VALIDATION OF A SELECTION BATTERY USED BY THE SOUTH AFRICAN MILITARY ACADEMY

Marlize Pretorius
Department of Industrial Psychology
University of Stellenbosch
South Africa

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

Supervisor
Prof CC Theron
Department of Industrial Psychology
Faculty of Economic and Management Sciences

December 2010

DECLARATION

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the owner of the copyright thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Copyright © 2010 Stellenbosch University
All rights reserved

ABSTRACT

The objective of this study is to determine whether the psychometric evaluation procedure, used by the South African Military Academy to make selection decisions, can validly predict academic performance of first year learners, whether this procedure is fair and whether the procedure is efficient. The sample used for this study consisted of three year groups (First Year Students of 2001, 2002 and 2003) enrolled at the Military Academy. In theory specific learning behaviours (learning competencies) are instrumental in attaining academic performance. These learning behaviours, in turn, depend on and are expressions of a complex nomological network of person-centered characteristics (learning competency potential). Differences in learning performance can be explained in terms of learning behaviours. Learning competencies are instrumental in achieving the learning outcomes for which the academic programme exists.

Learning competencies, in turn, can be explained in terms of learner characteristics. In order to differentiate between candidates who have better or poorer training prospects in terms of a construct orientated approach to selection, a performance hypothesis on the person-centered drivers of the learning competencies is used. It is argued that the degree of competence in: (1) the core cognitive processes/competencies that constitute learning (transfer and automatization) and are necessary to create meaningful structure in novel learning material, (2) the intellectual drivers of these learning competencies (fluid intelligence and information processing capacity), (3) proficiency in English and (4) past academic performance, should discriminate between better or poorer academic performance of learners attending the academic programmes at the SA Military Academy. The grade point average of the first year first semester academic results is used as a measure of the criterion construct.

Almost all of the results obtained in this study support the theory and propositions made by the performance hypothesis. Only one variable, accuracy of information processing, did not perform as predicted by the performance hypothesis. Prior learning explained the most variance in the criterion ($r=0,431^2$). The inter-correlation amongst the predictors is used to infer the proportion of unique variance each predictor accounts for in the composite criterion. A regression of the composite criterion on the array of predictors (X_2 – X_{12}) revealed that only memory and understanding (X_9) and prior learning (X_{12}) uncovered

relevant and unique information about determinants of performance on the criterion not conveyed by the remaining predictors in the model. The remaining predictors in the selection battery can consequently be considered redundant since they provide no new information not already conveyed by X_9 and X_{12}. When Y_{GPA} is regressed on the weighted combination of X_9 and X_{12}, only X_{12} significantly explains unique variance in Y_{GPA} when included in a regression model already containing X_9. In the light of the reported findings there is no need to create a combined weighted linear predictor composite (X_{comp}) which would form the basis of the actuarial mechanical decision rule that would guide selection decisions. Prior learning proved to be the only predictor that warrants inclusion in the actuarial mechanical prediction rule that will form the basis of selection decisions. In terms of the derived actuarial prediction rule the expected criterion performance of all applicants ($E[Y|X_{12}]$) could consequently be estimated by inserting the measures obtained during selection of prior learning into the derived regression equation. The use of this equation could be regarded as permissible to the extent to which $E[Y|X_{12}]$ correlates significantly with Y_{GPA}. Since $E[Y|X_{12}]$ correlates 0,431 and statistically significantly ($p<0,05$) with Y_{GPA}, the predictions derived from this equation are valid.

The findings of this research suggest that black and white students were sampled from the same population and therefore the use of the single, undifferentiated prediction rule would lead to fair selection decisions. To answer the question whether the selection procedure under investigation is adding any value to the organization, utility analysis is done based on the Taylor-Russell utility model as well as the Naylor-Shine interpretation of selection utility. A criterion-referenced norm table that expresses the risk of failure conditional on expected academic performance is derived from the use of only X_{12}. Recommendations for further research are put forward.

OPSOMMING

Die doel van hierdie studie is om te bepaal of die psigometriese evaluasie-prosedure wat deur die Suid Afrikaanse Militêre Akademie gebruik word vir keuringsbesluite, akademiese prestasie van eerstejaar leerders geldig voorspel, en of hierdie prosedure regverdig en effektief is. Die steekproef vir hierdie studie bestaan uit drie jaargroepe (eerstejaar studente van 2001, 2002 en 2003) wat ingeskryf was by die Militêre Akademie. Teoreties is daar spesifieke leergedrag (leerbevoegdhede) wat instrumenteel is in die bereiking van akademiese prestasie. Hierdie leergedrag hang af van en is weer ‘n uitdrukking van ‘n komplekse nomologiese netwerk van persoongesentreerde eienskappe (leerbevoegdheidspotensiaal). Verskille in leerprestasie kan verklaar word in terme van leergedrag. Leerbevoegdhede is instrumenteel in die bereiking van die leeruikomste waarvoor die akademiese program bestaan. Leerbevoegdhede, op sy beurt, kan weer verklaar word in terme van leerdereienskappe.

Ten einde ‘n onderskeid te kan tref tussen kandidate met beter of slegter opleidingsvooruitsigte, in terme van ‘n konstrukgeorienteerde benadering tot keuring, word ‘n prestasiehipotese gebruik wat gebaseer is op die persoongesentreerde drywers van die leerbevoegdhede. Dit word aangevoer dat die graad van bevoegdheid in: (1) die kern kognitiewe prosesse/bevoegdhede waaruit leer bestaan (oordrag en outomatisasie) en wat nodig is om sinvolle struktuur in nuwe leermateriaal te skep, (2) die intellektuele drywers van hierdie leerbevoegdhede (vloeibare intelligensie en informasieverwerkingskapasiteit), (3) bevoegdheid in Engels, en (4) vorige akademiese prestasie sal onderskei tussen beter of slegter akademiese prestasie van leerders wat akademiese programme by die SA Militêre Akademie bywoon. Die gemiddelde van eerstejaar eerste semester akademiese uitslae is gebruik as meting van die kriteriumkonstruk.

Byna al die resultate wat in hierdie studie verkry is ondersteun die teorie en proposisies soos aangevoer deur die prestasiehipotese. Slegs een veranderlike, akkuraatheid van informasie-prosessering, het nie gereageer soos voorspel deur die prestasiehipotese nie. Vorige leer het die meeste variansie in die kriterium verklaar ($r=0,431^2$). Die inter-korrelasie tussen die voorspellers is gebruik om die proporsie unieke variansie wat elke voorspeller in die saamgestelde kriterium verklaar te skat. ‘n Regressie van die saamgestelde kriterium op die reeks voorspellers ($X_2 – X_{12}$) toon aan dat slegs geheue en

begip (X_9) sowel as vorige leer (X_{12}) relevante en unieke informasie in verband met die determinante van prestasie in die kriterium weergee wat nie reeds weergegee word deur die oorblywende voorspellers in die model nie. Die oorblywende voorspellers in die keuringsbattery kan gevolglik as oorbodig beskou word aangesien hulle geen nuwe informasie verskaf wat nie reeds deur X_9 en X_{12} oorgedra word nie. Wanneer Y_{GPA} geregresseer word op die geweegde kombinasie van X_9 en X_{12}, verklaar slegs X_{12} unieke variasie in Y_{GPA} wanneer dit ingesluit word in ‘n regressiemodel wat alreeds X_9 bevat. In die lig van die gerapporteerde bevindinge is dit onnodig om ’n gekombineerde geweegde liniêre voorspellerkombinasie (X_{comp}) te skep om as basis van ‘n aktuariële meganiese besluitnemingsreël te dien aan hand waarvan keuringsbesluite geneem sal word. Vorige leer blyk die enigste voorspeller te wees wat insluiting regverdig in die aktuariële meganiese besluitnemingsreël wat die basis van keuringsbesluite sal vorm. In terme van die afgeleide aktuariële besluitnemingsreël sal die verwagte kriteriumprestasie van alle toekomstige aansoekers ($E[Y|X_{12}]$) geskat word deur die meting van vorige leer verkry tydens keuring in die afgeleide regressievergelyking in te stel. Die gebruik van hierdie vergelyking kan as toelaatbaar beskou word in die mate waartoe $E[Y|X_{12}]$ betekenisvol met Y_{GPA} korreleer. Aangesien $E[Y|X_{12}]$ statisties betekenisvol 0,431 ($p<0,05$) met Y_{GPA} korreleer, kan die voorspellings afgelei vanuit hierdie vergelyking as geldig beskou word.

Die bevindinge van hierdie navorsing dui daarop dat swart en wit studente van hierdie steekproef uit dieselfde populasie geneem is en daarom sal die gebruik van ‘n enkele, ongedifferensieerde voorspellingsreël lei tot regverdige keuringsbesluite. Om ‘n antwoord te verkry op die vraag of hierdie keuringsprosedure enige waarde tot die organisasie toevoeg is ‘n nutanaliese gedoen wat gebaseer is op Taylor-Russell se nutmodel so wel as die Naylor-Shine interpretasie van keuringsnut. ‘n Kriteriumgerigde normtabel, wat die voorwaardelike risiko op mislukking gebaseer op akademiese prestasie uitdruk, is afgelei deur die gebruik van slegs X_{12}. Aanbevelings vir verdere navorsing word voorgestel.

ACKNOWLEDGEMENTS

Thank you to every person who assisted me in realizing this project. Without you this end result would not have been possible, in particular thanks to Prof Theron and Dries.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION AND OBJECTIVE OF THE STUDY

1.1 INTRODUCTION

Organizations are man-made phenomena established for a definite reason and with a specific purpose. In order to reach the specific goal for which the organization was established, an organization has to combine and transform scarce resources into products and/or services with maximum utility. This is also true of the South African National Defence Force (SANDF). As a service organization, the SANDF is confronted with a choice of alternative utilisation possibilities regarding the limited resources of the Department of Defence. The SANDF is guided in this choice by the economic principle, which commands, on behalf of society, that this institution should strive to attain the highest possible output of quality services with the lowest possible input of resources. This institution should comply with the demand of the economic principle because such compliance would allow it to maximize its service delivery. In order to deliver optimal services to society, however, maximum utility must be designated as an important organizational goal. This objective of the SANDF therefore is the maximisation of the value of its services over a particular period relative to the resources used to deliver those services.

In order to achieve the above-mentioned objective in any service organization, a multitude of mutually coordinated activities need to be performed. These activities can be categorized as a system of inter-related organizational functions. The human resource function represents one of these organizational functions. The human resource function aspires to contribute towards organizational objectives through the acquisition and maintenance of a competent and motivated work force, as well as the effective and efficient utilisation of such a work force (Nel, Gerber, Van Dyk, Haasbroek, Schultz, Sono & Werner, 2001). Government, depending on the country's specific situation, prescribes the utilisation of its defence capability which in turn dictates the strategic goals of the SANDF. The strategic needs determine the acquisition, maintenance and utilisation of soldiers. The importance of human resource management therefore flows from the basic premise that combat readiness of soldiers and organizational success is significantly dependent on the quality of the SANDF's work force and the way the work force is utilized and managed.

Labour constitutes the most important resource of the SANDF due to the fact that this institution is managed, operated and run (i.e., commanded and lead) by people. Labour is therefore the heartbeat of this organization through which all other factors are combined and mobilized for service delivery. Evidently labour represents the factor which determines the cost effectiveness and efficiency with which the other factors of production are utilized (Milkovich & Boudreau, 1994).

The management of human resources is, however, complicated by the intricate, and to a certain extent enigmatic, nature of the working person as the carrier of labour as production factor. The behaviour of the working person is nonetheless not a random walk in the work place. The performance of the working person is rather the systematic expression of a complex nomological network of influencing variables characterising the individual and his or her working environment. This leads to the basic premise that credible and valid theoretical explanations for the different facets of the behaviour of the working person constitute a fundamental and indispensable, though not sufficient, prerequisite for efficient and equitable human resource management. Although a perfect understanding and complete certainty about the nature of the nomological network of variables governing the performance of the working person will probably never be possible, Industrial/Organizational Psychologists have, nonetheless succeeded to produce credible, and valid, albeit limited, (close fitting) theoretical explanations for the different facets of the behaviour of the working person. This in turn provides, through deductive inference, the opportunity to derive practical human resource interventions designed to affect either employee flows or employee stocks (Boudreau, 1991; Milkovich & Boudreau, 1994).

Interventions designed to affect the flow of employee attempt to change the composition of the work force by adding, removing or reassigning employees (e.g. through recruitment, selection, turnover, or internal staffing) with the expectation that such changes will manifest in improvements in employee performance and eventually organizational performance. In contrast, interventions designed to affect employee stock attempt to change the characteristics of the existing work force in their current positions or the work situation itself (e.g. through skills development[1], performance feedback, compensation or job redesign). The expectation is again that such changes will manifest in improvements in work performance (Boudreau, 1991; Milkovich & Boudreau, 1994). Improvements in work performance are affected through increases in work force quality, which in turn are brought about by the aforementioned two types of human resource interventions.

Personnel selection is probably the single most important human resource intervention aimed at affecting employee flows into, through and out of the organization. Selection normally implies a situation where there are more applicants than the number of available job or training and developmental vacancies. The objective of selection is to fill the available number of vacancies with those applicants who will eventually optimally succeed in the job or training. Selection is meant to be a value adding process. Effective selection adds value to an organization by ensuring that the right quality and quantity of employees are put in the right work or training positions at the right time in order to contribute towards the functioning of the organization (Nel et al., 2001). During the process applicants are put through a series of steps to determine which candidates are likely to be successful and eliminate those likely to fail (Nel *et al.*, 2001). The objective of personnel selection is to optimize employee work or training performance by appropriately assigning applicants to either an accept or a reject treatment (Cronbach & Gleser, 1965). Individuals who are selected are expected to perform better than rejected applicants (Guion, 1991; 1998).

Given this objective, the phenomenon of interest in personnel selection is the criterion construct job or training performance/success (η). If the selection process is to contribute to the organization's success, selection decisions should be focused on the comprehensive performance construct (Werther & Davis, 1993). The ultimate criterion (job or training performance/success) always remains the focus of interest in selection decision-making (Ghiselli, Campbell & Zedeck, 1981). This seemingly innocent but too often neglected fact has powerful implications for the interpretation and evaluation of information entering the selection decision.

Sufficient understanding and adequate detail of the work or training position is required to constitutively define the criterion construct in terms of the latent performance dimensions (η_i) comprising success. Sufficient understanding and adequate detail of the work or

[1] Education, training and development

training position is also required to operationalize the performance construct in terms of the behavioural denotations in which the ultimate criterion expresses itself. Measurement of job or training success should be based on these position-relevant behavioural denotations. The position-relevant criteria are identified and chosen through an understanding of the duties and responsibilities of the position (job description), as well as an understanding of the organization's strategic needs (Guion, 1991; 1998). A job analysis should thus be performed to establish the job description for a specific work or training position. A job analysis is a procedure used to develop insights into the components of a position. It should provide the necessary detailed information of the components of the position, as well as a sufficient understanding of the position. Job components include things such as the activities people perform in the position, resources they draw on when performing those activities, and the organizational implications of performing it well or poorly (Nel et al., 2001).

The ideal would therefore be to base selection decisions on valid and reliable measures (Y_i) of the criterion construct (η_i) personnel selection is meant to affect. This information is, however, not directly available at the time of the selection decision. Under these circumstances, and in the absence of any (relevant) information on the applicants, no possibility exists to improve the quality of the decision making over that which would have been obtained by chance. The only alternative to random decision-making (other than not to take any decision at all) would be to base the decision on predictions of the criterion rather than on direct measures of it. An accurate prediction of Y ($E[Y|X_i]$) will only be possible from information available at the time of the selection decision (X_i) if such information systematically correlates with a valid and reliable measure of the criterion and the nature of this relationship is known (Theron, 2001). An accurate understanding of this predictor-criterion relationship would enable 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. The selection decision would then be based on the expected criterion performance of applicants. It would be considered permissible to do so because of the systematic relationship existing between Y and ($E[Y|X_i]$).

Only two possible alternatives exist to obtain accurate predictions of the criterion (Binning & Barrett, 1989). The first option is to identify the dimensions of the performance construct or competencies required to successfully deliver the outputs for which the position[2] exists (inferred from the job description) and to operationalize these competencies in a simulated or natural work environment corresponding to the position in question in terms of the demands placed on the incumbent. This could be termed a content orientated approach to personnel selection (Binning & Barrett, 1989). The level of competence an applicant will eventually demonstrate on the competencies should he/she be placed in the position would not be a random event but rather an expression of a complex nomological network of person-centered attributes. Although their identity might not be known this complex nomological network of person-centered attributes will also determine the performance level achieved in the simulated work environment. The predictor and criterion scores are expected to correlate in a content orientated approach to selection because of the common source of variance the two measures. It should therefore be possible to reasonably accurately estimate the latter from the former provided that the manner in which the criterion and predictor is related is accurately understood.

The second option is to infer these critical incumbent attributes that determine the level of criterion performance that would be attained from the description of the position content and context. These critical attributes are unfortunately also sometimes referred to as competencies (Spangenberg, 1990) thereby creating considerable confusion, misunderstanding and discord in contemporary psychometric debate. The presumed interrelationship between these hypothesized determinants and the way they collectively combine in the criterion is postulated in a nomological network or latent structure (Campbell, 1991) as a complex performance hypothesis explaining criterion performance in the job in question. These hypothesized determinants of criterion performance, or a person centered subset thereof, could, to the extent that the tentative performance theory is indeed valid, be used in combined form to derive estimates of the, still to be realized, actual criterion scores. This could be termed a construct orientated approach to personnel selection (Binning & Barrett, 1989). The way these hypothesized determinants of performance should be combined is suggested by the way these determinants are linked in the postulated nomological network (Theron, 2001). The hypothesized nomological network of predictor and criterion constructs can be depicted as a structural model and tested by means of structural equation modelling (Diamantapoulos & Sigauw, 2001). If such a performance structural model would fit predictor and criterion data satisfactory the

---

[2] Position should here be interpreted to refer to either a job or a training position.

possibility of estimating latent criterion scores via the structural model parameter estimates creates a provocative alternative way of estimating criterion performance over the traditional multiple regression approach to personnel selection.

