theory. In keeping with the logic set out in the previous paragraphs this section will start off by arguing the need for the analysis and conceptualisation of learning performance. The learning performance construct will then be analysed and interpreted as the ultimate or final criterion construct in training and development selection models. Based on Taylor's theory (Taylor, 1989; 1994; 1997), the hypothesised human qualities needed by an affirmee trainee in order to demonstrate successful learning performance (as conceptualised), will then be discussed. The main aim of this section of the thesis could be summed up as an attempt to reconstruct Taylor's thoughts, arguments and hypotheses in order to generate a logical theoretical justification of his theory on learning potential and ultimately to explicate the learning performance structural model implicit in his views on learning potential. This would require the development of constitutive definitions for all major constructs contained in the model and the development of theoretical arguments justifying the proposed path influences between constructs. This would allow the empirical evaluation of the fit of the structural model, which would reflect (although not in any definitive sense) on the justifiability of the use of the APIL test battery as a selection instrument for affirmative development interventions.

2.2   THE NEED FOR THE ANALYSIS AND CONCEPTUALISATION OF LEARNING PERFORMANCE

In the introductory section of this thesis it was argued that valid predictors used fairly in strict top down criterion referenced selection would in all likelihood cause significant adverse impact against the previously disadvantaged groups in South Africa. In the final

analysis this is due not to bias in the predictor, nor unfairness in the inferences made from predictor data, but rather due to significant differences in the criterion distributions of minority and majority groups. In South Africa, it was argued, it does not seem unreasonable to attribute the systematic differences in criterion distributions to a socially engineered environment, which systematically denied members of the majority group access to developmental opportunities. It was further, perhaps somewhat optimistically argued that these past social injustices impacted directly on the attainments and dispositions required to perform successfully and not [so much] on psychological processes and structures that play a role in the development of the attributes required to succeed on the job. If past social injustices had the latter, more far reaching impact, rehabilitation of the psychological processes and structures through which critical attributes and competencies develop, would also have been required. If the systematic differences in criterion distributions, and consequently the adverse impact caused by strict top-down criterion referenced selection, can in fact be attributed to artificially created deficiencies in competency potential, the logical remedy seems to be to provide individuals who have been denied opportunities in the past with opportunities to develop the still lacking attainments and dispositions. The introductory argument, however, acknowledges that the resources needed to implement such developmental affirmative action opportunities are limited and that not everyone can be given access to costly developmental opportunities. To obtain the optimum return on the resources invested in developmental affirmative action interventions, it should be restricted to those individuals who would achieve the highest possible level of competence in the behaviours that constitute job performance; thus those individuals whose relevant job competency potential could be lifted to the highest possible level. A need thus exists in South Africa for a method to identify individuals who will gain maximum benefit from affirmative developmental opportunities, especially cognitively demanding development opportunities.

It was further argued that when candidates are being selected for a specific educational- or training programme that decision-makers (i.e. human resource managers) are faced with the dilemma of not having information at the time of the selection-decision, on the criterion variable they are trying to maximize, that is, on the learning performance that each

candidate would have achieved at the end of the programme. Through selection an attempt is made to predict the future learning success of candidates and select those with the highest level of predicted learning performance. This gives rise to a need for relevant substitute measures of the ultimate or final criterion (Ghiselli *et al.*, 1981), which can be interpreted criterion referenced at the time of the selection decision, or, stated differently, a need for valid predictors from which criterion performance can be predicted at the time of the selection decision. Finding relevant substitute measures for the criterion which can be interpreted criterion referenced or valid predictors from which criterion performance can be predicted would require (a) a comprehensive understanding of why differences in performance exist and (b) a sound understanding of the nature of the relationships between predictors and criterion latent variables (Society for Industrial Psychology, 1998). This, in turn, implies that a conceptualisation of the ultimate or final criterion is required so as to permit the building of a comprehensive learning performance structural model, which would explicate the aforementioned two prerequisites.

The final criterion (η) in the case of an educational or training and development selection procedure is learning performance and success should thus be conceptualised in terms of that which constitutes successful learning in a training and development or educational programme. This, in turn, requires an extensive analysis of the educational or training and development task (or "job") facing affirmative action candidates. Learning performance can only be appropriately conceptualised if it is based on a proper and correct task or job description, which in turn is dependent on a proper task or job analysis. It is essential that the fundamental nature of the key performance areas, comprising the specific task, be kept in mind. Keeping these in mind is important, because it is the objectives or desired training outcomes of a specific training programme that differentiate one programme from another. These differences in programme-specific objectives or outcomes would imply different conceptions of success, which only emphasise the importance of the initial analysis. A proper analysis would reveal different objectives or outcomes of different programmes, which would ultimately lead to appropriate programme-specific conceptions of success (Cross *et al.*, 2002; Anastasi & Urbina, 1997).

These detailed, programme-specific conceptions of success are of critical importance when evaluating the affirmative development intervention and, more importantly in this case, when validating the application of a selection procedure feeding affirmees into a specific intervention. This is important because affirmative development interventions are developed to build up specific (underdeveloped) attainments and dispositions relevant to success in specific jobs and because the competency potential required by different jobs differ.

In the development of a learning performance structural model that would explain variance in learning performance and that would form the theoretical foundation for a generally applicable learning potential selection battery, a more generic conceptualisation of the ultimate criterion is, however, required.

The effective selection of previously disadvantaged individuals into formal education, training and jobs would thus only be possible if there is a comprehensive understanding of that (the learning competencies and learning outcomes) which constitutes successful learning performance (including clarity about the fundamental nature of the generic key performance areas comprising the learning task).
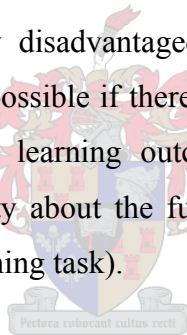
2.3   LEARNING PERFORMANCE

The key to the conceptualisation of the final criterion in selection for affirmative development seems to be the foregoing argument in terms of which the need for such interventions has been established. A dual competency model assists in clarifying this argument.

Competency modelling is a contentious topic in I/O Psychology (Schippmann *et al*, 2000). Nonetheless the competency model concept can serve as a powerful conceptual framework. Saville & Holdsworth (2001) proposed a conceptual model of performance at work, which captures the relationships between competency potential, competency requirements, competencies and outcomes in a manner, which allows for the integration and alignment of

the spectrum of human resource interventions. According to Saville & Holdsworth (2001, p. 6) the Performance@Work model represents:

> … a model of performance at work that defines the relationship between competency potential, competency requirements and competencies themselves. "Competencies" are defined as desired behaviours that support the attainment of organisational objectives. "Competency potential" is seen to derive from individual dispositions and attainments, and "competency requirements" involve both facilitators of and barriers to effective performance in the workplace. The framework points to ways in which people and work settings interact, and has implications for how performance in the workplace can be managed.

Jobs exist so that specific objectives can be realised through specific outcomes or output for which specific job behaviours (job competencies) are required. A complex nomological network of person-centred characteristics determines the level of competence achieved on these job competencies, some of which are relatively easily malleable (attainments) whilst others are more difficult to modify (dispositions). The legacy of Apartheid expresses itself in deficiencies in critical attainments and dispositions (job competency potential) that result in inferior job performance on the key performance areas (job competencies), which result in unsatisfactory outcomes and thus less than satisfactory attainment of the objectives for which the job exists. Affirmative development tries to salvage the situation by building up the specific attainments and dispositions relevant to success in a specific job with the expectation that this would eventually be reflected in the quality of the outputs for which the job exists.

In principle the same structure also applies to the affirmative training and development interventions. Individuals are assigned to affirmative development treatments with the aim of achieving specific learning objectives through specific learning outcomes. These learning outcomes are the exceedence of the minimum critical job competency potential (most likely, attainment) levels required to display the job competencies on a quality level sufficient to achieve the outcomes for which the job exists. Specific learning competencies are instrumental in attaining these desired learning outcomes. These learning behaviours, in turn, depend on and are expressions of a complex nomological network of person-centred

characteristics (learning competency potential), some of which are relatively malleable (attainments) and some of which are less easily altered (dispositions). A Performance@Learning competency model could thus be assumed, analogous to the Performance@Work model originally proposed by Saville & Holdsworth (2001). Moreover the Performance@Learning model should be sequentially linked to the Performance@Work competency model to provide a fertile conceptual model to explore the relationship between the characteristics of the learner required to exhibit the learning behaviours needed to develop the qualities necessary to exhibit the work behaviours that are instrumental in achieving the outcomes for which the job in question has been created. Figure 2.1 represents a schematic representation of the essence of the argument.
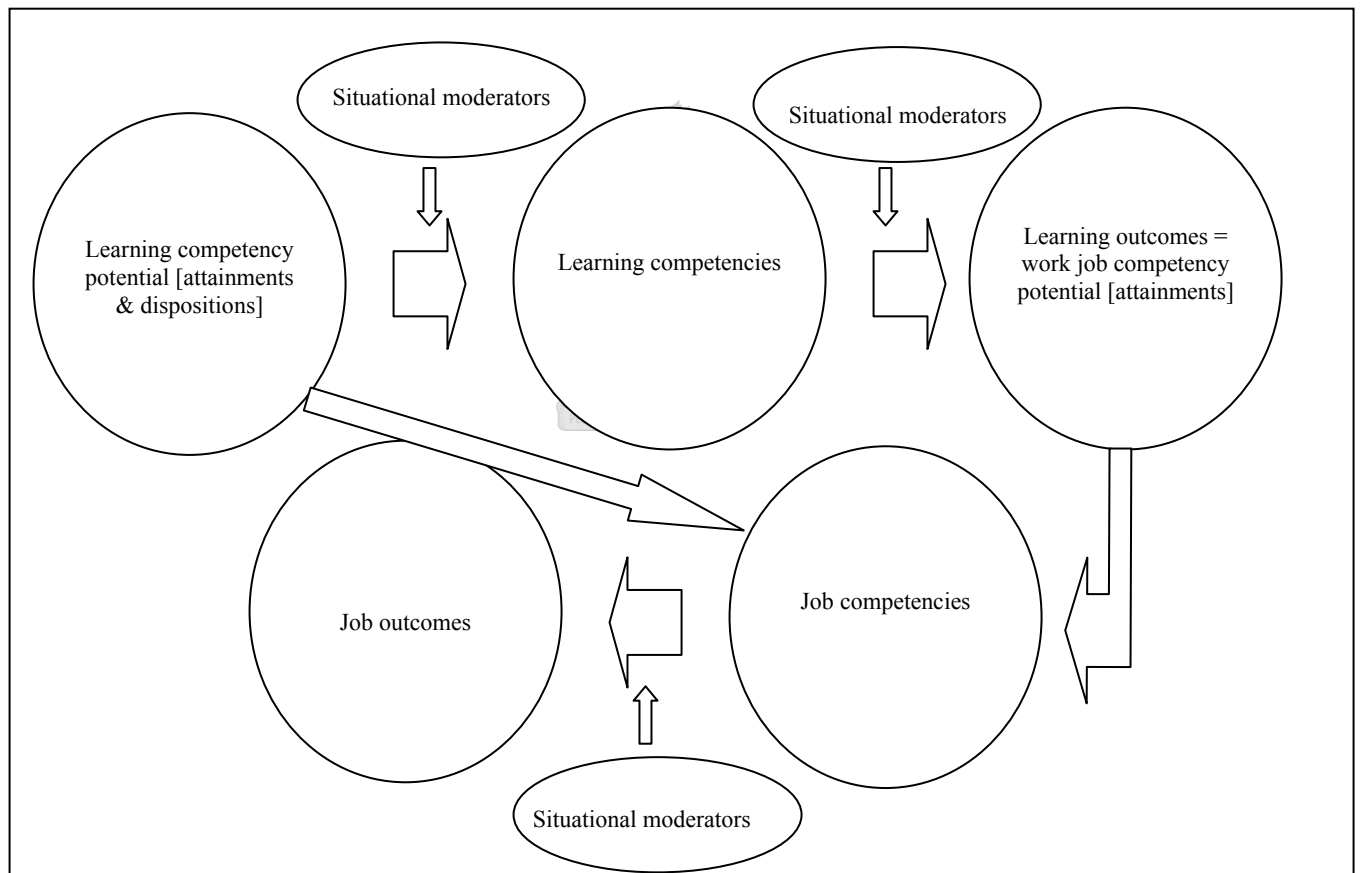
Figure 2.1

Integrated Performance@Learning & Performance@Work Model (adapted from Saville & Holdsworth, 2000, p. 7).

Learning performance, like job performance could be conceptualised (and therefore also assessed) on a behavioural (i.e., learning and/or job competency) level and/or on a learning outcome level. Learning performance thus should be defined in terms of the two core learning competencies (transfer of knowledge and automatization), in terms of the specific attainments and dispositions (to the extent that they could be modified through development interventions) required to succeed on the job in question and again on a behavioural level, in terms of the job competencies served by the development intervention. The logic underlying this position becomes apparent when considering the objective of affirmative training and development raised earlier. Affirmative development programmes are designed to empower employees with the job competency potential and job competencies required to deliver the outputs for which the job in question exists.

This would, however, hopefully mean more than simply retrieving previously transferred and automated (i.e., learned) responses to now familiar stimuli (although the application of newly acquired skills should not be dismissed altogether). The expectation rather would be that the affirmee would be able to apply the newly derived knowledge to novel stimuli not explicitly covered in the affirmative action development programme. The application of newly acquired knowledge in solving new work related problems is, however, again transfer at work and thus dependent on (a) fluid intelligence and, since fluid intelligence can not operate in a vacuum, (b) the extent to which previous relevant learning (transfer) has been successfully internalised (automated). No sharp division exists between learning and application. Class room learning occurs for the sake of world of work action learning. Learning performance should thus ultimately be (sequentially) assessed in terms of competence during training, in terms of the consequences or outcomes of learning (i.e., crystallized attainments/knowledge), and in terms of the ability to creatively utilise the newly derived knowledge in solving novel problems that could realistically be encountered in the work environment. This seems to have significant implications for the manner in which the criterion construct should be defined and operationalised in this validation study.

Taylor (1994) seems to focus his interpretation on the consequences or outcomes of learning. He, in addition, makes a very important distinction between learning performance and learning potential. Taylor (1994) writes:

> Learning performance is demonstrated when an individual acquires specialized skill through transfer from other fairly specialized skills or abilities. The more elaborated and developed a person's skill repertoire, the more effectively and swiftly he or she is likely to acquire the new skill. Learning potential is shown when a person comes to grips with a novel learning task involving unfamiliar stimulus material; in this case previously developed specific skills are of relatively little help to him or her, and the learner has to use very general transfer and skill acquisition strategies. (p. 190)

The distinction Taylor (1994) makes between learning potential and learning performance makes sense, and is especially important in understanding the underlying structural model. Learning performance should be understood as crystallised learning potential (acquired job competency potential) in action. Learning performance is the final criterion (η), not available at the time of the selection decision, which the selector is attempting to predict in a training and development scenario, while, according to Taylor's theory, learning potential should be understood as the substitute predictor construct (ξ) of learning performance.

Learning performance can be interpreted as the extent to which an individual has acquired a specific skill, knowledge or ability (job competency) and is the manifestation of that specific skill, ability or knowledge in action in a situation corresponding to the job for which the affirmative development is initiated. Learning potential, the individual's capacity to be modified and the capacity to acquire novel skills, is what needs to be assessed in disadvantaged individuals. It is learning potential that is crystallised through remedial intervention, and which allows an individual to demonstrate successful learning performance (Taylor, 1989).

The question subsequently arises as to what the learning competencies are that allow one individual to be more successful than another in acquiring a novel intellectually demanding skill (job competency). In other words, what learning competencies contribute to differences in learning performance between individuals?

To find an answer to these questions, Taylor (1994) reviewed the learning or dynamic approach to cognitive assessment, which focus on learning and modifiability, and found (a) transfer of knowledge and (b) automatization of information processes to be the two dimensions of learning (or learning competencies) to which successful learning performance could be attributed. Taylor (1994) makes the following comment:

> …it is clear that in the learning domain it is potential rather than achievement that should be assessed. The critical learning aspects to measure appear to be the implementation of general transfer strategies in dealing with novel material, and the early stages of proceduralization and automatization. (p. 190)

These two fundamental learning competencies that form the core of Taylor's (1989; 1994; 1997) theory on learning potential, will subsequently be discussed.

2.4 LEARNING COMPETENCIES

2.4.1 TRANSFER OF KNOWLEDGE

The acquisition of new job-specific knowledge, abilities and insight (job competency potential) can be described as a process during which new attainments have to be built on older ones and these have to be integrated into conceptual frameworks that subsequently become more general and elaborated (Taylor, 1994). According to Ferguson (1954; 1956), who attempts to relate learning and human abilities, and Taylor (1994), transfer forms the basis of this process of elaboration. Ferguson (1956) makes the following general formulation regarding transfer of knowledge:

> At any given point in time the organism may be said to be in a particular state. The concept of the state of a system is of importance in physics. It has a role in psychology also. This state undergoes continuous change because of a large number of circumstances both inside and outside the organism. One set of factors leading to a change in state is the behaviour of the organism in response to specific environmental circumstances, e.g. the performance of a task. Any change of state leads, theoretically, to changes in an indefinitely large number of other possible forms of performance. Any covariation which can be identified between any two or more forms of performance is conceptualised as a transfer function. (p. 126)

34

Ferguson (1954; 1956) clearly argues that the abilities of man emerge through a process of differential transfer. Taylor (1992; 1994) concurs with Ferguson (1954; 1956) and states that the abilities, which an individual already possesses, contribute to the development of new abilities. Transfer is the process through which crystallized abilities develop from the confrontation between fluid intelligence (Cattell, 1971) and novel stimuli (Taylor, 1994). Transfer is the application of that which an individual already knows to novel problems (McGeoch, 1946). Transfer can also be described as the effect previously learned behaviour has on the performance of new learning tasks (Gouws, Louw, Meyer & Plug, 1979). By implication it means that a task that is already learned or an ability that is already acquired makes it easier or more difficult to learn a new task or acquire a new ability (positive and negative transfer). Many learning theorists consider transfer as the most fundamental learning competency. Neither Ferguson (1954; 1956), nor Taylor, say anything about the psychological nature of transfer or why or how it occurs, instead, both of them rather try to clarify the concept. Taylor (1992) gives the following example to help clarify the concept of transfer:

> An example would be learning to program a computer in one's twenties or thirties (or even later in life). This ability may develop through transfer from verbal, numerical and reasoning skills, which in turn may have developed from the cognitive "engine" of fluid intelligence. (p. 6)
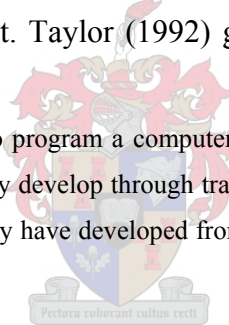
Ferguson (1956) tries to clarify the concept even further, by comparing it to the mathematical concept of a function. Ferguson (1956) writes:

> When two variables are so related that the values of one are dependent on the values of the other, they may be said to be functions of each other. It is customary to distinguish between dependent and independent variables, the value of the dependent variable being dependent on variation in the independent variable. The idea of function is descriptive of change in something with change in something else.
>
> The essence of the idea of transfer, also, is concomitant change, and in the simplest case implies change in performance on one task with change resulting from practice on another. (p. 124)

Ferguson (1954, 1956) believes that transfer is the more general phenomenon and learning is a particular formal case (Ferguson, 1954). An implicit condition for transfer to occur

seems to be that the prior task must differ in some respect from the subsequent task (Ferguson, 1956). In fact, Ferguson (1954; 1956) argues that if two tasks are presumed similar, based on superficial inspection, then the changes that occur in the "ability" of an individual to perform the specific task is functionally dependent on, or assignable to, repetition and not transfer *per se*.

In support of Ferguson (1954) Cook (1944) also argues that learning is a particular formal case of the general phenomenon of transfer. Cook (1944) writes:

> There is no separate problem of transfer of training. Or conversely, all learning (unless there exists a limiting case in which successive trials are identical on all counts) involves the problem posited in the transfer of training experiments: What identities and differences in successive trials affect what sort of learning? (p. 27)

It would seem as if a mass of prior experience is brought to bear on the learning of any task (Ferguson, 1954). McGeoch (1946, pp. 445-446) writes in support:

> After small amounts of learning early in the life of the individual every instance of learning is a function of the already learned organisation of the subject; that is all learning is influenced by transfer… The learning of complex, abstract, meaningful materials and the solution of problems by means of ideas (reasoning) are to a great extent a function of transfer. Where the subject "sees into" the fundamental relations of a problem or has insight, transfer seems to be a major contributing condition. It is, likewise, a basic factor in originality, the original and creative person having, among other things, unusual sensitivity to the applicability of the already known to new problem situations. Perceiving, at whatever complex level, is probably never free of its influence, and there is no complex psychological event which is not a function of it.

In discussing the same point Hebb (1949) writes:

> If the learning we know and can study, in the mature animal, is heavily loaded with transfer effects, what are the properties of the original learning from which those effects came? How can it be possible even to consider making a theory of learning in general from the data of maturity only? There must be a serious risk that what seems to be learning is really half transfer. We cannot assume that we know what learning transfers and what does not: for our knowledge of the extent of transfer is also derived from behavior at maturity, and the transfer from infant experiences may be much greater and more generalised. (p. 110)

Clearly, Cook (1944), Ferguson (1954, 1956), Hebb (1949), McGeoch (1946) and Taylor (1994) believe that transfer is a fundamental aspect of learning and cognitive development. It would seem as if they argue that individuals who are able to show superior learning performance would be those who are able to transfer better. In studies conducted by Campione, Brown, Ferrara, Jones and Steinberg (1985) and Ferretti and Butterfield (1992) it was found that slower students have the greatest difficulty with transfer. In other words, there is a definite difference in the ability to transfer between above average students and slower or lower ability students. These findings clearly support the arguments made by Cook (1944), Ferguson (1954, 1956), Hebb (1949), McGeoch (1946) and Taylor (1994).

Relating knowledge transfer to an educational and training situation, Taylor (1997) believes that a good student is one who is able to apply the knowledge that he or she has acquired from prior learning to other similar or related problems. Taylor (1997, p. 10) writes:

> In the work situation, the effective job incumbent is able to apply experience gained in one context to other related situations. Individuals low in this capacity are poor problem solvers and frequently have to refer to superiors or more competent peers for guidance.

In conclusion Taylor (1992) writes:

> Transfer is a phenomenon which is expressed when an individual comes to terms with novel or partially novel problems. Each subsequent set of problems in a transfer test differs from those that have come before, and is usually more complex than those that have come before. Therefore, the subject is continuously challenged, and the attainment of full understanding and correct answers is the pursuit of a shifting target. The stimulus material is "open-ended" in that new material is continuously being added. The educational process, as well as the process of acquiring new job skills, tends to be like this: new competencies are built on older ones and have to be integrated into conceptual frameworks that become more general and elaborate. Transfer lies at the heart of this process of elaboration. (p. 6)

Following the arguments put forward by the authors quoted in the previous paragraphs, it seems reasonable to argue that an individual would have to be able to transfer if he/she is to function successfully in a job (in the sense of solving novel problems via transfer from newly learned competency potential) and in a educational or training and development

environment. More so, it would make sense to include transfer of knowledge as a construct in a learning potential theory aiming to predict learning performance. Specifically, it should be viewed as a critical learning competency.

2.4.2 AUTOMATIZATION

Transfer of knowledge undoubtedly plays a role when learning tasks involve material that continuously change, but what about those situations where stimuli do not change dramatically over time? In such circumstances the challenge for the learner is rather to become more effective and efficient at what he or she is doing (Taylor, 1992). Moreover, learning tasks are not concluded once sense has been made out of novel stimuli. Unless an efficient cognitive algorithm can be written (Taylor, 1994) and stored for later retrieval that captures the insight/problem solving derived through transfer, the stimulus will remain a novel problem to be solved via transfer every time it is encountered. This would have greatly reduced the adaptive value of learning. Moreover, transfer would have been severely inhibited if newly derived insights did not accumulate in knowledge stations (Sternberg, 1984) to serve as the cognitive platforms from which subsequent problem solving/transfer occurs.
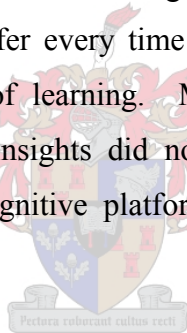
The only way in which an individual can become more effective and efficient in the execution of a task is if the individual automate many of the operations involved in performing the task. Sternberg (1984) agrees and argues that it is the automatization of a substantial proportion of the operations required to perform complex tasks that allows an individual to perform the task with minimal mental effort.

Sternberg's (1984) proposed model of automatization of information processes suggests that controlled information processing is under the conscious direction of the individual and that it is hierarchical in nature. Here he distinguishes between executive processes, which direct non-executive processes. Executive processes would be those that are used to plan, monitor, and revise strategies of information processing, while non-executive processes are

38

those used to actually carry out the strategies that the executive processes select, monitor, and revise (Sternberg, 1984).

On the other hand, Sternberg (1984) proposes that automatic information processing is pre-conscious, thus, not under the conscious direction of the individual and not hierarchical in nature. Here it is not possible to make the same functional distinction between executive and non-executive processes. Sternberg (1984) writes:

> Instead, production is in the mode of a production system, where all kinds of processes function at a single level of analysis. (p. 278)

When an individual is processing information from old domains or domains that are entrenched by nature, he or she primarily rely upon automatic, local processing (Sternberg, 1984). With regards to the processing of information from old domains or domains entrenched by nature, Sternberg (1984) writes:

> A central executive initially activates a system consisting of locally applicable processes and a locally applicable knowledge base. Multiple local systems can operate in parallel. Performance in these systems is automatic and of almost unlimited capacity; attention is not focussed upon the task at hand. Only knowledge that has been transferred to the local knowledge base is available for access by the processes utilized in a given task and situation. A critical point is that activation is by executive processes in the global system to the local system as a whole. The executive processes can instantiate themselves as part of this local system; when used in this instantiation, they do not differ functionally from processes of any other kind. (p. 278)

It would seem as if control is passed unto an already existing local system once an executive process has recognised a given situation as one for which a local system might be relevant. The local system would then act upon the given problem as a production system with a set of readily available productions. Functions in the production systems are both executive and non-executive in nature and integrated into a single non-hierarchical system (Sternberg, 1984).

When none of the productions in a system is able to satisfy a given present condition, the control is passed back to the global processing system. Sternberg (1984) writes:

> When the bottom of the production list is reached and no given condition is satisfied, global processing is necessary to decide how to handle the new task or situation. (p. 278)

Here, it would seem as if Sternberg (1984) acknowledges the role of transfer as the expression of an individual's fluid intelligence or abstract reasoning capacity operating on the content of a local processing system in solving novel problems. In addition, Sternberg (1984) writes:

> Once this task or situation is successfully handled, it is possible to pack what has been learned from global processing of the new experience into a given local processing system, so that the next time such a situation is encountered, there will be no need to exit from the local processing system. According to this view, the extent to which one develops expertise in a given domain largely depends on the ability of the individual to pack new information, in a useable way, into a given local processing system and on the ability to gain access to this information as needed. (p. 278)

In other words, when an individual faces a novel learning task he or she would first attempt to find a way of coping with the problem by "scanning" the already existing skills, knowledge and abilities. If a way of coping with a similar problem has been automated before then, the individual would use a learned response, as per Ferguson's (1954, 1956) theory, to deal with the new problem in a similar manner. However, if no directly applicable skills, knowledge or abilities exist, the individual would make use of fluid intelligence or abstract reasoning capacity to cope with the task by transferring existing relevant, but not directly applicable skills, knowledge and abilities onto a solution of the novel problem. Once the task is mastered the individual can add what has been learned to his or her already existing pool of skills, knowledge and abilities, thus, elaborating it. Once an individual is then again faced with a novel task he or she can now apply learned knowledge from a more elaborate pool of skills, knowledge and abilities, because of the addition of what has been learned, to master the new task.

Sternberg (1984) seem to have a similar understanding:

> In essence, a loop is set up whereby packing more information and processes into the local system enables them to automate more processing, and thus, to have global resources more available for what is new in a given task or situation. (p. 278)

Sternberg (1984) summarises his view on the ability to automate information processing as follows:

> … the present view essentially combines hierarchical and nonhierarchical viewpoints by suggesting that information processing is hierarchical and controlled in a global processing mode, and nonhierarchical and automatic in local processing modes. Expertise develops largely from the successively greater assumption of information processing by local resources. When these local resources are engaged, parallel processing of multiple kinds of tasks becomes possible. Global resources however, are serial and of very limited capacity in their problem-solving capabilities. (p. 278)

Following the arguments thus far, automatization seems to be an important dimension of learning and, therefore, Taylor includes automatization as a second learning competency in his theory of learning potential. Taylor (1997) writes:

> From both a practical (manpower utilisation) and ethical (fairness of opportunity) view it becomes important to assess dynamic aspects of the individual's cognitive endowment, especially his or her capacity to learn and acquire new competencies and ultimately to automize them. (p. 8)

2.5 LEARNING COMPETENCY POTENTIAL

In the formulation of his learning potential theory, Taylor (1992) also reviewed the conventional psychometric approach and the information processing approach to cognitive assessment. Taylor (1992) concluded that the capacity to form abstract concepts and information processing efficiency (speed, accuracy, flexibility) make up constituent parts of cognitive ability or intelligence. Moreover, Taylor seems to argue that these two facets of intelligence constitute the nucleus of the learning competency potential that drives the two learning competencies that constitute learning (transfer and automatization).

2.5.1   ABSTRACT THINKING CAPACITY

For many years there have been two paradigms present in psychology regarding intelligence. The first is that of Sir Francis Galton, who posited a unitary general cognitive ability as underlying all learning, problem solving, and other cognitive processing (Eysenck, 1986). Then there was Binet, who thought of intelligence as merely the average of a number of independent or semi-independent abilities. He argued that intelligence has no real existence, but that it is rather a statistical artefact (Eysenck, 1986). This dispute has now pretty much been resolved by psychometric studies, as Eysenck (1986) writes:

> …there clearly is a need for a general factor of intelligence to account for the "positive manifold" usually produced when IQ tests are intercorrelated and for the low rank of the matrices constituted of these intercorrelations. (p. 3)

Eysenck's (1986) quote links very well with the work of Spearman (1904, 1927) who proposed that the base of human intelligence lies in a unitary, general intelligence factor, which he dubbed the *g*-factor (*g*). However, Binet was not completely wrong in his theory of separate abilities either. Eysenck (1986) explains:

> …the evidence is now strong that in addition to general intelligence (*g*), we have a number of what English psychologists usually refer to as "group factors" and American psychologists as "primary abilities", independent of *g* and adding a certain amount to the total variance in cognitive testing. (p. 3)

But, commenting on the superior importance of *g*, Eysenck continues (1986):

> In the total variance, however, *g* is clearly much more important than any primary ability or even than all primary abilities taken together. (p. 4)

In 1971 Cattell proposed a theory in line with the above statements made by Eysenck (1986) in which he hypothesised that Spearman's (1904, 1927) general intelligence factor (*g*) is in fact not a unitary factor, but that it is made up of two distinct factors. He termed these fluid- (*Gf*) and crystallised (*Gc*) intelligence (Jensen, 1998). Here it could be argued that Cattell's *Gf* is probably very similar to Spearman's (1904, 1927) *g*, while *Gc* is the same as the "group factors" or "primary abilities" of which Eysenck (1986) speaks.