Both options obtain predictions of the criterion by measuring constructs through observable behaviour brought forward by a stimulus set (selection procedure). In the first option (content orientated approach) the selection method is designed to obtain the same response as actual facets of the work or training position would have brought forward. The same performance constructs that would have been measured in the work or training position during performance appraisal or evaluation of training are therefore measured but they are measured off the job during selection. Although a network of person constructs determines the applicant's reaction to the stimulus set, the nature of these constructs are not necessarily known or specified. In the second option (construct oriented approach) selection methods are designed so that applicants' responses to them are primarily a function of specific, defined person constructs, presumed to be determinants of the level of competence that would eventually be achieved on the job competencies (or performance constructs).

The basic question, from which a construct orientated selection procedure is ultimately conceived, asks with rather deceptive simplicity why differences in work or training performance exist. Inability to answer this question in terms of a valid performance hypothesis effectively eliminates the possibility to differentiate between better and poorer employment or training prospects in terms of a construct orientated approach to selection. The formation of a predictive hypothesis (also referred to as a performance hypothesis) is central to personnel selection and selection research under a construct-orientated approach. The performance hypothesis is based on an understanding of the position for which people are to be selected and on knowledge of the relevant background research because it consists of a specific, valued aspect of behaviour to be predicted (criterion), with one or more applicant traits hypothesized to predict it (predictors) (Guion, 1991). The outcomes for which the position exists, the competencies instrumental in achieving these desired outcomes, the person characteristics shaping the competencies and the situational characteristics moderating the effect of person characteristics on work or training behaviour, are all relevant in the formation of informed hypotheses about potential predictors of job or training success. To facilitate the prediction of success of potential employees or trainees, a comprehensive competency model is thus required. In other words, in order to predict possible job or training success it is important to know the individual competencies as well as the knowledge, skills, abilities, and other attributes that drive these competencies to obtain a comprehensive understanding of what makes employees successful on the job or to succeed in training (Sherman, Bohlander & Snell, 1998).

The performance hypothesis can be expressed in the form of a functional relationship, $\eta=f(\xi_i)$, in which η is the latent criterion variable and ξ_i is an array of latent predictor variables on which the criterion construct is dependent and thus can be used to predict the criterion phenomenon of interest at the time of decision making. The foregoing argument would concur with the point stressed by Guion (1991; 1998) that the hypothesis should be based on a clearly articulated reason to believe that the predictor set (ξ_i) is indeed relevant to, and would permit an accurate estimate of, the criterion (η) (Guion, 1991). The foregoing discussion would suggest that it probably would be more fruitful to express the performance hypothesis as a fully-fledged structural model of exogenous and endogenous latent variables (Diamantopoulos & Siguaw, 2000). This would, however, imply, as argued earlier, that multiple regression no longer would be the statistical estimation method of choice, but rather structural equation modelling (Diamantopoulos & Siguaw, 2000). This in turn leads to the question whether the latter, somewhat more involved, approach would enhance selection utility to an extent that would justify its use over the simpler, conventional approach.

If η and ξ_i can be operationalized in terms of valid and reliable indicator variables (Y and X_i) the latter can be used to obtain a prediction of the former. The relevance of the predictor measures from which the criterion estimates are derived is established through an extensive validation study as a form of applied explanatory research (Ellis & Blustein, 1991; Landy, 1986; Schmitt & Landy, 1993). What is being tested is in fact the performance hypothesis that variance the criterion construct is brought about by a network of predictor constructs. Landy (1986, pp. 1187-1188) supports this assertion, by stating:

> The validity analyst is carrying out traditional hypothesis testing. At least by implication, the hypothesis being considered is of the following form: People who do well on test X will do well on activity Y, or Y=f(X). Investigators should not lose sight of the fact that validity studies are attempts to develop a theory of

> performance that explains how an individual can (or will) meet the demands of a particular job.

Validity should be interpreted as the extent to which the inferences made from test scores are warranted; the extent to which the interpretation (i.e. meaning) assigned to test scores is justified (Guion, 1991; 1998). Strictly speaking, what is being validated is therefore not the measuring instrument, nor the measures obtained from the instrument, but rather the inferences made from the measures. Messick (1989, p. 13), in his influential and definitive treatment of the validity concept, stresses this when he states:

> Validity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on test scores or other modes of assessment. .... Broadly speaking, then, validity is an inductive summary of both the existing evidence for and the potential consequences of score interpretation and use. Hence what is to be validated is not the device as such but the inferences derived from test scores or other indicators - inferences about score meaning or interpretation and about the implications for action that the interpretation entails.

In the case of personnel selection the question a validation study needs to answer is whether inferences on the criterion may permissibly be made from the scores obtained on the predictors. In answering this question, however, more is involved than merely correlating the various predictors with the criterion. The different validity analysis strategies are not alternatives but rather form supplementary facets of a single unitary validity concept (Binning & Barrett, 1992; Ellis & Blustein, 1991; Guion, 1991; Messick, 1989; Schmitt & Landy, 1993) which all should come into play when validating a selection procedure. The data in terms of which the performance hypothesis is evaluated should be construct valid and reliable measures of the latent predictor and criterion variables comprising the hypothesis. Moreover, the data upon which selection decisions are based should be construct valid and reliable measures of those person characteristics that, according to the performance hypothesis, determine performance on the criterion in order to be useful as predictors of future job or training success (Sherman et al., 1998).

Once the case for the relevance of the predictor constructs has been successfully argued, the question on how to combine the information obtained from the various predictors to arrive at a selection decision arises. Two basic options exist in terms of which information can be combined for decision-making. Both options require that the nature of the relationship between the criterion and the substitute information should be accurately understood. The two options, however, differ in the way they express their understanding of the criterion - predictor relationship. The first option could be termed a judgmental, subjective or clinical mode of information combination since the decision outcome is derived from human judgement based on an inexplicit and unstandardized decision rule. The second option could be termed a mechanical, statistical or actuarial mode of information combination since an explicit and standardized rule or formula dictates the decision outcome (Gatewood & Feild, 1994; Grove & Meehl, 1996). An actuarial mode of information combination represents a mechanical prediction system to arrive at an overall inference about the expected criterion performance of an individual that was objectively derived via statistical or mathematical analysis from actual criterion and predictor data sets (Meehl, 1957; Murphy & Davidshofer, 1988). Within the mechanical option a number of different actuarial selection strategies can be distinguished.

A selection strategy in the current context refers to an explicit rule which determines, conditional on predictor measures, the assignment of applicants to one of two possible outcomes, namely terminal rejection or acceptation (Cronbach & Gleser, 1965; Gatewood & Feild, 1994). Reviews of the two approaches (Grove & Meehl, 1996; Gatewood & Feild, 1994; Murphy & Davidshofer, 1988) seem to suggest that clinicians very rarely make better predictions that can be made using actuarially derived prediction methods, that statistical methods are in many cases more accurate in predicting relevant criteria than are highly trained clinicians, and that clinical judgement should be replaced, wherever possible, by mechanical methods of integrating the information used in forming predictions (Murphy & Davidshofer, 1988). Gatewood & Feild (1994, p. 262) for example quite categorically argue in favour of the mechanical combination of selection data.

> The judgement of the decision maker can and should play an important part in data gathering (e.g., interview assessments), but should not play a major role in combining the various sources of information into a prediction about success. A mechanical formula/statistical model that is statistically derived and systematically applied is the best way to make accurate hiring decisions. When judgemental data are collected (e.g., interview assessments), it is better to convert those assessments to a rating and then enter the data into a statistical formula that combines the various data to make a prediction of job success.

Given the argument presented thus far it follows that effective selection will be possible under the construct orientated approach to the extent to which the nature of the relationship between the criterion construct and the person-centred variables influencing it can be accurately captured in a explicit mechanical prediction/decision rule. Stated differently, accurate predictions of the criterion construct are possible from measures of the predictor constructs only if the relationship between the criterion construct and the person-centred variables influencing it is understood accurately. One of the primary objectives of selection validation research is to actuarially derive a model/description of the relationship between the criterion construct and the person-centered variables influencing it so as to permit the accurate prediction of criterion performance on the basis of knowledge about predictor variables.

Multiple linear regression analysis would typically be used in the derivation of such a model. The objective of multiple linear regression analysis is to find a weighted linear combination of the individual information sources that minimizes the sum of the squared deviations between the linear combination and the actual criterion and thus that maximally correlates with the actual criterion (Tabachnick & Fidell, 1989). The multiple linear regression model assumes a linear relationship between the criterion Y and p predictor variables X_i that can, as a population model, be expressed as Equation 1.

$$E[Y \mid X_i] = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_p X_p \qquad 1$$

The development of a mechanical decision rule brings the question to the fore whether the criterion inference or prediction derived from Equation 1 is valid, in other words whether the criterion inference is permissible (Messick, 1989). Demonstrating that the predictor variables used in selection individually all correlate significantly with the criterion constitutes insufficient evidence to justify the use of the predictor variables for selection. It needs to be demonstrated that the manner in which the information obtained from the predictors is combined to infer/predict the degree of success applicants will achieve in a specific job correlates with the actual levels of success achieved. This important realisation often seems to be absent in validation studies, especially those, which combine selection information in accordance with a clinical or judgemental strategy.

If a high and significant $R(Y,E[Y|X_i])$ would be found, and all p regression parameter estimates would be significant, the selection decision rule can affect significant improvements in employee performance by controlling the quality of employees who enter the organization, move up in the organization or enter training and development interventions. The magnitude of the improvement in performance affected by the selection decision rule would increase linearly as $R(Y,E[Y|X_i])$ increases, (Brogden, 1946; 1949a; 1949b) but would also depend on situational characteristics like the selection ratio and the base rate (Cascio, 1991b). Demonstrating that a selection procedure affects significant improvement in work performance over that which would otherwise have been obtained would, however, not amount to sufficient evidence to justify that selection intervention. Given that the human resource function is included in the family of organizational functions based on the promise to contribute towards the primary organizational objective of maximizing the value of the organization for its owners, it logically follows that all interventions initiated by the human resource function should, in the final analysis, also be evaluated with the yardstick of profitability. The design, implementation and operation of a personnel selection procedure thus only make sense from an institutional perspective if a satisfactory (appropriately discounted) return on the capital invested in the selection procedure is achieved over the period in which the intervention generates its effect. There therefore rests a burden of persuasion on the human resource function to prove through appropriate financial indicators (Boudreau, 1991; Cronshaw & Alexander, 1985) that its selection procedures do add value to the organization (Cascio, 1991b). The burden of persuasion rest particularly heavy on the human resource function due to its general inability in the past to demonstrate its ability to contribute to bottom-line success (Cascio, 1991b).

Human resource selection constitutes a potent instrument enabling the human resource function to add value to the organization by virtue of its ability to regulate the quality and quantity of employees flowing into, through and out of the organization. Human resource selection procedures derive their ability to add value to the organization from their capability to discriminate between applicants in terms of attributes relevant to job performance. Selection measures are designed to discriminate and in order to accomplish its professed objective it must do so (Cascio, 1991a). Equal access to job or training opportunities for all current and aspirant employees would, from an institutional perspective, be considered irrational since it would nullify any institutional payoff that could otherwise have been derived from selection. However, due to the relative visibility of the selection mechanism's regulatory effect on the access to employment opportunities, the question readily arises whether the selection decision rule discriminates fairly. Section 6(1) of Chapter 2 of the Employment Equity Act (Republic of South Africa, 1998, p. 14) states that:

> No person may unfairly discriminate, directly or indirectly, against an employee in any employment policy or practice on one or more grounds, including race, gender, sex, pregnancy, marital status, family responsibility, ethnic or social orientation, colour, sexual orientation, age, disability, religion, HIV status, conscience, belief, political opinion, culture, language and birth.

There therefore also rests a burden of persuasion on the human resource function to prove through appropriate statistical indicators (Berenson, Levine & Goldstein, 1983) that its selection procedures are used to discriminate fairly between applicants (Cascio, 1991a).

1.2 SELECTION INTO THE ACADEMIC PROGRAMME OF THE SA MILITARY ACADEMY

The SA Military Academy is an educational institution of the South African National Defence Force (SANDF) and houses the Faculty of Military Science of the University of Stellenbosch. It provides university education and professional military development for young officers. This education aspires to equip these officers with knowledge, analytic skills and insight to be able to perform successfully as officers in the SANDF. Currently any junior officer can apply for studies at the SA Military Academy, given that they comply with certain requirements. All prospective learners have to go through a process of selection. However, the SANDF's Human Resource Strategy 2010 envisages a new system, which will regulate the manner in which members of the SANDF will be utilized in future (Defence Corporate Communication, 2003). This new strategy has a profound impact on the SA Military Academy's current resources. The SA Military Academy has limited capacity in terms of the total number of first year learners that can be accommodated.

In the new system there are three career stages, namely: (a) the Military Skills Development System (MSDS), (b) the Core Service System (CSS), and (c) the Senior Career System (SCS) (Department of Defence, 2003).

The MSDS represents the first career stage of members serving in the SANDF. Most new members who join the SANDF without any professional qualifications enter the organization through the MSDS. Candidates who wish to join the SANDF go through a selection process. Members who fit the profile of a soldier are enlisted into the Regular Force in the MSDS. MSDS members undergo full-time training and utilisation for a period of two years. During the first year of service in the MSDS, members undergo basic military training as well as basic functional training provided by the different functional training institutions of the SANDF. All MSDS members are assessed in terms of their leadership potential. Candidates identified for junior leader training undergo Field Section Leader training where after another selection board decides who will become officers and non commissioned officers. Since officers are expected to achieve a tertiary qualification, MSDS candidate officers without a degree undergo the Certificate in Military Studies at the SA Military Academy. Potential officers have to comply with the entry level requirements of the SA Military Academy. After this selection, the candidate officers and non-commissioned officers respectively receive officers' formative and junior leader training.

During the second year of MSDS service the junior non-commissioned officers and troops are utilized and deployed. MSDS candidate officers who have graduated at a university become officers and are utilized in accordance with the type of functional training they have received, whereas MSDS candidate officers without a tertiary qualification will undergo the Certificate in Military Studies at the SA Military Academy (Department of Defence, 2003). The Certificate in Military Studies is the first year of the B Mil Degree presented at this institution.

The CSS represents the second career stage of members serving in the SANDF (Regular Force). The goals of the CSS are to (a) provide the bulk of the junior and middle level leadership, (b) make provision for a contingent of enlisted members (privates) who have demonstrated the potential for further development, and (c) rejuvenate the HR component. Only those members who demonstrated development potential and performance proficiency to assume leadership and managerial positions will be selected for the CSS.
The first MSDS intake was in January 2003 and the first group of MSDS candidate officers reported at the SA Military Academy in January 2004. Candidate officers from the MSD System can, however, only enter the SA Military Academy's degree programmes if they comply with the relevant selection criteria imposed by the SA Military Academy (Minutes of

meeting on the MSDS intake at the Military Academy, 2003). In the current flow of events this would imply another selection process conducted by the SA Military Academy. A situation arose where candidates were selected to become officers but later on failed to comply with the relevant selection criteria set by the SA Military Academy and consequently were rejected by the SA Military Academy. This situation clearly was not tenable and posed tremendous problems to the system of identifying and developing military leadership potential. This selection process thus was far from optimal. An alternative system was utilized after realisation of the shortcomings of the selection system in use. Since the results of the current study were not yet available, an alternative selection process was decided on for the interim. The new system does not fall within the scope of the current study, and will therefore not be discussed.

Nonetheless, the aim of selection processes for studies at the SA Military Academy remains the same: in order to optimize the utilisation of limited resources, it is important that those members who are selected for studies at the SA Military Academy, should be academically successful. More specifically, to optimize the return achieved on the capital invested in the SA Military Academy programme, it is imperative to identify those individuals from the MSDS intake that would maximally benefit from the learning/development opportunity offered by the SA Military Academy. In order for a learner to be regarded as academically successful, he or she should complete their studies in the prescribed three-year time span. Moreover, the SA Military Academy would want to admit those learners who would deliver the highest possible academic performance. The performance criterion, for learners at this institution, therefore is academic success.

The question then arises if those learners who have performed well in their academic studies, will also perform well as officers in the SANDF once deployed? This question goes beyond the scope of this study, but should nevertheless be considered. Does good academic performance predict good performance of officers? Is the knowledge, skill and abilities developed by the academic programme of the SA Military Academy a necessary prerequisite in achieving successful officer performance in the SANDF? The implicit assumption seems to be that the knowledge, skill and abilities developed by the academic programme of the SA Military Academy are necessary but not sufficient conditions to achieve success as an officer in die SANDF. This line of reasoning could be interpreted to suggest that selection for admission into the Academy should take the form of a multiple hurdle strategy in which the first stage of selection occurs in terms of predicted success as an officer in the SANDF and the second stage occurs in terms of academic potential. The argument could also be interpreted the other way round. Both ways of looking at the problem, however, brings to the fore another troublesome question, namely what constitutes a good or successful officer?

To enter the SA Military Academy applicants have to comply with a number of specific selection criteria. These selection criteria only focus on an individual's ability to perform academically. During the SA Military Academy selection process, applicants are thus subjected to a number of specific assessment techniques. Selection decisions are based predominately on the results of a psychometric evaluation battery. In recent years the conventional psychometric tests, previously used during the selection process, were accused of being biased and under representing the cognitive capacity of individuals from historically disadvantaged backgrounds. These tests were subsequently replaced with a selection battery thought to be less susceptible to culture, race and gender bias. In as far as measurement bias would affect the validity of selection instruments (Millsap & Everson, 1993), the decision should be welcomed. This change, however, seems to have been motivated, at least in part, by the desire to comply with the Employment Equity Act's (Republic of South Africa, 1998) prohibition of unfair discrimination. To the extent that this had in fact been the case, the motivation behind the decision should be questioned. Selection fairness cannot be attained through the judicious choice of selection instruments. Neither can selection fairness be attained through the choice of unbiased selection instruments (Schmidt & Hunter, 1981; Theron, 2007). Nor can adverse impact be avoided through the judicious choice of selection instruments.

Up to 2003 the SA Military Academy administered the following predictors as part of a psychometric test battery to applicants to obtain information that is used to predict their future academic performance. The Ability, Processing of Information and Learning Battery (APIL-B) was used to measure the learning potential of candidates. To obtain an indication of whether a candidate commands the necessary English vocabulary and reading comprehension required to study in an environment where the medium of instruction is English, the Academic Aptitude Test (AAT) sub-test 3 (English Vocabulary) and sub-test 4 (English Reading Comprehension) was used. The Self-Directed Search (SDS) questionnaire was used to give an indication of a candidate's interests for the purpose of career counselling. Since the Self-Directed Search (SDS) questionnaire does not influence the selection decision itself, it is excluded from this study. Candidates' previous academic results (matriculation results) were also taken into account by the selection board.

The aforementioned selection procedure has an institutional as well as an individual impact. It firstly impacts on the academic success achieved by the SA Military Academy and eventually the performance of the officers of the SANDF. It, however, also has a significant impact on the personal lives of the individual applicants. The question is whether the selection procedure used to select candidate officers into the SA Military Academy can be justified in terms of its efficiency and fairness. With regards to the latter aspect, the Employment Equity Act (Republic of South Africa, 1998, p. 16) states that:

> Psychological testing and other similar assessments of an employee are prohibited unless the test or assessment being used (a) has been scientifically shown to be valid and reliable; (b) can be applied fairly to all employees; and (c) is not biased against any employee or group.

In addition paragraph 11 of the Employment Equity Act (Republic of South Africa, 1998, p. 16) requires that:

> Whenever unfair discrimination is alleged in terms of this Act, the employer against whom the allegation is made must establish that it is fair.

While it is true that the Employment Equity Act does not apply to members of the SANDF, it nonetheless remains possible for individual members of the SANDF to bring unfair discrimination claims before the Constitutional Court, or lodge complaints with the Human Rights Commission. Moreover, it could be argued that the Employment Equity Act (Republic of South Africa, 1998) simply formalized psychometric best practice; it forces organizations to perform the validation, fairness and utility studies they should have performed in their own self-interest anyway.

The aim of this research consequently, is to psychometrically evaluate the selection procedure used to select candidate officers into the academic programme of the SA Military Academy in a manner that would enable it to successfully meet the burden of proof implied by paragraph 11 of the Employment Equity Act.

1.3 RESEARCH OBJECTIVES

The foregoing argument would suggest that the selection procedure used to select candidate officers into the academic programme of the SA Military Academy would meet the burden of persuasion implied by paragraph 11 of the Employment Equity Act if it could be shown that:

- The learning performance hypothesis on which the selection procedure is based is true;
- The selection instruments offer reliable and construct valid measures of the exogenous and endogenous latent variables comprising the performance hypothesis;
- The learning performance inferences/predictions derived from the selection battery predictors correlate significantly with a reliable and valid measure of learning performance (or in terms of an alternative formulation the structural model corresponding to the performance hypothesis fits test and criterion data closely);
- The learning performance inferences/predictions are derived fairly from the measures obtained on the predictors;
- The fair use of the learning performance inferences affects an increase in the learning performance levels of selected candidate officers over that which would have resulted under random selection; and
- The value of the increase in learning performance of students exceeds the investment required to affect the improvement.

The specific objectives of the study consequently are:

- To test the propositions made by the performance hypothesis;
- To determine the predictive validity of the individual predictors of the selection battery;
- To derive a weighted linear prediction model actuarially from a set of predictor and criterion data;
- To determine the validity of the inferences derived from prediction model;
- To evaluate the fairness of the inferences/predictions derived from the prediction model and adapt the model if necessary;
- To evaluate the utility of the fair prediction model over random selection; and

- To develop a criterion-referenced norm table that expresses the risk of failure conditional on expected academic performance.

Ideally the study should also have investigated the fit of the structural model implied by the learning performance hypothesis on which the selection procedure is based. To accomplish this in addition to the above research objectives, however, seems to go beyond the scope of a study of this nature. The fit of the structural model implied by the learning performance hypothesis on which the selection procedure is based nonetheless remains an important concern that should be investigated empirically in subsequent research. De Goede (2007) has investigated the fit of the structural model underlying the APIL on a sample of student police officers. Reasonable model fit was obtained. Concerns about the adequacy of the measures used to operationalize the latent learning performance construct, however, necessitates further research on the model.

The study thus essentially aims at determining whether the psychometric evaluation procedure, used by the SA Military Academy to make selection decisions, can validly predict academic performance (success) of first year learners, whether this procedure is fair and whether the procedure is efficient.

CHAPTER 2
LEARNING PERFORMANCE HYPOTHESIS UNDERLYING THE SA MILITARY ACADEMY SELECTION PROCEDURE

2.1 EDUCATION, TRAINING AND DEVELOPMENT

Before a learning performance hypothesis is developed it is important to define learning in the context of this study. According to the Dictionary of Psychology (Plug, Louw, Gouws & Meyer, 1997), learning is an extensive term which encompasses a broad spectrum of connotative meaning. It can refer to the relative enduring change as the result of an experience, or to the processes from which these changes originated. All activities performed by a person who is learning are considered part of the learning process. Performance in learning can consequently be defined as the performance measured at the end of the learning process (experience) or it could also be defined in terms of the level of competence displayed in the behaviours that constitute learning. The former interpretation is typically used to determine (summatively) the success of the learning process. (Plug et al., 1997). The latter interpretation can, however, never be separated from the former interpretation. Crystallised abilities developed through learning has relevance for on the job performance largely because it serves as input in on-the-job action learning behaviours aimed at solving novel job-related problems.

Education, training and development are three related concepts with a central purpose - learning. Education can be formal or informal. Formal education is considered as the development of knowledge, attitudes, habits and personality characteristics (Plug et al., 1997). In addition, it can be described as the endeavour to transmit, evoke or acquire the above mentioned attributes as well as any learning that results from the attempt, intentional or unintentional (International Encyclopedia of Education, 1994). Furthermore, education is aimed at the development of cognitive processes to improve a person’s ability to understand and interpret knowledge. (De Cenzo & Robbins, 1994; Van Dyk, Nel, Loedolff & Haasbroek, 2001). Training is viewed as a systematic series of planned actions, such as instruction, exercise, revision, practical work as well as examinations, a person is exposed in order to change old or establish new knowledge, skills or behaviour in such a way that the organization’s objectives are achieved (Erasmus & Van Dyk, 1999; Plug et al.

1997). Training can therefore be regarded as a learning experience aimed at changing the individual to improve his or her ability to perform on the job (De Cenzo & Robbins, 1994).

Learning in the context of this study is regarded as any development activity that takes place in the process of developing good SANDF officers.

2.2 CREATING A CONCEPTUAL FRAMEWORK THROUGH THE CREATION OF A PERFORMANCE@LEARNING COMPETENCY MODEL

Competency modelling is seemingly a somewhat contentious topic in Industrial Psychology (Schippmann, Ash, Battista, Carr, Eyde, Hesketh, Kehoe, Pearlman, Prien & Sanchez, 2000). Nonetheless the competency model concept can serve as a powerful conceptual framework within which to develop a coherent performance hypothesis. Saville and Holdsworth (SHL) proposes a Universal Competency Framework which incorporates a model of performance at work that describes the relationships between competency potential, competency requirements as well as competencies (Bartram, 2006). According to Bartram (2006, p. 1) the SHL Universal Competency Framework represents:

> … a single underlying construct framework that provides a rational, consistent and practical basis for the purpose of understanding people’s behaviours at work and the likelihood of being able to succeed in certain roles and in certain environments.

As mentioned, the Universal Competency Framework incorporates the Performance@Work model. According to SHL (2001, p. 6):

> … “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.

In principle the same logic applies with regards to the education, training and development environment. Individuals are assigned to education, training or development treatments (opportunities) with the aim of achieving specific learning (education/training/development) objectives or outcomes (formulated in terms of performance competency potential attainments). These learning outcomes in the form of specific competency potential attainments are sought because they determine the level of competence achieved on job relevant competencies. Specific learning behaviours (learning competencies) are instrumental in attaining these desired outcomes. These learning behaviours, in turn, depend on and are expressions of a complex nomological network of person-centered 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 equivalent to the performance@work model originally proposed by Saville and Holdsworth (2001). Moreover the performance@learning model could be sequentially linked to the performance@work competency model. This provides a fertile conceptual model to explore the relationship between the characteristics required by a learner to be able to exhibit the learning behaviours needed to develop the qualities necessary to exhibit the work behaviours 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 this argument.

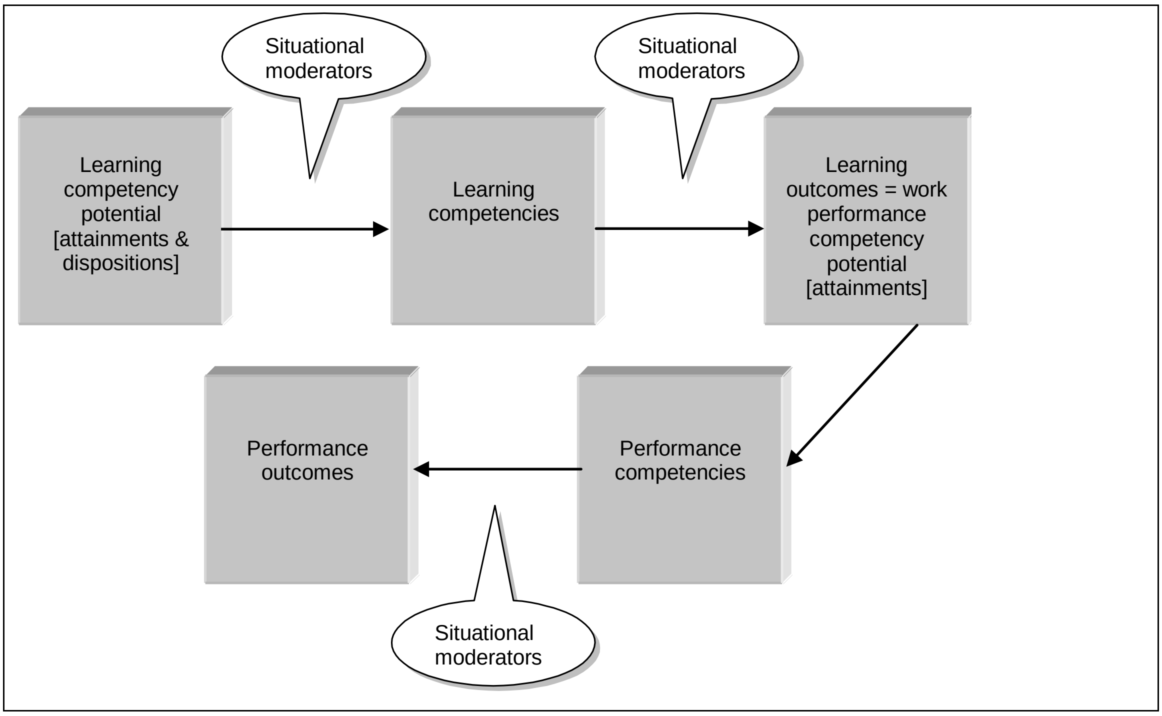

Figure 2.1. Performance@learning model (adapted from SHL, 2001, p. 7)

If training is to provide a worthwhile return on investment for the organization, training programmes have to be relevant to the job for which employees are selected and trained (Van Dyk et al., 2001). Training programmes are relevant to the extent to which they empower employees with the performance competency potential and performance competencies required to deliver the outputs for which the job in question exists.

Officers and candidate officers selected to study at the South African Military Academy come from different functional backgrounds – different Corps – and different Arms of Service in the South African National Defence Force (SANDF). The aim of their education at this institution is to acquire the necessary knowledge, analytic abilities and insight necessary to exhibit the work behaviours required of successful officers in the SANDF to deliver the expected output. Some of the knowledge, analytic abilities and insight required are function/discipline related and tends to be exclusive to specific occupational competencies within the SANDF. However not all of the development presented at the SA Military Academy, is job specific. At least some of the training is aimed at specific critical cross-field competencies (PEC, 2003), which officers in all the various branches of the SANDF should be able to exhibit if they are to achieve the objectives for which their specific positions exist. These generic, critical cross-field competencies are depicted in Table 2.1.

Table 2.1. Critical cross-field competencies (PEC, 2003)

Identifying and solving problems in which responses display responsible decisions using critical and creative thinking have been made
Working with others as a member of a team, group, organization, community
Organising and managing oneself and one's activities responsibly and effectively
Collecting, analysing, organising and critically evaluating information
Communicating effectively using visual, mathematical and/or language skills in the modes of oral and/or written persuasion
Using science and technology effectively and critically, showing responsibility towards the environment and health of others
Demonstrating an understanding of the world as a set of related systems by recognizing that problem-solving contexts do not exist in isolation
Contributing to the full personal development of each learner and the social and economic development of society at large, by making it the underlying intention of any programme of learning to make an individual aware of the importance of:
reflecting on and exploring a variety of strategies to learn more effectively
participating as responsible citizens in the life of local, national and global communities
being culturally and aesthetically sensitive across a range of social contexts
exploring education and career opportunities
developing entrepreneurial opportunities

The professional military development of young officers is a further mandate of the SA Military Academy. This function probably is not really served by the academic programmes of the Academy, but rather by the non-academic activities surrounding the academic programmes. Strictly speaking, however, the academic development and professional military development objectives cannot really be separated. The training and development presented at SA Military Academy thus serves three functions; it provides officers with discipline-specific knowledge, analytic abilities and insight, it equips officers with more generic cross-field competencies and it provides professional military development for young officers. The overarching function of the SA Military Academy, therefore, is to deliver tertiary educated officers to the SANDF.

From the aforementioned it can be deduced that the job of a learner is to actively partake in his or her education for the duration of their studies at this institution. Accordingly, it is expected of learners to attain a particular standard of achievement in this job; they have to be academically successful. In order to be regarded as academically successful, a learner has to complete his or her studies in the allotted time period. In other words, the execution of a learner's job will entail all the activities he or she needs to perform in pursuit of their academic goals and the completion his or her studies in the allotted time period. Despite its three-pronged objective, the decision on whether officers have successfully navigated the programme is, however, primarily based on the formal evaluation of discipline-specific knowledge, analytic abilities and insights and, to a lesser extent, by integrating it into the discipline-specific evaluations, the appraisal of the generic cross-field competencies.

In an attempt to attain the highest possible output of graduated officers at the lowest possible cost, this institution would consequently need to admit only those learners that would benefit the most from the training programmes, and who would achieve the highest possible academic performance. In order to differentiate between candidates in terms of their training or development prospects, it is imperative to determine why differences in training (academic) performance exist. Differences in **learning performance** (defined in terms of learning outcomes, i.e., discipline-specific knowledge, analytic abilities and insights) can be explained in terms of **learning behaviours** (or learning competencies, i.e., in terms of differences in what learners do). Moreover, learning competencies can be explained in terms of **learner characteristics** (or learning competency potential, i.e., in terms of the attributes of the learner). To successfully differentiate between candidates

who have better or poorer educational prospects in terms of a construct orientated approach to selection (Binning & Barrett, 1989), a valid performance hypothesis on the person-centered drivers of the learning competencies is required.
According to Guion (1991), hypothesizing the identity of the latent variables determining performance is a work of 'scientific imagination'. He is of the opinion that during hypothesis development the researcher has to introspectively 'imagine' the nature of the demands placed on a person by his or her job, as well as the characteristics a person needs to be able to meet those demands. The development of an informed hypothesis consequently depends on a proper understanding of the job for which people are to be selected (Guion, 1991). A performance@learning competency model indicating the learning competencies and learning competency potential learners need to be successful at the SA Military Academy thus needs to be developed, in order to formulate an informed performance hypothesis.

2.3 FLUID INTELLIGENCE AND TRANSFER

The job of a person occupying a particular position is comprised of a collection of duties that a person needs to perform which are instrumental in achieving specific outcomes. These duties (or functions), in turn consist of different tasks, and each task can be broken down in to a series of actions leading to a meaningful outcome (Van Dyk, *et al*., 2001). Information about these various tasks (or competencies) are used to infer the knowledge, abilities, and other person-centered prerequisites (performance competency potential) necessary to successfully execute the actions that collectively constitute the job (Van Dyk *et al*., 2001).

The job of a learner is to respond to a set of (novel) educational stimuli with specific behavioural (learning) actions that would allow the learner to create meaningful structure from the initially meaningless learning stimuli, and which would enable him or her to develop discipline-specific and generic performance competencies and the attainments and dispositions that underpin them. Moreover, the expectation is that the learner should attain the highest possible academic learning performance, in an endeavour to be regarded as academically successful.

During a learner's confrontation with a learning task he or she has to decide on the appropriate behaviour in response to these novel stimuli. Insight into the reason(s) for individual differences in response to novel situations might shed some light on why differences in learning performance exist.

The ability to learn and deal with novel situations is popularly labelled as intelligence (Kline, 1991). Cattell's (1971) investment theory distinguished two forms of intelligence, namely fluid and crystallized ability. According to this theory fluid ability develops as a single, general relation-perceiving ability, which is connected to the total associational neuronal development of the cortex. Fluid ability is regarded as the basic reasoning ability and it is mainly a function of the human's neural structures, and therefore highly heritable and less susceptible to the effects of environmental deprivation. Crystallized ability, on the other hand, develops as a result of investing fluid ability in particular learning experiences. In other words, crystallized ability consists of fluid ability as it is evidenced in the skills valued by the culture in which the individual lives. For example in a euro-centric environment, fluid ability may be vested in science and technology related competency potential and competencies whereas in a rural afro-centric environment it may be vested in hunting and tracking related knowledge and skills. Thus at an early age, say 2 or 3 years, fluid ability and crystallized ability are highly correlated. As children grow older and undergo different experiences at school and in the family, so, fluid ability and crystallized ability become less highly correlated as differences in learning opportunity affect additional difference in crystallized abilities over and above the difference explained by differences in fluid intelligence. The bright and well-adjusted individual who attends a good school and receives encouragement at home will invest most of his fluid ability in the crystallized skills of his culture. On the other hand, the equally bright individual from a home where education is not valued and who attends a school of inferior quality will be denied the opportunity to invest his fluid ability. His school performance consequently probably would be far worse than a child with moderate fluid intelligence who invests all his ability at school (Kline, 1991).