The two-factor model of fluid- and crystallised intelligence as proposed by Cattell (1971), taken in conjunction with the learning competency of transfer, offers an interesting explanation of why differences in abilities between individuals exist.

According to Cattell (1971) *Gf* is a fundamental, innate intelligence and can be related to all kinds of problem solving. *Gf* is related to how well an individual perceives complex relations, forms concepts and engages in abstract reasoning. It is the fundamental abstract reasoning and concept formation capacity or ability that an individual applies to novel problems (Cattell, 1971, Jensen, 1998). Furthermore, *Gf* is also applied in the development of new abilities and in the acquisition of new knowledge (Cattell, 1971). A very important point here is that G*f* is relatively formless and appears independent of experience and education. Therefore, it is *Gf* that is demonstrated in mental tests (e.g. Ravens Progressive Matrices or abstract reasoning tests) in which prior learned knowledge, skills, algorithms, or strategies offer little or no advantage (Jensen, 1998).

On the other hand, *Gc* refers to the acquired abilities and knowledge which arise from schooling, becoming competent with one's culture and mastering one's specific circumstances and could be called consolidated knowledge (Cattell, 1971; Jensen, 1998). Acquired abilities such as verbal and numerical comprehension could be categorised under *Gc*. *Gc* thus seem to have a scholastic and cultural foundation (Jensen, 1998).

The learning competency of transfer seems to link *Gf* with *Gc* in as far as transfer in essence is *Gf* in action in the solution of novel problems. Existing *Gc* is elaborated via transfer by *Gf* utilizing existing *Gc*. Moreover, the view that *Gf* operates on existing *Gc* in solving novel problems through transfer seems to have important practical implications for selection into affirmative development interventions. It seems to suggest that, apart from the static learning competency potential (*Gf* and information processing capacity), *Gc* also needs to be explicitly taken into account in a prediction model in as far as the ability to cope with novel, cognitively demanding learning material (i.e. transfer) will depend on the interaction between crystallized knowledge and abilities and the ability to transfer. Somehow it seems naive to assume that candidates for affirmative development will be able

to cope with novel, cognitively demanding learning material if a high *Gf* is present, irrespective of the extensiveness and level of crystallized knowledge and abilities in the domain on which the development intervention is focussed. In as far as *Gc* constitutes consolidated knowledge (Cattell, 1971, Jensen, 1998), it thus seemingly corresponds to what Sternberg (1984) refers to as the content of local processing systems or what is termed an attained (rather than dispositional) learning competency potential in the integrated Performance@Learning & Performance@Work model depicted in Figure 2.1. *Gf,* by contrast would correspond to a dispositional learning competency potential in the integrated Performance@Learning & Performance@Work model depicted in Figure 2.1.

There are quite a number of theorists arguing along the same lines as Cattell (1971), in that they also believe the core of intelligence to be fluid or abstract in nature. In fact, Snow, *et al*. (1984) represented cognitive abilities as points on a radex and found fluid intelligence to be right in the centre, while more specific abilities where situated at the periphery. Guttman (1965), as cited in Taylor (1994), represented test scores on a circumplex and also identified test scores at the centre as analytic or rule inferring and those further from the centre as more rule applying in nature.

Clearly, the ability to think abstractly and form concepts as described in Taylor's theory (Taylor, 1994), is the same as fluid intelligence proposed in the Cattell (1971) theory. As Taylor (1994) explains:

> The potentiality to think abstractly and form concepts develops as fluid intelligence. It consists of a set of general cognitive tools and strategies for application to novel problems. (p. 190)

Taylor (1992) further writes:

> Fluid intelligence is thus abstract thinking capacity, and it is best measured by confronting the testee with novel stimuli and asking him or her to find underlying concepts. (p. 5)

It would certainly seem as if an individual's abstract reasoning capacity plays an important role in both dealing with novel kinds of problems and learning. Therefore, an individual's

level of fluid intelligence or abstract reasoning capacity would (as a dispositional learning competency potential) either contribute or inhibit the individual's capacity to make sense of the learning task allowing the learning and acquisition of new knowledge, skills and abilities (via transfer), especially when the learning task becomes more complex in nature.

2.5.2   INFORMATION PROCESSING CAPACITY

The information processing approach towards the conceptualization of intelligence started to develop in the second half of the previous century and received a great deal of attention because of its more scientific approach to cognition. Psychometrics and cognition had a long alliance before the information processing approach came along, but during this period it was especially the development of theory that seemed to stagnate. The rise of a link between computer-systems and the understanding of human perception, thinking, and problem solving further fuelled the information processing approach. Thus, it was mainly the fact that man came to be seen primarily as an information-processor and the more scientific and theoretical nature of the information processing approach that contributed to its popularity (Estes, 1978; Taylor, 1992, 1994).

There exists a real danger that the constructs introduced thus far (transfer, automatization and abstract reasoning ability) can become conceptually confounded with the concept of information processing capacity. An unambiguous grasp of what is meant by information processing is thus first required. A clear understanding of what is meant by information processing is provided in the work of Jensen (1998). Jensen (1998) describes information processing, or more specifically information processes, as follows:

> Information processes are essentially hypothetical constructs used by cognitive theorists to describe how persons apprehend, discriminate, select, and attend to certain aspects of the vast welter of stimuli that impinge on the sensorium to form internal representations that can be mentally manipulated, transformed, stored in memory (short-term or long-term), and later retrieved from storage to govern the person's decisions and behaviour in a particular situation. (p. 205)

In the Concise Dictionary of Psychology information processing is defined as (Statt, 1998):

A key term in cognitive psychology used to denote what happens mentally between stimulus and response including perception, memory, thinking, problem-solving and decision-making. Information is usually taken to be any stimulus with a mental content- an image, idea, fact, opinion, etc. (p. 71)

First of all, it is important to clearly understand what is meant by the term information in this context. In the Concise Dictionary of Psychology definition, information is "taken to be any stimulus with a mental content (p. 71)". Jensen (1998) throws caution to the wind and conceives of "information" to have a more generalised and non-specific meaning than when it is commonly used. He goes on to write:

"Information" here does not refer to any specific fact or a particular item of acquired knowledge. It refers generally to any stimulus that *reduces uncertainty* in a given situation (p. 206).

Jensen (1998) subsequently concludes:

We use the term "information processing" here to describe the hypothetical processes that depend, presumably, on the structural and physiological properties of the brain that are activated whenever uncertainty is perceived and we work to reduce it. (p. 206)

Taylor (1994) believes that information processing does make up one of the constituent parts of cognitive ability. However, according to Taylor (1992), it would be "unwise" to conclude that information processing as widely measured with chronometric measures, which tap behaviour in very simple tasks, adequately accounts for human intelligence.

Hunt (1980) also goes on to write.

…it is unreasonable to expect that any one information-processing procedure would provide "the answer" to our questions about the nature of intelligence. … but the search for a "true" single information-processing function underlying intelligence is likely to be as successful as the search for the Holy Grail. (pp. 456-457)

In a learning context the learner is often faced with novel, intellectually challenging tasks. Such tasks cause the individual to experience a lot of uncertainty; thus, he or she would naturally try and reduce it. In order to reduce the uncertainty the individual has to firstly use executive processes (Sternberg, 1984) to process the bits of information or stimuli provided

in the task and select a strategy to follow and secondly, use non-executive processes (Sternberg, 1984) to actually carry out the strategy. The processing of bits of information through cognitive processes (executive and non-executive), which are activated in an uncertain situation in order to reduce the amount of uncertainty, could be termed information processing.

In an attempt to get to the core difference between information processing and information processing capacity Taylor (1992) writes:

> It is true that the information processing approach is more "dynamic" than the conventional structural approach. It does, after all, address itself to processes rather than merely to the resultant of these processes. But the information processing approach does not concern itself much with change over time. (p. 4)

Also, argues Underwood (1978):

> When information is presented to an individual the sequence of processing is not pre-determined, but the individual is able to select certain processes and reject others. The traditional flowchart of information processing suggests that once sensory data are entered into the system, then the response is structurally determined. The view here is that the response may be structurally limited, but that the strategies used by the processor- the individual- play a vital role. (p. 2)

The fact that differences exist in the choice of a problem-solving strategy is one of the reasons why cognitive psychologists often fail to relate information processing to intelligence scores. In fact, writes Hunt (1980):

> … it is shown that information-processing and psychometric measures are in much closer correspondence when account is taken of one's problem solving strategy. (p. 449)

The importance of distinguishing between information processing and information processing capacity now becomes clearer.

In clarifying the difference Taylor (1992) asks whether the elemental tasks in measures of information processing account for the "richness of more complex cognitive behaviours?" He also poses the question of whether they can "adequately account for the phenomena of learning?"

…in more complex behaviours, the person has to string together a large number of these processes. And as he or she becomes more adept with experience, he or she develops new and more efficient ways of assembling and employing the processes…

Information processing psychology as it is presently practised pays relatively little attention to individual differences in processing – "styles" of processing- and also ignores changes in styles with experience or learning… Various types of processing capacity, such as the rate and accuracy with which the stimuli of a problem can be taken in, the number of pieces of information that can be thought about at the same time, and the efficiency with which needed information can be retrieved from long-term memory would appear to have an impact on learning, particularly the rate of learning. (p. 4)

The strategy an individual selects to solve a given problem is one of the factors, which either contributes or impinges the capacity to solve the problem (Hunt, 1980; Underwood, 1978). Strategy, however, seems not to be the only factor that puts a boundary on an individual's capacity to process information (Taylor, 1992; Underwood, 1978). Underwood (1978) also writes:

Thus, our limitations in solving problems, given any one strategy, will be a composite of the speed of comprehension and assimilation of the information comprising the problem, of the storage limits of working memory, of the forgetting characteristics of the memory systems used, of the efficiency of the access code for retrieving information stored in permanent memory and which maybe relevant to the problem, and of the speed and efficiency of any other system used in the total activity. (p. 2)

Consequently, Taylor (1997) identifies three broad information processing capacity parameters. These are (a) the speed or quickness with which information of a moderate difficulty level is processed (processing speed). Taylor (1997) states that individuals who are slow information processors might fall behind in learning situations. The reason being, that they might not have enough time to investigate all the reasonable solution options to problems. (b) The accuracy with which information of a moderate difficulty level is processed (processing accuracy). Taylor (1997) argues that a person who is inaccurate in processing information often suffers from lapses in concentration accompanied by a failure to monitor and "quality control" processing performance.

Taylor (1997) clarifies:

> Information processing which is both quick and accurate is called efficient; that which is slow and accurate is called thorough; that which is quick and inaccurate is called impulsive; and that which is slow and inaccurate is called inefficient. (p. 7)

(c) The cognitive flexibility with which a problem-solving approach, which is appropriate to the problem, is selected. Cognitive flexibility, with which an individual selects a problem-solving approach, appropriate to the problem from a personal "toolkit" of cognitive strategies is a fundamental characteristic of intelligent behaviour (Hunt, 1980, Taylor, 1997). Individuals who keep on following an inappropriate strategy would be regarded as having a lesser capacity to process information.

At this stage it is very important to emphasise a point, which Taylor (1994) raises- that is that, just as with an individual's abstract reasoning capacity, an individual's capacity to process information is mostly genetically endowed, implying that an individual's capacity to process information is fairly free from the influence of culture and opportunities, but also that a certain capacity sets an upper limit to performance (Taylor, 1994). This might serve as another reason why attempts to relate information processing to intelligence scores in normal subjects, have not had much success, while it has been very successful in showing differences between extreme groups in information processing capacity (Hunt, 1980).

An individual's capacity to process information should play an important role in cognition and learning by which people become aware of and gain knowledge about the world. The argument would be that if one individual has a relatively better capacity to process information than another does, then that individual could be described as having relatively better cognitive ability. Thus, in a learning context it would seem as if the individual who can more efficiently and effectively (quickly, accurately and flexibly) process information would be the one who is able to acquire more, learn faster and perform better. For this reason Taylor (1994) includes information processing capacity as a (dispositional learning competency potential) construct in his theory.

2.6 TAYLOR'S THEORETICAL POSITION

In this section an attempt is made to clearly describe how Taylor (1992, 1989, 1994, 1997) integrates the ability, information processing, and learning traditions of cognitive psychology into a theory that accommodates all three traditions. Many of the points relating to such an integration most probably would have been discussed in the previous section but will nonetheless, for the sake of completeness, be repeated in this section which will serve as a final threading together of any loose ends.

The reason why Taylor wants to integrate all three approaches stems from an observation that the psychological tests that are widely available for use in industry and education are mostly designed to measure broad-based static psychological constructs such as abilities. Furthermore, it seems as if the two new independent approaches of information processing and learning and modifiability tend not to be widely used in industry, despite a need for assessment techniques of a more dynamic nature.

One of the reasons that could be contributing to this lack of usage might be the fact that the information processing and learning and modifiability approaches do not offer much in the form of practical measurement instruments, which can be used for selection or vocational guidance (Taylor, 1994). Thus, through his work Taylor (1992, 1994) attempts to relate his theoretical concepts to the mainstream of cognitive psychology while developing an assessment instrument that is suitable for practical application in an industrial or educational context.

Taylor (1992, 1994) believes that Ackerman's (1988) cylindrical cognitive model, an elaboration of Snow, *et al.* (1984) circular cognitive model, best accommodates his theoretical position of integrating all three traditions and explaining individual differences in skill acquisition (i.e. learning performance).

What follows is a brief account of the work done by Ackerman (1988). At the start of his article Ackerman (1988) states that what he is presenting is:

An integrative theory that links general models of skill acquisition with ability determinants of individual differences in performance…(p. 288)

Ackerman (1988) proposes that when an individual is initially faced with a skill-acquisition task (assuming that the information processing requirements are relatively novel) he or she places strong demand on the cognitive-attentional system. Ackerman (1988) writes:

During this phase performance is slow and error prone, as strategies (productions) are formulated and tested, and attention is primarily given to understanding and performing the task in question. (p. 289)

However, with consistent practice, performance speed and accuracy increase with a reduction in attentional demands (Ackerman, 1988). Here, the productions that are needed to perform the task become fully formulated. Ackerman (1988) explains:

During this second stage (Phase 2) the stimulus-response connections of the skill are refined and strengthened. (p. 290)

Lastly, the ultimate stage (Phase 3) of performance can best be characterised as autonomous or automatic. Ackerman (1988) writes:

Consistent practice results in fast and accurate performance; the task can often be completed competently even when attention is simultaneously devoted to other tasks. (p. 290)
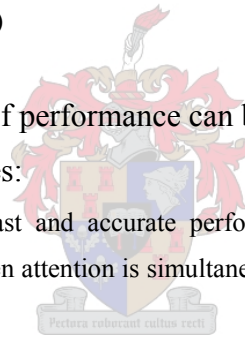
Ackerman (1988) also draws a link between three major ability factors and the three skill acquisition phases. The three ability factors are (a) general intelligence, (b) perceptual speed, and (c) psychomotor ability.

One point of concern that needs to be raised at this stage is the domain of the theory presented by Ackerman (1988). Ackerman (1988) explains:

In accordance with the definition of skills presented in the early part of this article, each experiment made use of tasks that depended to a substantial degree on motor behavior. … Skills such as chess mastery or physics problem solving do not depend to any significant degree on motor behavior and as such are not expected to follow the ability-performance transitions outlined in this theory. (p. 311)

51

From the above quotation it would seem as if Ackerman (1988) intended his theory to be more applicable to the acquisition of motor skills such as operating simple machinery, driving a car, playing musical instruments, and so forth. Even though this might be the case, Taylor (1992, 1994) obviously believes that Ackerman's (1988) theory is also relevant as a basis for understanding the acquisition of more cognitively oriented skills, such as playing chess. In a learning context where psychomotor ability would not play a significant role in final criterion performance (i.e. a university) the argument would, thus, be that it is, especially, the general ability- and perceptual speed factors that would cause variance in final performance. Therefore, the inclusion of general intelligence and perceptual speed in a theory pertaining to the described scenario would still make sense.

Ackerman (1988) regards general intelligence (general ability) as a broad construct that underlies non-specific information processing efficacy. Ackerman (1988) also states that the reasoning processes that account for individual differences across different content domains represent one component of a general intellectual ability. Clearly, Ackerman (1988) is a great supporter of the information processing approach to intelligence.

Taylor (1992, 1994) agrees with Ackerman (1988) in that information processing makes up one of the constituent parts of cognitive ability or general intelligence. However, in also reviewing the conventional psychometric approach to cognitive assessment, Taylor (1992, 1994) concludes that another constituent part of cognitive ability or general intelligence is the capacity to form abstract concepts. As Taylor (1992) explains:

> I would rather adopt a two-factor model of intelligence, the two factors being abstract thinking capacity and information processing capacity. The two factors are probably not totally separate: processing efficiency no doubt assists in success in forming abstract concepts, and abstract thinking no doubt helps in the formation of effective strategies to process information in the most economical way. (p. 5)

Taylor also writes (1994).

> Information processing speed and capacity are not the complete foundation of intelligence, although these form one of two main fundamentals. The other is the potential to infer concepts and thus think abstractly. This potentiality is not independent

of processing speed and capacity-the two factors are related; but processing variables do not fully account for the individual's potential to think abstractly. (p. 190)

The other ability factor that is mentioned in Ackerman's (1988) theory is perceptual speed. The core of this concept appears to involve speed of consistent encoding and comparing symbols. Ackerman (1988) goes on to represent the structure of human abilities as a cylinder. Ackerman (1988) explains why:

Problems exist in locating perceptual speed or psychomotor abilities in the Marshalek et al. (1983) model. Representation of these abilities can be rectified by explicitly segregating the complexity-specificity dimension from one of level-speed. With this modification a third dimension allows for both perceptual speed and psychomotor abilities. By using the basic two-dimensional surface at the extreme on the power (level)-speed dimension (i.e., a zero value for speed of information processing demands) and an arbitrary value for the extreme in speed (with the absence of cognitive processing, i.e., non cognitive motor speed), the structure of human abilities can be presented as a cylinder… Theoretically, as one moves down the cylinder, concentric sections represent the basic cognitive ability groups, with increasing demands on speed. (p. 291)

Figure 2.2 is a depiction of Ackerman's (1988) representation of human abilities as a cylinder. Taylor (1994) understands and explains Ackerman's (1988) cylindrical representation as follows:

Competencies near the core of the cylinder are more general and closely related to the genotypic potential. Progressively larger concentric rings contain skills which are ever more specific and remote from fundamental potential. These rings also reflect the process of transfer in development and learning (Ferguson, 1954; 1956). The vertical dimension of the cylinder is a speed dimension. Starting from the top, each successive 'slice' through the cylinder contains skills which are of an increasingly speeded nature.

As development proceeds, skills and knowledge accumulated in prior learning have a growing impact on the emergence of new skills.

Several authors (e.g. Anderson, 1983; Shiffrin & Schneider, 1977) have distinguished three phases of learning: conceptual understanding of the task, compilation of execution procedures, and automatization of processing. The boundary between the second and third phase is not distinct and the second phase merges into the third. It seems likely that the abstract thinking factor will play the major role in the first phase, whereas the processing speed and capacity factor will play an increasingly important role as learning

53

progresses to the phase of automatization. As automatization progresses, skills shift outwards and downwards in the cognitive cylinder.

Measures in the core of the cylinder provide the best estimate of the individual's fundamental potentiality. Those skills at the periphery are the product of a longer process of learning and transfer. (p. 190)
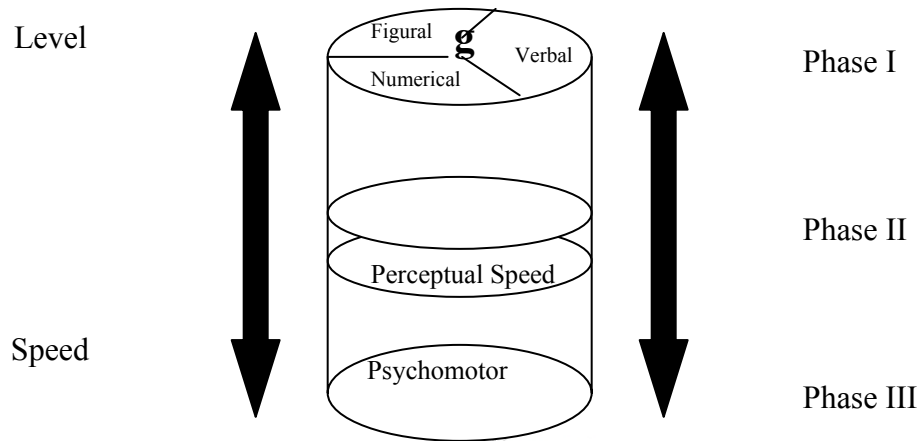
Figure 2.2

A Modified Radex-Based Model Of Cognitive Abilities **(Ackerman, 1988, p. 291)**

Ackerman (1988) explains individual differences in skill acquisition by providing three principles. Principle one, or phase one (refer to figure 2.2), corresponds to demands on general abilities. Here, individual differences in performance will be moderately to highly associated with general ability. In other words, abstract reasoning capacity and information processing capacity as in Taylor's theory. Ackerman (1988) writes:

> With practice, once production systems are formulated to accomplish the consistent components of the task, the influence of general and content abilities will diminish. (p. 293)

In other words, now the individual has started, through application of abstract reasoning, general information processing and the process of transferring bits and pieces of knowledge (relevant to the specific task or problem) from already existing local processing systems, to design a local processing station "written" specifically to deal with the type of task or problem at hand (refer back to automatization discussed earlier).

Ackerman's (1988) second principle states that skill acquisition phase two correspond to perceptual speed ability. Here it is important to remember that we are moving down- and outward, through practice, in the cylinder, thus, towards transfer from the new station and automatization. Ackerman (1988) explains this process:

> Early in practice the productions are still being formulated and tested; thus compilation and tuning are involved only to the degree that previously learned productions can be readily adapted for successful performance of the current task. Therefore, once the productions are formulated, there is an initially increasing association between perceptual speed ability and performance. Perceptual speed ability, that is, the facility and speed compilation of production systems that determine performance efficiency, is the essence of Phase 2. (p. 293)

Relating this back to Taylor's theory, at this stage, it is especially the learning dimensions of transfer of knowledge (now from the formulated production system) and automatization that will cause differences in individual performance i.e. the individual with a better capacity to transfer and automate will be the one who can better acquire or learn a new skill. Also, another way of understanding this is to refer back to Cattell's (1971) theory, in other words, the fluid abilities are now becoming crystallised abilities.

The third principle in Ackerman's (1988) theory relates more to individual differences stemming from motor abilities and will, thus, not be discussed here.

The argument thus far, seems to suggest that differences in skill acquisition (i.e. learning performance) between individuals could be explained in terms of four constructs, namely: abstract reasoning capacity, information processing capacity (speed, accuracy, flexibility), transfer of knowledge and automatization. But, the question still remains as to what the specific causal linkages between these constructs are, if indeed there are such causal linkages.

The preceding argument seems to suggest that information processing capacity and automatization should be causally linked, because it is the task- or role specific information processes that have to be automated. The individual's ability to store what has been learned

from global processing of a novel experience into a given local processing system (automatization), so that the next time such a situation is encountered, there will be no need to exit from the local processing system (Sternberg, 1984), depends on the speed, accuracy and flexibility with which information can be processed. But what about abstract reasoning capacity and transfer of knowledge?

Taylor (1992) argues that there is a direct causal link between abstract reasoning and transfer of knowledge:

> … the concept of fluid intelligence, which is seen by many cognitive psychologists as the fundamental or core ability, is related to the concept of transfer, which is seen by many learning theorists as the fundamental characteristic of learning. (p. 6)

In other words Taylor's theoretical argument is that an individual's capacity to transfer knowledge is causally linked to the individual's abstract reasoning capacity. Also, that an individual's ability to automate is causally linked to the individual's capacity to process information. Furthermore, that transfer of knowledge and automization is causally linked to learning performance. This theoretical argument culminates in a structural model (illustrated in Figure 2.3) that depicts the specific paths or hypothesised causal linkages between the constructs.
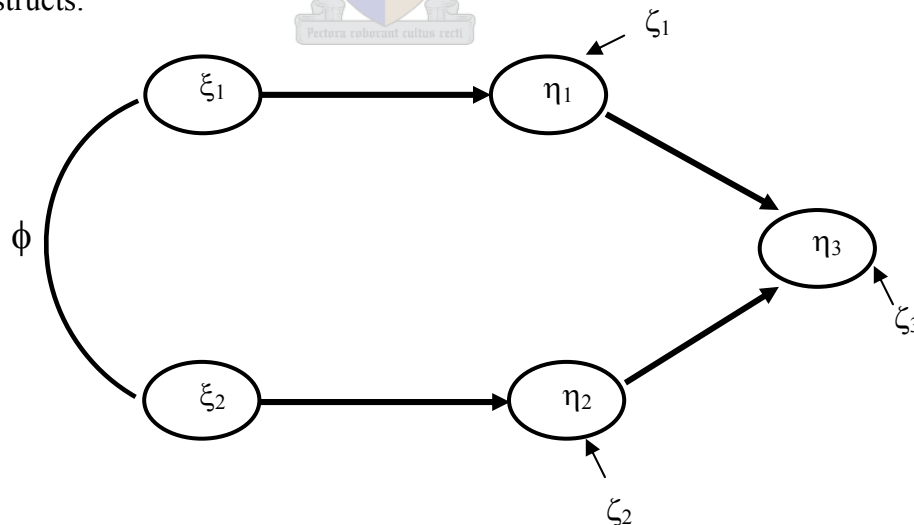
Where:
ξ_1 = *Abstract thinking capacity* η_1 = *Transfer of knowledge*
ξ_2 = *Information processing capacity* η_2 = *Automatization*
  η_3 = *Learning performance*

Figure 2.3

Graphical Portrayal Of Proposed Learning Potential Structural Model

The proposed structural model, which serves as the basis for this study, can be expressed as a set of structural equations (see equations one to three), representing the research problems that will be investigated:

$$\eta_1 = \gamma_{11}\xi_1 + \zeta_1 \text{--(1)}$$

$$\eta_2 = \gamma_{22}\xi_2 + \zeta_2 \text{--(2)}$$

$$\eta_3 = \beta_{31}\eta_1 + \beta_{32}\eta_2 + \zeta_3 \text{--(3)}$$

The structural model depicted in Figure 2.3 can also be portrayed mathematically in terms of a series of matrices. The structural model is defined by the following four matrices and three vectors:

- A 3 x 2 Γ (gamma)- matrix of path/regression coefficients γ describing the strength of the regression of η_i on ξ_i in the structural model;

- A 3 x 3 square B (beta)-matrix of regression/path coefficients (β) describing the strength of the regression of η_i on η_i in the structural model;

- A 2 x 2 symmetrical Φ (phi)-matrix of variance and covariance terms describing the variance in (φ_{ii}) and covariance between (φ_{ij}) the exogenous latent variables ξ_i and ξ_j (it is assumed that the exogenous latent variables are correlated and thus all off diagonal elements in Φ will be set free to be estimated);

- A 3x3 symmetrical Ψ (psi) matrix of variance and covariance terms describing the variance in (ψ_{ii}) and covariance between (ψ_{ij}) the structural error terms ζ_i and ζ_j (it is assumed that the structural error terms are uncorrelated and thus that Ψ is a diagonal matrix);

- A 2 x 1 ξ (ksi) column vector of exogenous latent variables;

- A 3 x 1 η (eta) column vector of endogenous latent variables;

- A 3 x 1 ζ (zeta) column vector of residual error terms.

More specifically, the hypothesised causal relationships depicted in Figure 2.3 can be expressed in matrix form as equations 4 and 5.

$$
\begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \beta_{31} & \beta_{32} & 0 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix} + \begin{pmatrix} \gamma_{11} & 0 \\ 0 & \gamma_{22} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{pmatrix} \quad \text{-------------(4)}
$$

$$
\eta = B\eta + \Gamma\xi + \zeta \quad \text{---(5)}
$$

The structural model depicted as Figure 2.3 and Equation 4 could possibly be extended by freeing the parameter β_{12}, thus making provision for a causal linkage between automatization and transfer. Automatization of the operations required to perform complex tasks allows an individual to perform the tasks with minimal mental effort (Sternberg, 1984), thus freeing cognitive capacity, specifically *Gf*, for novel problem solving (i.e., transfer) (Taylor, 1994). In addition the possibility should be considered to split the endogenous latent variable learning performance into a job competency potential latent variable (η_3) and a job competency latent variable (η_4) so as to align the structural model more closely with the argument underlying the competency model depicted in Figure 2.1.

If the endogenous latent variable learning performance would be interpreted to refer to the creative use of newly acquired knowledge (rather than the level to which job relevant knowledge and abilities have been developed), then the freeing of the parameter γ_{31} (in equation 4) should also be considered. Development programmes are designed to empower employees with the job competency potential and job competencies required to deliver the outputs for which the job in question exists. This should refer to more than simply the retrieving of previously transferred and automated (i.e. learned) responses to now familiar stimuli (again the application of newly acquired skills should not be dismissed altogether). The expectation rather would be that the affirmee would be able to apply the newly derived knowledge to novel stimuli not explicitly covered in the affirmative action development programme. The application of newly acquired knowledge in solving new work related problems is, however, again transfer at work and thus dependent on (a) fluid intelligence and, since fluid intelligence can not operate in a vacuum, (b) the extent to which previous relevant learning (transfer) has been successfully internalised (automated). By the same token information processing capacity should also affect the ability to apply newly derived

knowledge to novel stimuli not explicitly covered in the affirmative action development programme. The extended structural model can be depicted as Figure 2.4
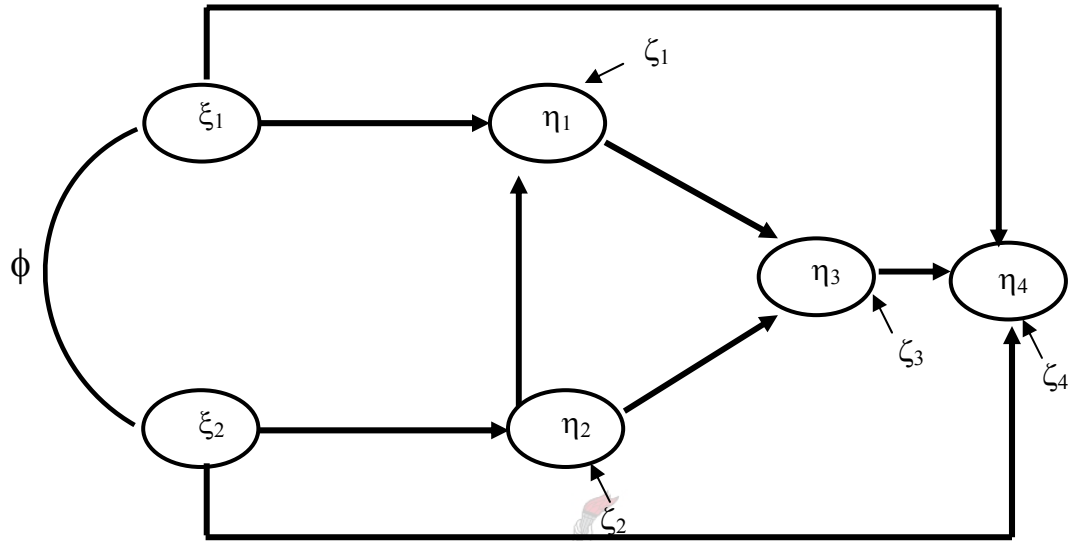
Figure 2.4

Graphical Portrayal Of Extended Learning Potential Structural Model

The revised structural model can again be expressed as a set of structural equations representing the research problems that will be investigated:

$\eta_1 = \gamma_{11}\xi_1 + \beta_{12}\eta_2 + \zeta_1$---(6)

$\eta_2 = \gamma_{22}\xi_2 + \zeta_2$--(7)

$\eta_3 = \beta_{31}\eta_1 + \beta_{32}\eta_2 + \zeta_3$--(8)

$\eta_4 = \beta_{43}\eta_3 + \gamma_{41}\xi_1 + \gamma_{42}\xi_2 + \zeta_4$---(9)

The revised structural model can again be portrayed mathematically in terms of a series of matrices:

- A 4 x 2 Γ (gamma)- matrix of path/regression coefficients γ describing the strength of the regression of η_i on ξ_i in the structural model;

- A 4 x 4 square B (beta)-matrix of regression/path coefficients (β) describing the strength of the regression of η_i on η_i in the structural model;

- A 2 x 2 symmetrical Φ (phi)-matrix of variance and covariance terms describing the variance in (Φ_{ii}) and covariance between (Φ_{ij}) the exogenous latent variables ξ_i and ξ_j (it is again assumed that the exogenous latent variables are correlated and thus all off diagonal elements in Φ will be set free to be estimated);

- A 4x4 symmetrical Ψ (psi) matrix of variance and covariance terms describing the variance in (ψ_{ii}) and covariance between (ψ_{ij}) the structural error terms ζ_i and ζ_j it is assumed that the structural error terms are uncorrelated and thus that Ψ is a diagonal matrix);

- A 2 x 1 ξ (ksi) column vector of exogenous latent variables;

- A 4 x 1 η (eta) column vector of endogenous latent variables;

- A 4 x 1 ζ (zeta) column vector of residual error terms.

More specifically, the hypothesised causal relationships depicted in Figure 2.4 can be expressed in matrix form as equations 10 and 11.

$$
\begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{pmatrix} = \begin{pmatrix} 0 & \beta_{12} & 0 \\ 0 & 0 & 0 \\ \beta_{31} & \beta_{32} & 0 \\ 0 & 0 & \beta_{43} \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{pmatrix} + \begin{pmatrix} \gamma_{11} & 0 \\ 0 & \gamma_{22} \\ 0 & 0 \\ \gamma_{41} & \gamma_{42} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \end{pmatrix} \quad \text{------------(10)}
$$

$$\eta = B\eta + \Gamma\xi + \zeta \text{--(11)}$$

CHAPTER 3
RESEARCH METHODOLOGY

3.1 INTRODUCTION

The literature study has culminated in a basic learning potential structural model in which learning performance has been treated as a job competency potential latent variable. The basic model has subsequently been expanded by making a distinction in the definition of learning performance between a job competency potential latent variable and a job competency latent variable. The expanded model introduces an additional latent variable and consequently the basic model is not nestled in the expanded model in a manner, which would allow one to statistically evaluate the merits of adding additional paths to the model.

The ideal, thus, would be to fit the expanded model since it corresponds more closely with the argument underlying the Performance@Learning competency model depicted in Figure 2.1. The ability to evaluate the fit of the expanded model is, however, contingent on the availability of suitable operational measures of the job competency latent variable.
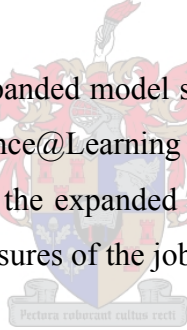
Unfortunately, the only data set that could be obtained for this study did not include both job competency potential and job competency as facets of learning performance. Only the basic model was consequently fitted in which learning performance is equated with the level of competence achieved in the job competency potential targeted by the affirmative training intervention. One modification was, however, made to Figure 2.3 in that automatization was permitted to exert a causal influence on transfer of knowledge as in the expanded model. Equation 4 thus can be expressed as equation 12.

$$
\begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix} = \begin{pmatrix} 0 & \beta_{12} & 0 \\ 0 & 0 & 0 \\ \beta_{31} & \beta_{32} & 0 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix} + \begin{pmatrix} \gamma_{11} & 0 \\ 0 & \gamma_{22} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{pmatrix} \quad \text{-------------(12)}
$$

The validity and credibility of the implicit claim of the study to have come to the correct/true verdict on the fit of the structural model depends on the methodology used to arrive at the verdict. Methodology is meant to serve the epistemic ideal of science. If very little of the methodology used is made explicit, there is no way of evaluating the merits of the researcher's conclusions, and the verdict therefore simply has to be accepted at face value (whilst the verdict might be inappropriate due to an inappropriate or wrong procedure for investigating the merits of the structural model). The rationality of science thereby suffers, as does ultimately the epistemic ideal of science (Babbie & Mouton, 2001). A comprehensive description of the research methodology is consequently presented below.

3.2 RESEARCH PROBLEMS

In determining the fit of the basic learning potential structural model the following questions will be investigated in this study:

- ❑ Does the basic learning potential structural model provide an adequate explanation of the covariance observed between the measures of learning performance, the learning competencies and learning potential?
- ❑ Is the extent to which transfer of knowledge occurs determined by the level of abstract thinking capacity?
- ❑ Is the extent to which automatization occurs determined by the level of information processing capacity?
- ❑ Is the extent to which transfer of knowledge occurs determined by the extent to which automatization occurs?
- ❑ Does transfer of knowledge determine job competency potential targeted by the affirmative training intervention?
- ❑ Does automatization determine job competency potential targeted by the affirmative training intervention?
- ❑ Is the influence of abstract thinking capacity on the job competencies targeted by the training intervention mediated by transfer of knowledge?

❑ Is the influence of information processing capacity on the job competencies targeted by the training intervention mediated by automatization?

The APIL test Battery provides dynamic measures of two latent learning competencies and static measures of two latent dispositions, which determine the learning competencies (Taylor, 1989, 1994, 1997). In estimating expected learning performance, these measures would typically be combined in a linear multiple regression model. Given the nature of the structural model underlying the APIL battery, the question, however, arises whether the static measures do not become redundant in a model that already includes the dynamic measures. The following research questions are thereby implied:

❑ Do the dynamic measures of the two latent learning competencies each explain unique variance in a composite measure of the job competency potential targeted by the affirmative training intervention?;

❑ Do the static measures of the two latent learning dispositions explain variance in a composite measure of the job competency potential targeted by the affirmative training intervention when added to a model already containing dynamic measures of the two latent learning competencies?; and

❑ Do the dynamic measures of the two latent learning competencies and the static measures of the two latent learning dispositions each explain unique variance in a composite measure of the job competency potential targeted by the affirmative training intervention?

3.3 MEASURING INSTRUMENTS/OPERATIONALISATION

To obtain empirical proof that the relationships postulated by the expanded learning potential structural model provides a plausible explanation for differences observed in learning performance, measures of the various exogenous and endogenous latent variables comprising the model are needed. In other words, the research hypothesis expressed as equation 12 should be operationalised by creating an exogenous and an endogenous measurement model.

The two measurement models describe how the exogenous and endogenous latent variables reflect themselves in manifest variables. However, to come to valid and credible conclusions on the ability of the expanded learning potential structural model to explain variance in learning performance, evidence is needed that the manifest indicators are indeed valid and reliable measures of the latent variables they are linked to in accordance with the measurement models. Diamantopoulos and Siguaw (2000) clarifies:

> Clearly, unless we can trust the quality of our measures, then any assessment of the substantive relations of interest (i.e., the links among the latent variables themselves) will be problematic. Thus an evaluation of the measurement part of the model should precede the detailed evaluation of the structural part of the model. (p. 89)

Viewed from the perspective of a traditional validation study more than a practically and statistically significant validity coefficient would be needed to justify the use of the APIL Test Battery for selection into affirmative development interventions. To justify the claim that inferences on learning performance (η) can be made from the observed scores obtained from the APIL test battery it needs to be shown that (a) Y is a valid and reliable measure of learning performance (η), (b) X_j are valid and reliable measures of the latent learning competencies and competency potential measured by the APIL test battery (ξ_j) (c) the valid and reliable measures (Y) of the conceptualised final criterion (η) is systematically related to valid and reliable substitute measures (X_i) of the latent variables measured by the APIL test battery (ξ_j), to ensure criterion-related validity (Guion, 1991; Theron 2002).

Part of the evidence needed to establish the psychometric integrity of the indicator variables, used to operationalise the latent variables comprising the expanded learning potential structural model, is presented below. The evaluation of the fit of the respective measurement models will, in addition, also reflect the extent of the success with which the indicator variables represent the latent variables to which they were linked.

3.3.1 ABSTRACT THINKING CAPACITY

Abstract thinking capacity (ξ_1) was measured with the Concept Formation Test, which is a sub-test of the test battery. This is a test that measures the individual's ability to form

abstract concepts, reason hypothetically, theorise, build scenarios and trace causes (Taylor, 1997).

The Concept Formation Test is a classificatory task where the testee is presented with sets of geometrical diagrams and then has to identify a diagram, which does not share a characteristic that all the others share (Taylor, 2006).

The reliability of the Concept Formation Test scores was calculated with Kuder-Richardson-type estimates. KR-20 coefficients (with correction applied under the assumption that the item difficulties are normally distributed) ranging between 0,78 and 0,87 were obtained for the Concept Formation Test (Taylor, 2006).

The nature of the APIL test battery allowed assigning each of the thirty items in the Concept Formation Test with either a 0 or 1 value. A score of 0 was assigned to each incorrect answer, while a score of 1 was assigned to each correct answer.

In order to fit the model through structural equation modelling (see paragraph 3.8.2.1) two operational measurement scores for abstract reasoning capacity were needed. Therefore, two parcels were made by assigning all the items with equal numbering (i.e. 2, 4, 6 etc.) to one parcel and assigning all items with unequal numbering (i.e. 1, 3, 5 etc.) to the other parcel. The total number of correct answers (i.e. answers assigned 1) in each individual parcel of items were then used as operational measurement scores (X1 & X2).

For the regression analysis (see paragraph 3.8.2.1) the total number of items correct was used as the measurement score for abstract thinking capacity.

3.3.2   TRANSFER OF KNOWLEDGE

Transfer of knowledge was measured with the Knowledge Transfer Test, which is a sub-test of the APIL test battery. The Knowledge Transfer Test measures knowledge transfer by exposing the testee to a number of related but increasingly complex problems. The

individual is given answers and feedback to example problems after he or she has completed each problem (Taylor, 2006).

Reliabilities for the Knowledge Transfer Test were estimated through the split-half method. Taylor (2006) explains:

> … the scores for problems 1 and 3 are summed and the scores for problems 2 and 4 are summed and the totals for these two halves are correlated and corrected for test shortening. (p. 63)

Split-half reliability coefficients ranging between 0,71 and 0,84 were obtained for the Knowledge Transfer Test (Taylor, 2006).

Each candidate completed four individual sub-tests in the Knowledge Transfer Test. The total number of items correct obtained by each candidate in each individual sub-test was used as operational measurement scores (Y1, Y2, Y3 and Y4). In other words, four operational measurement scores, expressing transfer of knowledge, were used to fit the model through structural equation modelling.

For the regression analysis the total number of items correct, across the four sub-tests, was used as the measurement score for transfer of knowledge.

3.3.3 INFORMATION PROCESSING CAPACITY

Information processing capacity was measured with the Flexibility-Accuracy-Speed-Tests. The Flexibility-Accuracy-Speed-Tests is a battery of four sub-tests within the APIL test battery that provides measures of the speed (quickness), the accuracy and the cognitive flexibility of information processing (Taylor, 2006).

Three scores obtained for speed (X3), accuracy (X4) and flexibility (X5) were used as operational measures for information processing capacity to fit the model through structural equation modelling.

The Processing Speed score was calculated by adding the total number of items attempted (whether correct or incorrect) over the first three sub-tests (the forth sub-test requires the testee to work with all three problem types presented in the first three subtests) (Taylor, 2006).

Taylor (2006) states that the reliability of the Information Processing Speed variable cannot be directly determined. He does, however, go on to say that some indication of the reliability can be obtained by inspecting the correlations between the three components that are added together to derive the speed score. These are the Series Number Attempted, Mirror Number Attempted and Transformations Number attempted. Correlation coefficients ranging between 0,45 and 0,72, with a mean of 0,61, have been obtained between the three sub-test scores across six samples (Taylor, 2006). Taylor (2006) states that, given the correlations, it is expected that the Processing Speed score comprising all three scores will have a reliability in the 0,80's.

The Accuracy score is a logarithmically transformed and inverted score of error rate. The formula that was used to calculate the Accuracy score is as follows (Taylor, 2006):

[5]$Accuracy = 100\text{-}30\log_{10}[(\text{Number of Errors/Number Attempted}) \times 200]$

The reliability of the Accuracy score was estimated by combining sub-tests 1 and 3 and also sub-tests 2 and 4 (Taylor, 2006). Separate accuracy indices for each of the two combined scores were calculated and corrections were made for test shortening. Reliability coefficients ranging between 0,70 and 0,86 were obtained across six samples (Taylor, 2006).

The Flexibility score is a function of the amount of work correctly done in the first three sub-tests in comparison with the amount of work correctly done in the final sub-test (Taylor, 2006). The following formula was used to calculate the Flexibility score:

$Flexibility = (\text{correct output in sub-test 4})^2/(\text{correct output in sub-test 1, 2 \& 3})$

---

[5] All four sub-tests were taken into account in calculating the number of errors and number attempted as per the formula.

Taylor (2006) argues that it is not possible to calculate the reliability of the Flexibility score. This is due to the fact that the learning/familiarity effect would corrupt the scores, unless the test-retest exercise is conducted many months apart. He does, however, go on to argue that the Flexibility score typically has large variance, which is a pre-requisite (but no guarantee) for good reliability (Taylor, 2006).

Once again the total sum of all three scores (speed, accuracy and flexibility) was used in combination as one measurement score for information processing capacity in the regression analysis.

3.3.4   AUTOMATIZATION

An indication of automatization would be if an individual becomes ever more adept and efficient at what he or she is doing. Such mastery is often expressed as a learning curve reflecting the number of units of work correctly done in successive time segments. The steeper the learning curve, the more rapid the process of automatization (Taylor, 1992).

Automatization was assessed with the Curve of Learning test, a sub-test of the APIL test battery, as the increase of work output over four sessions (Taylor, 2006).

Two operational measurement scores for automatization, a total output score and a memory and understanding score was used to fit the structural model. The output score was calculated as follows:

$$^6\text{COL1} + 1{,}75\text{COL2} + 2{,}33\text{COL3} + 2{,}8\text{COL4}$$

---

[6] The adjustment factors applied to COL2, COL3 and COL4 is to correct for time shortening

Reliability for the total output score was estimated by computing COL1 + COL3 and also COL2 + COL4 and then correlating these two scores and correcting for test shortening (Taylor, 2006). Reliability estimates, across six samples, ranging between 0,88 and 0,97 were obtained across six samples (Taylor, 2006). Correlations ranging between 0,43 and 0,91 were obtained between the four Curve of Learning sessions across the six samples (Taylor, 2006).

Taylor (2006) explains the memory test as follows:

> Immediately after they have finished the COL, testees are administered a Memory and Understanding Test based on COL material. COL problems involve transformations of symbols and determining the meaning of these symbols. While doing COL, testees have access to a Dictionary of this information, but are encouraged to learn as much of the meanings and transformations and are told that they will shortly have to do a test on the content of the Dictionary.

> Testees who internalise more of the information while doing the COL (as opposed to simply looking the material up in the Dictionary) will do better on the Memory and Understanding Test.

From the explanation above, it is clear that the Memory and Understanding score displays an individual's ability to automate responses. The reliabilities of the Memory and Understanding scores were calculated with KR-type estimates. The KR-20 coefficients (with correction applied under the assumption that the item difficulties are normally distributed) ranged between 0,70 and 0,82 over six samples (Taylor, 2006). The total number of items correct in the Memory and Understanding test was used in fitting the structural model.

The total number of all items correct, across all four Curve of Learning sub-tests were used as the operational measurement score in the regression analysis.

3.3.5   JOB COMPETENCY POTENTIAL

Job competency potential was determined by two measures as being used by the South African Police Service (SAPS) in the basic training learning programme. Scores obtained by entry level constables in the Specific Crimes (Y_1) and Statutory Law (Y_2) modules were used as an indication of the level of job competency potential, with the basic premises that a higher score obtained by a candidate indicates a higher level of job competency potential.

The two modules were selected because of the fairly broad distribution and variance in scores and seem to require the creative use of newly obtained knowledge in applied problem solving and also because performance on the examinations set in these two modules to a greater degree than the other modules.

Another concern, purely based on *prima facie* evidence and not on an in-depth investigation into the matter, is that the training institution in question seems to have fallen into the trap of designing evaluations or measurement instruments (i.e. examinations and tests) to merely measure the extent to which students are able to recollect information from memory rather than their ability to creatively use the newly obtained knowledge in problem solving. Differences in scores are, thus, not primarily determined by the extent to which real learning has taken place i.e. the ability of the student to automate responses, through automation of information processing procedures, given familiar problems (maybe previously dealt with or covered in training) and the extent to which a student can transfer knowledge obtained through training to solve similar (but not exactly the same) unique problems, previously dealt with or covered in training.

The argument that it is often impractical or not always possible to design measures in such a, maybe more complicated, but definitely more valid and credible manner, will always be posed. However, such an argument only serves as an 'easy way out'-type of argument and only aggravates the problem that the extent to which real learning took place is not effectively determined in many training institutions, where the main aim should be to ensure that students who qualify through the system are in fact truly competent and ready to

face the action learning challenges posed by the specific job, role or function that the training is aimed at.

Many students who qualify through training institutions are presented to the market as potentially ready, but in fact, if the measures used in the training institution are not valid and credible measures of the competencies needed to eventually perform successfully in the job, then training institutions are presenting candidates to the market who have no or very little real potential to perform in the job. Maybe, this is exactly part of the problem that lies at the core of the inability of South Africa to be a competitive global player as presented in the opening argument of this paper. However, even though this issue is a critical one that needs to be urgently addressed, it is not the purpose of this paper to address it.

However, the validity and credibility of conclusions on the ability of the expanded learning potential structural model to explain variance in learning performance, is highly dependant on evidence that the manifest indicators, in this case the measures used by the SAPS, are indeed valid and reliable measures of job competency potential. Such evidence could not be obtained from the SAPS and it is, therefore, acknowledged that the credibility of conclusions made in this study are severely crippled by this.

In this specific study, and in any study making use of data obtained from training institutions, given the foregoing argument, the measures used by these institutions to measure (Y) the conceptualised final criterion (job competency potential), should be approached with a fair degree of caution. Ideally, valid and reliable measures (Y) of the conceptualised final criterion should rather be designed and applied by the researcher him- or herself. However, given the already formidable magnitude of the study and due to various practical, and financial reasons, this was not possible.

3.4 SAMPLING

It is not always practical or possible to obtain measurements from every subject in a target population (N). The more practical and viable option is to focus on a representative sample (n) of the target population. The extent to which observations can or may be generalised to the target population is a function of the number of subjects in the chosen sample and the representativeness of the sample (SIP, 1998), while the power of inferential statistical tests also depends on sample size (Elmes, Kantowitz & Roediger, 1999; Theron, 2002). Given the nature of the study, the question of sample size should primarily be considered from the perspective of Structural Equation Modelling (SEM). Kelloway (1998) argues that SEM is very much a large sample technique and that tests of model fit are based on the assumption of large samples.

Determining the correct sample size is critical for power analysis purposes, especially the determination of both Type I and Type II errors. The detail concerning power analysis will, however, be discussed at a later stage (refer to section 4.10).

The MacCallum, Browne and Sugawara (1996) tables indicate that a sample size of 221 subjects is required to ensure a 0,80 probability of correctly rejecting an incorrect model with 59[7] degrees of freedom when actual model fit is close (i.e., ε_a=0,05), if the probability of a Type I error in testing the null hypothesis of exact fit (i.e., ε_a=0,0) is fixed at 0,05 [i.e., P(reject H_0: RMSEA=0|RMSEA=0,05)].

The tables further indicate that a sample size of 190 subjects is required to ensure a 0,80 probability of correctly rejecting an incorrect model with 59 degrees of freedom when actual model fit is mediocre (i.e., ε_a=0,08), if the probability of a Type I error in testing the null hypothesis of close fit is fixed at 0,05 (MacCullum *et al.*, 1996) [i.e., P(reject H_0: RMSEA=0,05|RMSEA=0,08)].

---

[7] The proposed learning potential structural model has 35 free parameters. However, only 32 parameters are effectively estimated since the Lisrel fixes one lambda-Y element for each of the three endogenous latent variables to unity (Personal communication, Gerhard Mels, 17 July 2006). The degrees of freedom are therefore the number of unique elements in the observed covariance matrix minus the number of model parameters to be estimated. In this case therefore 91-32=59.

For this study a non-probability sample of 434 new recruits from the South African Police Service Training College in Philippi, Cape Town was used. Even though the size of the selected sample is quite satisfactory, as in any study making use of a non-probability sample of the target population, caution should be taken when making generalisations of findings to the target population. Moreover caution should also be exercised when considering the generalizability of the study findings to other populations of learners.

3.5   MISSING VALUES

Multivariate data sets more often than not contain missing values, which may result from non-responses or absenteeism. (Mels, 2003). The issue of missing values had to be addressed in this study before the data could be analysed.

The missing data values were in part dealt with in the traditional way through a method of list wise deletion by only including cases in the analysis that had values on the two criterion measures Specific Crimes (Y_1) and Statutory Law (Y_2). The problem, however, as Mels (2003) warns, is that the dataset tends to be substantially reduced. In this case list-wise deletion resulted in an effective sample size of 130 subjects. Moreover the subset of 130 cases still included a limited number of missing values on the predictor variables.

To solve the remaining missing values problem the Multiple Imputation (MI) and Full Information Maximum Likelihood (FIML) procedures available in LISREL 8.54 (Jöreskog & Sörbom, 2003), could have been used.

Even though the Full Information Maximum Likelihood (FIML) estimation procedure is the more efficient method of the two (Du Toit & Mels, 2002; Mels, 2003), no separate imputed data set is created which thus would have prevented item and dimensionality analyses and the formation of item parcels (Du Toit & Du Toit, 2001; Mels, 2003), which is a requirement in this study. Another reason for not using the Full Information Maximum Likelihood (FIML) estimation procedure is the fact that FIML assume that the values are missing at random and that the observed variables are continuous and follow a multivariate

normal distribution (Du Toit & Du Toit, 2001). However, the variables most probably do not follow a multivariate normal distribution. Moreover, due to the missing value problem it is not possible to test the multivariate normality assumption. In addition the individual items could probably not legitimately be considered continuous variables but rather should be regarded as ordinal variables.

The most satisfactory solution of the two would thus have been to use a multiple imputation procedure (Du Toit & Du Toit, 2001; Mels, 2003). The biggest advantage of both the two multiple imputation procedures available in LISREL 8.54 is that estimates of missing values are derived for all cases in the initial sample (i.e., no cases with missing values are deleted) and the full data set is available for subsequent item and dimensionality analyses, and the formation of item parcels (Du Toit & Du Toit, 2001; Mels, 2003).

The problem, however, was that the multiple imputation procedures available in LISREL 8.54, assume that the values are missing at random and that the observed variables are continuous and follow a multivariate normal distribution (Du Toit & Du Toit, 2001). Especially the latter two prerequisites were seen as problematic in this case.

Therefore, to solve the remaining missing value problem, imputation by matching was used, specifically because the assumption of multivariate normality was not met.

Imputation by matching refers to a process of substituting real values for missing values. The substitute values replaced for a case are derived from one or more other cases that have a similar response pattern over a set of matching variables (Jöreskog & Sörbom, 1996a).

The ideal is to use matching variables that will not be utilized in the confirmatory factor analysis. This, however, was not possible in this case. Four variables, Speed and Flexibility (two measures of information processing capacity) and Curve of Learning Total and Curve of Learning Adjusted (two measures of Automatization) were identified, through statistics output options in PRELIS, as being least plagued by missing values and, thus, served as

matching variables. PRELIS succeeded in successfully imputing missing values for 119 cases.

The race and age frequency distributions across the final effective sample of 119 respondents are displayed in Table 3.1 and Table 3.2 below:

Table 3.1

Race Frequency Distribution Across The Sample Population.

Race

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		5	4.2	4.2	4.2
	African	61	51.3	51.3	55.5
	Coloured	50	42.0	42.0	97.5
	European	3	2.5	2.5	100.0
	Total	119	100.0	100.0	

Table 3.2

Age Statistics And Frequency Distribution Across Sample Population.

Age

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	18	1	.8	.8	.8
	19	2	1.7	1.7	2.5
	20	6	5.0	5.0	7.6
	21	9	7.6	7.6	15.1
	22	13	10.9	10.9	26.1
	23	10	8.4	8.4	34.5
	24	14	11.8	11.8	46.2
	25	8	6.7	6.7	52.9
	26	17	14.3	14.3	67.2
	27	14	11.8	11.8	79.0
	28	10	8.4	8.4	87.4
	29	11	9.2	9.2	96.6
	30	1	.8	.8	97.5
	31	2	1.7	1.7	99.2
	33	1	.8	.8	100.0
	Total	119	100.0	100.0	

3.6   RESEARCH DESIGN

To empirically investigate the hypothesis that variance in learning performance can be explained in terms of the learning competencies and learning competency potential as measured by the APIL test battery, a strategy was required that will provide *unambiguous* empirical evidence in terms of which to evaluate the operational hypotheses. This empirical evidence providing strategy is known as the research design (Kerlinger, 1973; Theron, 2002).

The research design is the plan and structure of the investigation which is set up to firstly, procure answers to the research question and secondly, to control variance (Kerlinger, 1973). The ability of the research design to maximise systematic variance, minimise error variance and control extraneous variance (Kerlinger, 1973; Kerlinger & Lee, 2000) will determine the unambiguousness with which the empirical evidence can be interpreted for or against the learning potential hypothesis.

An ex post facto correlational design was used in this study. Ex post facto research is a form of systematic empirical enquiry in which the researcher does not have direct control over the independent variables. Their manifestations have, thus, either already occurred or they are not inherently manipulable (Kerlinger & Lee, 2000). Inferences about the hypothesised relation between the latent variables ξ_j and η_i are made from concomitant variation in independent and dependent variables (Kerlinger & Lee, 2000).

Kerlinger and Lee (2000) mention the following three major limitations regarding ex post facto research. The limitations are:

- ❑ The inability to manipulate independent variables;
- ❑ The lack of power to randomise; and
- ❑ The risk of improper interpretation.

Ex post facto research lacks control (especially when compared to experimental designs) and erroneous interpretations may originate due to the possibility of more than one explanation for the obtained difference or correlation (Kerlinger & Lee, 2000). Thus, Kerlinger and Lee (2000) warn that results obtained from ex post facto research should be treated with caution. Although it is possible to enhance the degree of control achieved in any given ex post facto design (Kerlinger & Lee, 2000), these extensions to the basic correlational design could not be practically utilized in this particular study.

The objective of this study was to establish whether specific causal linkages exist between learning competency potential, the learning competencies and learning performance as proposed by the expanded learning potential structural model(Equation 12). The ex post facto nature of the research design, however, precluded the drawing of causal inferences from significant path coefficients.

If unsatisfactory absolute model fit would be found (Byrne, 1989; Kelloway, 1998) the conclusion would inevitably follow that the comprehensive model does not provide an acceptable explanation for the observed covariance matrix and thus that the structural model expressed as equation 12 does not satisfactorily explain variance in learning performance (assuming acceptable measurement model fit).

The converse, however, is not true. If the covariance matrix derived from the estimated model parameters closely correspond to the observed covariance matrix it would not imply that the processes postulated by the structural model necessarily *must have* produced the observed covariance matrix and that the theory underlying the APIL test battery must be valid. A high degree of fit between the observed and estimated covariance matrices would only imply that the processes portrayed in the structural model provide one plausible explanation for the observed covariance/correlation matrix. The structural model could, under such an outcome, be considered corroborated in the sense that it survived an opportunity to be refuted (Popper, 1972).

3.7 HYPOTHESES

In accordance with the literature study, the proposed research problems and the derived structural model (expressed as equation 12) the following substantive research hypotheses and associated statistical hypotheses are formulated (See Appendices A and B).