Sternberg (1985) developed a triachic theory of intelligence based on three cornerstones. According to his theory, intelligence cannot be understood outside of a socio-cultural context. In other words, what is "intelligent" in one environment may be irrelevant or even unintelligent in another. Secondly, intelligence is purposeful, goal-oriented, relevant behaviour consisting of two general skills; the ability to deal with novel tasks and the ability to develop expertise, that is to learn from experience to perform mental tasks effortlessly or automatically. Thirdly, intelligence depends on acquiring information processing skills and strategies (Weinberg, 1989).

The ability to deal with novel tasks is what was earlier referred to as fluid intelligence. Fluid intelligence is a function of the cognitive strategies available to the individual, and consists of a set of general cognitive tools and strategies (Taylor, 1994). Because fluid ability is a fundamental capacity that can be directed at novel and unusual problems (Taylor, 2001), crystallized abilities can be regarded as a product of this process. Crystallized abilities develop with repeated practice in a particular domain, which was initially unfamiliar to the individual. Stated differently, crystallized ability is specialized insights and knowledge that result from the use of fluid ability, via transfer. Transfer in this context is described as the process through which crystallized abilities develop from the confrontation between fluid intelligence (Cattell, 1971) and novel stimuli (Taylor, 1994). In other words, transfer is the adaptation of knowledge and skill to address problems somewhat different from those already encountered. Transfer can also be described as the phenomenon observed in terms of the effect that previously learned behaviour has on the performance in another situation or new learning tasks (Ormrod, 1990; Plug et al., 1997), meaning that a task that was already learned made it easier (or even possible) to learn a new task or solve an intellectually more challenging subsequent novel problem. The one pillar of academic learning is therefore the transfer of existing knowledge and skills on to novel learning material presented in class in an attempt to create meaningful structure in the learning material.

Through a process of transfer the individual’s structure of abilities and skills are elaborated over a period of time (Ferguson, 1956). At early ages the structure is simple, possibly dominated by the fluid ability. The individual’s fluid ability is responsible for the development of the first specific abilities. After the first crystallized abilities were developed, these specific abilities assist, through a process of transfer of skill, in the emergence of yet more specific skills. Fluid intelligence continues to be important in this process. This process may continue on into adulthood. No sharp division thus exists between (academic) learning and application; in as far as application is in fact further (action) learning. The intellectual territory captured by fluid intelligence becomes the intellectual launching pad for subsequent intellectual conquests. For this reason, transfer is regarded as a fundamental aspect of learning and cognitive development.

Furthermore, Ferguson (1956) regards learning as a process during which some attributes of a person's behaviour differentiate and attain relative stability (or invariance). He termed these attributes abilities, and added that the learning process is facilitated by the abilities a person already possesses. Stated differently, existing abilities contribute to the development of new abilities via intellectual transfer driven by fluid intelligence. The essence of transfer then, is perpetual concomitant change, and in the simplest case implies change in performance on one task with change resulting from practice on another. It can thus be said that performance on one task is some unspecified function of performance on another task, and measures of the amount of practice on the two tasks. There seems to be a close relation between Ferguson's (1956) notion of transfer and Sternberg's (1985) mastery of novelty, as stated earlier.

It becomes clear that the job of a learner is in essence that of problem solving where new competencies are built on existing ones and have to be integrated into conceptual frameworks (or knowledge stations) of the learner that become progressively more general and elaborate. It seems as if transfer lies at the heart of this process of elaboration (Taylor, 1994).

Transfer can be "near" or "far" – in other words, the new problem may be quite similar to previously encountered and solved problems, or it may be rather different. The nearness or farness of the transfer required has a bearing on the difficulty of the problem (Taylor, 2001). The difficulty of the problem is, however, also determined by the abstract reasoning/fluid ability of the learner.

2.4 INFORMATION PROCESSING AND AUTOMATIZATION

Taylor (1994) regards fluid intelligence as the potentiality to think abstractly and to infer concepts or rules. However, according to Taylor (1994) intelligence comprises more than just abstract thinking. He divides intelligence into abstract thinking and information processing factors. The information-processing approach to understanding intelligence describes how people gather, store and use information to solve problems and to acquire knowledge (Weinberg, 1989). Taylor (1994) asserts that although abstract reasoning ability is not independent of information processing efficiency, but rather related to it, the processing variables nonetheless do not fully account for the individual's abstract reasoning ability. Together these two factors set the natural upper limits of an individuals' performance because they are biologically or genetically endowed (Taylor, 1994). About 65 percent of the population variance in intelligence is attributed to genetic factors (Kline, 1991). Findings indicate that both genetic and environmental factors are important in the development of crystallized intelligence (Kline, 1991).

If the stimulus material does not change dramatically over time, the learning challenge lies in becoming more skilled and efficient in the performance of the newly derived response. Becoming more skilled and efficient in such a task can be referred to as automatization (Taylor, 1994). The rate at which newly acquired insight, derived through transfer, is automated is expressed in a learning curve reflecting the rate of work output correctly done over a period of time. Ferguson (1956) noted that a learning curve is a portrayal of change in ability with repetition. It can be expected that the steepness of learning curve will be influenced by fluid ability and transfer in the beginning of a closed-ended learning task. However, information processing variables are expected to have a more significant influence on learning performance during the later stages of the learning task, where learning progresses to the phase of automatization (Taylor, 2001; Taylor, 1994) and gains in performance are the results of practice (Jensen, 1980). This is congruent with Ferguson’s (1956) statement that abilities involved at one phase of learning differ from the abilities involved at another phase. It appears as if automatization and processing capacity may be related in the same way as transfer and fluid intelligence are related.

Ackerman (1988) indicates that the correlation between measures of fluid intelligence and performance tend to decline as learning progress (on relatively simple tasks that do not change substantially during the evaluation period). In other words, the correlation between performance early in the process of learning something new and a measure of fluid intelligence (e.g., Raven) at that time, is stronger than a correlation later during the learning process. This is consistent with aspects of Sternberg's (1985) theory in which he states that learning tasks measure "novelty-coping skills earlier during practice and automatization skills later during practice" (p. 77). It is evident that learning tasks are not only about continuously making sense out of novel stimuli. In many cases, the stimuli do not change dramatically over time, and the challenge for the learner is to become ever more adept and efficient at what he or she is doing. Moreover, if automatization did not successfully write the insights derived through transfer to knowledge stations where it can be subsequently retrieved to be transferred on to new problems, learning would be caught in a futile repetitive cycle of endlessly solving the same problem every time it is encountered.

Kyllonen and Christal (cited in Taylor, 1994, p. 190) distinguished four sources of individual differences in learning: Knowledge and skills (the enablers) and processing speed and processing capacity (the mediators). As discussed, knowledge and skills are crystallized ability that result from the application of fluid ability. These crystallized abilities thus *enable* the transfer required for development of new knowledge or skills, suggesting that transfer could not occur in a vacuum. The processing speed and processing capacity in turn play a critical role in *mediating* the automatization of processing.

Sternberg (1985) states that an individual's ability to deal with novelty and to automate performance is facilitated by components and their interactions. He describes a component as an elementary information process that serves (at least) three kinds of functions. Furthermore, a distinction is made between three kinds of components: meta-components (i.e., higher-order executive processes used in planning, monitoring, and decision making in task performance), performance components (i.e., processes used in the execution of a task), and knowledge-acquisition components (i.e., processes used in learning new information). The knowledge acquisition components provide the mechanisms for the steady development of an individual's knowledge base (knowledge stations). Increments in the knowledge base, in turn, allow for more sophisticated forms of knowledge acquisition and possibly for greater ease in executing processes used in task execution. In other words, as the base of old knowledge becomes deeper and broader, the possibility for relating new knowledge to old knowledge increase, and consequently, the new knowledge is incorporated into the existing knowledge base (Sternberg, 1985). Taylor (1994) contends that the skills and knowledge accumulated in prior learning have a growing impact on the emergence of new skills. The acquisition of new discipline-specific knowledge, abilities and insight can therefore be described as a process during which the learner has to build new attainments on older ones and these have to be integrated into a knowledge base that subsequently become more general and elaborated.

Efficient cognitive algorithms can be written (Taylor, 1994) to capture new intellectual insights derived through transfer of previously existing insights. The use of insight (previously learned information) to resolve a problem situation may also be referred to as problem-solving, which is a form of transfer. According to information processing theory, problem-solving success is influenced by cognitive factors such as short-term memory capacity, encoding, and retrieval (Ormrod, 1990). Unless efficient cognitive algorithms are formed and stored in memory for later retrieval, the stimulus will remain a novel problem to be solved via transfer every time it is encountered. It can thus be said that transfer is inhibited if newly derived insights cannot be captured and accumulate in knowledge stations (Sternberg, 1985) which serve as the cognitive platforms from which subsequent problem solving/transfer occurs. In some instances lack of knowledge can block successful execution of the performance components needed for intelligent functioning (Sternberg, 1985).

Learning performance, particularly the rate of learning, appears to be influenced by 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 (Taylor, 2001).

Since all ongoing intellectual activity is carried out in the short-term or working memory (Vernon, 1990), it may probably be the most important index of information processing capacity (Taylor, 1994). Working memory has been defined as short-term memory 'in action' as an alternative to just the temporary storage of information (Baddeley, cited in Taylor, 1994, p. 187). This may be because working memory consists of two aspects namely storage and processing. The processing draws on data stored in short-term memory, and once having manipulated it, returns it to short-term memory (Taylor, 1994). However, the working memory system is constrained by a limited storage capacity. Because of an inability to store information for an extended period of time in the absence of continued rehearsal, and a trade-off between the amount of information that it can hold at any one time and the amount of other information that it can process at the same time, even a fairly simple problem-solving task might place sufficient storage and processing demands on working-memory to reach (or exceed) its threshold. Were it not for a speed-of-processing property, the working memory system would exceed its limited storage and processing capacities. This fourth property allows short-term memory to cope with information-processing demands of complex problem-solving tasks (Vernon, 1990). As information is initially encoded into working memory, the more quickly it can be recorded and broken down into a small number of chunks, the less likely it is that the system's storage capacity will be exceeded. The more quickly the information necessary to solve a problem can be searched for, accessed, and retrieved from long-term memory, the more likely that the earlier encoded information will not have been lost due to decay. At each stage in the solution of a problem, the more quickly the required cognitive processes can be executed, the higher the probability that the system will not reach its threshold and, hence, the higher the probability that the problem will be solved correctly. Working memory has been shown to correlate strongly with fluid intelligence measures (Larson & Saccuzzo, 1989). The reason might be because those subjects who obtain higher IQ scores (i.e., who can successfully solve and answer more items on an intelligence test) are those who can process information more quickly, as measured by their performance on reaction time tests (Vernon, 1990).

IQ alone does not capture the full range of human cognition (Weinberg, 1989). In studies of differences between experts and novices in problem solving, it was found that the expert's knowledge is organised around central principles of his field of expertise, whereas the knowledge of the novice is organized around the physical entities or objects directly indicated in the problem description. Both expert and novice solve the problem, but the way the problem is initially represented determines different problem-solving procedures that result in differences in efficiency and ability to handle difficult situations (Glaser, 1981). Glaser concludes that the learning and experience of the competent individual results in knowledge, and in the organization of that knowledge into a fast-access pattern recognition or encoding system that greatly reduces the mental processing load. Understanding of the results from these acquired knowledge patterns enables the expert to form a particular representation of a problem situation. Novices also have systematic knowledge structures, but at a qualitatively different level than do experts. Novices' initial representation of the situation (which is determined by acquired knowledge structures) appears to be an index of developing competence (Glaser, 1981). In general, novices in a domain tend to focus on those features of examples with which they are most familiar and to miss the underlying concepts that the examples are supposed to demonstrate (Catrambone & Holyoak, 1989). In addition, research indicate that the phrasing of the transfer problem is important, suggesting that learners may need to learn how to encode novel problems in terms of features that are shared by prior examples. Two factors thus seem to be implicated in mediating transfer. One involves the quality of the representation of the commonalities among multiple source analogs; the other involves the presence of cues in the target problem that activate relevant features of the source analogs. In both cases the effect is to increase the likelihood that relevant features from prior training examples will be recalled and applied to the target problem (Catrambone & Holyoak, 1989).

2.5 LEARNING PERFORMANCE

In summary it can thus be said that transfer and automatization capacities are determined by the intelligence of the learner. Furthermore, it became evident that the transfer facet of learning is in effect fluid intelligence in action, and that the ability to automate newly acquired solutions/insights is determined by the information processing efficiency.

The job of a learner can now be rephrased as the solving of (cognitively demanding) novel learning problems by transferring current knowledge and competencies to the unfamiliar educational stimuli so as to acquire knowledge and understanding of the learning material and to internalize this insight through a process of automatization. It is clear that a learner needs the capacity to form effective cognitive strategies and the capacity to process information efficiently to succeed in these learning competencies or behaviours (of transfer and automatization).

Ideally a learner should also display insight into the possible application of this knowledge and skills during the rest of their military careers. This would, however, hopefully mean more than simply the retrieving of previously transferred and automated (i.e., learned) responses to now familiar stimuli. The expectation rather would be that the learner would be able to apply the newly derived knowledge to novel stimuli not explicitly covered in the academic programme. The application of newly acquired knowledge in solving new work related problems is, however, again transfer at work and thus dependent on fluid intelligence and, since fluid intelligence cannot operate in a vacuum, also the extent to which previous relevant learning (transfer) has been successfully internalized (automated).

Performance can be considered as something that a person does, in other words, the actions of that person. Work or job performance is therefore related only to the actions, behaviours or performance competencies relevant to the organization’s goals. Furthermore, performance pertains to the actual action, and not the consequence or result of action. Because behaviour is not always observable, this distinction is difficult. It is sometimes easier to observe the effects of behaviour than the behaviour itself. For example, the cognitive behaviour used in solving an algebra problem is difficult to observe and it is easier to see the effect of this behaviour - the production of a solution after the application of the mind. However, the result of covert cognitive behaviour, such as “solutions,” “statements,” or “answers” are regarded as actions, and is therefore also considered as performance (Campbell, 1991). Learning performance thus should be defined in terms of the two core learning competencies (transfer and automatization). However, when considering the conceptualization and operationalization/assessment of learning performance, the objective of training and development raised earlier should be kept in mind.

Training programmes are designed to empower employees with the performance competency potential and performance competencies required to deliver the outputs for which the job in question exists. Learning performance should thus ultimately not be assessed in terms of the consequences or outcomes of learning (i.e., crystallized knowledge), nor in terms of competence during training, but rather in terms of the ability to creatively utilize the newly derived knowledge in solving novel problems that could realistically be encountered in the work environment. This should, however, not be construed to mean that the assessment of the change in crystallized abilities affected by training interventions and the assessment of transfer and automatization during training is of no value. This seems to have significant implications for the manner in which the criterion construct should ideally be defined and operationalized in this validation study. De Goede (2007, pp. 70-71) expresses the concern that training institution too often fall into the trap of designing evaluations 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:

> 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.

2.6 LANGUAGE PROFICIENCY

Vocabulary is acquired incidentally throughout one's lifetime as a result of knowledge-acquisition components (Sternberg, 1985), and for this reason vocabulary tests measure acquired knowledge. A vocabulary test will provide quite a good predictor of academic achievement, because academic achievement is strongly dependent upon knowledge acquisition and upon the meta-components that control the components of knowledge acquisition (Sternberg, 1985). For this reason vocabulary tests may be regarded as a measure of crystallized ability.

Ferguson (1956) indicated that cultural factors prescribe what shall be learned and consequently different patterns of abilities emerge in different cultures. Therefore an individual's vocabulary can be regarded as an indication of his fluid ability as it is evidenced in his crystallized language ability, influenced by the values of his culture.

Glaser (1981) has reported that low levels of reading performance result from the interfering effects of slow, inefficient word decoding on the execution of higher-level comprehension components. Less-skilled readers are less efficient in elementary word processing tasks; this takes up time and memory space that is necessary for efficient sentence comprehension, since the latter depends on the availability of relevant knowledge stored in memory to which the new information can be related. This general hypothesis of the interference effects between basic and higher-level component processes raises a point for consideration here. Apparently, word-decoding processes need to attain a certain level of efficiency before more advanced processes can be carried out. Language proficiency thus seems to play a significant role in learning performance.

The foregoing argument would therefore suggest that learning performance on the academic programmes offered by the SA Military Academy will not only depend on the ability to transfer and automate newly derived insight (i.e., fluid intelligence and information processing efficiency). Since the medium of communication in the SANDF is English, the medium of education at the SA Military Academy is consequently also English. Consequently the presentation of novel learning stimuli occurs in English. It therefore stands to reason that a lack of proficiency in English will significantly constrain learner’s ability to master such learning material. A fluid intelligence by language proficiency interaction effect on transfer thus seems to be suggested. In addition the question could be asked whether efficient cognitive algorithms can be written and stored for later retrieval that captures the insight/problem solving derived through transfer (i.e., automatization) if an insufficient language proficiency would exist? If so, a language proficiency main effect on learning performance would also be implied. The latter argument is however somewhat more tenuous in as far as insight gained in novel learning material presented in English could possibly be written and stored in cognitive algorithms that captures the insight/problem solving derived through transfer for later retrieval in the mother tongue language.