Hypothesis 1a:

The structural model expressed as equation 12 exactly fits the data in the parameter. There is therefore no significant discrepancy between the reproduced covariance matrix implied by the model ($\Sigma(\Theta)$; see Figure 3.1) and the observed population covariance matrix (Σ).

$$H_{01a}: \Sigma = \Sigma(\Theta)$$
$$H_{a1a}: \Sigma \neq \Sigma(\Theta)$$

The exact fit hypothesis could alternatively be formulated as:

$$H_{01a}: \text{RMSEA}=0$$
$$H_{a1a}: \text{RMSEA}>0$$

Hypothesis 1b:

The structural model expressed as equation 12 fits the data in the parameter closely. The reproduced covariance matrix implied by the model ($\Sigma(\Theta)$) closely approximates the observed population covariance matrix (Σ).

$$H_{01b}: \text{RMSEA} \leq 0,05$$
$$H_{a1b}: \text{RMSEA} > 0,05$$

Hypothesis 2:

Abstract thinking capacity (ξ_1) has a statistically significant positive effect on transfer of knowledge (η_1).

$$H_{02}: \gamma_{11} = 0$$
$$H_{a2}: \gamma_{11} > 0$$

Hypotheses 3:

Information processing capacity (ξ_2) has a statistically significant positive effect on automatization (η_2).

$$H_{03}: \gamma_{22} = 0$$
$$H_{a3}: \gamma_{22} > 0$$

Hypotheses 4:

The extent to which transfer of knowledge (η_1) occurs is positively determined by the extent to which automatization occurs (η_2).

$$H_{04}: \beta_{12} = 0$$
$$H_{a4}: \beta_{12} > 0$$

Hypotheses 5:

Transfer of knowledge (η_1) has a statistically significant positive effect on job competency potential targeted by the affirmative training intervention (η_3).

$$H_{05}: \beta_{31} = 0$$
$$H_{a5}: \beta_{31} > 0$$

Hypotheses 6:

Automatization (η_2) has a statistically significant positive effect on job competency potential targeted by the affirmative training intervention (η_3).

$$H_{06}: \beta_{32} = 0$$
$$H_{a6}: \beta_{32} > 0$$

Hypothesis 7:

The influence of abstract thinking capacity (ξ_1) on the job competencies targeted by the training intervention (η_3) is mediated by transfer of knowledge (η_1).

$$H_{07}: \gamma_{11}\beta_{31} = 0$$
$$H_{a7}: \gamma_{11}\beta_{31} > 0$$

Hypothesis 8:

The influence of information processing capacity (ξ_2) on the job competencies targeted by the training intervention (η_3) is mediated by automatization (η_2).

$$H_{08}: \gamma_{22}\beta_{32} = 0$$
$$H_{a8}: \gamma_{22}\beta_{32} > 0$$

Hypothesis 9:

The dynamic measures of the two latent learning competencies, transfer of knowledge (X_3) and automatization (X_4), each explain unique variance in a composite measure of the job competency potential targeted by the affirmative training intervention (Y).

$$H_{09a}: \beta[X_3] = 0|\beta[X_4] \neq 0$$
$$H_{a9a}: \beta[X_3] > 0|\beta[X_4] \neq 0$$
$$H_{09b}: \beta[X_4] = 0|\beta[X_3] \neq 0$$
$$H_{a9b}: \beta[X_4] > 0|\beta[X_3] \neq 0$$

Hypothesis10[8]:

The static measures of the two latent learning dispositions (X_1 & X_2) explain variance in a composite measure of the job competency potential targeted by the affirmative training intervention (Y) when added to a model already containing dynamic measures of the two latent learning competencies (X_3 & X_4).

$$H_{010}: \beta[X_1] = \beta[X_2] = 0|\beta[X_3] \neq 0, \beta[X_4] \neq 0$$
$$H_{a10}: \beta[X_1] \neq \beta[X_2] \neq 0|\beta[X_3] \neq 0, \beta[X_4] \neq 0$$

Hypothesis 11a[9]:

Abstract reasoning capacity (X_1) produces unique variance in job competency potential targeted by the affirmative training intervention (Y) not attributable to transfer of knowledge (X_3), automatization (X_4) or information processing capacity (X_2).

$$H_{011a}: \beta[X_1] = 0 \mid \beta[X_2] \neq 0, \beta[X_3] \neq 0, \beta[X_4] \neq 0$$

---

[8] H_{010} will only be tested if either H_{09a} or H_{09b} or both are rejected. If only H_{09a} or H_{09b} would be rejected H_{010} would be amended accordingly.
[9] The family of null hypotheses under hypothesis 11 will only be tested if H_{010} is rejected. Hypotheses might be amended depending on the results obtained in the preceding hypothesis tests.

$$H_{a11a}: \beta[X_1] \neq 0 \mid \beta[X_2] \neq 0, \beta[X_3] \neq 0, \beta[X_4] \neq 0$$

Hypothesis 11b:

Transfer of knowledge (X_3) produces unique variance in job competency potential targeted by the affirmative training intervention (Y) not attributable to abstract reasoning capacity (X_1) automatization (X_4) or information processing capacity (X_2).

$$H_{011b}: \beta[X_3] = 0 \mid \beta[X_1] \neq 0, \beta[X_2] \neq 0, \beta[X_4] \neq 0$$
$$H_{a11b}: \beta[X_3] \neq 0 \mid \beta[X_1] \neq 0, \beta[X_2] \neq 0, \beta[X_4] \neq 0$$

Hypothesis 11c:

Information processing capacity (X_2) produces unique variance in learning performance (Y) not attributable to transfer of knowledge (X_3), abstract reasoning capacity (X_1) or automatization (X_4).

$$H_{011c}: \beta[X_2] = 0 \mid \beta[X_1] \neq 0, \beta[X_3] \neq 0, \beta[X_4] \neq 0$$
$$H_{a11c}: \beta[X_2] \neq 0 \mid \beta[X_1] \neq 0, \beta[X_3] \neq 0, \beta[X_4] \neq 0$$

Hypothesis 11d:

Automatization (X_4) produces unique variance in learning performance (Y) not attributable to transfer of knowledge (X_3), abstract reasoning capacity (X_1) or information processing capacity (X_2).

$$H_{011d}: \beta[X_4] = 0 \mid \beta[X_1] \neq 0, \beta[X_2] \neq 0, \beta[X_3] \neq 0$$
$$H_{a11d}: \beta[X_4] \neq 0 \mid \beta[X_1] \neq 0, \beta[X_2] \neq 0, \beta[X_3] \neq 0$$

3.8 STATISTICAL ANALYSIS TECHNIQUES AND STATISTICAL PACKAGE

3.8.1 ITEM- AND DIMENSIONALITY ANALYSIS

Item analysis is a technique that is generally used to identify and eliminate items from a measure that do not contribute to an internally consistent description of the sub-scale in question. Therefore, high validity and reliability can be built into tests in advance through

item analysis, thus, improving tests through the selection, substitution, or revision of items (Anastasi & Urbina, 1997).

The architecture of the APIL test battery reflects the intention to construct essentially one-dimensional sets of items to reflect variance in each of the latent variables collectively comprising learning potential. The items are meant to function as homogenous stimulus sets to which raters respond with behaviour, which is primarily a relatively uncontaminated expression of a specific underlying latent variable. The same argument would apply to multiple indicator learning performance measures. The objective of dimensionality analysis is to confirm the uni-dimensionality of each sub-scale and to remove items with inadequate factor loadings or to split heterogeneous sub-scales into two or more homogeneous subsets of items (and revise the structural model).

Unfortunately the nature of most of the individual items of the measures used in this study did not allow item- and dimensionality analysis to be done. It is, nonetheless, believed that Taylor (1989, 1994, 1997) did conduct item- and dimensionality analysis when he originally developed the APIL test battery although no detailed account of the results could be traced. For the purpose of this study, only the individual items as obtained through the Concept Formation Test allowed for item- and dimensionality analysis.

3.8.2   STRUCTURAL EQUATION MODELLING (SEM)

Structural equation modelling (SEM) was used as the statistical analysis technique to test the proposed model's absolute fit. Kelloway (1998) gives three arguments in favour of SEM as an analysis technique

Firstly, Kelloway (1998) argues that in the social sciences, measures are often used to represent constructs. SEM allows the researcher to determine how well these measures reflect the intended constructs. Kelloway (1998, p. 2) argues:

> Confirmatory factor analysis, an application of structural equation modelling, is both more rigorous and more parsimonious than the "more traditional" techniques of exploratory factor analysis.

Furthermore, factor analysis as *per* SEM is based on the testing of hypotheses, with explicit tests of both the overall quality of the factor solution and the specific parameters (e.g. factor loadings) composing the model (Kelloway, 1998).

Secondly, Kelloway (1998) argues that social scientists are mostly interested in the question of prediction. He argues that predictive models have become very complex and that SEM allows the testing and specification of these more complex "path" models as an entity in addition to testing the components comprising the model.

Lastly, Kelloway (1998) argues that SEM provides a flexible, yet powerful, method by which the quality of measurement can be taken into account when evaluating the predictive relationships existing amongst the underlying latent variables. Unlike more traditional analysis techniques, SEM permits estimates of the strength of the relationship existing between latent variables unattenuated by measurement error.

Also in favour of SEM Bollen and Long (1993) writes:
> Structural equation models (SEMs) are a well-known component of the methodological arsenal of social sciences. Much of their attractiveness stems from their generality. Like econometric methods, SEMs allow consideration of simultaneous equations with many endogenous variables. Unlike most econometric methods, SEMs allow measurement error in the exogenous and endogenous variables. As with factor analysis developed in psychometrics and related procedures in sociometrics, SEMs permit multiple indicators of latent constructs and estimation of reliability and validity. In addition, SEMs allow more general measurement models than traditional factor-analytic structures and enable the researcher to specify structural relationships among the latent variables. Thus structural equation models are a synthesis of procedures developed in econometrics, sociometrics, and psychometrics. (p. 1)

The arguments provided by Kelloway (1998) and Bollen and Long (1993) serve as the reason behind the selection of SEM as the statistical analysis technique used in this study. The following five distinct but interrelated steps, which characterise most applications of SEM, were adhered to (Bollen and Long, 1993; Diamantopoulos & Siguaw, 2000):

- ❑ Model specification;
- ❑ Evaluation of model identification;
- ❑ Estimation of model parameters;
- ❑ Testing model fit; and
- ❑ Model respecification.

Model specification involves describing the nature and number of model parameters to be estimated in the initial comprehensive model. It would also include the construction of a comprehensive path diagram depicting the substantive hypotheses and measurement system. Evaluation of model identification involves an examination of the data to determine whether it is possible to find unique values for the freed parameters of the specified model. Once a model is identified the researcher has to select an estimation technique. This process is often determined by the nature and distributional properties of the variables that are being analysed. After parameter estimates are obtained, the researcher has to test whether the model is consistent with the data, in other words, does the model fit the data. If the model does indeed fit the data, the process can stop. However, the fit of the model can more than often be improved through respesification of the model either by fixing currently free parameters, constraining parameters or freeing additional parameters, whereupon steps 2-5 can be repeated (Bollen & Long, 1993).

Ideally, should satisfactory model fit be achieved, the model should be cross-validated by fitting the model with parameters constrained to the estimated values found during the initial study on a fresh data set from the same population. This aspect is especially important if the initial data set has been used to modify the original model (Diamantopoulos & Siguaw, 2000). Cross-validation did not form part of this study. Another, independent study will investigate the stability of the model parameter estimates by fitting the model to a validation sample.

The statistical package that was used in the analysis is LISREL 8.54 for Windows (Du Toit & Du Toit, 2001; Jöreskog & Sörbom, 2003; Mels, 2003).

3.8.2.1 Specification of the Full LISREL Model

In its most general form, the LISREL model consists of a set of linear structural and measurement equations. The variables in the model are either directly observed or latent (theoretical) variables. The general assumption is that there is a causal structure among the latent variables and that the observed variables are indicators of the latent variables (Jöreskog, 2003).The comprehensive model is made up of two parts, the measurement model and the structural model. The measurement model specifies how latent variables are indicated by the observed variables; in other words it describes the reliabilities and validities (measurement properties) of the observed variables. The structural equation model, on the other hand, specifies the causal relationships among the latent variables, describes the causal effects, and assigns the explained variance (Jöreskog, 2003).

Structural model

The revised basic learning potential structural model is schematically depicted in Figure 3.1.

Where:
ξ_1 = Abstract reasoning capacity
ξ_2 = Information processing capacity

η_1 = Transfer of knowledge
η_2 = Automatization
η_3 = Job competency potential targeted by the affirmative training intervention
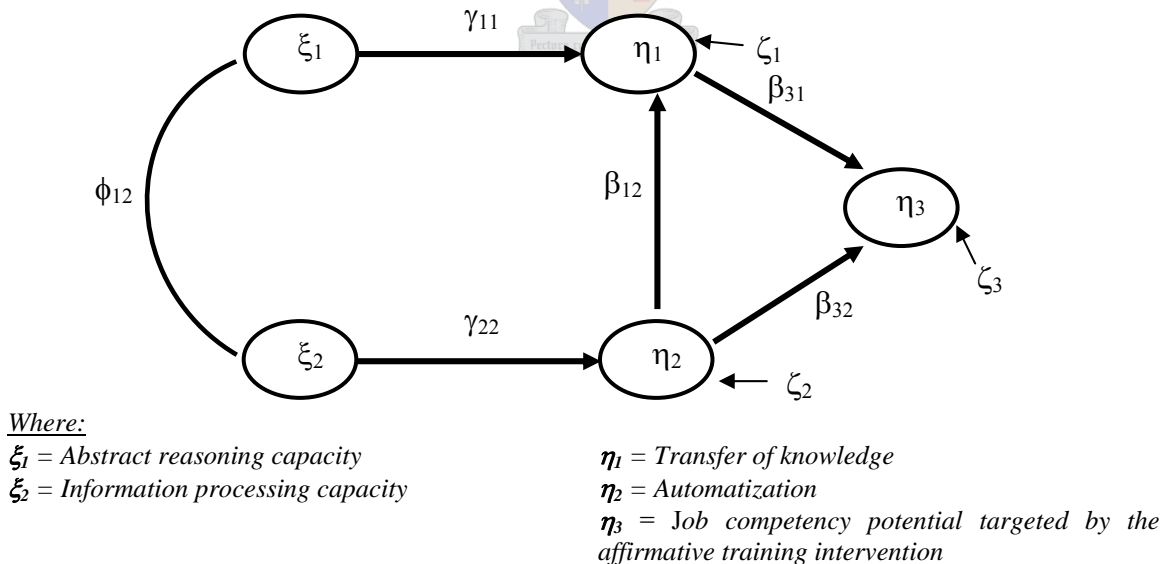
Figure 3.1

Graphical Portrayal Of Fitted Learning Potential Structural Model

The proposed structural model, which serves as the basis of this study, can be expressed as a set of structural equations representing the research hypotheses that will be investigated.

$$\eta_1 = \gamma_{11}\xi_1 + \beta_{12}\eta_2 + \zeta_1 \text{-- ---------------------------(13)}$$

$$\eta_2 = \gamma_{22}\xi_2 + \zeta_2 \text{--- (14)}$$

$$\eta_3 = \beta_{31}\eta_1 + \beta_{32}\eta_2 + \zeta_3 \text{---(15)}$$

The structural model can also be portrayed mathematically in terms of a series of matrices. The structural model is defined by the following three matrices and three vectors:

- ❑ A 3 x 2 Γ (gamma)- matrix of path/ regression coefficients γ_{ij} describing the strength of the regression of η_i on ξ_j) in the structural model;

- ❑ A 3 x 3 square B (beta)-matrix of regression/ path coefficients (β_{ij}) describing the strength of the regression of η_i on η_j in the structural model;

- ❑ A 2 x 2 symmetrical matrix Φ (phi)-matrix of variance and covariance terms describing the variance in (Φ_{ii}) and covariance between (Φ_{ij}) the exogenous latent variables ξ_i and ξ_j (it is again assumed that the exogenous latent variables are correlated and thus all off diagonal elements in Φ will be set free to be estimated);;

- ❑ A 3x3 symmetrical Ψ (psi) matrix of variance and covariance terms describing the variance in (ψ_{ii}) and covariance between (ψ_{ij}) the structural error terms ζ_i and ζ_j it is assumed that the structural error terms are uncorrelated and thus that Ψ is a diagonal matrix);

- ❑ A 2 x 1 ξ (ksi) column vector of exogenous latent variables;

- ❑ A 3 x 1 η (eta) column vector of endogenous latent variables;

- ❑ A 3 x 1 ζ (zeta) column vector of residual error terms.

More specifically, the hypothesised causal relationships depicted in Figure 3.1 can again be expressed in matrix form as equation 12 and more succinctly as equation16.

$$
\begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix} = \begin{pmatrix} 0 & \beta_{12} & 0 \\ 0 & 0 & 0 \\ \beta_{31} & \beta_{32} & 0 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix} + \begin{pmatrix} \gamma_{11} & 0 \\ 0 & \gamma_{22} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{pmatrix} \text{------------------(12)}
$$

$$\eta = B\eta + \Gamma\xi + \zeta \text{--(16)}$$

Endogenous measurement model

The endogenous measurement model (depicted in Appendix A) can be expressed in terms of the following set of measurement equations:

$$y_1 = \lambda^y_{11}\eta_1 + \varepsilon_1 \text{---(17)}$$

$$y_2 = \lambda^y_{21}\eta_1 + \varepsilon_2 \text{---(18)}$$

$$y_3 = \lambda^y_{31}\eta_1 + \varepsilon_3 \text{---(19)}$$

$$y_4 = \lambda^y_{41}\eta_1 + \varepsilon_4 \text{---(20)}$$

$$y_5 = \lambda^y_{52}\eta_2 + \varepsilon_5 \text{---(21)}$$

$$y_6 = \lambda^y_{62}\eta_2 + \varepsilon_6 \text{---(22)}$$

$$y_7 = \lambda^y_{73}\eta_3 + \varepsilon_7 \text{---(23)}$$

$$y_8 = \lambda^y_{83}\eta_3 + \varepsilon_8 \text{---(24)}$$

More specifically, the set of endogenous measurement equations can be summarized as matrix equations 25 and 26. The loadings (λ^y_{ji}) of the observed Y_j variables on the endogenous latent variables (η_i), as depicted in Appendix A are reflected in Λ_y.

$$
\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \end{pmatrix}
=
\begin{pmatrix} \lambda^y_{11} & 0 & 0 \\ \lambda^y_{21} & 0 & 0 \\ \lambda^y_{31} & 0 & 0 \\ \lambda^y_{41} & 0 & 0 \\ 0 & \lambda^y_{52} & 0 \\ 0 & \lambda^y_{62} & 0 \\ 0 & 0 & \lambda^y_{73} \\ 0 & 0 & \lambda^y_{83} \end{pmatrix}
\begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix}
+
\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \end{pmatrix}
\text{--------------------------(25)}
$$

$$Y = \Lambda_y\eta + \varepsilon \text{…………………………………………………………………….(26)}$$

Exogenous measurement model

The exogenous measurement model (schematically depicted in Appendix A) can be expressed as the following set of equations:

$$x_1 = \lambda^x_{11}\xi_1 + \delta_1 \text{---(27)}$$

$$x_2 = \lambda^x_{21}\xi_1 + \delta_2 \text{---(28)}$$

$$x_3 = \lambda^x_{32}\xi_2 + \delta_3 \text{---(29)}$$

$$x_4 = \lambda^x_{42}\xi_2 + \delta_4 \text{---(30)}$$
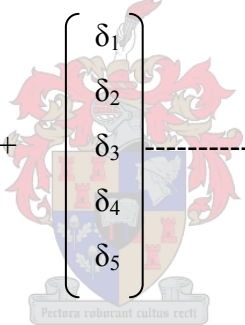
$$x_5 = \lambda^x_{52}\xi_2 + \delta_5 \text{---(31)}$$

More specifically, the set of exogenous measurement equations can be summarized as matrix equation 32 and 33. The loadings (λ^x_{ji}) of the observed X_j variables on the exogenous latent variables (ξ_i), as depicted in Appendix A are reflected in Λ_x.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} \lambda^x_{11} & 0 \\ \lambda^x_{21} & 0 \\ 0 & \lambda^x_{32} \\ 0 & \lambda^y_{42} \\ 0 & \lambda^x_{52} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \end{bmatrix} \text{--(32)}$$

$$X = \Lambda_x \xi + \delta \text{ ....(33)}$$

Full LISREL model

The full LISREL model for single samples, for deviations about the means, can be expressed mathematically as the following three equations (Jöreskog, 2003):

❏ The structural model:

$$\eta = B\eta + \Gamma\xi + \zeta \text{--(16)}$$

❏ The measurement model for y:

$$Y = \Lambda_y\eta + \varepsilon \text{--(26)}$$

❏ The measurement model for x:

$$X = \Lambda_x\xi + \delta \text{---(33)}$$

The terms in these models can be defined as follows (Jöreskog, 2003):

- η (eta) is a 3 x 1 random vector of latent dependant or endogenous variables;
- ξ (ksi) is a 2 x 1 random vector of latent independent or exogenous variables;
- Y is a 8 x 1 vector of observed indicators of the dependent latent variables η;
- X is a 5x 1 vector of observed indicators of the independent latent variables ξ;
- ε is a 8 x 1 vector of measurement errors in y;
- δ is a 5 x 1 vector of measurement errors in x;
- Λ_y is a 8 x 3 matrix of regression coefficients of the regression of y on η;
- Λ_x is a 5 x 2 matrix of regression coefficients of the regression of x on ξ;
- Γ (gamma) is a 2 x 2 matrix of path/regression coefficients γ describing the strength of the regression of η_i on ξ_i in the structural model;
- B (beta) is a 3 x 3 symmetrical matrix of regression/path coefficients (β) describing the strength of the regression of η_i on η_i in the structural model;
- 3 x 1 ζ (zeta) vector of equation errors (random disturbances) in the structural relationship between η and ξ;
- Φ (phi) is a 2 x 2 symmetrical matrix of variance and covariance terms describing the variance in (Φ_{ii}) and covariance between (Φ_{ij}) the exogenous latent variables ξ_i and ξ_j [Cov (ξ) = Φ(2x2)].

Jöreskog (2003) suggests that the random components in the LISREL model are assumed to satisfy the following minimum assumptions:

- ε is uncorrelated with η;
- δ is uncorrelated with ξ;
- ζ is uncorrelated with ξ; and
- ζ is uncorrelated with ε and δ.

3.8.2.2 Model Identification

Model identification is an important but difficult question that needs to be examined prior to confronting the structural model with data (MacCallum, 1995). The essential issue is whether the nature of the model and the data would permit the determination of unique estimates for the freed parameters in the model. This would be possible if for each free

parameter there would exist at least one algebraic function that expresses that parameter as a function of sample variances/covariance terms (MacCallum, 1995).

Unfortunately, there is no uncomplicated set of necessary and sufficient conditions that if satisfied, would ensure that the model is identified (Diamantopoulos & Siguaw., 2000; MacCallum, 1995). There are, however, two critical and necessary conditions that have to be met. It is firstly essential that a definite scale should be established for each latent variable. The second requirement is that the number of model parameters to be estimated should not exceed the number of unique variance/covariance terms in the observed sample covariance matrix (Diamantopoulos & Siguaw., 2000; MacCallum, 1995). The model depicted in Appendix A satisfies both these necessary conditions for identification. The first requirement will be met by treating each latent variable as a (0; 1) standardized variable (MacCallum, 1995). The number of model parameters that are set free to be estimated (t = 35) are less than the number of non-redundant elements in the observed sample covariance matrix ($[(p+q)(p+q+1)]/2=91$)[10] (Diamantopoulos & Siguaw., 2000). The degrees of freedom for the structural model are therefore 91-35=56.

---
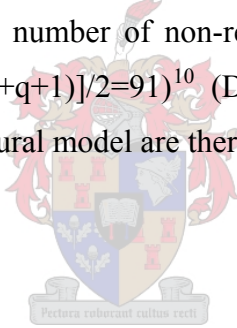
[10] p=the number of y-variables; q=the number of x-variables.

CHAPTER 4
RESEARCH RESULTS

The theoretical model derived from the literature study hypothesizes specific structural relationships between the latent variables. In accordance with the proposed relationships among the latent variables as depicted in Figure 3.1 specific statistical hypotheses were formulated. The purpose of this chapter is to report the results of the statistical analyses aimed at testing the stated null hypotheses.

4.1   PARAMETER ESTIMATION METHOD

The purpose of parameter estimation is to find numerical values for the freed parameters of the structural model that would minimize the difference between the observed and estimated/reproduced sample variance/covariance matrices (Diamantopoulos & Siguaw., 2000). LISREL 8.54 offers a number of different estimation methods (Jöreskog & Sörbom, 1996a; Mels, 2003). The appropriate method to use depends on the nature of the variables to be analysed and the distributional properties of the data.

The composite indicator variables were treated as continuous variables. The analysis of the covariance matrix instead of the polychoric correlation matrix was therefore permissible (Jöreskog & Sörbom, 1996a; Mels, 2003). The default method of estimation when fitting models to continuous data (maximum likelihood), however, requires multivariate normality (Kaplan, 2000). This is also true for the generalized least squares (GLS) and full information maximum likelihood (FIML) methods for structural equation modeling (Mels, 2003). The analysis of continuous non-normal variables in structural equation models can result in incorrect standard errors and chi-square estimates (Du Toit et al., 2001; Mels, 2003). The univariate and multivariate normality of the composite indicator variables was consequently evaluated via PRELIS (Jöreskog & Sörbom, 1996b).

Table 4.1

Test Of Univariate Normality For Continuous Variables Before Normalisation

```
          Skewness          Kurtosis      Skewness and Kurtosis

Variable Z-Score P-Value   Z-Score P-Value   Chi-Square P-Value

 ZABSTR1   2.186   0.029     0.290   0.772        4.865   0.088
 ZABSTR2   1.919   0.055    -0.621   0.535        4.067   0.131
  ZSPEED  -0.983   0.326    -4.112   0.000       17.872   0.000
   ZACCU   1.771   0.077     0.121   0.904        3.150   0.207
   ZFLEX   5.470   0.000     3.704   0.000       43.636   0.000
 ZTRANS1   2.043   0.041   -10.784   0.000      120.470   0.000
 ZTRANS2   3.754   0.000    -0.954   0.340       15.004   0.001
 ZTRANS3   4.061   0.000     1.454   0.146       18.608   0.000
 ZTRANS4   5.256   0.000     2.421   0.015       33.491   0.000
  ZAUTO1   4.171   0.000     3.309   0.001       28.346   0.000
  ZAUTO2   0.701   0.484    -1.054   0.292        1.601   0.449
ZSPECCRI   1.332   0.183     1.127   0.260        3.044   0.218
 ZSTATUT  -4.405   0.000     4.130   0.000       36.466   0.000

Relative Multivariate Kurtosis = 1.132
```

Table 4.2

Test Of Multivariate Normality For Continuous Variables Before Normalisation

```
       Skewness                  Kurtosis            Skewness and Kurtosis

  Value  Z-Score P-Value    Value  Z-Score P-Value     Chi-Square P-Value
 ------  ------- -------   ------  ------- -------     ---------- -------
 46.152  11.895   0.000   220.663   5.487   0.000        171.594   0.000
```

Table 4.1 indicates that eight of the indicator variables failed the test of univariate normality ($p < 0,05$) . Table 4.2 indicates that the null hypothesis that the data follows a multivariate normal distribution also had to be rejected ($\chi^2 = 171,598$; $p < 0,05$). The data was subsequently normalised through PRELIS.

Table 4.3

Test Of Univariate Normality For Continuous Variables After Normalisation

```
          Skewness          Kurtosis      Skewness and Kurtosis

Variable Z-Score P-Value   Z-Score P-Value   Chi-Square P-Value

 ZABSTR1   0.027   0.978    -0.009   0.993        0.001   1.000
 ZABSTR2  -0.012   0.991    -0.094   0.925        0.009   0.996
  ZSPEED   0.008   0.994     0.105   0.917        0.011   0.994
   ZACCU   0.000   1.000     0.116   0.907        0.014   0.993
   ZFLEX   0.000   1.000     0.116   0.907        0.014   0.993
 ZTRANS1   0.238   0.812    -8.710   0.000       75.913   0.000
 ZTRANS2   0.934   0.351    -3.055   0.002       10.204   0.006
```

```
 ZTRANS3   0.710   0.478   -1.041   0.298      1.588   0.452
 ZTRANS4   1.844   0.065   -3.028   0.002     12.568   0.002
  ZAUTO1   0.000   1.000    0.117   0.907      0.014   0.993
  ZAUTO2   0.043   0.966    0.039   0.969      0.003   0.998
ZSPECCRI   0.153   0.878    0.030   0.976      0.024   0.988
 ZSTATUT   0.277   0.781    0.111   0.912      0.089   0.956


Relative Multivariate Kurtosis = 0.985


Test of Multivariate Normality for Continuous Variables
```

Table 4.4

Test Of Multivariate Normality For Continuous Variables After Normalisation

Skewness			Kurtosis			Skewness and Kurtosis	
Value	Z-Score	P-Value	Value	Z-Score	P-Value	Chi-Square	P-Value
------	-------	-------	-------	-------	-------	----------	-------
23.413	0.330	0.741	192.002	0.134	0.894	0.127	0.939

Table 4.3 indicates that the normalisation procedure succeeded in rectifying the univariate normality problem on most indicator variables and Table 4.4 indicates that the normalized data now also pass the test of multivariate normality ($\chi^2 = 0,127$; p > 0,05).

A covariance matrix was subsequently computed from the transformed data set to serve as input for the LISREL analysis. Maximum likelihood estimation was used to estimate the parameters set free in the model. Instead of defining the origin and unit of the latent variable scales in terms of observable reference variables, the latent variables were standardised (Jöreskog & Sörbom, 1996a). All factor loadings of each latent unit performance variable were set free to be estimated, but only with regard to its designated observed variables. All remaining elements of Λ_x were fixed at zero loadings to reflect the assumed factorial simplicity of the indicator variables (Tabachnick and Fidell, 1989). The elements of the covariance/correlation matrix (Φ) and the diagonal elements of the variance/covariance matrix (θ_δ) were treated by default as free.

4.2   ASSESSING THE OVERALL GOODNESS-OF-FIT OF THE MEASUREMENT MODEL

The main aim of SEM is to explain the patterns of covariances observed amongst the observed variables in terms of the relationships hypothesized by the measurement and

93

structural models. The two measurement models describe how the exogenous and endogenous latent variables reflect themselves in manifest variables. To come to valid and credible conclusions on the ability of the structural model to explain the pattern of covariance in learning performance, evidence is needed that the manifest indicators are indeed valid and reliable measures of the latent variables they are linked to. Unless the operational measures can be trusted to validly represent the latent variables they have been tasked to reflect, any assessment of the substantive relations of interest will be problematic in as far as the meaning of poor or good structural model fit would become ambiguous. Therefore the evaluation of the measurement part of the model should precede the detailed evaluation of the structural part of the model (Diamantopoulos & Siguaw, 2000). Instead of fitting the endogenous and exogenous measurement models separately a single measurement model had been fitted to evaluate the success with which the learning potential latent variables had been operationalised.

An admissible final solution of parameter estimates for the APIL measurement model was obtained after 10 iterations.

The χ^2 test statistic is the measure that is traditionally used to test the hypothesis that there is no significant discrepancy between the reproduced covariance matrix implied by the model ($\Sigma(\Theta)$) and the observed population covariance matrix (Σ). The exceedence probability reported by LISREL is the probability of obtaining a χ^2 value larger than the calculated value, given that the exact fit null hypothesis is true. Jöreskog and Sörbom (1993, p. 122) writes:

> Chi-square is a badness of fit measure in the sense that a small chi-square corresponds to good fit and a large chi-square to bad fit. Zero chi-square corresponds to perfect fit.

Thus, contrary to traditional hypothesis testing, a non-significant χ^2 would imply that there is no significant discrepancy between the covariance matrix implied by the measurement model and the observed covariance matrix. Therefore, a non-significant χ^2 would indicate that the model fits the data in that the model can exactly reproduce the population covariance matrix (Kelloway, 1998).

The p-value associated with the Normal Theory χ^2 value in Table 4.5 [0,0029] indicates a significant test statistic (p < 0,05). This suggests that there is a significant discrepancy between the covariance matrix implied by the measurement model and the observed covariance matrix, thus rejecting the exact fit null hypothesis (Kelloway, 1998). The measurement model is, therefore, not able to reproduce the observed covariance matrix (Kelloway, 1998).

The problem; however is that the χ^2 measure is distributed asymptotically as a χ^2 distribution. This causes the frustrating dilemma that just at the point where the distributional assumptions of the test statistic become tenable the statistical power of the test also becomes extremely high. It thus becomes extremely unlikely to obtain the desired insignificant χ^2 statistic in a large sample even when the model fits the empirical data quite well (Hu et al., 1995). In addition Browne and Cudeck (1993) argue:

> In applications of the analysis of covariance structures in the social sciences it is implausible that any model that we use is anything more than an approximation to reality. Since a null hypothesis that a model fits exactly in some population is known a priori to be false, it seems pointless even to try to test whether it is true. (p. 137)

The evaluation of the fit on the basis of the normed chi-square statistic χ^2/df ($\chi^2/df = 1,606$) for the measurement model suggests that the model fits the data well. Ratios less than 2 have, however, been interpreted as indicating over-fitting. Judged by these standards the model could, when viewed optimistically, be seen to fit the data well, or, when viewed somewhat pessimistically, be seen to have been over-fitted. Kelloway (1998), however, comments that the guidelines indicative of good fit (ratios between 2 and 5) have very little justification other than the researcher's personal modelling experience, and does not advise a strong reliance on the normed chi-square.

Numerous alternative descriptive goodness-of-fit measures of model fit have been developed and should be interpreted as well. The version of LISREL used in this study reports 18 indices of model fit, of which four address the question of absolute fit (Kelloway, 1998). A verdict on model fit would be more credible if derived from a synthesis of a variety of measures of fit. A problem complicating the evaluation of model

fit is that quite often the multitude of fit measures are not in agreement thus introducing a certain degree of ambiguity (Byrne, 1989; Diamantopoulos & Siguaw, 2000)

The simplest fix index provided by LISREL is Root Mean Squared Residual (RMR). Kelloway (1998) writes:

> This is the square root of the mean of the squared discrepancies between the implied and observed covariance matrices. The lower bound of the index is 0, and low values are taken to indicate good fit. (p. 27)

The RMR (0,058) value indicates that the measurement model fits the data reasonably well.

A problem with interpreting this index is the fact that it is sensitive to the scale of measurement of the model variables and as a result it is difficult to determine what a low value actually is (Diamantopoulos & Siguaw, 2000; Kelloway, 1998). Therefore, LISREL also provides the standardised RMR, which has a lower bound of 0 and an upper bound of 1. Values less than 0,05 are generally regarded as indicating good fit to the data (Kelloway, 1998; Diamantopoulos & Siguaw, 2000).

The standardized RMR (0,058), as per table 4.5, also indicates that the measurement model fits the data reasonably well.

LISREL also reports the root mean squared error of approximation (RMSEA), which is also based on the analysis of residuals, with smaller values indicating a better fit to the data. The RMSEA is generally regarded as one of the most informative fit indices (Diamantopoulos & Siguaw, 2000). Kelloway (1998) writes:

> Unlike all other fit indices discussed in this chapter, the RMSEA has the important advantage of going beyond the point estimates to the provision of 90% confidence intervals for the point estimate. Moreover, LISREL also provides a test of the significance of the RMSEA by testing whether the value obtained is significantly different from 0,05. (p. 27)

Diamantopoulos & Siguaw (2000) and MacCallum *et al*. (1996) suggest that values <0,05 are indicative of good fit and values of between 0,05 and under 0,08 indicates reasonable

fit, while values between 0,08 and 0,1 indicates mediocre fit and values >0,1 indicates poor fit.

Therefore, according to Diamantopoulos & Siguaw (2000) and MacCallum *et al*. (1996), the RMSEA (0,072) value indicates reasonably good fit. The 90% confidence interval for RMSEA shown in table 4.5 (0.042 - 0.099) contains the critical 0,05 value. A test of the significance of the obtained value is performed by LISREL by testing H_0: RMSEA $\leq 0,05$ against H_a: RMSEA $> 0,05$. Table 4.5 indicates that the obtained RMSEA value of 0,072 is not significantly different from the target value of 0,05 (i.e. the close fit null hypothesis is not rejected; $p > 0,05$) and since the confidence interval does include the target value of 0,05, a good fit seems to have been achieved.

The goodness-of-fit index (GFI) measures are 'based on a ratio of the sum of the squared discrepancies to the observed variances (for generalised least squares, the maximum likelihood version is somewhat more complicated)' (Kelloway, 1998, p. 27). The GFI ranges from 0 to 1, with values exceeding 0,9 indicating good fit to the data (Kelloway, 1998). However, Kelloway (1998) warns:

> It should be noted that this guideline is based on experience. Like many of the fit indices that will be presented, the GFI has no known sampling distribution. As a result, "rules" about when an index indicates a good fit to the data are highly arbitrary and should be treated with caution. (p. 27)

The Goodness of Fit Index (GFI) = 0.90 indicates that the measurement model fits the data reasonably well.

The adjusted GFI (AGFI), adjusts the GFI for degrees of freedom in the model (Diamantopoulos & Siguaw, 2000; Kelloway, 1998). The AGFI also ranges from 0 to 1, with values above 0,9 indicating a good fit to the data (Kelloway, 1998). Therefore, the Adjusted Goodness of Fit Index (AGFI) = 0.83 indicates reasonable, but not good fit.

After interpreting all the fit indices, the conclusion would have to be drawn that the measurement model fit the data reasonably well, but not perfectly.

To ensure a thorough assessment of fit and especially because the overall measures of fit indicates that the measurement model fits the data only reasonably well, it is necessary to investigate the standardised residuals and modification indices to further determine the success with which the model explains the observed covariances amongst the manifest variables (Jöreskog & Sörbom, 1993).

Table 4.5
Goodness-Of-Fit Statistics For The Measurement Model

Degrees of Freedom	55
Minimum Fit Function Chi-Square =	92.10 (P = 0.0013)
Normal Theory Weighted Least Squares Chi-Square =	88.34 (P = 0.0029)
Estimated Non-centrality Parameter (NCP)	33.34
90 Percent Confidence Interval for NCP	(11.56 ; 63.04)
Minimum Fit Function Value	0.78
Population Discrepancy Function Value (F0)	0.28
90 Percent Confidence Interval for F0	(0.098 ; 0.53)
Root Mean Square Error of Approximation (RMSEA)	0.072
90 Percent Confidence Interval for RMSEA	(0.042 ; 0.099)
P-Value for Test of Close Fit (RMSEA < 0.05)	0.10
Expected Cross-Validation Index (ECVI)	1.36
90 Percent Confidence Interval for ECVI	(1.17 ; 1.61)
ECVI for Saturated Model	1.54
ECVI for Independence Model	12.97
Chi-Square for Independence Model with 78 Degrees of Freedom	1505.02
Independence AIC	1531.02
Model AIC	160.34
Saturated AIC	182.00
Independence CAIC	1580.15
Model CAIC	296.39
Saturated CAIC	525.90
Normed Fit Index (NFI)	0.94
Non-Normed Fit Index (NNFI)	0.96
Parsimony Normed Fit Index (PNFI)	0.66
Comparative Fit Index (CFI)	0.97
Incremental Fit Index (IFI)	0.97
Relative Fit Index (RFI)	0.91
Critical N (CN)	106.44
Root Mean Square Residual (RMR)	0.058
Standardized RMR	0.058
Goodness of Fit Index (GFI)	0.90
Adjusted Goodness of Fit Index (AGFI)	0.83
Parsimony Goodness of Fit Index (PGFI)	0.54

4.3 EXAMINATION OF MEASUREMENT MODEL RESIDUALS

Residuals refer to the differences between corresponding cells in the observed and fitted covariance/correlation matrices (Jöreskog & Sörbom, 1993). Residuals, and especially standardised residuals, provide diagnostic information on sources of lack of fit in models (Jöreskog & Sörbom, 1993; Kelloway, 1998). Jöreskog & Sörbom (1993) explain that a standardised residual refers to a residual that is divided by its estimated standard error. The standardized residuals are depicted in Table 4.6.

Table 4.6

Standardized Residuals

	ZABSTR1	ZABSTR2	ZSPEED	ZACCU	ZFLEX	ZTRANS1
ZABSTR1	- -					
ZABSTR2	- -	- -				
ZSPEED	1.17	1.05	- -			
ZACCU	1.00	-0.74	-2.87	- -		
ZFLEX	1.05	-1.34	-1.42	4.26	- -	
ZTRANS1	0.39	-1.43	1.69	0.34	-1.08	- -
ZTRANS2	0.77	-0.26	1.10	-1.26	-2.33	1.55
ZTRANS3	0.44	0.55	2.06	-0.30	-1.30	-1.20
ZTRANS4	0.57	0.95	2.25	1.62	2.21	-1.66
ZAUTO1	-0.55	0.18	5.58	-2.58	-1.22	0.30
ZAUTO2	-2.97	1.36	0.52	0.68	-1.68	1.05
ZSPECCRI	-0.03	0.75	0.37	0.22	-0.65	1.27
ZSTATUT	-0.14	-0.98	-0.06	0.15	0.08	-0.49

Standardized Residuals

	ZTRANS2	ZTRANS3	ZTRANS4	ZAUTO1	ZAUTO2	ZSPECCRI
ZTRANS2	- -					
ZTRANS3	0.67	- -				
ZTRANS4	-0.07	-0.17	- -			
ZAUTO1	-1.69	0.17	2.34	- -		
ZAUTO2	-1.30	0.06	1.36	- -	- -	
ZSPECCRI	-1.22	0.89	-1.45	-0.70	0.43	- -
ZSTATUT	0.07	-0.26	0.72	0.86	-0.14	- -

Standardized Residuals

	ZSTATUT
ZSTATUT	- -

Standardised residuals can be interpreted as z-scores (i.e. number of standard deviations above or below the mean). Standardised residuals are considered to be large if they exceed +2.58 or –2.58 (Diamantopoulos & Siguaw, 2000). The specific covariance terms that were

poorly estimated as judged by this criterion are listed in Table 4.7. A large positive residual would indicate that the model underestimates the covariance between two variables, while a large negative residual would indicate that the model overestimates the covariance between variables. Underestimation indicates that the model needs to be modified by adding additional explanatory paths, which could better account for the covariance between the variables. On the other hand, if the model overestimates the covariance between the variables, the model should be modified by trimming paths that are associated with the particular covariance term (Jöreskog & Sörbom, 1993).

Table 4.7
Summary Statistics for Standardized Residuals

Smallest Standardized Residual = -2.97
Median Standardized Residual = 0.00
Largest Standardized Residual = 5.58
Largest Negative Standardized Residuals
Residual for ZACCU and ZSPEED -2.87
Residual for ZAUTO1 and ZACCU -2.58
Residual for ZAUTO2 and ZABSTR1 -2.97
Largest Positive Standardized Residuals
Residual for ZFLEX and ZACCU 4.26
Residual for ZAUTO1 and ZSPEED 5.58

The two large positive residuals (> 2,58) and three large negative residuals (< -2,58) in Table 4.7 indicates five observed covariance terms in the observed sample covariance matrix (out of 78 covariance terms) being poorly estimated by the derived model parameter estimates. This would indicate reasonable model fit.

Jöreskog & Sörbom (1993) state that all the standardised residuals may be examined collectively in a stem-and-leaf plot and a Q-plot. The stem-and-leaf plot is depicted in Figure 4.1. A good model would be characterised by a stem-and-leaf plot in which the residuals are distributed approximately symmetrical around zero. An excess of residuals on the positive or negative side would indicate that the covariance terms are systematically under- or overestimated.

```
 - 3|0
 - 2|963
 - 1|777544333322210
 - 0|77765333211110000000000000000000
   0|11122223344445667778999
   1|000112344567
   2|1223
   3|
   4|3
   5|6
```

Figure 4.1

Stem-And-Leaf Plot Of Standardized Residuals

From the stem-and-leaf plot depicted in Figure 4.1, the distribution of standardised residuals appears only slightly positively skewed, but not overly so. This indicates that there is a slightly stronger tendency for the model to overestimate the observed covariance terms.

The Q-plot is depicted in Figure 4.2. When interpreting the Q-plot it is important to note whether the data points fall on the 45-degree reference line or not. If the points fall on the 45-degree reference line, it would be indicative of a good model fit. (Jöreskog & Sörbom, 1993). To the extent that the data points swivel away from the 45-degree reference line the model fit is less than satisfactory.
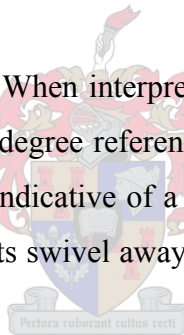
The Q-plot in Figure 4.2 clearly indicates less than perfect model fit because the standardised residuals for all pairs of observed variables tend to deviate from the 45^0-reference line in both the lower- and upper region of the X-axis. Subsequently, given the examination of the residuals, it is important to also evaluate the model modification indices.
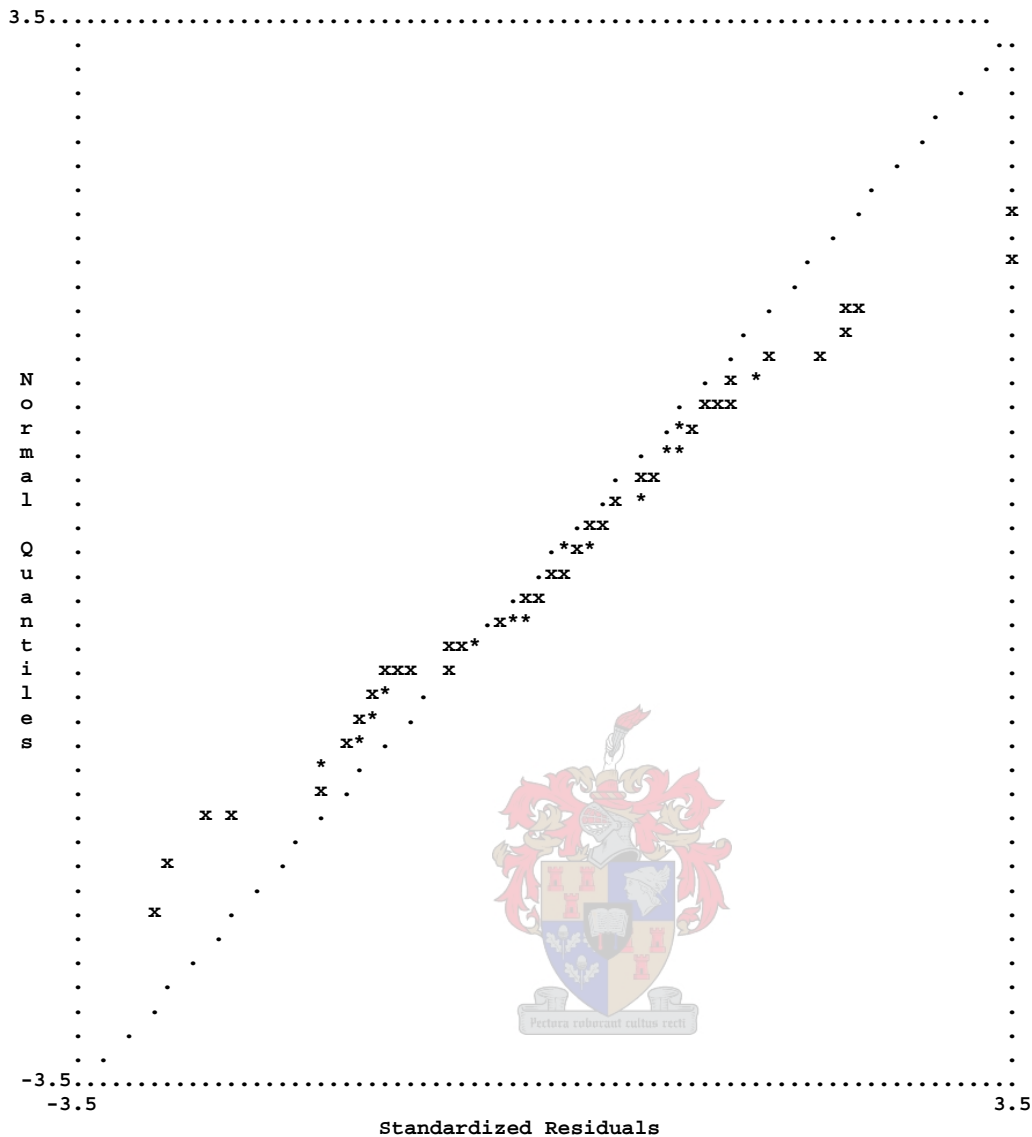
Figure 4.2

Q-plot of Standardized Residuals

4.4   MEASUREMENT MODEL MODIFICATION INDICES

Model modification indices are aimed at answering the question whether any of the currently fixed parameters, when freed in the model, would significantly improve the parsimonious fit of the model. Modification indices (MI) indicate the extent to which the χ^2 fit statistic will decrease if a currently fixed parameter in the model is freed and the model re-estimated (Jöreskog & Sörbom, 1993). Large modification index values (> 6,6349)

would be indicative of parameters that, if set free, would improve the fit of the model significantly (p<0,01) (Diamantopoulos & Siguaw, 2000; Jöreskog & Sörbom, 1993). It is important to note that parameters with high MI values should only be freed if it makes substantive sense to do so (Kelloway, 1998). The expected change for the parameter is the expected value of the parameter if it were freed. The standardised and completed standardised expected changes are the expected values in the standardised and completely standardised solution if the parameter were freed.

Jöreskog & Sörbom (1993) argue that modification indices should be used in the following way in the process of model evaluation and modification:

> If chi-square is large relative to the degrees of freedom, one examines the modification indices and relaxes the parameter with the largest modification index *if this parameter can be interpreted substantively*. If it does not make sense to relax the parameter with the largest modification index, one considers the second largest modification index etc. If the signs of certain parameters are specified a priori, positive or negative, the expected parameter changes associated with the modification indices for these parameters can be used to exclude models with parameters having the wrong sign. (p. 127)

Table 4.8

Lambda-X Modification Indices for Measurement Model

	ABSTRACT	INFOPROC	TRANSFOR	AUTOMAT	LEARNPER
ZABSTR1	– –	2.40	0.65	7.96	0.04
ZABSTR2	– –	2.40	0.65	7.96	0.04
ZSPEED	3.08	– –	10.10	18.47	0.18
ZACCU	0.19	– –	0.32	2.53	0.07
ZFLEX	1.34	– –	5.31	5.91	0.45
ZTRANS1	0.90	0.23	– –	0.45	1.12
ZTRANS2	0.12	3.52	– –	4.91	1.69
ZTRANS3	0.35	0.03	– –	0.05	0.50
ZTRANS4	1.90	5.99	– –	6.57	0.56
ZAUTO1	0.00	0.02	0.07	– –	0.12
ZAUTO2	0.00	0.02	0.07	– –	0.12
ZSPECCRI	0.67	0.00	0.05	0.04	– –
ZSTATUT	0.67	0.00	0.05	0.04	– –

Examination of the modification values calculated for the Λ_X matrix (see Table 4.8 above) indicates that the Speed measure of Information Processing also load on the automatization and transfer of knowledge latent variables. The fact that only two additional paths would

103

significantly improve the fit of the model should be interpreted as a positive and favourable comment on the merits of the measurement model.

4.5   INTERPRETATION OF THE MEASUREMENT MODEL

Through the examination of the magnitude and the significance of the slope of the regression of the observed variables on their respective latent variables an indication of the validity of the measures is obtained. In other words, if a measure is designed to provide a valid reflection of a specific latent variable, then the slope of the regression of X_i on ξ_j in the fitted measurement model has to be substantial and significant (Diamantopoulos & Siguaw, 2000).

The unstandardized Λ_x (see Table 4.9 below) matrix contains the regression coefficients of the regression of the manifest variables on the latent variables they were linked to. The regression coefficients/loadings of the manifest variables on the latent variables are significant (p < 0,05) if the t-values, as indicated in the matrices, exceed |1,96|. Significant indicator loadings provide validity evidence in favour of the indicators (Diamantopoulos & Siguaw, 2000).
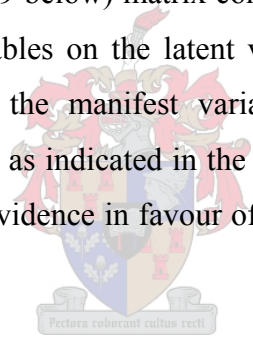
Table 4.9

Unstandardized Lambda-X Matrix

	ABSTRACT	INFOPROC	TRANSFOR	AUTOMAT	LEARNPER
	--------	--------	--------	--------	--------
ZABSTR1	0.68	- -	- -	- -	- -
	(0.09)				
	7.70				
ZABSTR2	0.91	- -	- -	- -	- -
	(0.09)				
	10.57				
ZSPEED	- -	0.71	- -	- -	- -
		(0.08)			
		8.51			
ZACCU	- -	0.87	- -	- -	- -
		(0.08)			
		11.23			
ZFLEX	- -	0.76	- -	- -	- -
		(0.08)			
		9.32			
ZTRANS1	- -	- -	0.73	- -	- -

	– –	– –	(0.09) 8.37		
ZTRANS2	– –	– –	0.77 (0.09) 9.07	– –	– –
ZTRANS3	– –	– –	0.72 (0.09) 8.23	– –	– –
ZTRANS4	– –	– –	0.41 (0.10) 4.24	– –	– –
ZAUTO1	– –	– –	– –	0.81 (0.08) 9.95	– –
ZAUTO2	– –	– –	– –	0.83 (0.08) 10.37	– –
ZSPECCRI	– –	– –	– –	– –	0.83 (0.15) 5.70
ZSTATUT	– –	– –	– –	– –	0.46 (0.11) 4.16

All the factor loadings, indicated in the Lambda-X matrix, are significant with t > 1,96. However, Diamantopoulos & Siguaw (2000) warn that there is indeed a problem with relying on unstandardised loadings and their associated t-values. The problem is that it might be hard to compare the validity of different indicators measuring a particular construct. Diamantopoulos & Siguaw (2000) explains:

> This problem arises because indicators of the same construct may be measured on very different scales; if this is the case, then direct comparisons of the magnitudes of the loadings are clearly inappropriate. In addition, bearing in mind that each latent variable has to be assigned a scale by fixing the loading of one of its indicators to unity, the loadings of the other indicators for the latent variable are only interpretable relative to the unit of the reference indicator. Clearly, if a different indicator is used as the reference variable, the magnitudes of the loadings will change. (p. 89)

As a result, Diamantopoulos & Siguaw (2000) recommend that the magnitudes of the standardised loadings should also be investigated. This is done by examining the *Completely Standardised Solution*, in which both latent and manifest variables have been standardized, available as part of the LISREL output. The completely standardized factor loading matrix is presented in Table 4.10. The values shown in Table 4.10 could be

interpreted as the regression slopes of the regression of the standardized indicator variables on the standardized latent variables. The completely standardized factor loadings therefore indicate the average change expressed in standard deviation units in the indicator variable associated with one standard deviation change in the latent variable. Interpreted in this sense, the loading of the first abstract reasoning indicator variable on the abstract reasoning latent variable, the fourth transfer indicator variable on the transfer latent variable and the second learning performance indicator variable on the learning performance indicator variable could be regarded as somewhat problematic. The square of the completely standardized factor loadings indicate the proportion of indicator variance explained in terms of the latent variable it is meant to express (Diamantopoulos & Siguaw, 2000). Since each indicator only loads on a single latent variable the squared completely standardized loadings equal the R^2 values shown below in Table 4.11.

Table 4.10

Completely Standardized Lambda-X Matrix

	ABSTRACT	INFOPROC	TRANSFOR	AUTOMAT	LEARNPER
ZABSTR1	0.68	- -	- -	- -	- -
ZABSTR2	0.91	- -	- -	- -	- -
ZSPEED	- -	0.71	- -	- -	- -
ZACCU	- -	0.87	- -	- -	- -
ZFLEX	- -	0.76	- -	- -	- -
ZTRANS1	- -	- -	0.73	- -	- -
ZTRANS2	- -	- -	0.77	- -	- -
ZTRANS3	- -	- -	0.72	- -	- -
ZTRANS4	- -	- -	0.41	- -	- -
ZAUTO1	- -	- -	- -	0.81	- -
ZAUTO2	- -	- -	- -	0.83	- -
ZSPECCRI	- -	- -	- -	- -	0.83
ZSTATUT	- -	- -	- -	- -	0.46

The squared multiple correlations (R^2) of the indicators depicted in Table 4.11 show the proportion of variance in an indicator that is explained by its underlying latent variable. A high R^2 value would indicate that variance in the indicator in question to a large degree reflects variance in the latent variable to which it has been linked. The rest of the variance, not explained by the latent variable can be ascribed to systematic and random measurement error (Diamantopoulos & Siguaw, 2000).

Table 4.11

106

Squared Multiple Correlations for X – Variables

ZABSTR1	ZABSTR2	ZSPEED	ZACCU	ZFLEX	ZTRANS1
--------	--------	--------	--------	--------	--------
0.46	0.82	0.51	0.76	0.58	0.53

ZTRANS2	ZTRANS3	ZTRANS4	ZAUTO1	ZAUTO2	ZSPECCRI
--------	--------	--------	--------	--------	--------
0.60	0.51	0.17	0.65	0.70	0.68

ZSTATUT

0.21

The total variance in the ith indicator variable (X_i) could be decomposed into variance due to variance in the latent variable the indicator variable was meant to reflect (ξ_j), variance due to variance in other systematic latent effects the indicator variable was not designed to reflect and random measurement error. The latter two sources of variance in the indicator variable are acknowledged in equations 26 and 33 through the measurement error terms ε_i and δ_i. The measurement error terms ε and δ thus do not differentiate between systematic and random sources of error or non-relevant variance. The square of the completely standardized factor loading λ (see Table 4.10) could be interpreted as the proportion systematic-relevant indicator variable variance. This corresponds to the information provided in Table 4.11. The diagonal of the completely standardized theta-delta (θ_δ) matrix (shown as Table 4.12) reflect the proportion of non-relevant item parcel variance.

Table 4.12

Completely Standardized Theta-Delta Matrix

ZABSTR1	ZABSTR2	ZSPEED	ZACCU	ZFLEX	ZTRANS1
--------	--------	--------	--------	--------	--------
0.54	0.18	0.49	0.24	0.42	0.47

ZTRANS2	ZTRANS3	ZTRANS4	ZAUTO1	ZAUTO2	ZSPECCRI
--------	--------	--------	--------	--------	--------
0.40	0.49	0.83	0.35	0.30	0.32

ZSTATUT

0.79

The completely standardized error variance of the ith indicator variable ($\theta_{\delta ii}$) in Table 4.12 thus consists of systematic non-relevant variance and random error variance. The values shown in Table 4.11 could therefore be interpreted as indicator variable validity coefficients, $\rho(X_i,\xi_j)$. Since ($\lambda_{ij}^2 + \theta_{\delta ii}$) are equal to unity in the completely standardized solution, the validity coefficients, $\rho(X_i,\xi_j)$ can be defined as follows:

$\rho(X_i,\xi_j) = \sigma^2_{\text{systematic-relevant}}/(\sigma^2_{\text{systematic-relevant}} + \sigma^2_{\text{non-relevant}})$

$\qquad = \lambda_{ij}^2/[\lambda_{ij}^2 + \theta_{\delta ii}]$

$\qquad = 1 - (\theta_{\delta i}/[\lambda_{ij}^2 + \theta_{\delta ii}])$

$\qquad = 1 - \theta_{\delta ii}$

$\qquad = \lambda_{ij}^2$ ---34

Since reliability could be defined as the extent to which variance in indicator variables can be attributed to systematic sources, irrespective of whether the source of variance is relevant to the measurement intention or not, the values shown in Table 4.11 could simultaneously be interpreted as lower bound estimates of the item reliabilities ρ_{ii} (Diamantopoulos & Siguaw, 2000; Jöreskog & Sörbom, 1996a). The extent to which the true item reliabilities would be under-estimated would be determined by the extent to which δ_{ii} contains the effect of the systematic non-relevant latent influences.
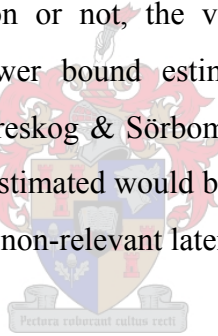
In terms of the foregoing argument the values of the squared multiple correlations for the indicator variables shown in Table 4.11 cause some concern as there are various indicators (as highlighted) that fail to adequately reflect variance in the latent variables they are meant to reflect. Especially the TRANS4 measure and the STATUT measure seem to have failed to represent the transfer of knowledge and job competency potential latent variables respectively. Only 17% of the variance in TRANS4 and 21% of the variance in STATUT can be explained in terms of their respective underlying latent variables. ABSTRA1 to a somewhat lesser extent (46%) fails to represent the abstract reasoning latent variable successfully. But for ABSTRA2 and AUTO2, the underlying latent variables at best only explain modest proportions of the variance in the indicator variables in which they are meant to express themselves. This tends to erode the confidence with which any definite conclusion on the merits of the learning potential structural model will be possible.

Diamantopoulos & Siguaw (2000) further suggest that it is useful to also calculate a composite reliability value for each latent variable. This can be done by using the information on the indicator loadings end error variances from the *Completely Standardised Solution* (Tables 4.11 and 4.12). The following formula will be used to calculate composite reliability values (Diamantopoulos & Siguaw, 2000, p. 90):

$$\rho_c = (\Sigma\lambda)^2/[(\Sigma\lambda)^2 + \Sigma(\theta)] \text{--}(35)$$

Where:

ρ_c = composite reliability;

λ = completely standardized indicator loadings;

θ = completely standardized indicator error variances (i.e. variances of the δ's and ε's);

Σ = summation over the indicators of the latent variable.

Diamantopoulos & Siguaw (2000) report that $\rho_c > 0,6$ would be indicative that the composite indicators linked to a given latent variable provides a satisfactory reliable measurement of the construct. The composite reliability scores for the composite indicators linked to the latent variables are displayed in Table 4.13 below:

Table 4.13

Composite Reliability Scores For Composite Indicators

Latent Variable	Composite Reliability Score
Abstract reasoning capacity	**0,778**
Information processing capacity	**0,826**
Transfer of knowledge	**0,760**
Automatization	**0,805**
Job competency potential	**0,600**

As indicated in Table 4.13, the two indicator variables, Specific Crimes and Statutory Law, designed to represent the job competency potential latent variable, collectively failed to

successfully achieve this since the composite reliability score is equal to 0,60. Tables 4.10 and 4.11 suggest this is primarily due to the failure of the STATUT indicator variable. Table 4.13 further indicates that the composite reliability score calculated for the composite indicators of transfer of knowledge is also somewhat less favourable, but still exceeding the desired composite reliability score of 0,60. This again suggests that a question mark hangs over the success with which at least some of the latent variables comprising the learning potential structural model had been operationalized, thereby jeopardizing an unambiguous verdict on the merits of the learning potential structural model.