2.7 PRIOR LEARNING

Previously it had been argued that transfer could not occur in a vacuum. Fluid intelligence needs a cognitive platform of existing, automated knowledge from which it creatively assembles solutions to the novel problems confronting it. If this is true, then an additional determinant of academic/learning performance at the SA Military Academy will be the level of the relevant crystallized abilities with which the learner arrives at the Academy. This in turn will dependent on how well learners performed at school. A school performance (or prior learning) main effect on transfer is thereby suggested as well as a fluid intelligence by school performance interaction effect on transfer. Significant differences unfortunately existed in the past (and some of which still persist today despite fifteen years of democracy) in the learning and developmental opportunities afforded to members from different backgrounds. Moreover, due to the consequence of the discrepancies in learning and developmental opportunities in the history of South Africa, strict top-down selection decisions based on predicted future academic performance derived from previous academic results, or any other valid predictor for that matter, would create significant adverse impact against members of historically disadvantaged groups. The logical, rational response to this dilemma would be to treat the cause of the problem by augmenting the deficiencies in the crystallized abilities through bridging programmes. No magic psychometric assessment wand will be able to transform the legacy of discrepancies in learning opportunities; it can only reveal its unfortunate consequences.

2.8 PERFORMANCE HYPOTHESIS UNDERLYING THE SELECTION PROCEDURE OF THE SA MILITARY ACADEMY

From the aforementioned it is apparent that the presence, or absence of the necessary cognitive competencies that would assist in the understanding and interpretation of the learning material, the intellectual drivers of these learning competencies, proficiency in English and past academic performance should discriminate between better or poorer academic performance of learners attending the academic programmes of the SA Military Academy. The performance hypothesis for success at the SA Military Academy unfolded thus far is schematically depicted as Figure 2.2.

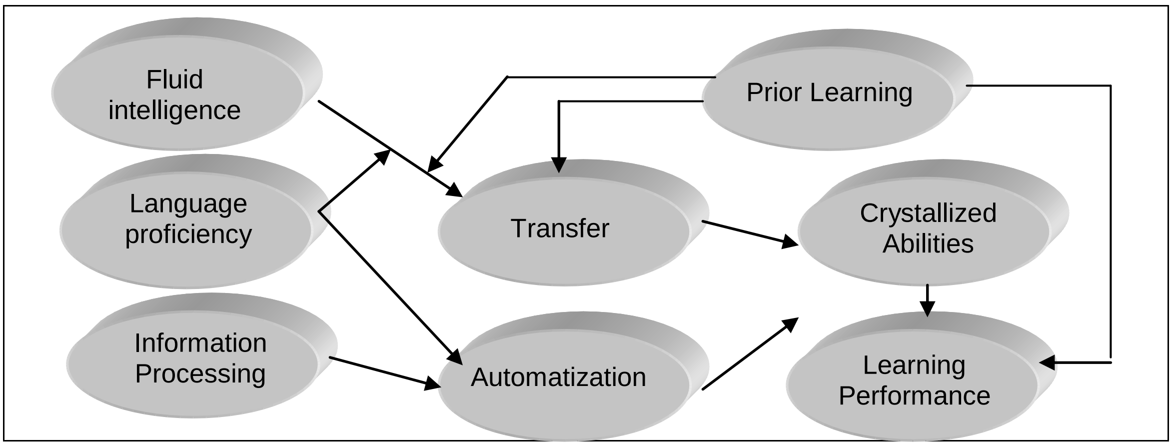

Figure 2.2. Performance hypothesis underlying the selection procedure of the SA Military Academy

2.9 CONCEPTUAL DEFICIENCIES IN THE PERFORMANCE HYPOTHESIS UNDERLYING THE SELECTION PROCEDURE OF THE SA MILITARY ACADEMY

One level on which HR interventions should be evaluated is in terms of the theoretical model on which it is based (Babbie & Mouton, 2001). Logic and literature, seems to suggest that learning/academic performance is also shaped by a number of additional factors not taken into account by the existing selection procedure of the SA Military Academy and thus not reflected in Figure 2.2. To the extent that the current selection procedure fails to accurately reflect the manner in which important influential determinants of performance combine to affect learning performance it should be regarded as deficient.

Individual differences in performance is a function of various abilities, not just intelligence according to the psychometric model of performance. Such other abilities include personality, motivational, and mood factors (Kline, 1991). Factors affecting academic performance on tertiary level are reported consistently in relevant research performed locally and internationally (Engelbrecht, 2000; Harackiewicz, Barron, Tauer & Elliot, 2002; Nel, 1997; Rademeyer & Schepers, 1998; Vosloo, 1987). These factors can be divided in two primary categories, namely cognitive as well as non-cognitive factors (Engelbrecht, 2000; Nel, 1997). Cognitive factors typically include factors such as school achievement, cognitive aptitude, and cognitive learning potential (as defined above). However, apart from the cognitive variables, it is important to also take non-cognitive factors into consideration (Engelbrecht, 2000; Nel, 1997; Lavin, 1965; Pienaar, 1991). Non-cognitive factors include motivation, study habits, socio-economic background, personality, interest, locus of control, self-esteem, career maturity, learning- and study strategies, and biographical data (Nel, 1997; Rademeyer & Schepers, 1998). The best prediction of academic performance would probably be achieved with a combination of cognitive and non-cognitive factors (Engelbrecht, 2000; Nel, 1997). To successfully elaborate the performance@learning structural model depicted in Figure 2.2 would however also require that the conceptualization of learning should be broadened to include more learning competencies than transfer and automatization. Learning behaviourally involves more than this. Possible behavioural learning performance dimensions that should be included in the performance@learning structural model over and above transfer and automatization would be time at task, self motivation and management of resources.

There is a magnitude of studies published with regards to the prediction of academic achievement of learners. The results of a limited sample of such research will be briefly discussed below in order to indicate the vast amount of variables influencing academic achievement or learning performance.

Different opinions exist with regards to the validity of high school grades as predictor of academic performance on tertiary level (Rademeyer & Schepers, 1998). According to various researchers (Behr, 1985; Ting, & Robinson, 1998; Van Eeden, De Beer & Coetzee, 2001) high school grades are the most significant predictors of academic performance on this level. However, Shochet (1994) argues that in the South African context, universities must find selection criteria, other than high school grades, that are both fair and valid for disadvantaged Black applicants with inferior school education. The merit of this argument, however, seems questionable. Fairness firstly does not reside in the predictor. Selection fairness cannot be assured through the careful development or judicious choice of selection instruments. Fair selection can only be assured by determining whether group membership systematically affects any of the parameters defining the regression of the criterion on the predictors and appropriately accounting for the group effect in the selection decision rule (Theron, 2007). Moreover if previously disadvantaged Black students find it difficult to gain access to universities or to succeed at university because of inferior school education the intellectually honest solution lies in remedying the deficiencies in prior learning. Crystallized abilities, developed through prior learning to a specific level, is a necessary condition to be successfully solve the learning problems posed by tertiary education given a specific level of fluid intelligence. There is no point in sweeping this uncomfortable reality under the rug.

Research conducted by Van Eeden, De Beer and Coetzee (2001) seems to suggest that cognitive predictor variables in their study reflect a disadvantage of having English as a second language when studying. This study indicates that English was the only significant predictor of academic performance for African language speaking learners compared to English-speaking learners. No relationship was found between English and the other variables in this study for the English home language group. It was concluded that better results might be obtained if the role of language is controlled for.

In a longitudinal study Harackiewicz, Barron, Tauer and Elliot (2002) explored the role of achievement goals, ability, and high school performance in predicting academic success. Their findings suggest that mastery and performance goals have positive and complementary consequences for motivation and performance in college courses over the period of learners' academic careers.

A significant and persistent relationship between self-esteem and academic achievement was confirmed in a sample of Black and Coloured first-year university learners. Academic success or failure appears to be as deep-rooted in self-concept as in measured mental ability, suggesting that, especially amongst black university learners, motivational and personality variables moderate the effect of academic aptitude scores on actual academic performance (Howcroft, 1991).

Research conducted to gain more clarity on the role of personality traits, personality types and learning strategies related to unsatisfactory academic achievement by some first year learners, suggest that these factors can be used to predict (with statistical significance) whether a first year learner will be an achiever or a non-achiever (Pelser, 1992). Mpofu and Oakland (2001) found that learners who use a surface approach to learning achieve lower levels of academic achievement compared to those who use deep and achieving approaches to learning.

At the South African Military Academy, nearly twenty years ago, academically unsuccessful learners (somewhat surprisingly) measured significantly higher with regard to participation in, commitment to and value expectations of their role as worker (officer in the SANDF). These learners also presented as more career mature with regard to decision-making, world-of-work information and career planning (Kotze, 1993), implying that their attitude towards their studies (or role as learner) might have had a significant influence on their learning performance. Two possible explanations were presented for this phenomenon. Firstly, it might be that academically unsuccessful students manipulated their response, indicating a bigger focus on their role as worker, as a result of their rationalizing for poor academic performance. A second possibility is that during the process of academic failure, the students matured psychologically, realizing that they need to change their focus to the role as worker, implying a turning point in their career development because they need to return to the line function earlier than their initial plan.

Ting and Robinson (1998) assessed the effectiveness of cognitive and psychosocial variables in predicting grade point average (GPA) and retention at university. High school GPA was the most significant predictor for first year GPA. Other factors that also significantly correlate includes: educational level of parents, course load, and extracurricular activities. The predictive ability of GPA was greatly influenced by race, gender and by factors within genders, including science skills and financial aid. The findings indicated that multivariate models that include race/gender as a dummy variable to predict academic performance across gender and race are more effective than a general model.

Rademeyer and Schepers (1998) concluded that the best approach to selection would probably be to compute the canonical discriminant function for each faculty, and to classify prospective learners accordingly. It was further suggested that if a learner is not selected for a specific faculty he or she should be considered of another faculty. However, the learner's interests should also be considered.

The selection battery employed by the SA Military Academy for the selection of students into the academic programmes of the Academy evaluated in this study only takes into account the cognitive factors that have an influence on academic performance. There are, however, as indicated by the preceding discussion, numerous other non-cognitive variables that may influence performance of students at tertiary institutions that the selection procedure under consideration does not account for. To the extent that the battery under review fails to reflect critical learner attributes that do affect learning performance the selection battery should be regarded as theoretically deficient.

CHAPTER 3
RESEARCH METHODOLOGY

3.1 RESEARCH OBJECTIVE

The general, overarching aim of this study is to psychometrically evaluate the selection battery used by the SA Military Academy for the selection of learners. The following more specific objectives were formulated earlier for the study:

- To test the propositions made by the performance hypothesis depicted as a structural model in Figure 2.2;
- To determine the predictive validity of the individual predictors of the selection battery;
- To derive a weighted linear prediction model actuarially from a set of predictor and criterion data;
- To determine the validity of the inferences derived from prediction model;
- To evaluate the fairness of the inferences/predictions derived from the prediction model and adapt the model if necessary;
- To evaluate the utility of the fair prediction model over random selection; and
- To develop a criterion-referenced norm table that expresses the risk of failure conditional on expected academic performance.

3.2 RESEARCH DESIGN

The performance hypothesis derived from the literature study hypothesizes specific structural relationships between the criterion construct and the latent variables being assessed by the selection battery (see Figure 2.2). The validity of these hypothesized relationships is to be investigated empirically. The research design constitutes the formal logic in terms of which the validity of the hypothesized relations amongst the variables will be examined. The function of the research design is to ensure empirical evidence that can be interpreted unambiguously for or against the stated hypotheses. The research design achieves this through control of variance in the measures of the exogenous and endogenous latent variables. More specifically the primary function of a research design is to maximize systematic variance, to minimize error variance and to control systematic non-relevant variance (Kerlinger & Lee, 2000).

An *ex post facto* correlational design is used in this study. The study aims at testing the empirical validity of hypotheses of the basic form "if ξ then η" as proposed by the performance hypothesis depicted in Figure 2.2. The predictor variables in this study are, however, inherently not manipulable. Testing the empirical validity of the relational claims made by the performance hypothesis is therefore not possible by inducing variance in the independent variable through experimental manipulation to determine whether the dependent variable responds with concomitant variation. Random assignment to treatment conditions is consequently also not feasible. As a result, *ex post facto* correlational designs suffer from a relative lack of control of variance in the dependent variable. The logical confidence with which the finding of significant correlations could be interpreted as corroboration of the relational claims made by the performance hypothesis is thereby eroded (Babbie & Mouton, 2001; Kerlinger & Lee, 2000). Although the argument underlying the performance hypothesis depicted in Figure 2.2 is distinctly causal in its thinking, the *ex post facto* nature of the research design will preclude the drawing of causal inferences from significant correlation coefficients. The observed correlation matrix (or alternatively the observed covariance matrix) reflecting the strength of the relationship between measures of the various latent variables still begs the question what caused the latent variables to correlate the way they do? Various possible structural models serve as alternative possible explanations for the observed correlations. The model depicted in Figure 2.2 is only one of a variety of plausible models.

3.3 MEASURING INSTRUMENTS

The performance hypothesis should be operationalized to obtain empirical proof that the relationships suggested by the performance hypothesis (as depicted in Figure 2.2) developed in this research provide a plausible explanation for differences observed in learning performance. To justify the claim that inferences on the learning performance (η) can be made from the observed scores obtained from the SA Military Academy Selection Battery[3] it is imperative to demonstrate that the selection battery has predictive validity for this specific criterion. To convincingly achieve this would require that it should be demonstrated that:

---

[3] Ability, Processing of Information and Learning Battery (APIL-B), the Academic Aptitude Test (subtest 3 and 4) (AAT), and Matrix M-scores.

- Composite first year first semester academic results (Y) is a valid and reliable measure of learning performance (η),
- X_j as measured by the SA Military Academy Selection Battery are valid and reliable measures of the latent learning competencies and competency potential variables (ξ_j), and
- the valid and reliable measure (Y) of the conceptualised final criterion (η) is systematically related to valid and reliable substitute measures (X_j) of the latent variables measured by the SA Military Academy Selection Battery (ξ_j), (Binning & Barrett, 1989; Guion, 1991; Theron, 2002).

Psychometric evidence is therefore needed to establish the psychometric integrity of the indicator variables used to operationalize the latent variables comprising the learning performance hypothesis. The foregoing argument is schematically depicted as Figure 3.1. Without empirical evidence supporting inferences 1 and 2 in Figure 3.1, inference 3 can not be justified from finding a significant ($p<0,05$) correlation between a weighted composite of predictor measures and the criterion measure.

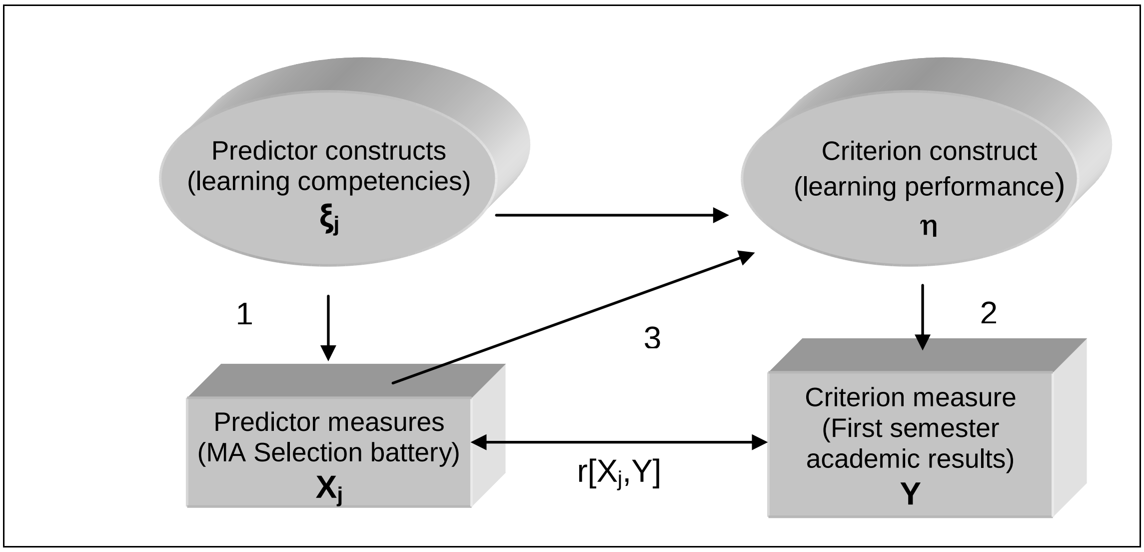

Figure 3.1. The nature of the evidence required to justify the use of the substitute measures (X_j)

The manner in which the performance hypothesis was operationalized is visually represented as Figure 3.2.