Another measure, complementary to the composite reliability measure, is the average variance extracted (ρ_v). This measure indicates the amount of variance ascribed to the construct in relation to the amount of variance due to measurement error. ρ_v values less than 0,5 would indicate that the measurement error accounts for a greater amount of variance in the indicators than the underlying variable does. If this is indeed the case then serious doubts arise regarding the soundness of the indicators and/or the latent variable itself (Diamantopoulos & Siguaw, 2000). The following formula was used to calculate the average variance extracted (Diamantopoulos & Siguaw, 2000, p. 91):

$$P_v = (\Sigma\lambda^2)/[(\Sigma\lambda^2) + \Sigma(\theta)] \text{--(36)}$$

Where λ, θ and Σ are defined as in equation 35.

The results achieved by solving equation 36 for each of the latent variable indicator variable sets are displayed in Table 4.14.

Table 4.14

Average Variance Extracted For Composite Indicators

Latent Variable:	Average variance extracted
Abstract reasoning capacity	**0,642**
Information processing capacity	**0,615**
Transfer of knowledge	**0,453**

Automatization	**0,674**
Job competency potential	**0,450**

Table 4.14 essentially corroborates the conclusion that emerged from the results presented thus far.

The measurement model fit could be described as reasonable. The claim that specific indicator variables reflect specific latent variables and not others does therefore, not seem unreasonable. However, the success with which at least two of the indicator variables represent the latent variables they were meant to reflect seems limited. As such, the integrity of the analysis of the hypothesized structural relations is threatened. As was argued earlier, unless there is evidence to suggest that the operational measures do in fact reflect the latent variables of interest, the usefulness of using such data to investigate hypotheses on the assumed nature of relationships between the latent variables becomes contentious. If poor structural model fit would be obtained it would not be possible to unequivocally rule out the possibility that it was not due to inherent structural flaws but rather to shortcomings in the operationalization of specific latent variables (specifically the focal job competency potential latent variable and the transfer of knowledge latent variable).

4.6   ASSESSING THE OVERALL GOODNESS-OF-FIT OF THE STRUCTURAL MODEL

An admissible final solution of parameter estimates for the expanded Learning Potential structural model was obtained after 14 iterations. The full spectrum of fit indices provided by LISREL to assess the absolute fit of the model is presented in Table 4.15.

Table 4.15

Goodness-Of-Fit Statistics For The Structural Model

Degrees of Freedom	59
Minimum Fit Function Chi-Square	101.24 (P = 0.00052)
Normal Theory Weighted Least Squares Chi-Square	97.97 (P = 0.0011)
Estimated Non-centrality Parameter (NCP)	38.97
90 Percent Confidence Interval for NCP	(15.62 ; 70.21)
Minimum Fit Function Value	0.86
Population Discrepancy Function Value (F0)	0.33
90 Percent Confidence Interval for F0	(0.13 ; 0.60)
Root Mean Square Error of Approximation (RMSEA)	0.075
90 Percent Confidence Interval for RMSEA	(0.047 ; 0.10)
P-Value for Test of Close Fit (RMSEA < 0.05)	0.066
Expected Cross-Validation Index (ECVI)	1.37
90 Percent Confidence Interval for ECVI	(1.17 ; 1.64)
ECVI for Saturated Model	1.54
ECVI for Independence Model	12.97
Chi-Square for Independence Model with 78 Degrees of Freedom	1505.02
Independence AIC	1531.02
Model AIC	161.97
Saturated AIC	182.00
Independence CAIC	1580.15
Model CAIC	282.90
Saturated CAIC	525.90
Normed Fit Index (NFI)	0.93
Non-Normed Fit Index (NNFI)	0.96
Parsimony Normed Fit Index (PNFI)	0.71
Comparative Fit Index (CFI)	0.97
Incremental Fit Index (IFI)	0.97
Relative Fit Index (RFI)	0.91
Critical N (CN)	102.60
Root Mean Square Residual (RMR)	0.057
Standardized RMR	0.057
Goodness of Fit Index (GFI)	0.89
Adjusted Goodness of Fit Index (AGFI)	0.83
Parsimony Goodness of Fit Index (PGFI)	0.57

The p-value associated with the χ^2 value in Table 4.15 clearly indicates significant test statistics. A non-significant χ^2 indicates model fit in that the model can reproduce the observed covariance matrix. (Bollen and Long, 1993; Kelloway, 1998). In this case, the model is not able to reproduce the observed covariance matrix to a degree of accuracy that can be attributed to sampling error only, in other words, H_{01a}: $\Sigma = \Sigma(\Theta)$ is rejected in favour of H_{a1a}: $\Sigma \neq \Sigma(\Theta)$ (Kelloway, 1998). By implication H_{01a}: RMSEA=0 is also rejected in favour of H_{a1a}: RMSEA>0.

The evaluation of fit on the basis of the normed chi-square statistic χ^2/df ($\chi^2/df = 1,66$) for the structural model suggest that the model fits the data well (refer to paragraph 4.2. for a more in-depth interpretation of this ratio).

The RMSEA value of 0,075 indicates reasonable, but not good fit, as values less than 0,05 indicates good fit. The RMR (0,057) and standardised RMR (0,057) also indicate reasonable, but not good fit. Values of less than 0,05 on the latter indices are regarded as indicative of a model that fits the data well (Kelloway, 1998). The 90% confidence interval for RMSEA shown in Table 4.15 (0.047 ; 0.10) includes the critical 0,05 value, indicating reasonable to good fit. A test of the significance of the obtained value is performed by LISREL by testing H_{01b}: RMSEA $\leq$ 0,05 against H_{a1b}: RMSEA > 0,05. Table 4.15 indicates that the obtained RMSEA value of 0,075 is not significantly different from the target value of 0,05 (i.e. H_{01b} is not rejected; p > 0,05) and since the confidence interval does include the target value of 0,05, a close fit seems to have been achieved.

The goodness-of-fit index (GFI) and the adjusted GFI (AGFI) both indicate reasonable, but not good fit. Values exceeding 0,9 indicates good fit to the data (Jöreskog & Sörbom, 1993; Kelloway, 1998).

The assessment of parsimonious fit acknowledges that model fit can always be improved by adding more paths to the model and estimating more parameters until perfect fit is achieved in the form of a saturated or just-identified model with no degrees of freedom (Kelloway, 1998). The objective in model building is, however, to achieve satisfactory fit with as few model parameters as possible (Jöreskog & Sörbom, 1993). The objective is therefore to find, in this sense, the most parsimonious model.

Indices of parsimonious fit relate the benefit that accrues in terms of improved fit to the cost incurred (in terms of degrees of freedom lost) to affect the improvement in fit (Jöreskog & Sörbom, 1993). The values for the Aiken information criterion (AIC = 161,97) shown in Table 4.15 suggest that the fitted structural model provides a more parsimonious fit than the independent/null model (1531,02) as well as the saturated model (182.00) since smaller

values on these indices indicate a more parsimonious model (Kelloway, 1998). The values for the consistent Aiken information criterion (CAIC = 282.90) also suggest that the fitted structural model provides a more parsimonious fit than both the independent/null model (1580.15) and the saturated model (525.90).

The expected cross-validation index (ECVI) expresses the difference between the reproduced sample covariance matrix $(\hat{\Sigma})$ derived from fitting the model on the sample at hand and the expected covariance matrix that would be obtained in an independent sample of the same size from the same population (Byrne, 1989; Diamantopoulos & Siguaw, 2000). Since the model ECVI (1.37) is smaller than the value obtained for the independence model (12.97) and smaller than the ECVI value associated with the saturated model (1.54), a model resembling the fitted model seems to have a better chance of being replicated in a cross-validation sample than the independence model or the saturated model. This finding is echoed by the Aiken information criterion and the consistent Aiken information criterion results. The proposed learning potential structural model therefore does not seem to be overly elaborate in how it conceptualizes the causal processes amongst the learning potential latent variables, nor does the proposed model seem to under-represent the causal processes.
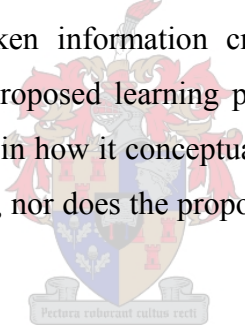
After interpreting all the fit indices, the conclusion would have to be drawn that the structural model fit the data reasonably well. Integrating the results obtained on the full spectrum of fit statistics depicted in Table 4.15 seems to suggest a reasonable fitting model that clearly outperforms the independence model and that seems to acknowledge the true complexity of the processes underlying the APIL test battery.

However, to ensure a thorough assessment of the fit of the structural model and especially because the structural model only fits the data reasonably well, it is necessary to also investigate the standardised residuals and modification indices to determine the extent of success with which the model explains the observed covariances amongst the manifest variables (Jöreskog & Sörbom, 1993).

114

4.7 EXAMINATION OF STRUCTURAL MODEL RESIDUALS

Residuals refer to the differences between corresponding cells in the observed and fitted covariance/correlation matrices (Jöreskog & Sörbom, 1993). Residuals, and especially standardised residuals, provide diagnostic information on sources of lack of fit in models (Jöreskog & Sörbom., 1993; Kelloway, 1998). Jöreskog & Sörbom (1993) explain that a standardised residual refers to a residual that is divided by its estimated standard error.

Standardised residuals can be interpreted as z-scores (i.e. number of standard deviations above or below the mean). Standardised residuals are considered to be large if they exceed +2.58 or –2.58 (Diamantopoulos & Siguaw, 2000). A large positive residual would indicate that the model underestimates the covariance between two variables, while a large negative residual would indicate that the model overestimates the covariance between variables. Underestimation indicates that the model needs to be modified by adding additional explanatory paths, which could better account for the covariance between the variables. On the other hand, if the model overestimates the covariance between the variables, the model should be modified by trimming paths that are associated with the particular covariance term (Jöreskog & Sörbom, 1993). The standardized residuals resulting from the covariance estimates derived from the estimated comprehensive model parameters are shown in Table 4.16 and summarized in Table 4.17.

Table 4.16

Standardized Residuals

	ZTRANS1	ZTRANS2	ZTRANS3	ZTRANS4	ZAUTO1	ZAUTO2
ZTRANS1	1.59					
ZTRANS2	1.68	1.59				
ZTRANS3	-1.09	0.79	1.59			
ZTRANS4	-1.64	-0.05	-0.16	1.59		
ZAUTO1	0.29	-1.63	0.15	2.32	- -	
ZAUTO2	1.19	-0.93	0.26	1.41	1.40	- -
ZSPECCRI	1.29	-1.19	0.91	-1.45	-1.70	-0.50
ZSTATUT	-0.56	0.03	-0.31	0.71	0.55	-0.61
ZABSTR1	0.14	0.49	0.20	0.46	-0.23	-1.87
ZABSTR2	-1.08	0.08	0.77	1.01	1.44	2.44
ZSPEED	1.19	0.52	1.57	2.07	4.93	-0.88
ZACCU	0.29	-1.04	-0.27	1.57	-3.47	0.22
ZFLEX	-1.02	-2.13	-1.23	2.17	-1.89	-1.98

```
        Standardized Residuals

          ZSPECCRI   ZSTATUT   ZABSTR1   ZABSTR2    ZSPEED    ZACCU
          --------  --------  --------  --------  --------  --------
ZSPECCRI      1.59
 ZSTATUT      1.59      1.59
 ZABSTR1     -0.13     -0.22      - -
 ZABSTR2      0.43     -0.82      0.33      - -
  ZSPEED      1.20      0.34      0.01     -0.05      - -
   ZACCU      1.94      0.93      0.00     -0.80     -2.17      - -
   ZFLEX      0.84      0.69      0.39     -1.44     -1.37      4.86


        Standardized Residuals

            ZFLEX

          --------
   ZFLEX      - -
```

Table 4.17

Summary Statistics For Standardized Residuals

Smallest Standardized Residual = -3.47
Median Standardized Residual = 0.15
Largest Standardized Residual = 4.93
Largest Negative Standardized Residuals
Residual for ZACCU and ZAUTO1 -3.47
Largest Positive Standardized Residuals
Residual for ZSPEED and ZAUTO1 4.93
Residual for ZFLEX and ZACCU 4.86

Two large positive residuals and one large negative residual indicate three observed covariance terms in the observed sample covariance matrix being poorly estimated by the derived model parameter estimates. Inspection of the variables associated with these standardised residuals reveals no clear specific suggestions for possible model modification. The small number of covariance terms poorly reproduced by the fitted model parameter corroborates the earlier conclusion that the model succeeds reasonably well in explaining the observed data.

Jöreskog & Sörbom (1993) states that all the standardised residuals may be examined collectively in a stem-and-leaf plot and a Q-plot. A good model would be characterised by a stem-and-leaf plot in which the residuals are distributed approximately symmetrical around zero. An excess of residuals on the positive or negative side would indicate that the

residuals are systematically under- or overestimated. The stem-and-leaf plot of the structural model standardized residuals is depicted in Figure 4.3.

```
- 3|5
- 2|210
- 1|99766544221100
- 0|998866533222110000000000
  0|1122233333445555778889
  1|0222344466666666679
  2|1234
  3|
  4|99
```

Figure 4.3

Stem-And-Leaf Plot Of Standardized Residuals

From the stem-and-leaf plot depicted in Figure 4.3, the distribution of the standardised residuals appears to be slightly positively skewed. This is supported by the fact that the median standardized residual is 0,15. The estimated model parameters therefore, on average, tend to under-estimate the observed covariance terms. This would suggest that the model fails to account for one or more influential paths. Moreover the distribution of the standarized residuals seem to be somewhat less leptokurtic than would be typical of good model fit.
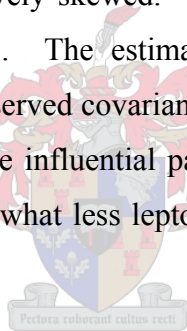
The Q-plot is depicted in Figure 4.4.
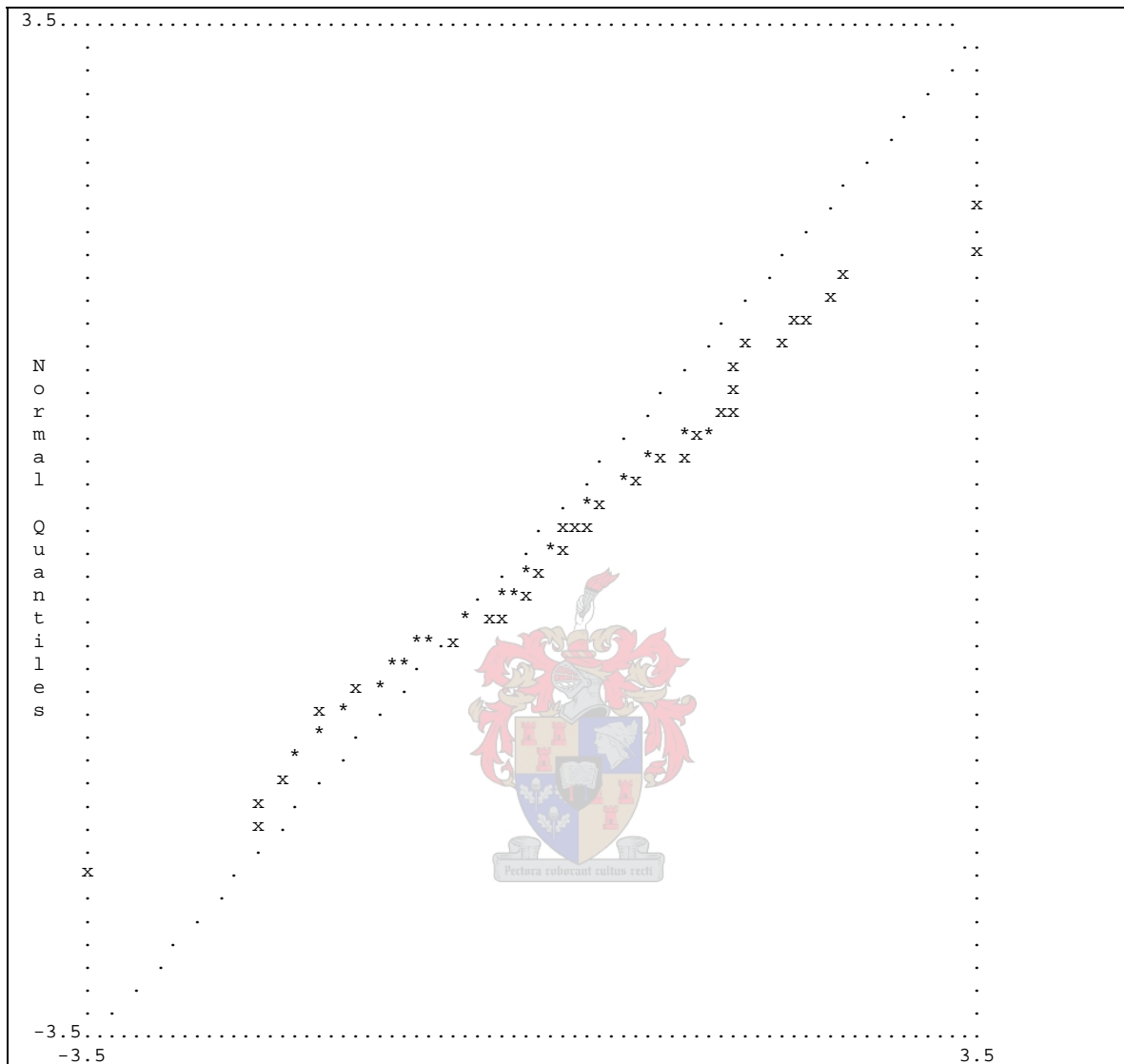
Figure 4.4
Q-Plot Of Standardized Residuals

When interpreting the Q-plot it is important to note whether the data points fall on the 45-degree reference line or not. If the points fall on the 45-degree reference line, it would be indicative of a good model fit (Jöreskog & Sörbom, 1993). The model fit would be less than satisfactory if the data points swivel away from the 45-degree reference line.

Less than perfect model fit is indicated by the fact that the standardised residuals for all pairs of observed variables tend to deviate slightly from the 45^0 – reference line in the Q-plot in both the lower and upper region of the X-axis. The deviation is, however, not pronounced and less severe than in the case of the measurement model.

4.8   FURTHER ASSESSMENT OF THE STRUCTURAL MODEL

Since the structural model adequately fits the data as judged by the overall goodness-of-fit measures, the structural model will be evaluated further. The aim of further assessing the structural model is to determine whether each of the hypothesized theoretical relationships is supported by the data (Diamantopoulos & Siguaw, 2000).

Diamantopoulos & Siguaw (2000) identify four issues relevant to assessing the structural model. First, it is important to assess whether the signs of the parameters representing the paths between latent variables are in agreement with the nature of the causal effect hypothesised to exist between the latent variables. Secondly, it is important to assess whether the parameter estimates are significant ($p < 0,05$). Thirdly (assuming significance), it is important to assess the magnitude of the parameter estimates indicating the strength of the hypothesized relationships. Lastly, it is important to evaluate the squared multiple correlations (R^2), indicating the amount of variance in each endogenous latent variable that is explained by the latent variables linked to it in terms of the hypothesized structural model.

The parameters of interest in assessing the structural model are the freed elements of the gamma (Γ) and beta (B) matrices.

The unstandardised Γ matrix (Table 4.18) is used to assess the significance of the estimated path coefficients γ_{ij}, expressing the strength of the influence of ξ_j on η_i. These parameters are significant ($p < 0,05$) if $t > |1,96|$ (Diamantopoulos & Siguaw, 2000). A significant γ estimate would imply that the corresponding H_0-hypothesis will be rejected in favour of the

relevant H_a-hypothesis. The hypotheses which are relevant to the Γ matrix in this study are H_{02} and H_{03}.

Table 4.18

Unstandardized Gamma (Γ) Matrix

```
            ABSTRACT    INFORMAT
            --------    --------
TRANSFER      0.27        - -
            (0.15)
             1.72

 AUTOMAT      - -         0.89
                        (0.10)
                          8.64

LEARNPER      - -         - -
```

The values in the matrix (Table 4.18) indicate that, the null hypothesis, that information processing capacity (ξ_2) has no statistically significant effect on automatization (η_2) (hypothesis 3, H_{03}: $\gamma_{22} = 0$), can be rejected in favour of H_{a3} (p < 0,05). Thus, the relationship postulated between information processing capacity (ξ_2) and automatization (η_2) in the structural model, is corroborated. In addition, the sign associated with this significant γ parameter estimate is consistent with the nature of the relationship hypothesised to exist between these latent unit performance dimensions.

Table 4.18 further indicates that the null hypothesis, that abstract thinking capacity (ξ_1) has no statistically significant positive effect on transfer of knowledge (η_1) (hypothesis 2, H_{02}: $\gamma_{11} = 0$), cannot be rejected. An insignificant (p > 0,05) relationship is, therefore, evident between abstract thinking capacity and transfer of knowledge. The causal relationship hypothesized between abstract thinking capacity (ξ_1) and transfer of knowledge (η_1) is therefore not corroborated. The question invariably arises to what extent this is due to the inability to successfully operationalize the transfer of knowledge latent variable.

The unstandardised B matrix (Table 4.19) is used to assess the significance of the estimated path coefficients β_{ij}, expressing the strength of the influence of η_j on η_i. Unstandardised β_{ij} estimates are also significant (p<0,05) if t > │1,96│ (Diamantopoulos & Siguaw, 2000). A

significant β estimate would imply that the corresponding H_0-hypothesis will be rejected in favour of the relevant H_a-hypothesis. The hypotheses which are relevant to the B matrix in this study are H_{04}, H_{05} and H_{06}.

Table 4.19

Unstandardized Beta (B) Matrix

	TRANSFER	AUTOMAT	LEARNPER
TRANSFER	– –	0.54 (0.16) 3.29	– –
AUTOMAT	– –	– –	– –
LEARNPER	0.26 (0.19) 1.33	0.33 (0.19) 1.75	– –

The values in Table 4.19 indicate that the null hypothesis, that the extent to which transfer of knowledge (η_1) occurs is not determined by the extent to which automatization occurs (η_2), H_{04}: $\beta_{12} = 0$, can be rejected in favour of H_{a4} (p < 0,05). Thus, the relationship postulated between transfer of knowledge (η_1) and automatization (η_2) in the structural model are corroborated. In addition, the sign associated with this significant β parameter estimate is consistent with the nature of the relationship hypothesised to exist between these latent unit performance dimensions.

Table 4.19 further indicates that the null hypothesis, that transfer of knowledge (η_1) has no statistically significant effect on job competency potential targeted by the affirmative training intervention (η_3), H_{05}: $\beta_{31} = 0$, cannot be rejected. Table 4.19 moreover indicates that the null hypothesis, that automatization (η_2) has no statistically significant effect on job competency potential targeted by the affirmative training intervention (η_3), H_{06}:$\beta_{32} = 0$, also cannot be rejected. An insignificant (p>0,05) relationship is, therefore, evident between transfer of knowledge and job competency potential and between automatization and job competency potential. The causal relationships hypothesized between transfer of knowledge and learning performance and between automatization and learning performance are therefore not corroborated. Again the question invariably arises whether

these finding is due to a conceptual flaw in Taylor's original theorizing or whether it is due to the inability of this study to successfully operationalize the job competency potential latent variable.

LISREL has the ability to decompose total effects between latent variables into direct and indirect effects. Indirect effects refer to the influence of ξ_j or η_i on η_j as mediated by η_k. Indirect effects are derived by multiplying the unstandardized parameter estimates of the paths comprising the indirect effect. LISREL also computes an estimated standard error and an accompanying t-value for each indirect effect in the model (Diamantopoulos & Siguaw, 2000; Kaplan, 2000). The matrix of indirect effects of ksi on eta (Table 4.20) will be used to test the mediation null hypotheses H_{07} and H_{08}. Diamantopoulos & Siguaw (2000) however warns that the indirect effect statistics need to be interpreted with caution when any of contributing parameter estimates is insignificant.

Table 4.20

Unstandardized Indirect Effects Of Ksi On Eta

	ABSTRACT	INFOPROC
TRANSFER	- -	0.48
		(0.15)
		3.22
AUTOMAT	- -	- -
LEARNPER	0.07	0.42
	(0.07)	(0.12)
	1.04	3.55

Table 4.20 indicates that the null hypothesis, that the influence of abstract thinking capacity (ξ_1) on the job competencies targeted by the training intervention (η_3) is not mediated by transfer of knowledge (η_1), H_{07}: $\gamma_{11}\beta_{31} = 0$, can not be rejected ($p > 0,05$). Table 4.20 moreover indicates that the null hypothesis, that the influence of information processing capacity (ξ_2) on the job competencies targeted by the training intervention (η_3) is not mediated by automatization (η_2), H_{08}: $\gamma_{22}\beta_{32} = 0$, can be rejected ($p < 0,05$). Table 4.20 in addition indicates that the indirect effect of Information Processing on Transfer is significant ($p < 0,05$). No formal mediation hypothesis was formulated in this regard.

Diamantopoulos & Siguaw (2000) suggest that additional insights can be obtained by looking at the completely standardised B and Γ parameter estimates. The completely standardised B and Γ parameter estimates are not affected by differences in the unit of measurement of the independent variables and can, thus, be compared across equations. The completely standardised B and Γ parameter estimates reflect the average change, expressed in standard deviation units, in the endogenous latent variable directly resulting from a one standard deviation change in an endogenous or exogenous latent variable to which it has been linked, holding the effect of all other variables constant (Diamantopoulos & Siguaw, 2000). The completely standardised B and Γ parameter estimates are depicted in Table 4.21.

Table 4.21

Completely Standardized Gamma (Γ) and Beta (B) Estimates

```
GAMMA

                ABSTRACT    INFOPROC
                --------    --------
    TRANSFER       0.27        - -
     AUTOMAT        - -        0.89
    LEARNPER        - -         - -


BETA

                TRANSFER    AUTOMAT    LEARNPER
                --------    --------   --------
    TRANSFER        - -       0.54       - -
     AUTOMAT        - -        - -        - -
    LEARNPER       0.26       0.33       - -
```

Table 4.21 indicates that of the two significant effects, the effect of information processing on automatization is more pronounced than the effect of automatization on transfer of knowledge.

4.9   STRUCTURAL MODEL MODIFICATION INDICES

Model modification indices are aimed at answering the question whether any of the currently fixed parameters, when freed in the model, would significantly improve the parsimonious fit of the model. Modification indices (MI) indicate the extent to which the χ^2 fit statistic will decrease if a currently fixed parameter in the model is freed and the model

re-estimated (Jöreskog & Sörbom, 1993). Large modification index values (> 6,6349) would be indicative of parameters that, if set free, would improve the fit of the model significantly (p<0,01) (Diamantopoulos & Siguaw, 2000; Jöreskog & Sörbom, 1993). It is important to note that parameters with high MI values should only be freed if it makes substantive sense to do so (Kelloway, 1998). The expected change for the parameter is the expected value of the parameter if it were freed (i.e., the extent to which it would change from its currently fixed value of zero). The standardised and completed standardised expected changes are the expected values in the standardised and completely standardised solution if the parameter were freed.

Jöreskog & Sörbom (1993) argue that modification indices should be used in the following way in the process of model evaluation and modification:

> If chi-square is large relative to the degrees of freedom, one examines the modification indices and relaxes the parameter with the largest modification index *if this parameter can be interpreted substantively*. If it does not make sense to relax the parameter with the largest modification index, one considers the second largest modification index etc. If the signs of certain parameters are specified a priori, positive or negative, the expected parameter changes associated with the modification indices for these parameters can be used to exclude models with parameters having the wrong sign. (p. 127)

The proposed structural model depicted in Figure 3.1 seems to fit the data reasonably well. The foregoing analysis of the standardised residuals, however, suggests that the addition of one or more paths would probably improve the fit of the model. Examination of the modification indices calculated for the B matrix, depicted in Table 4.22 suggest that there exists no additional paths between any endogenous latent variables that would significantly improve the fit of the proposed learning potential structural model.

Table 4.22

Modification Indices And Expected Change Calculated For The B Matrix

```
Modification Indices for BETA

           TRANSFER    AUTOMAT    LEARNPER
           --------    --------    --------
 TRANSFER    - -         - -         0.07
  AUTOMAT    0.03        - -         3.95
 LEARNPER    - -         - -         - -

      Expected Change for BETA

           TRANSFER    AUTOMAT    LEARNPER
           --------    --------    --------
 TRANSFER    - -         - -        -0.12
  AUTOMAT    0.03        - -        -0.24
 LEARNPER    - -         - -         - -

   Standardized Expected Change for BETA

           TRANSFER    AUTOMAT    LEARNPER
           --------    --------    --------
 TRANSFER    - -         - -        -0.12
  AUTOMAT    0.03        - -        -0.24
 LEARNPER    - -         - -         - -
```

Examination of the modification indices calculated for the Γ matrix indicates depicted in Table 4.23 suggest that there also exists no additional paths between any exogenous latent variable and any endogenous latent variable that would significantly improve the fit of the proposed learning potential structural model.

Table 4.23

Modification Indices And Expected Change Calculated For The Γ Matrix

```
           ABSTRACT   INFOPROC
           --------   --------
 TRANSFER    - -        0.11
  AUTOMAT    2.52       - -
 LEARNPER    0.07       4.17
```

The inability of this study to successfully operationalize the job competency potential latent variable should, however, be taken into account when considering the modification index findings with regards to the Γ matrix.

4.10 POWER ASSESSMENT

When evaluating the findings on the fit of a model it is very important to investigate the statistical power associated with testing the model. Statistical power refers to the conditional probability of rejecting the null hypothesis given that it is false (P[reject H_0: $\Sigma = \Sigma(\Theta)|H_0$ false]). In the context of SEM statistical power therefore refers to the probability of rejecting an incorrect model. Diamantopoulos & Siguaw (2000) explain:

> When we test a model's fit by, say, the chi-square test, we emphasize the probability of making a **Type I error**, i.e. rejecting a correct model; this probability is captured by the **significance level**, α which is usually set at 0,05. A significant chi-square result indicates that *if* the null hypothesis is true (i.e. the model is correct in the population), then the probability of incorrectly rejecting it is low (i.e. less than five times out of 100 if α= 0,05). However, another error that can occur is *not* to reject an incorrect model. This type of error is known as **Type II error** and the probability associated with it is denoted as β. The probability of avoiding a Type II error is, therefore, 1-β and it is this probability that indicates the power of our test; thus the power of the test tells us how likely it is that a false null hypothesis (i.e. incorrect model) will be rejected. (p. 93)

Unfortunately, this issue is more often than not neglected, but it is important to understand that any model evaluation would be incomplete if power considerations were ignored. The importance of conducting a power analysis stems from the critical role that sample size plays in the decisions made in model testing (Diamantopoulos & Siguaw, 2000). Specifically in large samples (i.e., high power) the decision to reject a null hypothesis of exact fit (or a null hypothesis of close fit) becomes problematic because it is not clear whether the model was rejected because of severe misspecifications in the model or to the (too) high sensitivity of the test to detect even minor flaws in the model. Conversely in small samples (i.e., low power) the decision not to reject the null hypothesis of exact/close fit results in ambiguity because it is not clear whether the decision was due to the accuracy of the model or to the insensitivity of the test to detect specification errors in the model. When the chi-square test is applied only Type I errors are explicitly taken into account, thus, a power analysis must be undertaken to also account for the probability of Type II errors (Diamantopoulos & Siguaw, 2000).

Two types of power calculations were performed. First, the power associated with a test of exact fit [i.e. testing the null hypothesis that the model fits perfectly in the population (as done by the conventional chi-square test)] was estimated. However, as argued earlier, this test is very limited since models are only approximations of reality and, therefore, rarely do they fit exactly in the population. The power associated with a test of close fit was consequently also estimated. Here the null hypothesis states that the model has a close, but imperfect fit in the population. The stated null hypothesis takes the error of approximation [i.e. the discrepancy between Σ and $\Sigma(\theta)$] into account and is, therefore, more realistic (Diamantopoulos & Siguaw, 2000).

Both the test of exact fit and the test of close fit make use of the RMSEA statistic. If a model fits perfectly in the population the error due to approximation is set at 0 and the null hypothesis formulated earlier as H_{01a} is consequently tested against H_{a1a} (Diamantopoulos & Siguaw, 2000).

To determine the power of a test of the exact fit hypothesis, a specific value for the parameter needs to be assumed under H_a, because there are as many power estimates, as there are possible values for the parameter under H_a. A value that makes good sense to use in this instance is RMSEA = 0,05, as RMSEA < 0,05 is indicative of good fit. If a model achieves close fit in the population the error due to approximation will be set equal to or less than 0,05 (Diamantopoulos & Siguaw, 2000).

If a model fits only approximately in the population the error due to approximation is set at 0,05 and the null hypothesis formulated earlier as H_{01b} is consequently tested against H_{a1b} (Diamantopoulos & Siguaw, 2000). To determine the power of a test of the close fit hypothesis a specific value for the parameter again needs to be assumed. A reasonable value to assume is RMSEA = 0,08, since RMSEA = 0,08 is the upper limit of reasonable fit. Diamantopoulos & Siguaw (2000) explains:

> The choice of the values for ε_0 and ε_0 (ε represents RMSEA) reflect the recommendations in the literature regarding RMSEA thresholds for close and mediocre fit respectively. Here we are asking the question: if the true fit of the model was

mediocre (i.e. if H_{a1b} is correct), what is the power of the test that H_{014}: RMSEA $\leq 0,05$

(i.e. that fit is close)? (p. 95)

With the information on H_0 and H_a and given a significance level (α) of 0,05 and a sample size N, the power of the test becomes a function of the degrees of freedom (v) in the model ($v=\frac{1}{2}[(p+q)][p+q+1]-t)=91-35=56^{11}$). All other things being equal, the higher the degrees of freedom, the greater the power of the test (Diamantopoulos & Siguaw, 2000). Power tables compiled by MacCallum *et al.* (1996) only make provision for degrees of freedom ≤ 100 and N ≤ 500. A SPSS translation of the SAS syntax provided by MacCallum *et al.* (1996) was consequently used to derive power estimates for the tests of exact and close fit, given the effect sizes assumed above, a significance level (α) of 0,05 and a sample size of 119. The degrees of freedom (v) in the model is ($\frac{1}{2}[(p+q)][p+q+1]-t)=91-35=56$.

Power values of 0,433 were obtained for the test of exact fit. The probability of rejecting the null hypothesis of exact fit under the true condition of close fit are thus low enough to have provided reason for concern at the outset of the study. However, since the null hypothesis of exact fit has been rejected this need no longer be a reason for concern. The probability of rejecting the null hypothesis of close fit if the true model fit was mediocre at 0,558. The latter power estimate, taken in conjunction with the decision not to reject the null hypotheses of close fit, suggest that the conclusion of close model fit could be somewhat contentious in that the tests were not highly sensitive to misspecifications in the model.

4.11 REGRESSION ANALYSES

The APIL test battery provides dynamic measures of two latent learning competencies and static measures of two latent dispositions, which determine the learning competencies (Taylor, 1989, 1994, 1997). In estimating expected learning performance, these measures would typically be combined in a linear multiple regression model. Given the nature of the structural model underlying the APIL battery, the question however arises, whether the

---

[11] t represents the number of parameters to be estimated in the fitted model.

static measures do not become redundant in a regression model that already includes the dynamic measures. To a certain extent the inability of the preceding analysis to corroborate all of the hypothesized causal linkages in the learning potential structural model makes this somewhat of a less pressing question.

The matrix of zero-order Pearson correlation coefficients between the five learning potential measures and the corresponding conditional probabilities is portrayed in Table 4.25. The convention proposed by Guilford (cited in Tredoux & Durrheim, 2002, p. 184) and depicted in Table 4.24 was used to interpret sample correlation coefficients. Although somewhat arbitrary and although it ignores the normative question about the magnitude of values typically encountered in a particular context, it nonetheless fosters consistency in interpretation.

Table 4.24

Guilford's Interpretation Of The Magnitude Of Significant r

Absolute value of r	Interpretation
< 0,19	Slight; almost no relationship
0,20 – 0,39	Low correlation; definite but small relationship
0,40 – 0,69	Moderate correlation; substantial relationship
0,70 – 0,89	High correlation; strong relationship
0,90 – 1,00	Very high correlation; very dependable relationship

Table 4.25 indicates that the two learning competencies measures (transfer of knowledge and automatization) as well as the two latent learning competency potential measures (abstract reasoning capacity and information processing capacity) correlate low but statistically significantly ($p < 0,05$) with the learning performance measures (job competency potential). Of concern, however, is the fact that the learning competencies measures and the learning competency potential measures correlate moderately and statistically significantly ($p < 0,05$) amongst themselves.

Table 4.25

Learning Potential Correlation Matrix

Correlations

		Abstract Reasoning Capacity	Information Processing Capacity	Transfer of Knowledge	Automatiz ation	Job Competency Potential
Abstract Reasoning Capacity	Pearson Correlation	1	.615**	.542**	.510**	.266**
	Sig. (1-tailed)		.000	.000	.000	.002
	N		119	119	119	119
Information Processing Capacity	Pearson Correlation	.615**	1	.618**	.680**	.385**
	Sig. (1-tailed)	.000		.000	.000	.000
	N	119	119	119	119	119
Transfer of Knowledge	Pearson Correlation	.542**	.618**	1	.512**	.303**
	Sig. (1-tailed)	.000	.000		.000	.000
	N	119	119	119	119	119
Automatization	Pearson Correlation	.510**	.680**	.512**	1	.328**
	Sig. (1-tailed)	.000	.000	.000		.000
	N	119	119	119	119	119
Job Competency Potential	Pearson Correlation	.266**	.385**	.303**	.328**	1
	Sig. (1-tailed)	.002	.000	.000	.000	
	N	119	119	119	119	119

**. Correlation is significant at the 0.01 level (1-tailed).

4.11.1 TESTING HYPOTHESIS 9

Standard linear multiple regression analysis was used to determine whether the dynamic measures of the two latent learning competencies (X_3 & X_4) each explain unique variance in a composite measure of the job competency potential targeted by the affirmative training intervention (Y). More specifically to test H_{09a}: $\beta[X_3] = 0|\beta[X_4] \neq 0$ and H_{09b}: $\beta[X_4] = 0|\beta[X_3] \neq 0$ against directional alternative hypotheses the following linear regression model was fitted to the data using standard multiple regression.

$$E[Y \mid X_i] = \alpha + \beta_1[X_3] + \beta_2[X_4]\text{---}(37)$$

Where:

$E[Y \mid X_i]$ = Expected learning performance;

α = Y intercept

β_1 = slope of Y with variable X_3, holding variables X_4 constant;

β_2 = slope of Y with variable X_4, holding variables X_3 constant;

X_3 = Transfer of knowledge

X_4 = Automatization

4.11.1.1   Testing Hypothesis H_{09a}

Evaluating the contribution of each independent variable to the multiple regression model is important because only those independent variables that are useful in predicting the value of the dependent variable should be included in the regression model (Berenson, Levine & Goldstein, 1983).

The partial F test criterion was used for determining the contribution of each independent variable. This method involves determining the contribution to the regression sum of squares (SSR) made by each independent variable after all other independent variables have been included in the model. The new independent variable will only be included if it significantly improves the model (Berenson *et al.*, 1983).

To test H_{09a} a univariate analysis of variance was done through general linear modelling (GLM) on SPSS (2006). The GLM Univariate procedure allows you to model the value of a dependent scale variable based on its relationship to categorical and scale predictors (SPSS, 2006). The GLM Univariate procedure is based on the General Linear Model procedure, in which factors and covariates are assumed to have a linear relationship to the dependent variable (SPSS, 2006).

H_{09a}: $\beta[X_3] = 0 | \beta[X_4] \neq 0$ was tested by calculating the following test statistic from Table 4.26:

F = {(SSR[b3,b4] – SSR[b4]/[p-1]}/MSE[b3,b4]--(38)

   = {(803,084 – 654.541/[2-1]}/45,408

   = 3,27

Where:

F~F[p-1, n-p-1]

Using a 0,05 level of significance, H_{09a} may be rejected if $F \geq F_{1-\alpha;\ p-1,\ n-p-1}$.

F = 3,27 < F(1;116) = 3,9201

Table 4.26

Regression Of Job Competency Potential On Transfer Of Knowledge (X_3) And Automatization (X_4)

Tests of Between-Subjects Effects

Dependent Variable: Job Competency Potential

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	803.084[a]	2	401.542	8.843	.000
Intercept	58784.767	1	58784.767	1294.602	.000
X3	148.543	1	148.543	3.271	.073
X4	247.367	1	247.367	5.448	.021
Error	5267.281	116	45.408		
Total	521600.250	119			
Corrected Total	6070.366	118			

a. R Squared = .132 (Adjusted R Squared = .117)

Tests of Between-Subjects Effects

Dependent Variable: Job Competency Potential

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	654.541[a]	1	654.541	14.140	.000
Intercept	59152.436	1	59152.436	1277.891	.000
X4	654.541	1	654.541	14.140	.000
Error	5415.824	117	46.289		
Total	521600.250	119			
Corrected Total	6070.366	118			

a. R Squared = .108 (Adjusted R Squared = .100)

Tests of Between-Subjects Effects

Dependent Variable: Job Competency Potential

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	555.717[a]	1	555.717	11.790	.001
Intercept	161895.867	1	161895.867	3434.818	.000
X3	555.717	1	555.717	11.790	.001
Error	5514.648	117	47.134		
Total	521600.250	119			
Corrected Total	6070.366	118			

a. R Squared = .092 (Adjusted R Squared = .084)

H_{09a}: $\beta[X_3] = 0|\beta[X_4] \neq 0$ is, therefore, not rejected, implying that transfer of knowledge (X_3) does not significantly ($p > 0,05$) explain unique variance in job competency potential when included in a model already containing automatization (X_4). The observed exceedence probability associated with the calculated F statistic (0,073) does, however, not exceed the critical value of 0,05 by a great margin.

4.11.1.2   Testing Hypothesis H_{09b}:

H_{09b}: $\beta[X_4] = 0|\beta[X_3] \neq 0$ was tested by calculating the following test statistic from Table 4.26:

F = {(SSR[b3,b4] – SSR[b3]/[p-1]}/MSE[b3,b4]---(39)

  = {(803.084 – 555.717)/[2-1]}/45.408

  = 5,448
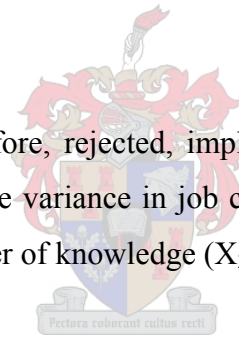
Where:

F~F[p-1, n-p-1]

Using a 0,05 level of significance, H_{09b} may be rejected if $F \geq F_{1-\alpha; p-1, n-p-1}$.

F = 5,54 > F(1;116) = 3,9201

H_{09b}: $\beta[X_4] = 0|\beta[X_3] \neq 0$ is, therefore, rejected, implying that automatization (X_4) does significantly (p<0,05) explain unique variance in job competency potential when included in a model already containing transfer of knowledge (X_3).

The regression equation, $E[Y \mid X_i] = \alpha + \beta_1[X_3] + \beta_2[X_4]$, is, thus, reduced to:

$E[Y \mid X_i] = \alpha + \beta_2[X_4]$

Where:

      $E[Y \mid X_i]$ = Expected learning performance;

      α = Y intercept

      β_2 = slope of Y with variable X_4;

      X_4 = Automatization

4.11.2 TESTING HYPOTHESIS 10

Hierarchical multiple regression analysis was used to determine whether the static measures of the two learning dispositions (X_1 & X_2) would significantly explain variance in learning performance when added to a linear regression model already containing automatization (X_4). More specifically to test H_{010}: $\beta[X_1] = \beta[X_2] = 0|\beta[X_4] \neq 0$ the following two observed variable linear multiple regression models will be fitted to the data using standard multiple regression:

$$E[Y \mid X_i] = \alpha + \beta_1[X_1] + \beta_2[X_2] + \beta_4[X_4] \text{---(40)}$$
$$E[Y \mid X_i] = \alpha\ \beta_4[X_4] \text{---(41)}$$

Where:

   $E[Y \mid X_i]$ = Expected learning performance;

   α = Y intercept

   β_1 = slope of Y with variable X_1, holding variables X_2 and X_4 constant;

   β_2 = slope of Y with variable X_2, holding variables X_1 and X_4 constant;

   β_4 = slope of Y with variable X_1, holding variables X_1 and X_2 constant;

   X_1 = Abstract reasoning capacity

   X_2 = Information processing capacity

   X_4 = Automatization

4.11.2.1   Testing Hypothesis H_{010}:

Evaluating the contribution of each independent variable to the multiple regression model is important because only those independent variables that are useful in predicting the value of the dependent variable should be included in the regression model (Berenson *et al.*, 1983). The partial F test criterion will be used for determining the contribution of each independent variable. This method involves determining the contribution to the regression sum of squares (SSR) made by each independent variable after all other independent variables have been included in the model. The new independent variable will only be included if it significantly improves the model (Berenson *et al.*, 1983). To test H_{010} a

univariate analysis of variance was done through general linear modelling (GLM) on SPSS (2006). The GLM Univariate procedure allows you to model the value of a dependent scale variable based on its relationship to categorical and scale predictors (SPSS, 2006). The GLM Univariate procedure is based on the General Linear Model procedure, in which factors and covariates are assumed to have a linear relationship to the dependent variable (SPSS, 2006).

H_{010}: $\beta[X_1] = \beta[X_2] = 0|$ $\beta[X_4] \neq 0$ was tested by calculating the following test statistic from Table 4.27:

Table 4.27

Regression Of Job Competency Potential On Abstract Reasoning Capacity (X_1), Information Processing Capacity (X_2) And Automatization (X_4)

Tests of Between-Subjects Effects

Dependent Variable: Job Competency Potential

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	953.884[a]	3	317.961	7.147	.000
Intercept	27503.653	1	27503.653	618.183	.000
X1	3.093	1	3.093	.070	.793
X2	219.913	1	219.913	4.943	.028
X4	44.804	1	44.804	1.007	.318
Error	5116.482	115	44.491		
Total	521600.250	119			
Corrected Total	6070.366	118			

a. R Squared = .157 (Adjusted R Squared = .135)

Tests of Between-Subjects Effects

Dependent Variable: Job Competency Potential

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	654.541[a]	1	654.541	14.140	.000
Intercept	59152.436	1	59152.436	1277.891	.000
X4	654.541	1	654.541	14.140	.000
Error	5415.824	117	46.289		
Total	521600.250	119			
Corrected Total	6070.366	118			

a. R Squared = .108 (Adjusted R Squared = .100)

$$F = \{(SSR[b1,b2,b4] - SSR[b4]/[p-1]\}/MSE[b1,b2,b4] \text{------------------------}(42)$$

= {(953,884 – 654,541)/[3-1]}/44,491

= 3,36

Where:

F~F[p-1, n-p-1]

Using a 0,05 level of significance, H_{010} may be rejected if $F \geq F_{1-\alpha; p-1, n-p-1}$.

F = 3,36 > F(2;115) = 3.0718

In other words, H_{010} is rejected (p<0,05), indicating that X_1 (abstract reasoning capacity) and/or X_2 (information processing capacity) should be included in a model already containing automatization. This finding that at least one or both static measures of learning potential do add incremental validity to a regression model already containing a dynamic measure of learning potential challenges the earlier argument that the inclusion of the person-centred drivers of the learning competencies in a prediction model that already contains measures of the learning competencies will be redundant. Logically the redundancy argument would imply that the information processing capacity (X_2) variable should be redundant but that X_1 (abstract reasoning capacity) should explain unique variance in learning performance when included in a model already containing automatization (X_4).

The results of testing H_{011a}: $\beta[X_1] = 0 \mid \beta[X_2] \neq 0$, $\beta[X_4] \neq 0$ and H_{011b}: $\beta[X_2] = 0 \mid \beta[X_1] \neq 0$, $\beta[X_4] \neq 0$ will determine whether both static learning potential measures should be added to a model already containing automatization (X4).

4.11.3 TESTING HYPOTHESIS 11

Standard multiple regression analysis was used to investigate the predictive ability of an observed variable linear multiple regression model, regressing learning performance on a weighted linear combination of the two learning dispositions and automatization (X4). More specifically to test H_{011a}: $\beta[X_1] = 0 \mid \beta[X_2] \neq 0$, $\beta[X_4] \neq 0$, H_{011b}: $\beta[X_2] = 0 \mid \beta[X_1] \neq$

0, $\beta[X_4] \neq 0$ and H_{011c}: $\beta[X_4] = 0 \mid \beta[X_1] \neq 0, \beta[X_2] \neq 0$[12] the following observed variable linear multiple regression models was fitted to the data using standard multiple regression:

$$E[Y \mid X_i] = \alpha + \beta_1[X_1] + \beta_2[X_2] + \beta_4[X_4] \text{---}(43)$$

$$E[Y \mid X_i] = \alpha + \beta_1[X_2] + \beta_4[X_4] \text{---}(44)$$

$$E[Y \mid X_i] = \alpha + \beta_1[X_1] + \beta_4[X_4] \text{---}(45)$$

Where:

$\quad$ $E[Y \mid X_i]$ = Expected learning performance;

$\quad$ α = Y intercept

$\quad$ β_1 = slope of Y with variable X_1, holding variables X_2, X_3, and X_4 constant;

$\quad$ β_2 = slope of Y with variable X_2, holding variables X_1, X_3, and X_4 constant;

$\quad$ β_3 = slope of Y with variable X_3, holding variables X_1, X_3, and X_4 constant;

$\quad$ β_4 = slope of Y with variable X_1, holding variables X_1, X_2, and X_3 constant;

$\quad$ X_1 = Abstract reasoning capacity

$\quad$ X_2 = Information processing capacity

$\quad$ X_3 = Transfer of knowledge

$\quad$ X_4 = Automatization

Evaluating the contribution of each independent variable to the multiple regression model is important because only those independent variables that are useful in predicting the value of the dependent variable should be included in the regression model (Berenson, Levine & Goldstein, 1983). The partial F test criterion will be used for determining the contribution of each independent variable. This method involves determining the contribution to the regression sum of squares (SSR) made by each independent variable after all other independent variables have been included in the model. The new independent variable will only be included if it significantly improves the model (Berenson *et al.*, 1983).

To test H_{011I} a univariate analysis of variance was done through general linear modelling (GLM) on SPSS (2006). The GLM Univariate procedure allows you to model the value of a dependent scale variable based on its relationship to categorical and scale predictors (SPSS,

---

[12] These statistical hypotheses may have to be revised depending on the outcome of the analyses.

2006). The GLM Univariate procedure is based on the General Linear Model procedure, in which factors and covariates are assumed to have a linear relationship to the dependent variable (SPSS, 2006).

H_{011a}: $\beta[X_1] = 0 \mid \beta[X_2] \neq 0$, $\beta[X_4] \neq 0$ was tested by calculating the following test statistic from Table 4.28:

Table 4.28

Regression Of Job Competency Potential On Abstract Reasoning Capacity (X_1), Information Processing Capacity (X_2) And Automatization (X_4)

Tests of Between-Subjects Effects

Dependent Variable: Job Competency Potential

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	953.884[a]	3	317.961	7.147	.000
Intercept	27503.653	1	27503.653	618.183	.000
X1	3.093	1	3.093	.070	.793
X2	219.913	1	219.913	4.943	.028
X4	44.804	1	44.804	1.007	.318
Error	5116.482	115	44.491		
Total	521600.250	119			
Corrected Total	6070.366	118			

a. R Squared = .157 (Adjusted R Squared = .135)

Tests of Between-Subjects Effects

Dependent Variable: Job Competency Potential

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	950.791[a]	2	475.395	10.772	.000
Intercept	27959.627	1	27959.627	633.513	.000
X2	296.249	1	296.249	6.712	.011
X4	49.897	1	49.897	1.131	.290
Error	5119.575	116	44.134		
Total	521600.250	119			
Corrected Total	6070.366	118			

a. R Squared = .157 (Adjusted R Squared = .142)

F = {(SSR[b1,b2,b4] – SSR[b2,b4]/[p-2]}/MSE[b1,b2,b4]------------------(46)

= {(953,884 – 950.791)/[3-2]}/44,491

= 0,07

Where:

$F \sim F[p-2, n-p-1]$

Using a 0,05 level of significance, H_{010} may be rejected if $F \geq F_{1-\alpha;\ p-2,\ n-p-1}$.

$F = 0,07 < F(1;115) = 3,9201$

In other words, H_{011a} is not rejected (p>0,05), implying that X_1 (abstract reasoning capacity) should not be included in a model already containing automatization and information processing capacity. The model is, thus, reduced to:

$E[Y \mid X_i] = \alpha + \beta_2[X_2] + \beta_4[X_4]$--(47)

The revised hypothesis H_{011b}: $\beta[X_2] = 0 \mid \beta[X_4] \neq 0$, was subsequently tested by calculating the following test statistic from Table 4.28:

$F = \{(SSR[b2,b4] - SSR[b4]/[p-1]\}/MSE[b2,b4]$--(48)
$= \{(950.791 - 654,541)/[2-1]\}/44,134$
$= 6,71$

Where:

$F \sim F[p-1, n-p-1]$

Using a 0,05 level of significance, H_{011b} may be rejected if $F \geq F_{1-\alpha;\ p-1,\ n-p-1}$.

$F = 6,71 > F(1;116) = 3,9201$

Thus, H_{011b} is rejected (p<0,05), and information processing capacity (X_2) significantly explains unique variance in learning performance in a model already containing automatization.

The question, however, now arises whether automatization (X_4) still significantly explains variance in learning performance when retained in a model already including information processing capacity (X_2)? To investigate this question, the revised hypothesis H_{011c}: $\beta[X_4]$ = 0 │ $\beta[X2_4] \neq 0$, was subsequently tested by calculating the following test statistic from Table 4.28

$F = \{(SSR[b2,b4] – SSR[b2]/[p-1]\}/MSE[b2,b4]$--(49)
$= \{(950.791 - 900,894)/ [2-1]\}/44,134$
$= 1,13$
Where:
$F \sim F[p-1, n-p-1]$

Using a 0,05 level of significance, H_{011c} may be rejected if $F \geq F_{1-\alpha; \, p-1, \, n-p-1}$.

$F = 1,13 < F(1;116) = 3,9201$

H_{011C} is not rejected (p>0,05), implying that X_4 (automatization) does not significantly explain unique variance in learning performance when included in a model already containing information processing capacity (X_2). Automatization should therefore be eliminated from the regression model and the model consequently is, reduced to:

$E[Y │ X_i] = \alpha + \beta_2[X_2]$ --(50)

The finding that the addition of a measure of abstract reasoning capacity (X_1) does not add incremental validity to a model already including a measure of automatization (X_4) is rather surprising. Based on the logic of the proposed learning potential structural model one would have expected abstract reasoning capacity to have served as a substitute for transfer of knowledge and to significantly explain unique variance in learning performance in a model that already includes the dynamic automatization measure. The finding that the inclusion of both information processing capacity (X_2) and automatization (X_4) in the same prediction model serves little purpose does agree with the proposed learning potential

structural model (Figure 3.1). Logically one would, however, have expected that the more direct causal determinant of learning performance should have been the more influential predictor. This line of reasoning, however, does ignore the reliability and validity of the operational measures.

4.12 COMPARISON OF PREDICTIVE POWER

To compare the predictive power of the structural model to that of the observed variable multiple regression model, the R^2 of the regression model will be contrasted to the proportion of the variance in job competency potential targeted by the affirmative training intervention (η_3) explained by the structural model linked to it.

The squared multiple correlations for the endogenous latent variables in the learning potential structural model are shown in Table 4.29.

TABLE 4.29

Squared Multiple Correlations For Structural Equations

TRANSFER	AUTOMAT	LEARNPER
0.56	0.79	0.30

The proposed structural model satisfactorily explains variance in the endogenous latent variable automatization. The proposed structural model, moreover, succeeds modestly in explaining variance in transfer of knowledge. The proposed structural model, however does not really succeed in explaining variance in learning performance. The model's failure to account for the variance in the primary endogenous latent variable, Learning Performance, creates some reason for concern.

Table 4.25 indicates that information processing capacity (X_2) explains 14,8% of the variance in learning performance in the observed variable regression model. Inclusion of all four predictor variables in the regression model explains 16,1% of the variance in learning performance (See Table 4.30).

TABLE 4.30

Regression Of Job Competency Potential On Abstract Reasoning Capacity (X_1),

Information Processing Capacity (X_2) And Transfer (X_3) and Automatization (X_4)

Tests of Between-Subjects Effects

Dependent Variable: Job Competency Potential

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	979.520[a]	4	244.880	5.484	.000
Intercept	25419.822	1	25419.822	569.230	.000
X1	.231	1	.231	.005	.943
X2	155.021	1	155.021	3.471	.065
X3	25.636	1	25.636	.574	.450
X4	36.137	1	36.137	.809	.370
Error	5090.846	114	44.657		
Total	521600.250	119			
Corrected Total	6070.366	118			

a. R Squared = .161 (Adjusted R Squared = .132)

The structural model therefore outperformed the regression model in terms of its ability to explain variance in learning performance.

It could, however be argued that the comparison should rather be made between the structural model with all insignificant paths pruned away and the regression model with only effects included that significantly explain unique variance in learning performance. When fitting the structural model without the abstract reasoning exogenous latent variable and its associated indicator variables the fit of the model deteriorates (RMSEA=,087) although the proportion of variance in learning performance increases marginally to 0,31.

CHAPTER 5

CONCLUSIONS, RECOMMENDATION AND SUGGESTIONS FOR FUTURE RESEARCH

5.1   INTRODUCTION

In South Africa, specific groups had and still have easier and more access to opportunities that allow them to develop an array of coping strategies, knowledge, skills and abilities. Access to such opportunities often has the resultant effect that such individuals perform better in conventional assessment situations, in the workplace and in training programmes or educational institutions (Boeyens, 1989; Guthke, 1993; Hamers & Resing, 1993; Taylor, 1989; Taylor, 1992). It is clear that a need exists in South Africa for a method to identify individuals who will gain maximum benefit from affirmative developmental opportunities, especially cognitively demanding development opportunities.

Ideally, such measures would assess an individual's core or fundamental cognitive abilities and potentialities and not specific job skills that are strongly influenced by past opportunities (Taylor, 1997).

Vygotsky (1978) originally proposed the measurement of learning potential as a method of assessing an individual's core or fundamental cognitive abilities and potentialities. Drawing on ideas developed, amongst others, by Vygotsky (1978), Sternberg (1984), Snow, *et al.* (1984) and Ackerman (1988), Taylor (1989, 1994, 1997) developed a learning potential model, which explicates the latent variables collectively constituting learning potential.

Based on this learning potential model, a learning potential measure, specifically assessing an individual's hidden latent and reserve potential, reducing the influence of verbal abilities, cultural meanings and educational qualifications has been proposed and developed by Taylor (1989, 1992, 1994, 1997) in the form of the APIL test battery. Taylor (1997) claims that this learning potential measure is especially suited for application in the following two practical settings. Firstly, it can serve as a useful tool in making fair decisions when job applicants are selected. Allied to this, is the fact that it can also help

identify candidates who are likely to cope or master more demanding work roles. Secondly, it can be applied in the educational arena and will help identify candidates who are likely to master new cognitively demanding material in a formal educational or training context.

However, such an approach would imply that effective selection of previously disadvantaged individuals into formal educational or training is possible to the extent to which there exists a comprehensive understanding of the reasons underlying training performance and the manner in which they combine to determine learning performance in addition to clarity on the fundamental nature of the key performance areas comprising the learning task. The APIL test battery will thus result in effective selection to the extent to which the explanatory model on which it is based successfully explains variance in learning performance.

The primary objectives of this research were to (a) explicate the structural model underlying the APIL test battery and (b) evaluate the fit of the model on empirical data.

The APIL test battery provides dynamic measures of two latent learning competencies and static measures of two latent dispositions, which determine the learning competencies (Taylor, 1989, 1994, 1997). In estimating expected learning performance, these measures would typically be combined in a linear multiple regression model. Given the nature of the structural model underlying the APIL test battery, the question, however, arises whether the static measures do not become redundant in a model that already includes the dynamic measures.

The secondary objectives of this research consequently were to determine whether the static measures of the two latent learning dispositions would significantly explain variance in learning performance when added to a model already containing dynamic measures of the two latent learning competencies.

If the structural model was indeed found to be valid, and if the APIL test battery does succeed in selecting those who show a greater probability of succeeding in cognitively

demanding developmental opportunities aimed at enhancing the required knowledge, skills, and abilities needed to succeed on the job, and the development programmes do succeed in reducing the differences in the criterion distributions, then adverse impact in job selection should be reduced. Previously disadvantaged individuals should now be significantly less disadvantaged in terms of the required knowledge, skills and abilities. Theoretically, over time, this approach should work towards levelling the playing field so that success or failure in personnel selection can be attributed to previous opportunities or lack thereof to a lesser degree than is currently typically the case in South Africa, without even temporarily relinquishing on the utility objective.

The specific objectives of this research were:
- To explicate the underlying structural model upon which the APIL test battery was developed, explaining learning performance;
- To test the model's absolute fit;
- To evaluate the significance of the hypothesised paths in the model;
- To investigate the predictive ability of an observed variable linear multiple regression model, regressing learning performance on a weighted linear combination of the two learning dispositions and the two learning competencies;
- To determine whether the static measures of the two latent learning dispositions would significantly explain variance in learning performance when added to a linear regression model already containing dynamic measures of the two latent learning competencies;
- To compare the predictive power of the structural model to that of the observed variable multiple regression model;
- To modify the structural model if necessary; and
- To compare the fit of the revised structural model to that of the original model.