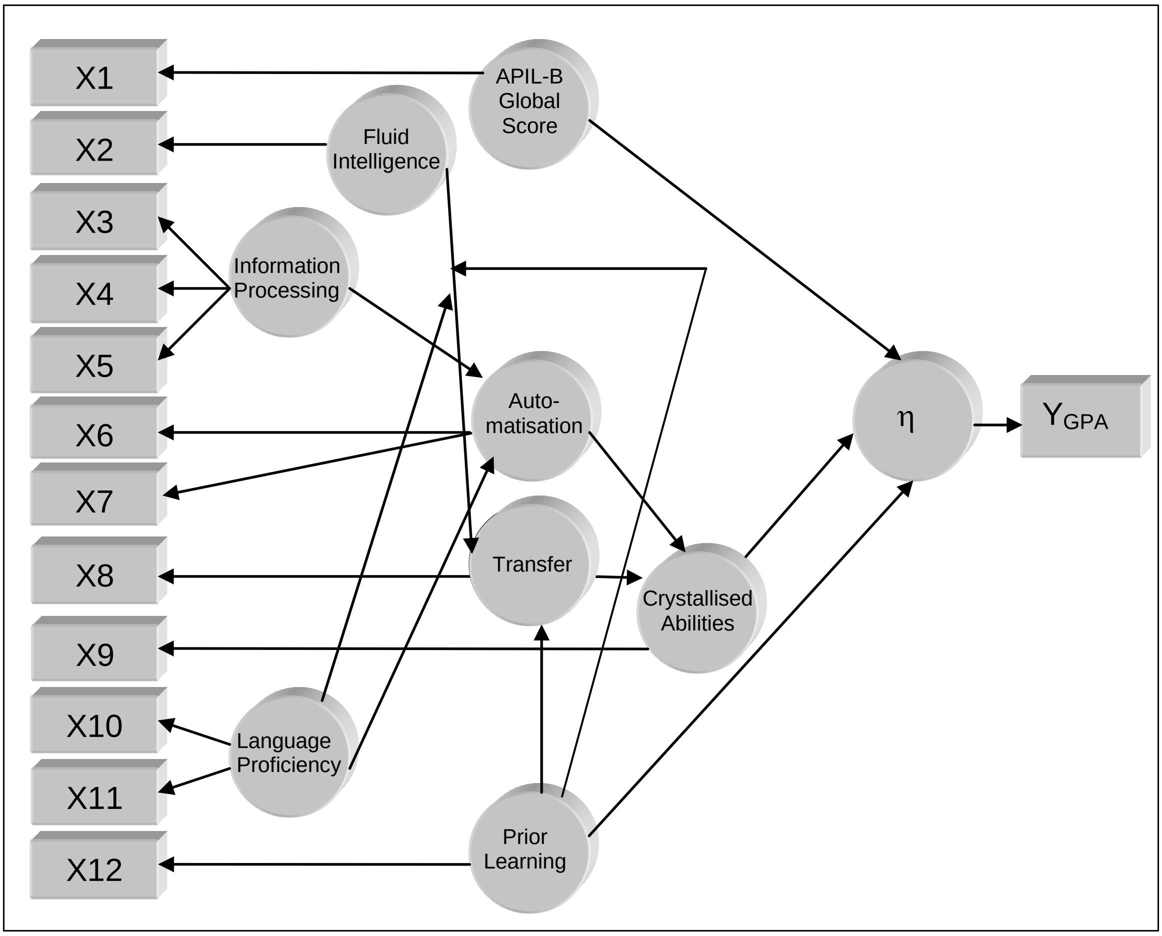

Figure 3.2. Operationalized performance hypothesis dimensions

In order to get optimum return on investment, the SA Military Academy needs to select, from all the applicants, the grouping which has the highest probability to be successful in their studies. This study focuses on the learning performance of students after their first semester at this institution. The average of the first year first semester academic results will be used as the measure of this criterion construct. As indicated in the performance@learning model (Figure 1.1), specific learning competencies are instrumental in attaining this desired performance outcome. Valid measures of these learning competencies (predictor constructs) can be used to predict the learning performance of applicants at the time when the selection decision needs to be made provided that a known systematic relationship exists between the measures of the predictor constructs and a measure of the criterion measure. The operationalization of the predictor constructs and the criterion construct will now be discussed.

3.3.1 Language Proficiency

Proficiency in the language of instruction (English) should have a significant influence on a student's ability to master novel learning tasks. Language proficiency is hypothesized to moderate the impact of fluid intelligence on transfer. Fluid intelligence can create meaningful structure in learning material presented in English only if a reasonable mastery of English exists. A language proficiency main effect on automatization moreover is implied. As indicated earlier this argument seems somewhat more tenuous. As stated previously, basic word-decoding processes need to attain a certain level of efficiency before more advanced cognitive processes can be carried out. The reason being that elementary word processing tasks takes up time as well as memory space necessary for efficient sentence comprehension, since reading comprehension depends on the availability of relevant knowledge stored in memory to relate new information with (Glaser, 1981). A distinction is made between basic vocabulary and reading comprehension as components of language proficiency for the purpose of this study. These two components (English vocabulary and reading comprehension) will be measured with Sub-Tests 3 and 4 of the Academic Aptitude Test (AAT) (University version).

Vocabulary (Sub-Test 3)

The purpose of this sub-test is to obtain an indication whether a testee commands the necessary vocabulary required for tertiary studies. This test is based on the assumption that the ability to recognize and select the best word from amongst a number of possibilities to fit in a specific context provides a valid indication of a testee's knowledge of words. The test consists of 30 sentences each. A word has been omitted in every sentence. The testee has to select the correct word from five given words. The reliability coefficients of the AAT battery were calculated with Kuder-Richardson formula 20. For the English Vocabulary test a coefficient of internal consistency value of 0,85 had been obtained (Owen & De Beer, 1981).

Reading Comprehension (Sub-test 4)

The purpose of this sub-test is to obtain an indication whether a testee commands the necessary reading comprehension required for further study. This test assesses a testee's ability to understand and put into practice that which he/she reads. This test is based on the assumption that a testee's ability to form the correct concept in reading words, sentences and paragraphs provides a valid indication of his ability to understand and apply that which he/she reads. The test consists of a number of passages which have to be read, with questions on every passage. There are thirty questions in every test. The English Reading Comprehension test returned a reliability coefficient value of 0,81, calculated with the Kuder-Richardson formula 20. (Owen & De Beer, 1981).

3.3.2 Fluid intelligence

The construct fluid intelligence is sometimes also referred to as the capacity to think abstractly (Cattell, 1971) or to think conceptually (Taylor, 1997). This ability is used to solve all new or unusual problems for which no predetermined solution exists, and drives the development of new skills and abilities. Students' ability to solve new or unusual problems, form abstract concepts, reason hypothetically, theorize, build scenarios, trace causes, etc. will be measured with the Concept Formation Test (CFT). This test measures the individual's capacity to think abstractly and conceptually, and forms part of the Ability, Processing, of Information and Learning Battery (APIL-B). Non-verbal tests of fluid intelligence have been used successfully for many years in cross-cultural research and assessment (Taylor, 1997).

The Concept Formation Test comprises a series of classificatory tasks where the testee is presented with sets of geometrical diagrams and must identify a diagram, which does not share a characteristic that all the other diagrams share (Taylor, 1997).

The reliability of the Concept Formation Test scores was also calculated with Kuder-Richardson formula 20. 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 (Taylor, 1997).

3.3.3 Information Processing Capacity

Information processing refers to the way individuals gather and use information to solve problems and to acquire knowledge (Weinberg, 1989). This capability comprises three latent variables namely speed of information processing, accuracy of information processing and flexibility of information processing. Speed of information processing refers to the quickness (speed) with which information is processed. Accuracy of information processing refers to the incidence of errors in work done in a given time segment. In combination the speed and accuracy measures reflect efficiency of information processing. Flexibility of information processing refers to the capacity to quickly choose the appropriate problem-solving strategy to solve the problem at hand, also referred to as cognitive flexibility.

Information processing capacity will be measured with the Flexibility-Accuracy-Speed-Tests (FAST). This sub-test of the APIL-B is a compendium of four short subtests. These four sub-tests yield primary scores that are combined and reworked to yield three secondary scores. The secondary scores provide measures of the speed of information processing, accuracy of information processing and capacity to cope with multiple problem formats under time pressure (Taylor, 1997).

Taylor (1997) points out that the reliability of the variable speed of information processing cannot be directly determined. However, some indication of the reliability was obtained by inspecting the correlations between the three components that were added together to derive the speed score. These are the number of items attempted on the Series component, the number of items attempted on the Mirror component and the number of items attempted on the Transformations component of the FAST sub-test of the APIL-B. Correlation coefficients ranging between 0,45 and 0,72 have been obtained for six samples (Taylor, 1997).

The accuracy score’s reliability was estimated by splitting the Flexibility-Accuracy-Speed-Tests into two, calculating separate accuracy indices on each score, correlating these two accuracy indices and correcting the obtained correlation for the shortening of the original scale. Correlation coefficients ranging between 0,70 and 0,86 have been obtained for six samples (Taylor, 1997).

According to Taylor (1997) it is impossible to estimate the reliability of the Flexibility score. He nonetheless concludes that the Flexibility scores have a large variance which, according to him, is a prerequisite for good reliability. This is, however, rather tenuous evidence since the variance could in fact also be an expression of a large random error component in the observed flexibility scores.

3.3.4 Automatization

Automatization is the capacity to become fast and efficient on a cognitive task with practice (Sternberg, 1985). Stated differently, automatization is the extent to which an individual becomes ever more skilled and efficient at what he is doing, and is often expressed as a learning curve reflecting the number of units of work correctly done over a period of time. In other words, learning rate can be regarded as a function of the improvement in performance of an individual expressed in terms of the number of units of work correctly done per unit time. The steeper the learning curve, the more rapid the rate of learning or process of automatization (Taylor, date unknown; Taylor, 1997). There are thus two components to automatization, the steepness of the learning curve and the total amount of work done during the process of automatization.

Automatization is assessed as the increase of work output over four sessions with the Curve of Learning sub-test of the APIL-B (Taylor, date unknown). During this subtest, testees are subjected to the same task (symbol-symbol and symbol-word translations making use of a special dictionary) on four occasions and testees are also given three study periods (after the first, second and third exposures to the learning material). Two scores are derived from this repeated-exposure exercise namely COLdiff and COLtot. These measures give an indication of performance gain in a learning task (difference in output between the fourth and first sessions) and overall work output on this task (total amount of work done in all four sessions) respectively (Taylor, 1997).

Taylor (1997, p. 63) explains how the reliability coefficients were determined for the Curve of Learning:

> COLdiff's reliability was estimated by subtracting number correct in COL3 from number correct in COL1 to produce one score and subtracting number correct in COL4 from number correct in COL2 to produce a second score, and then correlating these. No correction for test shortening was applied because the COLdiff score is not twice the above scores. However, COLdiff is likely to be appreciably more reliable than the indices quoted below because the difference between COL4 and COL1 is much larger than the difference between COL3 and COL1 or between COL4 and COL2. COLtot's reliability was estimated by computing COL1 + COL3 and also COL2 + COL4 and then correlating these two scores correcting for test shortening.

The reliability estimates for six samples ranged between 0.62 and 0.70 for COLdiff and for COL tot 0,88 and 0,97.

3.3.5 Transfer of Knowledge

Transfer was described earlier as the phenomenon observed in terms of the effect previously learned behaviour has on the performance in another situation. As previously stated, the relationship between an individual's abilities and performance is attributed to the process of transfer. Testee's capacity to transfer knowledge or skill from one problem situation to different but related problem situations will be measured with the Knowledge Transfer Test (KTT) sub-test of the APIL-B. This sub-test is a learning exercise that measures knowledge transfer by exposing a testee to a number of related but increasingly complex problems. The testee is given answers to two example problems and also feedback on his/her performance of the test problems (Taylor, 1997).

Reliabilities for the Knowledge Transfer Test were estimated through the split-half method. Taylor (1997, p.63) 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.

Split-half reliability coefficients ranging between 0,71 and 0,84 were obtained for the Knowledge Transfer Test (Taylor, 1997).

3.3.6 Crystallized Abilities

As previously stated, an individual's crystallized abilities develop with repeated practice in a particular domain where initially no such abilities existed. Crystallized abilities are therefore regarded as specialized insight or understanding and knowledge that emerge via transfer from existing knowledge and that is subsequently, successfully stored in memory. Memory content and understanding can thus be considered as crystallized ability.

The Memory and Understanding sub-test of the APIL-B will be used to measure the testee's crystallized ability. This test is a sequel to the Curve of Learning test, and measure the retention of the information in the Curve of Learning Dictionary (testees are encouraged to learn as much of the special Dictionary as possible during the administration of the Curve of Learning test). Learning of this material is difficult, and therefore high scores on this test indicate that the testee has conceptually mastered the Dictionary, and thus processed and understood it deeply (Taylor, 1997).

The crystallized abilities depicted in Figure 2.2 that have to be transferred by fluid intelligence onto the problems posed by the subject evaluations are the knowledge and abilities that the various academic modules have set out to develop. The level to which crystallized abilities would be developed would depend on the degree of competence on the learning competencies, which in turn would depend on the level of fluid intelligence and information processing capacity, prior learning and language proficiency. The learning competencies in action on the academic programme and the crystallized academic abilities can, however, clearly not be measured at the time of selection into the academic programme. In the APIL-B the learning competencies and the crystallized ability that emerges from them are measured in a simulated learning situation in which the role of prior learning and language proficiency is minimized. The question therefore is whether one could expect the level to which crystallized abilities are developed in a simulated learning scenario to correlate with the level to which crystallized abilities are developed in academic learning at the SA Military Academy? It could be argued that these two ability measures should correlate positively because they are both determined by a common fluid intelligence and information processing capacity. The correlation should, however, be attenuated by the fact that prior learning and language proficiency plays a significant role in academic learning but significantly less of a role in the artificial, simulated learning scenario created in the APIL-B and that prior learning and language proficiency varies across learners.

3.3.7 Prior Learning

As theorized, fluid intelligence needs a reservoir of knowledge to delve into in order to creatively solve novel problems and generate solutions. It was suggested that a student's academic performance would therefore also be dependent on the level of his relevant crystallized abilities obtained prior to this learning opportunity.

Performance at school (average of matriculation examination results) will be used as indication of the level of prior learning. The average of matriculation examination results as indicated on the matriculation certificate, issued by the Department of Education, will be used. A battery of subject specific knowledge tests that assess the degree to which the knowledge prerequisites assumed by the various modules have been met would, however, have been preferable.

3.3.8 Learning Performance

Grade Point Average

The Learning Performance of a student will be expressed as a Grade Point Average. The Grade Point Average of a student is the average weighted score for all subjects the student has taken that semester. The weighted average will be calculated for each student by multiplying the credits associated with each subject taken with the subject average achieved for that particular subject. The result will then be divided by the sum of all the credits.

The assumption is that the subject evaluations the student has to sit for do not evaluate the ability to recall the newly developed crystallized abilities but rather the ability to creatively utilize the newly derived knowledge in solving novel problems that could realistically be encountered in the military work environment.

3.4 STATISTICAL HYPOTHESES

Given the research objective outlined above and the proposed relationships among the latent variables as depicted in the basic performance hypothesis (Figure 2.2) the following research and statistical hypotheses are formulated.

The notational system used in the formulation of the hypotheses follows the practice typically employed in selection validation studies rather than the structural equation modelling convention. The symbols used to represent the indicator variables operationalizing various latent variables comprising the performance hypothesis are depicted in Figure 3.2 above. The first semester, first year weighted grade point average is represented by the symbol Y_{GPA}. Conceptual reasoning ability or fluid intelligence is represented by the symbol X_2, speed of information processing by X_3, accuracy of information processing by X_4, flexibility of information processing by X_5, steepness of learning curve in automatization by X_6, total amount of work done in automatization by X_7, transfer by X_8. Memory and understanding depicted as X_9, is interpreted to represent the crystallized ability latent variable in Figure 3.2. The two subtests of the Academic Aptitude Test used in the selection battery, sub-test 3 (English Vocabulary) and sub-test 4 (English Reading Comprehension) are represented by X_{10} and X_{11} respectively. Prior learning operationalized in terms of candidates' matriculation results is represented as X_{12}. In addition the Ability, Processing of Information and Learning Battery (APIL-B) provide a global score of overall learning potential, which was indicated as X_1. Race will be treated as a dichotomous dummy variable (X_{13}) in the fairness analysis with $X_{13} = 0$ representing Black learners and $X_{13} =1$ White learners.

The nature of the envisaged statistical analyses will necessarily affect the format in which the statistical hypotheses will be formulated. The possibility of utilizing structural equation modelling to evaluate the performance hypothesis was initially considered. The fitting of a structural model, which contains one or more interaction effects between continuous latent variables, however, is substantially more complicated than the fitting of a model where the relationships between all latent variables can be expressed by linear equations (Schumacker & Lomax, 1996). Although Kenny and Judd (1984) developed a procedure to estimate non-linear and interaction effects of latent variables in structural equation models, the implementation of their procedure via LISREL nonetheless remains extremely cumbersome (Schumacker & Lomax, 1996). Multi-group structural equation modelling could was also considered as another possibility (Hair, Anderson, Tatham, & Black, 1995). This option would have been an ambitious but nonetheless realistic possibility if only a single moderator variable would have been hypothesized. In the model depicted in Figure 2.2 both prior learning and language proficiency are hypothesized to act as moderator variables. Moreover the sample size would not allow for a multi-group analysis. It was consequently decided to rather restrict the evaluation of the causal linkages proposed by the performance hypothesis to correlation and regression analysis. This has the advantage of aligning the analyses used in the evaluation of the performance hypothesis with those used to evaluate the selection decision-making in terms of validity, fairness and utility. A more detailed account of the statistical analyses performed will be outlined below.

The following substantive research hypotheses and associated statistical hypotheses were formulated in pursuit of the research objective and based on the performance hypothesis (Figure 2.2). The hypothesized effect of prior learning on transfer (as a main effect and in interaction with fluid intelligence), the hypothesized language proficiency main effect on automatization and the hypothesized language proficiency x fluid intelligence interaction effects on transfer were, however, excluded from the empirical evaluation of the performance hypothesis. The APIL purposefully uses geometric test stimuli with which all testees are equally unfamiliar, irrespective of the educational opportunities they might have had in life. Although prior learning could logically be expected to play a significant role in solving the type of novel academic problems learners might encounter in their studies at the SA Military Academy, it should play no role in the solving of novel problems (transfer) in the simulated world created by the APIL-B. Likewise in a world where problems to be solved are presented in English, English language proficiency can logically be expected to play a significant role in solving academic problems and automating those solutions. But the same is not true in the contrived and largely non-verbal reality created by the APIL-B. Admittedly this largely nullifies the argument offered earlier to justify the decision not to make use of structural equation modelling to evaluate the performance hypothesis. The argument that the use of correlation and regression analysis would better align the analyses used in the evaluation of the performance hypothesis with those used to evaluate the selection decision-making in terms of validity, fairness and utility, however, remains true.

The simplified performance hypothesis is depicted in Figure 3.3. The performance hypothesis is simplified in a path diagram. X_1, X_2, X_3, … , X_{12} are the observed variables, indicated by boxes carrying the same meaning as indicated earlier. The unobserved, latent variables, depicted in Figure 3.2 as circles were not included in the simplified path diagram because of the decision to not use structural equation modelling to test the hypothesized model. They are assumed to influence the X's, thus the arrows run from latent variables to indicator variables. Single headed arrows are used to indicate causal influences, and double-headed arrows to indicate correlations. The symbol η represents the First Year first semester academic success latent variable.