Overall, It was expected that the structural model would fit the data reasonably well although it was expected that the null hypothesis of exact fit would be rejected. It was furthermore expected that all paths hypothesized in the model would be significant.

It was also expected that the dynamic measures of learning potential will each explain unique variance in a composite measure of the job competency potential targeted by the affirmative training intervention. It was, however, expected that the static measures of the two latent learning dispositions would not significantly explain variance in learning performance when added to a linear regression model already containing dynamic measures of the two latent learning competencies.

5.2   RESULTS

5.2.1 Evaluation of the Measurement Model

The overall goodness-of-fit of the measurement model was tested through structural equation modelling. Various indices were interpreted to assess the goodness-of-fit of the measurement model and it was found that the measurement model fits the data reasonably well, but not perfectly. After examination of the measurement model residuals it was found that five observed covariance terms in the observed sample covariance matrix (out of 78 covariance terms) was being poorly estimated by the derived model parameter estimates, which is also indicative of reasonable model fit. By examining the stem-and-leaf plot, the distribution of standardised residuals appeared only slightly positively skewed, but not overly so. This indicated that there was a slightly stronger tendency for the model to overestimate the observed covariance terms. However, the Q-plot clearly indicated less than perfect model fit. Subsequently, given the examination of the residuals, it was also important to evaluate the model modification indices. After examining the modification indices it was found that only two additional paths would significantly improve the fit of the measurement model, which was interpreted as a positive and favourable comment on the merits of the measurement model. The values of the squared multiple correlations for the indicators and the calculated composite reliability values for each latent variable caused concern and left a question mark hanging over the success with which at least some of the latent variables comprising the learning potential structural model had been operationalized, thereby jeopardizing an unambiguous verdict on the merits of the learning potential structural model. Overall, the measurement model fit could be described as reasonable. The

claim that specific indicator variables reflect specific latent variables and not others did therefore, not seem unreasonable. However, the success with which at least two of the indicator variables represented the latent variables they were meant to reflect seemed limited. As such, the integrity of the analysis of the hypothesized structural relations was threatened.

5.2.2 Evaluation of Structural Model

After interpreting all the fit indices, the conclusion was drawn that the structural model also fit the data reasonably well. Integrating the results obtained on the full spectrum of fit statistics seemed to suggest a reasonable fitting model that clearly outperforms the independence model and that seemed to acknowledge the true complexity of the processes underlying the APIL test battery. However, to ensure that a thorough assessment of the fit of the structural model was done and especially because it was found that the structural model only fits the data reasonably well, it was necessary to investigate the standardised residuals and modification indices to determine the extent of success with which the model explained the observed covariances amongst the manifest variables. Two large positive residuals and one large negative residual indicated three observed covariance terms in the observed sample covariance matrix being poorly estimated by the derived model parameter estimates. Inspection of the variables associated with these standardised residuals revealed no clear specific suggestions for possible model modification. The small number of covariance terms poorly reproduced by the fitted model parameter corroborated the earlier conclusion that the model succeeded reasonably well in explaining the observed data. From the stem-and-leaf plot the distribution of the standardised residuals appeared to be slightly positively skewed. The estimated model parameters therefore, on average, tend to under-estimate the observed covariance terms. This would suggest that the model failed to account for one or more influential paths. Moreover, the distribution of the standarized residuals seemed to be somewhat less leptokurtic than would be typical of good model fit. Less than perfect model fit was indicated by the fact that the standardised residuals for all pairs of observed variables tended to deviate slightly from the 45^0 – reference line in the Q-

plot in both the lower and upper region of the X-axis. The deviation was, however, not pronounced and less severe than in the case of the measurement model.

Upon further examination, the null hypothesis, that information processing capacity has no statistically significant positive effect on automatization was rejected. Thus, the relationship postulated between information processing capacity (ξ_2) and automatization (η_2) in the structural model, was corroborated. In addition, the sign associated with this significant γ parameter estimate was consistent with the nature of the relationship hypothesised to exist between these latent unit performance dimensions. The null hypothesis, that abstract thinking capacity (ξ_1) has no statistically significant positive effect on transfer of knowledge (η_1) was not rejected. An insignificant relationship was, therefore, evident between abstract thinking capacity and transfer of knowledge. The causal relationship hypothesized between abstract thinking capacity (ξ_1) and transfer of knowledge (η_1) was therefore not corroborated. The question invariably arose as to what extent this was due to the inability to successfully operationalize the transfer of knowledge latent variable. The null hypothesis, that the extent to which transfer of knowledge (η_1) occurs is not positively determined by the extent to which automatization occurs was also rejected. Thus, the relationship postulated between transfer of knowledge (η_1) and automatization (η_2) in the structural model was corroborated. In addition, the sign associated with this significant β parameter estimate was consistent with the nature of the relationship hypothesised to exist between these latent unit performance dimensions. The null hypothesis, that transfer of knowledge (η_1) has no statistically significant positive effect on job competency potential targeted by the affirmative training intervention (η_3), was not rejected. The null hypothesis, that automatization (η_2) has no statistically significant positive effect on job competency potential targeted by the affirmative training intervention was also not rejected. An insignificant relationship was, therefore, evident between transfer of knowledge and job competency potential and between automatization and job competency potential. The causal relationships hypothesized between transfer of knowledge and learning performance and between automatization and learning performance were not corroborated. Again the question invariably arose as to whether these finding was due to a conceptual flaw in

Taylors's original theorizing or whether it was due to the inability of this study to successfully operationalize the job competency potential latent variable.

The null hypothesis, that the influence of abstract thinking capacity (ξ_1) on the job competencies targeted by the training intervention (η_3) is not mediated by transfer of knowledge was not rejected. The null hypothesis, that the influence of information processing capacity (ξ_2) on the job competencies targeted by the training intervention (η_3) is not mediated by automatization was on the other hand rejected. It was further found that the indirect effect of information processing on transfer was significant. The effect of information processing on automatization was found to be more pronounced than the effect of automatization on transfer of knowledge.

Overall, it was found that the proposed structural model fit the data reasonably well. However, the analysis of the standardised residuals, suggested that the addition of one or more paths would probably improve the fit of the model. Examination of the modification indices suggested that there exist no additional paths between any endogenous latent variables or any exogenous latent variables that would significantly improve the fit of the proposed learning potential structural model. However, the inability of this study to successfully operationalize the job competency potential latent variable should be taken into account when considering the modification index findings. The probability of rejecting the null hypothesis of exact fit under the true condition of close fit was low enough to have provided reason for concern at the outset of the study. However, since the null hypothesis of exact fit was rejected, this no longer was a reason for concern. The probability of rejecting the null hypothesis of close fit if the true model fit was mediocre was also low enough to provide reason for concern. The latter power estimate, taken in conjunction with the decision not to reject the null hypotheses of close fit, suggested that the conclusion of close model fit could be somewhat contentious in that the tests were not highly sensitive to misspecifications in the model.

5.2.3 Regression Analysis

Further analysis were performed in the form of a correlation analysis, simple linear regression analysis and standard multiple linear regression analyses. It was found that the two learning competency measures (transfer of knowledge and automatization) as well as the two latent learning competencies potential measures (abstract reasoning capacity and information processing capacity) correlate low but statistically significantly with the learning performance measures (job competency potential). Of concern, however, was the fact that the learning competencies measures and the learning competency potential measures correlated moderately and statistically significantly amongst themselves.

The following was found through the regression analysis:

- Transfer of knowledge does not significantly explain unique variance in job competency potential when included in a model already containing automatization. The observed exceedence probability associated with the calculated F statistic did, however, not exceed the critical value of 0,05 by a great margin The problem seems to be the relatively high correlation existing between the two learning competency measures.;

- Automatization does significantly explain unique variance in job competency potential when included in a model already containing transfer of knowledge;

- Abstract reasoning capacity and/or information processing capacity should be included in a model already containing automatization. This finding that at least one or both static measures of learning potential do add incremental validity to a regression model already containing a dynamic measure of learning potential challenged the earlier argument that the inclusion of the person-centred drivers of the learning competencies in a prediction model that already contains measures of the learning competencies will be redundant. ;

- Abstract reasoning capacity should not be included in a model already containing automatization and information processing capacity;

- Information processing capacity significantly explains unique variance in learning performance in a model already containing automatization.

- Automatization does not significantly explain unique variance in learning performance when included in a model already containing information processing capacity.

The finding that the addition of a measure of abstract reasoning capacity does not add incremental validity to a model already including a measure of automatization was rather surprising. Based on the logic of the proposed learning potential structural model one would have expected abstract reasoning capacity to have served as a substitute for transfer of knowledge and to significantly explain unique variance in learning performance in a model that already includes the dynamic automatization measure. The finding that the inclusion of both information processing capacity and automatization in the same prediction model serves little purpose, does agree with the proposed learning potential structural model. It was, however, expected that the more direct causal determinant of learning performance should have been the more influential predictor. This line of reasoning, however, does ignore the reliability and validity of the operational measures.

5.2.4 Comparing the predictive power between the structural model and regression model
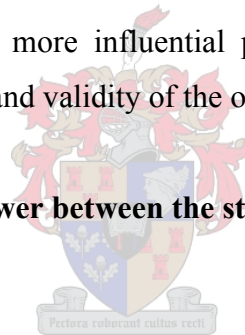
In comparing the predictive power of the structural model versus the regression model, it was found that the structural model outperformed the regression model in terms of its ability to explain variance in learning performance. However it could be argued that the comparison should rather have been made between the structural model with all insignificant paths pruned away and the regression model with only effects included that significantly explain unique variance in learning performance.

5.3 SUGGESTIONS FOR FUTURE RESEARCH

The results of this study partially justify the use of the APIL battery for affirmative development selection. Evidence on the fairness and the utility of the procedure would have to be examined to come to a definitive verdict on the usefulness of the battery. The

151

results would suggest that all the scores of the APIL need not be considered when estimating future training performance.

The stability of the model needs to be examined in a cross-validation study on a fresh sample of respondents taken from the same population. In addition, future research should investigate the possibility of expanding the model by formally incorporating latent variables like existing knowledge level, conscientiousness, tenacity, learning motivation, and learning support/infrastructure to explain additional variance in job competency potential targeted by the affirmative training intervention. It is extremely unlikely that differences in learning performance can be attributed to differences in intellectual ability only. It is furthermore unlikely that a mastery of learning material will necessarily mean that the newly acquired insights will be used in finding solutions to novel problems on the job. Incorporating latent variables like self-efficacy, performance motivation and mentoring/perceived support should therefore also be considered.

The degree of measurement model fit achieved could be described as reasonable. The claim that the specific indicator variables used to reflect the specific latent variables comprising the learning potential structural model does therefore, not seem altogether unreasonable. However, the success with which at least two of the indicator variables represent the latent variables they were meant to reflect seems limited. As such, the integrity of the analysis of the hypothesized structural relations is threatened. Especially the validity of the job competency potential and transfer of knowledge measures seems to have been questionable. To do something about the transfer of knowledge measures is not that easy since it forms an integral part of the APIL battery. The job competency potential measure could, however, be improved and the study repeated.

Earlier (paragraph 3.3.5) the concern was raised that the job competency potential measure did not really reflect the ability to creatively use newly obtained knowledge in problem solving (i.e. did not reflect action learning). The fact that information processing capacity turned out to be the best predictor of learning performance reinforces this concern. To the extent that this was the case the need to replicate this study with a learning performance

measure that does reflect problem solving based on newly acquired knowledge becomes even more important.
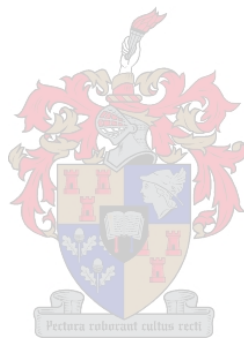
6.   REFERENCES

Ackerman, P.L. (1988). Determinants of individual differences during skill acquisition: cognitive abilities and information processing. *Journal of Experimental Psychology*. 117 (3), 288-318.

Anastasi, A., & Urbina, A. (1997). *Psychological Testing*. United States of America. Prentice-Hall.

Arvey, R.D., Faley, R.H. (1988). *Fairness in selecting employees*. United States: Addison-Wesley Publishing Company, Inc.

Austin, J.T., & Villanova, P. (1992). The criterion problem: 1917-1992. *Journal of Applied Psychology*, 77 (6), 836-874.

Babbie, E., & Mouton, J. (2001). *The Practice of Social Research*. Cape Town: Oxford University Press.

Berenson, M.L., Levine, D.M., Goldstein, M. (1983). *Intermediate Statistical Methods and Applications: a Computer Package Approach*. United Stated of America: Prentice-Hall, Inc.

Boeyens, J. (1989). *Learning Potential: an empirical investigation*. NIPR Special Report Pers 435. Pretoria: Human Sciences Research Council.

Bollen, K.A., & Long, J.S. (1993). Introduction. In K. A. Bollen, & J.S.Long (Eds.). *Testing Structural Equation Models* (pp.1-9). United States of America: SAGE Publications, Inc.

Boudreau, J.W. (1991). Utility Analysis for decisions in Human Resource Management. In M.D. Dunette & L.M. Hough, *Handbook of Industrial and Organisational Psychology*. California: Consulting Psychologists Press.

Browne, M.W., & Cudeck, R. (1993). *Alternative ways of assessing model fit*. In K.A. Bollen & J.S. Long (Eds.): Testing Structural Equation Models. Newbury Park: Sage Publications.

Byrne, B.M. (1989). *A primer of LISREL: Basic applications and programming for confirmatory factor analytic models*. New York: Springer Verlag.

Campbell, J.P. (1991). Modeling the performance prediction problem in industrial and organisational psychology. In M.D. Dunnette, & L.M. Hough (Eds.). *Handbook of*

Industrial and Organisational Psychology (Second edition; Volume 1). Palo Alto, California: Consulting Psychologists Press, Inc.

Carrell, M.R., Elbert, N.F., Hatfield, R.D., Grobler, P.A., Marx, M., van der Schyf, S. (1998). *Human Resource Management in South Africa*. Republic of South Africa: Prentice Hall.

Cascio, W.F., (1998). *Applied Psychology in Human Resource Management* (5th ed.). United States of America: Prentice-Hall.

Cattell, R.B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin.

Campione, J.C., Brown, A.L., Ferrara, R.A., Jones, R.S. & Steinberg, E. (1985). Breakdowns in flexible use of information: Intelligence-related differences in transfer following equivalent learning performance. *Intelligence, 9*(4), 297-315.

Cleary, T.A. (1968). Test Bias: Prediction of Grades of Negro and White Students in Integrated Colleges. *Journal of Educational Measurement, 5*(2), 115-124.

Cook, M. (1998). *Personnel Selection: Adding Value Through People* (3rd ed.). Chichester: John Wiley & Sons.

Cook, T.W. (1944). Repetition and Learning: I. Stimulus and Response. *Psychology Review*, *51*, 25-36.

Cross, E., Marais, W., Steel, H. & Theron, C.C. (2002). 'n Psigometriese oudit op 'n bestaande keuringsprosedure. *S.A. Journal of Industrial Psychology, 28*(2), 1-9.

De Beer, M. (2000). *Learning Potential Computerised Adaptive Test: User's Manual*. South Africa: Production Printers.

De Beer, M. (2003). *A Comparison of Learning Potential Results At Various Educational Levels*. Paper presented at the 6th Annual Society for Industrial and Organisational Psychology of South Africa (SIOPSA) conference, Johannesburg.

De Corte W. (1999). Weighing job performance predictors to both maximize the quality of the selected workforce and control the level of adverse impact. *Journal of Applied Psychology*, *84*(5), 695-702.

Diamantopoulos, A., & Siguaw, J.A. (2000). *Introducing LISREL*. London: Sage Publications.

Du Toit, M., & Du Toit, S.H.C. (2001). *Interactive LISREL: user's guide*. Lincolnwood, IL: Scientific Software International.

Du Toit, S.H.C. & Mels, G. (2002). Supplementary notes on multiple imputation. Retrieved June 2005 from http://www.ssicentral.com/other/impute.htm.

Dunnette, M.D. (1966). *Personnel Selection and Placement*. United States of America: Wadsworth Publishing Company.

Elmes, D.G., Kantowitz, B.H., Roediger, H.L. (1999). *Research Methods in Psychology* (6th ed.). United States of America: Brookes/ Cole.

Estes, W.K. (1978). The Information-Processing Approach to Cognition: A Confluence of Metaphors and Methods. In W.K. Estes (Ed.). *Handbook of Learning and Cognitive Processes: Human Information Processing* (Vol. 5)(pp. 1-18). New Jersey: Lawrence Erlbaum Associates, Publishers.

Eysenck, H.J. (1986). The Theory of Intelligence and the Psychophysiology of Cognition. In R.J. Sternberg (Ed.), *Advances in the Psychology of Human Intelligence¸ Vol. 3*. Hillsdale (NJ): Lawrence Erlbaum.

Ferguson, G.A. (1954). On Learning and Human Ability. *Canadian Journal of Psychology*, *8*(2), 95-112.

Ferguson, G.A. (1956). On Transfer and the Abilities of Man. *Canadian Journal of Psychology*, *10*(3), 121-131.

Ferretti, R.P., & Butterfield, E.C. (1992). Intelligence-related differences in the learning, maintenance, and transfer of problem-solving strategies. *Intelligence, 16*, 207-224.

Gatewood, R.D., & Feild, H.S. (1990). *Human Resource Selection* (2nd ed.). United States of America: The Dreyden Press.

Gatewood, R.D., & Feild, H.S. (1994). *HumanRresource Selection*. Fort Worth: The Dryden Press.

Ghiselli, E.E., Campbell, J.P. & Zedeck, S. (1981). *Measurement Theory for the Behavioural Sciences*. San Francisco: W.H. Freeman and Company.

Gibson, J.L., Ivancevich, J.M. (Jr.), Donnelly, J.H. (1997). *Organisations: Behavior, Structure, Processes*. United States of America: Irwin/ McGraw-Hill.

Gouws, L.A., Louw, D.A., Meyer, W.F. & Plug, C. (1979). *Psigologiewoordeboek*. Johannesburg: McGraw-Hill.

Guion, R.M. (1991). Personnel Assessment, Selection and Placement. In M.D. Dunnette & L.M. Hough (Eds.). *Handbook of Industrial and Organisational Psychology*. California: Consulting Psychologists Press.

Guthke, J. (1993). 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. 13-18). Lisse: Swets & Zeitlinger.

Guttman, L. (1965) A faceted definition of intelligence. In R. Eiferman (Ed.), *Scripta Hierosolymitana Studies in Psychology*, *14*, (pp.166-181). Jerusalem: Hebrew University.

Hamers, J.H.M., Resing, W.C.M. (1993). Current Trends in Theories and Assessment of Intelligence. In J.H.M. Hamers, K. Sijtsma & A.J.J.M. Ruijssenaars, *Learning Potential Assessment: Theoretical, Methodological and Practical Issues* (pp. 13-18). Lisse: Swets & Zeitlinger.

Hattrup, K., Rock, J. & Scalia, C. (1997). The effects of varying conceptualisations of job performance on adverse impact, minority hiring and predictor performance. *Journal of Applied Psychology*, *82*, 656-664.

Hebb, D.O. (1949). *The Organisation of Behavior*. New York: Wiley.

Holborn, P.T. (1991).Selection decisions: The quest for fairness. NIPR Special Report Pers 424. Pretoria: Human Sciences Research Council.

Hough, L.M. (2000). Personnel Selection: Looking Toward the Future-Remembering the past. *Annual Review of Psychology (Annual 2000)*. Retrieved July 2005 from http://www.findarticles.com/cf_0/m0961/2000_Annual/61855639/print.jhtml).

Hough, L.M., Oswald, F.L. & Ployhart (2001). Determinants, detection and amelioration of adverse impact in personnel selection procedures: issues evidence and lessons learned. *International Journal of Selection and Assessment*, *9*(1), 152-194.

Hu, L.T., & Bentler P.M. (1995). Evaluating model fit. In R.H. Hoyle (Ed.) *Structural Equation Modelling; Concepts, Issues and Applications.* Thousand Oaks, California: Sage Publications.

Hunt, E. (1980). Intelligence as an Information Processing Concept. *British Journal of Psychology*, *71*, 449-474.

Jensen, A.R. (1998). *The g Factor: The Science of Mental Ability*. Connecticut: Praeger Publishers.

Jones, G.R. (2001). *Organisational Theory: Texts and Cases* (3rd edition). New Jersey: Prentice Hall.

Jöreskog, K.G., & Sörbom, D (1993). *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language*. United States of America: Scientific Software International, Inc.

Jöreskog, K.G., & Sörbom, D. (1996a.) *LISREL 8: User's reference guide*. Chicago: Scientific Software International.

Jöreskog, K.G., & Sörbom, D. (1996b). *PRELIS 2: User's reference guide*. Chicago: Scientific Software International.

Jöreskog, K.G., & Sörbom, D. (2003). *LISREL 8: New Statistical Features (Third Printing with Revisions*. Lincolnwood, IL: Software International, Inc.

Jöreskog, K.G. (2003). *Description of the LISREL model*. Retrieved March 2006 from http://www.ssicentral.com/lisrel/define.htm.

Kaplan, D. (2000). *Structural Equation Modeling; Foundations and Extensions*. Thousand Oaks, California: Sage Publications.

Kelloway, E.K. (1998). *Using LISREL for Structural Equation Modelling: A Researcher's Guide*. United States of America: SAGE.

Kerlinger, F.N. (1973). *Foundations of Behavioural Research* (2nd ed.). United States of America: Holt, Rinehart and Winston, Inc.

Kerlinger, F.N. (1986). *Foundations of Behavioural Research*. New York: CBS Publishing.

Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of Behavioural Research.* (4th ed.). Fort Worth: Harcourt College.

Learning Potential Assessment (2003). Retrieved 5 July 2005 from http://www.ads.za.net/cuture1.htm.

MacCallum, R.C. (1995) Model specification: procedures, strategies and related issues. In R.H. Hoyle (Ed.). *Structural Equation Modelling; Concepts, Issues and Applications*. Thousand Oaks, California: Sage Publications.

MacCallum, R.C., Browne, M.W. & Sugawara, H.M. (1996). Power analysis and determination of sample size for covariance structure modelling. *Psychological Methods*, 1(2), 130-149.

McGeoch, J. A. (1946). *The Psychology of Human Learning: An Introduction*. New York: Longmans Green and Co.

Mels, G. (2003). *A workshop on structural equation modeling with LISREL 8.54 for Windows*. University of Port Elizabeth.

Milkovich, G.T., & Boudreau, J.W. (1994). *Human Resource Management* (7th edition). Boston: Irwin.

Noe, R.A., Hollenbeck, J.R., Gerhart, B., Wright, P.M. (2000). *Human Resource Management: Gaining a Competitive Advantage* (3rd ed.). United States of America: McGraw-Hill.

Petersen, N.S. & Novick, M.R. (1976). An evaluation of some models for culture-fair selection. *Journal of Educational Measurement*. *13* (1), p. 3-29.

Popper, K.R. (1972). *Conjectures and Refutations; the Growth of Scientific Knowledge*. London: Routledge and Paul.

Sackett, P.R. & Roth, L. (1996). Multi-stage selection strategies: a monte carlo investigation of effects on performance and minority hiring. *Personnel Psychology*, *49*, 549-572.

Sackett, P.R., Schmitt, N., Ellingson, J.E. & Kabin, M.B. (2001). High-stakes testing in employment, credentialing, and higher education; prospects in a post-affirmative-action world. *American Psychologist*, *56*(4), 302-318.

Saville & Holdsworth (SHL). (2000). Competency design; towards an integrated human resource management system. *SHL Newsline*, March, 7-8.

Saville & Holdsworth (SHL). (2001). Competencies and performance@work, *SHL Newsline*, May,6.

Schippmann, J.S., Ash, R.A., Battista, M., Carr, L., Eyde, L.D., Hesketh, B., Kehoe, J., Pearlman, K., Prien, E.P. & Sanchez, J.I. (2000). The practice of competency modelling. *Personnel Psychology*, *53*, 703-740.

Schmidt, F.L. & Hunter, J.E. (1981). Employment testing; old theories and new research findings. *American Psychologist*, *36*(10), 1128-1137.

Snow, R.E., Kyllonen, P.C., Marshalek, B. (1984). The topography of ability and learning correlations. In R.J. Sternberg (Ed.). *Advances in the Psychology of Human Intelligence, Vol. 2*. Hillsdale (NJ): Lawrence Erlbaum.

Society for Industrial Psychology (SIP). (1998). *Guidelines for the Validation and Use of Assessment Procedures for the Workplace*. Johannesburg: SIP.

Spearman, C.E. (1904). "General Intelligence" Objectively Determined and Measured. *American Journal of Psychology*, *15*, 201-293.

Spearman, C.E. (1927). *The Abilities of Man, their Nature and Measurement*. London: Macmillan.

SPSS (2006). Retrieved January 2006 from (http://www.spss.com).

Statt, D.A. (1998). *The Concise Dictionary of Psychology* (3rd edition.). New York: Routledge.

Sternberg, R.J. (1984). Toward a triarchic theory of human intelligence. *The Behavioral and Brain Sciences*, *7*, 269-271.

Tabachnick, B.G. and Fidell, L.S. (1989). *Using Multivariate Statistics*. New York: Harper Collins Publishers.

Taylor, T.R. (1989). International Developments in Psychological Assessment. *Congress on Psychometrics for Psychologists*. Megawatt Park, Sandton: Eskom and the Institute of Industrial Psychology of South Africa.

Taylor T.R. (1992). Beyond competence: measuring potential in a cross-cultural situation fairly: potential in psychometrics: Part two. *Congress on Psychometrics for Psychologists*. Megawatt Park, Sandton: Eskom and the Society of Industrial Psychology of South Africa.

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

Taylor, T.R. (1997). *Administrator's Manual for APIL BATTERY*. Auckland Park: AProLAB.

Taylor, T.R. (2006). *Technical Manual for APIL BATTERY*. Auckland Park: AProLAB.

Theron, C.C. (2001). Unpublished class notes (Research Methodology and Psychometrics 778/ 788, part 2), University of Stellenbosch.

Theron, C.C. (2002). Learning Potential Measurement. Unpublished class notes (Psychometrics and Research Methodology, 244), University of Stellenbosch.

Thorndike, R.L. (1982). *Applied Psychometrics*. United States of America: Houghton Mifflin Company.

Tredoux, C.G., & Durrheim, K.L. (2002). *Numbers, Hypotheses and Conclusions.* Cape Town: Juta.

Underwood, G. (1978) *Strategies of Information Processing.* London: Academic Press.

United States. Equal Employment Opportunity Commission, Civil Service Commission, Department of Labor, & Department of Justice. (1978). Uniform guidelines on employee selection procedures. *Federal Register*, August 25, *43* (No. 166), 38290-38315.

Vygotsky, L.S. (1978). *Mind in Society: The Development of Higher Psychological Processes.* United States of America: Harvard University Press.

Appendix A

Path diagram for full LISREL model
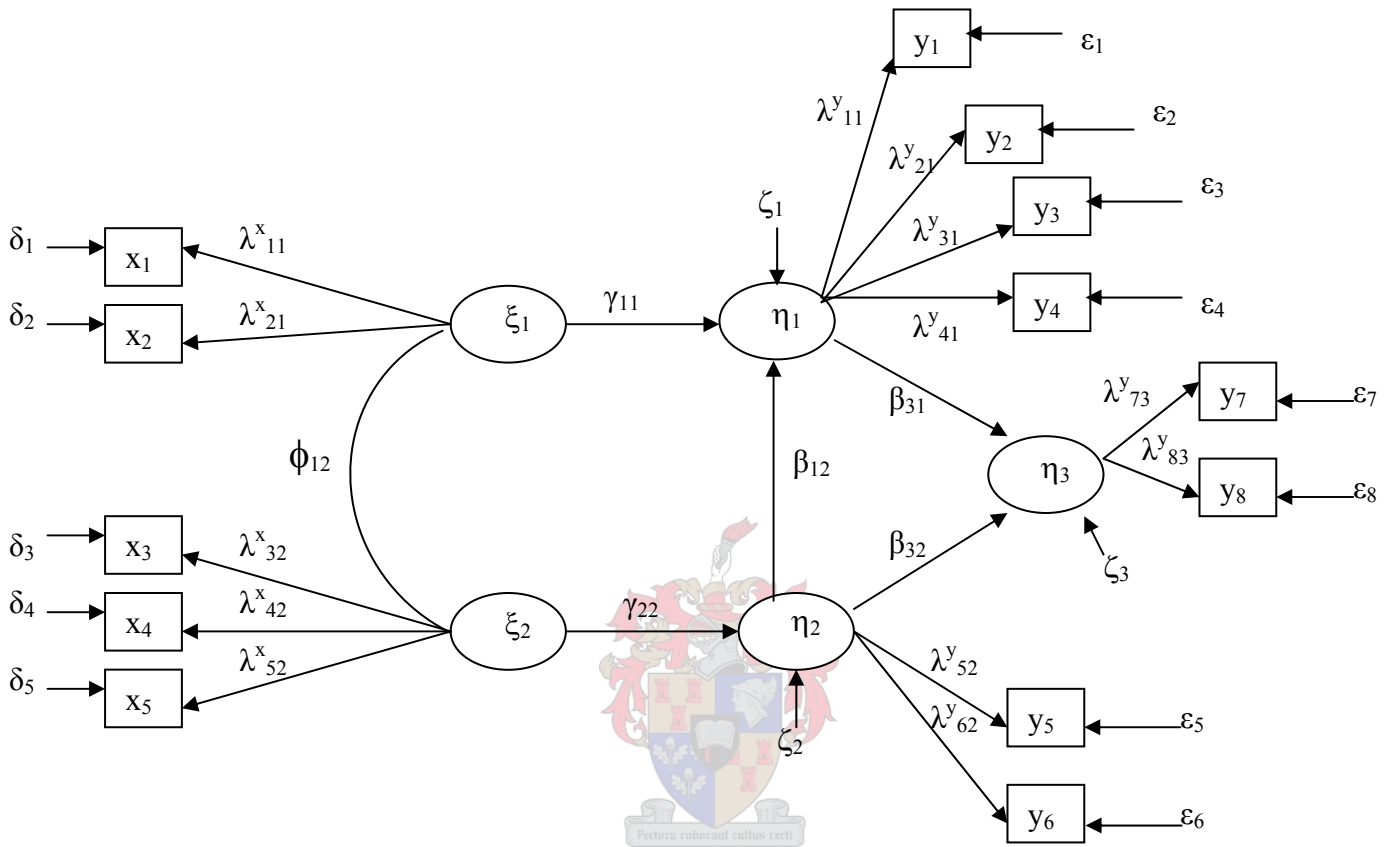
Symbol	Interpretation
ξ_1	Abstract reasoning capacity
ξ_2	Information processing capacity
η_1	Transfer of knowledge
η_2	Automatization
η_3	Job competency potential

Appendix B

Path diagram for observed variable regression model