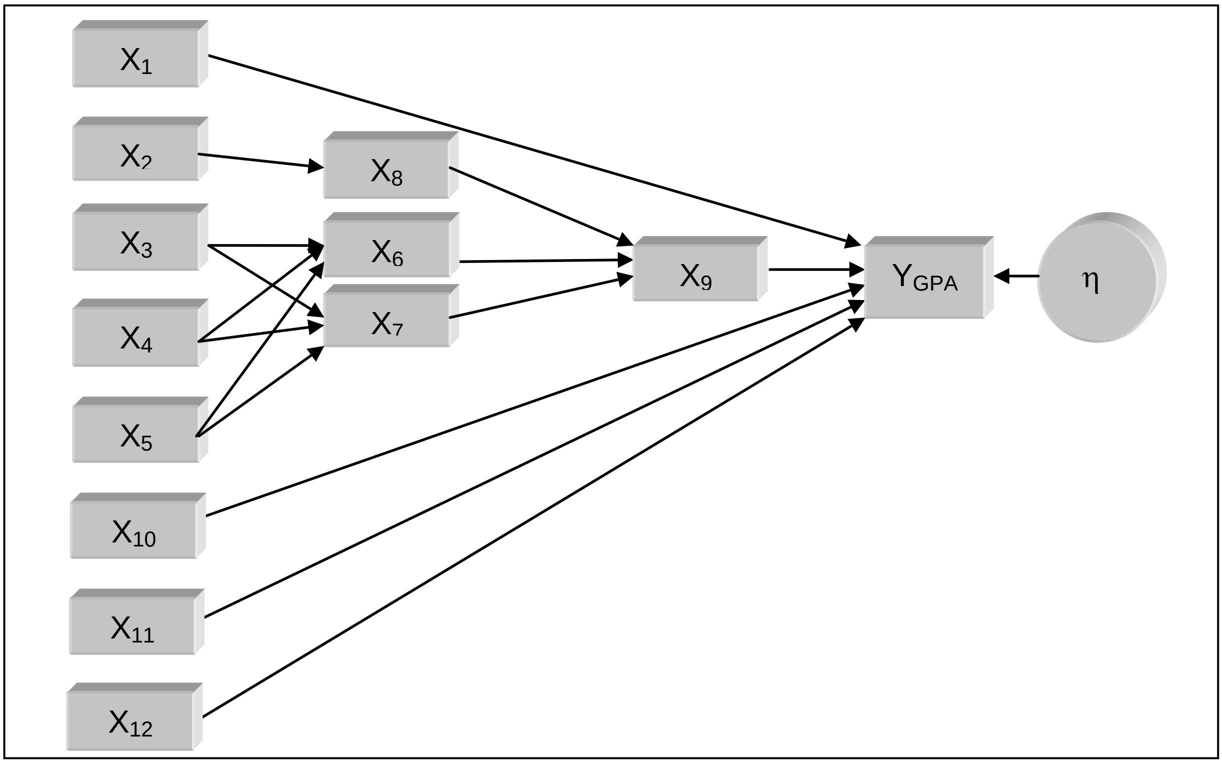

Figure 3.3. The simplified performance hypothesis

Hypothesis 1

Fluid intelligence (X_2) has a positive effect on transfer (X_8).

H_{01}: $\rho[X_2,X_8] = 0$

H_{a1}: $\rho[X_2,X_8] > 0$

Hypothesis 2

Speed (X_3), accuracy (X_4) and flexibility (X_5) of information processing each has a positive effect on the steepness of the learning curve in automatization (X_6) and the total amount of work done in automatization (X_7).

H_{02}: $\rho[X_3,X_6] = 0$

H_{a2}: $\rho[X_3,X_6] > 0$

H_{03}: $\rho[X_4,X_6] = 0$

H_{a3}: $\rho[X_4,X_6] > 0$

H_{04}: $\rho[X_5,X_6] = 0$

H_{a4}: $\rho[X_5,X_6] > 0$

H_{05}: $\rho[X_3,X_7] = 0$

H_{a5}: $\rho[X_3,X_7] > 0$

H_{06}: $\rho[X_4,X_7] = 0$

H_{a6}: $\rho[X_4,X_7] > 0$

H_{07}: $\rho[X_5,X_7] = 0$

H_{a7}: $\rho[X_5,X_7] > 0$

Hypothesis 3

The steepness of the learning curve in automatization (X_6) and the total amount of work done in automatization (X_7) both have a positive effect on the level to which crystallized abilities develop (X_9).

H_{08}: $\rho[X_6,X_9] = 0$

H_{a8}: $\rho[X_6,X_9] > 0$

H_{09}: $\rho[X_7,X_9] = 0$

H_{a9}: $\rho[X_7,X_9] > 0$

Hypothesis 4

Transfer (X_8) has a positive effect on the level to which crystallized abilities develop (X_9).

H_{010}: $\rho[X_8,X_9] = 0$

H_{a10}: $\rho[X_8,X_9] > 0$

Hypothesis 5

The level to which crystallized abilities develop (X_9) has a positive effect on learning performance (Y_{GPA})

H_{011}: $\rho[X_9,Y_{GPA}] = 0$

H_{a11}: $\rho[X_9,Y_{GPA}] > 0$

Hypothesis 6

Prior learning (X_{12}) has a positive effect on learning performance (Y_{GPA})

H_{012}: $\rho[X_{12},Y_{GPA}] = 0$

H_{a12}: $\rho[X_{12},Y_{GPA}] > 0$

Hypothesis 7

Speed (X_3), accuracy (X_4) and flexibility (X_5) of information processing each significantly explains unique variance in the steepness of the learning curve in automatization (X_6).

H_{013}: $\beta_1[X_3]=0|\beta_2[X_4]\neq0;\beta_3[X_5]\neq0$

H_{a13}: $\beta_1[X_3]>0|\beta_2[X_4]\neq0;\beta_3[X_5]\neq0$

H_{014}: $\beta_1[X_4]=0|\beta_2[X_3]\neq0;\beta_3[X_5]\neq0$

H_{a14}: $\beta_1[X_4]>0|\beta_2[X_3]\neq0;\beta_3[X_5]\neq0$

H_{015}: $\beta_1[X_5]=0|\beta_2[X_4]\neq0;\beta_3[X_3]\neq0$

H_{a15}: $\beta_1[X_5]>0|\beta_2[X_4]\neq0;\beta_3[X_3]\neq0$

Hypothesis 8

Speed (X_3), accuracy (X_4) and flexibility (X_5) of information processing each significantly explain unique variance in the total amount of work done in automatization (X_7).

H_{016}: $\beta_1[X_3]=0|\beta_2[X_4]\neq0;\beta_3[X_5]\neq0$

H_{a16}: $\beta_1[X_3]>0|\beta_2[X_4]\neq0;\beta_3[X_5]\neq0$

H_{017}: $\beta_1[X_4]=0|\beta_2[X_3]\neq0;\beta_3[X_5]\neq0$

H_{a17}: $\beta_1[X_4]>0|\beta_2[X_3]\neq0;\beta_3[X_5]\neq0$

H_{018}: $\beta_1[X_5]=0|\beta_2[X_4]\neq0;\beta_3[X_3]\neq0$

H_{a18}: $\beta_1[X_5]>0|\beta_2[X_4]\neq0;\beta_3[X_3]\neq0$

Hypothesis 9

The steepness of the learning curve in automatization (X_6) and the total amount of work done in automatization (X_7) each significantly explain unique variance in the level to which crystallized abilities develop (X_9)

H_{019}: $\beta_1[X_6]=0|\beta_2[X_7]\neq0$

H_{a19}: $\beta_1[X_6]>0|\beta_2[X_7]\neq0$

H_{020}: $\beta_1[X_7]=0|\beta_2[X_6]\neq0$

H_{a20}: $\beta_1[X_7]>0|\beta_2[X_6]\neq0$

Language proficiency and prior learning were, however, included in the psychometric evaluation of the selection decision-making since these variables in reality do influence the accept and reject decisions made with regards to learners applying for admission to the SA Military Academy.

Hypothesis 10[4]

Conceptual reasoning ability (X_2), speed of information processing (X_3), accuracy of information processing (X_4), flexibility of information processing (X_5), steepness of learning curve in automatization (X_6), total amount of work done in automatization (X_7), transfer (X_8), memory and understanding (X_9), English Vocabulary (X_{10}), English Reading Comprehension (X_{11}) and prior learning (X_{12}) each significantly explain variance in the composite criterion, first semester, first year weighted grade point average (Y_{GPA}).

H_{0j}: $\rho[X_i, Y_{GPA}] = 0$; j=21, 22, …, 31; i=2, 3, …, 12

H_{aj}: $\rho[X_i, Y_{GPA}] > 0$; j=21, 22, …, 31; i=2, 3, …, 12

Hypothesis 11

Global learning potential (X_1) significantly explains variance in the composite criterion, first semester, first year weighted grade point average (Y_{GPA}).

H_{032}: $\rho[X_1, Y_{GPA}] = 0$

H_{a32}: $\rho[X_1, Y_{GPA}] > 0$

Hypothesis 12

Conceptual reasoning ability (X_2), speed of information processing (X_3), accuracy of information processing (X_4), flexibility of information processing (X_5), steepness of learning curve in automatization (X_6), total amount of work done in automatization (X_7), transfer (X_8), memory and understanding (X_9), English Vocabulary (X_{10}), English Reading Comprehension (X_{11}) and prior learning (X_{12}) each significantly explain unique variance in the composite criterion, first semester, first year weighted grade point average (Y_{GPA}), not explained by the other variables included in the prediction model.

H_{033}: $\beta_1[X_2]=0|\beta_2[X_j]\neq 0$; j=3, 4, …, 12

H_{a33}: $\beta_1[X_2]>0|\beta_2[X_j]\neq 0$; j=3, 4, …, 12

H_{034}: $\beta_1[X_3]=0|\beta_2[X_j]\neq 0$; j=2, 4, …, 12

H_{a34}: $\beta_1[X_3]>0|\beta_2[X_j]\neq 0$; j=2, 4, …, 12

H_{035}: $\beta_1[X_4]=0|\beta_2[X_j]\neq 0$; j=2, 3, 5, …, 12

---

[4] Strictly speaking only X_1, X_9, X_{10}, X_{11} and X_{12} should be considered for inclusion in a selection battery since, according to the performance hypothesis depicted in Figure 3.3, only these five variables explain variance in Y_{GPA}.

H_{a35}: $\beta_1[X_4]>0|\beta_2[X_j]\neq0$; j=2, 3, 5, …, 12

H_{036}: $\beta_1[X_5]=0|\beta_2[X_j]\neq0$; j=2, 3, 4, 6, …, 12

H_{a36}: $\beta_1[X_5]>0|\beta_2[X_j]\neq0$; j=2, 3, 4, 6, …, 12

H_{037}: $\beta_1[X_6]=0|\beta_2[X_j]\neq0$; j=2, …, 5, 7, …, 12

H_{a37}: $\beta_1[X_6]>0|\beta_2[X_j]\neq0$; j=2, …, 5, 7, …, 12

H_{038}: $\beta_1[X_7]=0|\beta_2[X_j]\neq0$; j=2, …, 6, 8, …, 12

H_{a38}: $\beta_1[X_7]>0|\beta_2[X_j]\neq0$; j=2, …, 6, 8, …, 12

H_{039}: $\beta_1[X_8]=0|\beta_2[X_j]\neq0$; j=2, …, 7, 9, …, 12

H_{a39}: $\beta_1[X_8]>0|\beta_2[X_j]\neq0$; j=2, …, 7, 9, …, 12

H_{040}: $\beta_1[X_9]=0|\beta_2[X_j]\neq0$; j=2, …, 8, 10, …, 12

H_{a40}: $\beta_1[X_9]>0|\beta_2[X_j]\neq0$; j=2, …, 8, 10, …, 12

H_{041}: $\beta_1[X_{10}]=0|\beta_2[X_j]\neq0$; j=2, …, 9, 11, 12

H_{a41}: $\beta_1[X_{10}]>0|\beta_2[X_j]\neq0$; j=2, …, 9, 11, 12

H_{042}: $\beta_1[X_{11}]=0|\beta_2[X_j]\neq0$; j=2, …, 10, 12

H_{a42}: $\beta_1[X_{11}]>0|\beta_2[X_j]\neq0$; j=2, …, 10, 12

H_{043}: $\beta_1[X_{12}]=0|\beta_2[X_j]\neq0$;j=2, …, 11

H_{a43}: $\beta_1[X_{12}]>0|\beta_2[X_j]\neq0$; j=2, …, 11

All the predictors significantly explaining unique variance in the composite criterion will be combined in a weighted linear predictor composite (X_{comp}) in accordance with their partial regression weights in a multiple regression model. The following hypotheses will be tested to evaluate the presence of predictive bias in the weighted composite of significant predictors.

Hypothesis 13

The error variance of the regression of the composite criterion (Y_{GPA}) on the weighted linear predictor composite (X_{comp}) is the same for black and white learners.

H_{044}: $\sigma^2[Y_{GPA}|X_{comp};X_{13}=0]=\sigma^2[Y_{GPA}|X_{comp};X_{13}=1]$

H_{a44}: $\sigma^2[Y_{GPA}|X_{comp};X_{13}=0]\neq\sigma^2[Y_{GPA}|X_{comp};X_{13}=1]$

Hypothesis 14

The regression of the composite criterion (Y_{GPA}) on the weighted linear predictor composite (X_{comp}) does not coincide for black and white learners.

H_{045}: $\beta_2[X_{13}]=\beta_3[X_{comp}*X_{13}]=0|\beta_1[X_{comp}]\neq 0$

H_{a45}: $\beta_2[X_{13}]\neq\beta_3[X_{comp}*X_{13}]\neq 0|\beta_1[X_{comp}]\neq 0$

Hypothesis 15[5]

The regression of the composite criterion (Y_{GPA}) on the weighted linear predictor composite (X_{comp}) differs in terms of slope between black and white learners.

H_{046}: $\beta_3[X_{comp}*X_{13}]=0|\beta_1[X_{comp}]\neq 0$; $\beta_2[X_{13}]\neq 0$

H_{a46}: $\beta_3[X_{comp}*X_{13}]\neq 0|\beta_1[X_{comp}]\neq 0$; $\beta_2[X_{13}]\neq 0$

Hypothesis 16[6]

The regression of the composite criterion (Y_{GPA}) on the weighted linear predictor composite (X_{comp}) differs in terms of intercept between black and white learners.

H_{047a}: $\beta_2[X_{13}]=0|\beta_1[X_{comp}]\neq 0$; $\beta_3[X_{comp}*X_{13}]\neq 0$[7]

H_{a47a}: $\beta_2[X_{13}]\neq 0|\beta_1[X_{comp}]\neq 0$; $\beta_3[X_{comp}*X_{13}]\neq 0$

H_{047b}: $\beta_2[X_{13}]=0|\beta_1[X_{comp}]\neq 0$; $\beta_3[X_{comp}*X_{13}]=0$

H_{a47b}: $\beta_2[X_{13}]\neq 0|\beta_1[X_{comp}]\neq 0$; $\beta_3[X_{comp}*X_{13}]=0$

No statistical hypothesis was formulated with regards to the utility of the fair use of the weighted linear predictor composite (X_{comp}).

3.5 STATISTICAL ANALYSES

H_{01} to H_{012} were tested by calculating a matrix of zero-order Pearson correlation coefficients and the corresponding conditional probabilities $P[|r_{ij}| \geq r_c|H_0:\rho_{ij}=0]$. Given a 5% significance level and directional alternative hypotheses, $H_0:\rho_{ij}=0$ were rejected if $P[|r_{ij}| \geq r_c|H_0:\rho_{ij}=0] < 0,05$. The convention proposed by Guilford (cited in Tredoux & Durrheim, 2002, p. 184) and depicted in Table 3.1 was used to interpret the magnitude of the obtained sample correlation coefficients.

---

[5] H_{046} will be tested only if H_{045} is rejected ($p<0,05$).

[6] H_{047a} will be tested if H_{045} and H_{046} are rejected ($p<0,05$). H_{047b} will be tested if H_{045} is rejected ($p<0,05$) but H_{046} is not rejected ($p>0,05$).

[7] If H_{047a} is not rejected ($p>0,05$), H_{048}: $\beta_3[X_{comp}*X_{13}]=0|\beta_1[X_{comp}]\neq 0$; $\beta_2[X_{13}]=0$ could be tested. This would, however, be redundant since logically H_{048} must be rejected if H_{045} and H_{047a} are rejected ($p<0,05$).

Table 3.1 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

Hypothesis 7 (H_{013} – H_{015}) was tested by fitting the following multiple regression model to the data using standard multiple regression:

$$E[X_6|X_3,X_4, X_5] = \alpha + \beta_1[X_3] + \beta_2[X_4] + \beta_3[X_5]$$

Hypothesis 8 (H_{016} – H_{018}) was tested by fitting the following multiple regression model to the data using standard multiple regression:

$$E[X_7|X_3,X_4, X_5] = \alpha + \beta_1[X_3] + \beta_2[X_4] + \beta_3[X_5]$$

Hypothesis 9 (H_{019} & H_{020}) was tested by fitting the following multiple regression model to the data using standard multiple regression:

$$E[X_9|X_6,X_7] = \alpha + \beta_1[X_6] + \beta_2[X_7]$$

Hypothesis 10 and 11 (H_{021} to H_{032}) were tested by calculating a matrix of zero-order Pearson correlation coefficients and the corresponding conditional probabilities $P[|r_{ij}| \geq r_c|H_0{:}\rho_{ij}=0]$. Given a 5% significance level and directional alternative hypotheses, $H_0{:}\rho_{ij}=0$ were rejected if $P[|r_{ij}| \geq r_c|H_0{:}\rho_{ij}=0] < 0,05$.

Hypothesis 12 (H_{033} – H_{043}) was tested by fitting the following multiple regression model to the data using standard multiple regression:

$$E[Y_{GPA}|X_2,X_3, \ldots, X_{12}] = \alpha + \beta_1[X_2] + \beta_2[X_3] + \ldots + \beta_{11}[X_{12}]$$

Hypothesis 13 (H_{044}) was tested by regressing Y_{GPA} on X_{comp} for white and black learners separately. The following test statistic was subsequently calculated:

$$F= s^2[Y_{GPA}|X_{comp};X_{13}=0]/s^2[Y_{GPA}|X_{comp};X_{13}= 1]$$

In the calculation of the F-ratio it is assumed that $s^2[Y_{GPA}|X_{comp};X_{13}=0] > s^2[Y_{GPA}|X_{comp};X_{13}=1]$ from the output obtained from two separate standard simple regression analyses:

Hypothesis 14 (H_{045}) was tested by fitting the following two regression models to the data via standard multiple regression analysis:

$$E[Y_{GPA}|X_{comp}] = \alpha + \beta_1[X_{comp}]$$

$$E[Y_{GPA}|X_{comp},X_{13},X_{13}{*}\ X_{comp}] = \alpha + \beta_1[X_{comp}] + \beta_2[X_{13}] + \beta_3[X_{comp}{*}X_{13}]$$

If H_{045} is not rejected the fairness analysis will terminate since the regression of the criterion on the weighted composite will coincide in the two groups.

If H_{045} is rejected, hypothesis 15 (H_{046}) will be tested by fitting the following two multiple regression models to the data using standard multiple regression analysis:

$$E[Y_{GPA}|X_{comp}] = \alpha + \beta_1[X_{comp}] + \beta_2[X_{13}]$$

$$E[Y_{GPA}|X_{comp},X_{13},X_{13}{*}\ X_{comp}] = \alpha + \beta_1[X_{comp}] + \beta_2[X_{13}] + \beta_3[X_{comp}{*}X_{13}]$$

If H_{046} is not rejected, hypothesis 16 (H_{047b}) will be tested by fitting the following two multiple regression models to the data using standard multiple regression analysis:

$$E[Y_{GPA}|X_{comp}] = \alpha + \beta_1[X_{comp}]$$

$$E[Y_{GPA}|X_{comp},X_{13},X_{13}{*}\ X_{comp}] = \alpha + \beta_1[X_{comp}] + \beta_2[X_{13}]$$

If H_{046} rejected, hypothesis 16 (H_{047a}) will be tested by fitting the following two multiple regression models to the data using standard multiple regression analysis:

$$E[Y_{GPA}|X_{comp}] = \alpha + \beta_1[X_{comp}] + \beta_3[X_{comp}{*}X_{13}]$$

$$E[Y_{GPA}|X_{comp},X_{13},X_{13}{*}\ X_{comp}] = \alpha + \beta_1[X_{comp}] + \beta_2[X_{13}] + \beta_3[X_{comp}{*}X_{13}]$$

The utility of the fair use of the weighted linear predictor composite (X_{comp}) was examined in terms of the Taylor-Russell interpretation of selection utility (Taylor & Russell, 1939). The base rate (BR) was consequently calculated for the validation sample, given a minimum acceptable grade point average of 50%. The success ratio for various possible selection ratio's (SR) were subsequently calculated as well as the increase in the proportion selectees successful when selecting the best SR fairly with the weighted linear predictor composite (X_{comp}) rather than randomly.

The utility of the fair use of the weighted linear predictor composite (X_{comp}) was also examined in terms of the Naylor-Shine interpretation of selection utility (Naylor & Shine, 1965). The expected grade point average of the selected group achieved when using the weighted linear predictor composite (X_{comp}) fairly was consequently calculated under various selection ratio’s and compared to the expected grade point average of the selected group achieved under random selection.

A criterion-referenced norm table that expresses the risk of failure conditional on expected academic performance derived from the fair use of the weighted linear predictor composite (X_{comp}) was calculated by transforming the critical grade point average of 50% in the conditional Y_{GPA} distribution (conditional on X_{comp}) to a standard normal score (i.e., a $z_{critical}$-score) and determining the probability of obtaining a standard normal score or less than equal to $z_{critical}$ at selected X_{comp} values.

3.6 SAMPLING

The sample used for this study consisted of three year groups (First Year Students of 2001, 2002 and 2003) enrolled at the SA Military Academy for six different study directions. Table 3.3 – Table 3.5 reflects the distribution of the sample regarding the year group, gender, and race.

Table 3.2. Frequency Distribution of the Year Groups

Year Group					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2001	58	29.4	29.4	29.4
	2002	70	35.5	35.5	65.0
	2003	69	35.0	35.0	100.0
	Total	197	100.0	100.0	

Table 3.3. Frequency Distribution of Gender

Gender					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Female	40	20.3	20.4	20.4
	Male	156	79.2	79.6	100.0
	Total	196	99.5	100.0	
Missing	System	1	.5		
Total		197	100.0		

Table 3.4. Frequency Distribution of Race

Race					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Black	127	64.5	64.8	64.8
	White	69	35.0	35.2	100.0
	Total	196	99.5	100.0	
Missing	System	1	.5		
Total		197	100.0		

From Tables 3.2 and 3.3 it is apparent that gender and race is not uniformly distributed in the sample although the three year groups are somewhat more evenly distributed. The majority of learners were Black males

3.7 EXTERNAL VALIDITY OF THE VALIDATION DESIGN

Traditionally two validation designs are distinguished, namely a current employee design and a follow-up design. This simple dichotomous distinction, however, does not provide a satisfactorily comprehensive coverage of the different design possibilities. Sussmann and Robertson (1986) propose a more comprehensive classification comprising 11 different designs. In this particular study a variation on the Sussmann and Robertson design 5 is used. It is essentially a follow-up design in which all applicants are tested on the experimental battery, selection decisions are based on the experimental battery and criterion data is obtained after a short tenure. In evaluating the chosen design two aspects (at least) should be considered, namely the practical feasibility of the design, and the transportability of the study findings. The latter aspect would depend on:

- The extent to which the conditions under which the selection procedure is examined differs from the conditions under which the procedure will be used.
- The extent to which those aspects on which the two sets of conditions differ affect the quantities which are examined/described in the validation study.

The following aspects could differ across the three sets of conditions:

- The homogeneity of the validation sample versus the homogeneity of the applicant sample [restriction of range]
- The test motivation of testees
- Attributes which are affected by job experience

To explicitly consider these differences the exact nature of the actual selection design should be spelled out. The crucial aspects of the selection design that affect transportability is whether the selection procedure under evaluation will:

- Replace the existing selection procedure;
- Follow on the existing selection procedure in a multiple hurdle fashion; or
- Combine with the existing selection procedure in a linear composite as a single step selection stage.

It would therefore be wrong, although this is quite often the case, to show unqualified preference (from a theoretical perspective) for a predictive design.

The external validity of the validation design is, given the applied nature of the research, of critical importance since it affects the validity and credibility of (implicit) claims made with regards to the selection procedure in actual operation. In this case the selection procedure under evaluation is the existing procedure. The data provided in the sample only reflects the first year first semester results of three different intake groups or year groups at the SA Military Academy. These groups were selected based on their results on the battery being psychometrically evaluated. Although the test results of learners who did not make the selection are available, no criterion data is available for these learners. This results in data with a restricted range and interpretation of results should take this into consideration. Restriction of range will attenuate the obtained validity coefficients (Guion, 1991). Although the validation design therefore realistically simulates the actual selection design in terms of most characteristics it nonetheless fails to mirror the conditions under which selection eventually will occur in as far as it assumes a too homogenous applicant group. It is, however, possible to correct validity coefficients for restriction of range. The appropriate correction formula depends on the type of restriction of range (Guion, 1991; Thorndike, 1982).

Case 2 [Case A] exists if the correlation to be corrected is between two variables X and Y, selection occurred [directly/explicitly] on the variable X [or Y] through complete truncation on X at X_C [or on Y at Y_C] and both restricted and unrestricted variances are known only for the explicit selection variable X [or Y]. The validity coefficient corrected for Case 2 [Case A] restriction of range is given by (Thorndike, 1982):

$$r[X,Y] = Kr[x,y]/(K^2r^2[x,y]+1-r^2[x,y])^{1/2}$$

Where: $K = s[X]/s[x]$

Hypothesis 12 could have been tested via correlation and regression analysis for specific modules by using the module performance mark as the criterion measure instead of the GPA. Since selection decisions will never be based on expected module performance, formal statistical hypotheses were not formulated for these analyses in advance. One should, however, probably explore the possibility that study direction might explain variance in GPA as a main effect and/or in interaction with specific predictors or even X_{comp}.

When considering the various study directions the sample size decreases dramatically, which prevents the use of regression analysis for determination of the predictive power of the selection test battery for those specific groups.

CHAPTER 4
RESULTS AND CONCLUSIONS

4.1 INTRODUCTION

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 3.3 specific statistical hypotheses were formulated. Specific statistical hypotheses were moreover formulated to examine the predictive validity of the individual sub-tests of the APIL battery, to examine the merits of combining the various sub-tests in a selection battery and to examine the fairness of the battery. The purpose of this chapter is to report the results of the statistical analyses aimed at testing the stated null hypotheses.

4.2 THE PROPOSITIONS MADE BY THE PERFORMANCE HYPOTHESIS

4.2.1 The relationship between predictors and the composite criterion

The following discussion concentrates on the results obtained from the statistical analysis as it relates to Hypotheses 1 to 6 (H_{01} to H_{012}). Pearson correlation coefficients were calculated to determine the nature of the relationships between each of the predictors (APIL-B Global Score (X_1), conceptual reasoning ability (X_2), speed of information processing (X_3), accuracy of information processing (X_4), flexibility of information processing (X_5), steepness of learning curve in automatization (X_6), total amount of work done in automatization (X_7), transfer (X_8), memory and understanding (X_9), English Vocabulary (X_{10}), English Reading Comprehension (X_{11}) and prior learning (X_{12}) and the composite criterion, first semester, first year weighted grade point average (Y_{GPA}). The calculated correlation coefficients are displayed in Table 4.1 and will be referred to in the following paragraphs.

Table 4.1. Matrix of zero-order Pearson correlation coefficients and the corresponding conditional probabilities

Correlations

		X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	Y
X2	Pearson Correlation	1	.571**	.456**	.550**	.498**	.566**	.578**	.591**	.358**	.503**	.493**	.209**
	Sig. (1-tailed)		.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.002
	N	197	197	193	197	197	197	196	197	193	193	192	197
X3	Pearson Correlation	.571**	1	.360**	.683**	.546**	.729**	.683**	.501**	.355**	.550**	.488**	.197**
	Sig. (1-tailed)	.000		.000	.000	.000	.000	.000	.000	.000	.000	.000	.003
	N	197	197	193	197	197	197	196	197	193	193	192	197
X4	Pearson Correlation	.456**	.360**	1	.517**	.392**	.465**	.453**	.386**	.225**	.366**	.235**	.029
	Sig. (1-tailed)	.000	.000		.000	.000	.000	.000	.000	.001	.000	.001	.346
	N	193	193	193	193	193	193	192	193	189	189	188	193
X5	Pearson Correlation	.550**	.683**	.517**	1	.491**	.658**	.667**	.456**	.314**	.523**	.475**	.165*
	Sig. (1-tailed)	.000	.000	.000		.000	.000	.000	.000	.000	.000	.000	.010
	N	197	197	193	197	197	197	196	197	193	193	192	197
X6	Pearson Correlation	.498**	.546**	.392**	.491**	1	.841**	.588**	.627**	.279**	.473**	.465**	.164*
	Sig. (1-tailed)	.000	.000	.000	.000		.000	.000	.000	.000	.000	.000	.011
	N	197	197	193	197	197	197	196	197	193	193	192	197
X7	Pearson Correlation	.566**	.729**	.465**	.658**	.841**	1	.716**	.674**	.390**	.620**	.523**	.161*
	Sig. (1-tailed)	.000	.000	.000	.000	.000		.000	.000	.000	.000	.000	.012
	N	197	197	193	197	197	197	196	197	193	193	192	197
X8	Pearson Correlation	.578**	.683**	.453**	.667**	.588**	.716**	1	.515**	.426**	.605**	.432**	.141*
	Sig. (1-tailed)	.000	.000	.000	.000	.000	.000		.000	.000	.000	.000	.025
	N	196	196	192	196	196	196	196	196	192	192	191	196

Table 4.1. Matrix of zero-order Pearson correlation coefficients and the corresponding conditional probabilities (continued)

		X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	Y
X9	Pearson Correlation	.591**	.501**	.386**	.456**	.627**	.674**	.515**	1	.268**	.497**	.446**	.233**
	Sig. (1-tailed)	.000	.000	.000	.000	.000	.000	.000		.000	.000	.000	.000
	N	197	197	193	197	197	197	196	197	193	193	192	197
X10	Pearson Correlation	.358**	.355**	.225**	.314**	.279**	.390**	.426**	.268**	1	.734**	.450**	.314**
	Sig. (1-tailed)	.000	.000	.001	.000	.000	.000	.000	.000		.000	.000	.000
	N	193	193	189	193	193	193	192	193	193	193	188	193
X11	Pearson Correlation	.503**	.550**	.366**	.523**	.473**	.620**	.605**	.497**	.734**	1	.574**	.295**
	Sig. (1-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000		.000	.000
	N	193	193	189	193	193	193	192	193	193	193	188	193
X12	Pearson Correlation	.493**	.488**	.235**	.475**	.465**	.523**	.432**	.446**	.450**	.574**	1	.431**
	Sig. (1-tailed)	.000	.000	.001	.000	.000	.000	.000	.000	.000	.000		.000
	N	192	192	188	192	192	192	191	192	188	188	192	192
Y	Pearson Correlation	.209**	.197**	.029	.165*	.164*	.161*	.141*	.233**	.314**	.295**	.431**	1
	Sig. (1-tailed)	.002	.003	.346	.010	.011	.012	.025	.000	.000	.000	.000	
	N	197	197	193	197	197	197	196	197	193	193	192	197

** Correlation is significant at the 0,01 level (1-tailed) if $P[|r_{ij}| \geq r_c|H_0{:}\rho_{ij}=0] < 0{,}01$
* Correlation is significant at the 0,05 level (1-tailed) if $P[|r_{ij}| \geq r_c|H_0{:}\rho_{ij}=0] < 0{,}05$

4.2.2 The relationship between fluid intelligence and transfer

Hypothesis 1 postulates that fluid intelligence (X_2) has a positive directional effect on transfer (X_8). Table 4.1 indicates a substantial relationship (r=0,578), and the probability for this moderate correlation between fluid intelligence and transfer under H_0 was significant (p=0,000). H_{01} can therefore be rejected, and Hypothesis 1, stating fluid intelligence has a positive effect on transfer could not be refuted. This result is consistent with the theory presented in this study.

4.2.3 The relationship between speed, accuracy and flexibility of information processing each with the steepness of the learning curve in automatization and the total amount of work done in automatization

Hypothesis 2 proposes that speed (X_3), accuracy (X_4) and flexibility (X_5) of information processing each has a positive directional effect on the steepness of the learning curve in automatization (X_6) and the total amount of work done in automatization (X_7). Table 4.1 specifies the following:

- a substantial relationship (r=0,546) was found between speed of information processing (X_3) and steepness of the learning curve in automatization (X_6), the probability for this moderate correlation under H_0 was significant (p=0,000). H_{02} can therefore be rejected;
- a definite but small relationship (r=0,392) exists between accuracy of information processing (X_4) and steepness of the learning curve in automatization (X_6), the probability for this low correlation under H_0 was significant (p=0,000). Even though only a low correlation exists, H_{03} can be rejected in the favour of the alternative hypothesis;
- a substantial relationship (r=0,491) was found between flexibility of information processing (X_5) and steepness of the learning curve in automatization (X_6), the probability for this moderate correlation under H_0 was significant (p=0,000). H_{04} can comfortably be rejected in favour of the alternative hypothesis;
- a strong relationship (r=0,729) is evident between speed of information processing (X_3) and the total amount of work done in automatization (X_7), the probability for this high correlation under H_0 was also significant (p=0,000). H_{05} can be rejected with confidence in favour of the alternative hypothesis;

- a substantial relationship ($r=0,465$) was established between accuracy of information processing (X_4) and the total amount of work done in automatization (X_7), the probability for this moderate correlation under H_0 was significant with $p=0,000$. H_{06} can therefore be rejected in support of H_{a6};
- a substantial relationship ($r=0,658$) was observed between flexibility of information processing (X_5) and the total amount of work done in automatization (X_7) the probability for this moderate correlation under H_0 was significant ($p=0,000$) and H_{07} can therefore be rejected;

In conclusion, the results obtained implies that the relationships hypothesized between speed, accuracy and flexibility of information processing respectively and the steepness of the learning curve in automatization and the total amount of work done in automatization could not be refuted. The results however suggest that accuracy of information processing has a less pronounced effect on the steepness of the learning curve in automatization.

This result is consistent with the theory presented in this study, suggesting that all three components of information processing have an influence on the steepness of the learning curve in automatization. These results indicate though that the accuracy with which information is processed only has a modest influence on the rate at which an individual becomes more skilled and efficient in performing a new task.

4.2.4 The relationship between both the steepness of the learning curve in automatization and the total amount of work done in automatization with crystallized abilities

The proposition made by Hypothesis 3 is that the steepness of the learning curve in automatization (X_6) and the total amount of work done in automatization (X_7) both have a positive directional effect on the level of crystallized ability (X_9) development. Table 4.1 lists the following:

- a substantial relationship ($r=0,627$) between steepness of the learning curve in automatization (X_6) and crystallized abilities (X_9) was found, the probability for this moderate correlation under H_0 was significant ($p=0,000$) and H_{08} can therefore be rejected;

- a substantial relationship ($r=0,674$) was established between work done in automatization (X_7) and crystallized abilities (X_9), the probability for this moderate correlation under H_0 was significant ($p=0,000$) and H_{09} can therefore be rejected;

Hypothesis 3 could therefore not be proven false. This result is consistent with the theory presented in this study.

4.2.5 The relationship between transfer and crystallized abilities

According to Hypothesis 4 a positive directional relationship exists between transfer (X_8) and crystallized abilities (X_9). From the information in Table 4.1, a substantial relationship ($r=0,515$) is evident between transfer (X_8) and crystallized abilities (X_9). Again, the probability for this moderate correlation under H_0 was significant ($p=0,000$) and H_{010} can therefore be rejected. Hypothesis 4 can also not be contested.

4.2.6 The relationship between crystallized abilities and learning performance

Hypothesis 5 postulates a positive directional effect crystallized abilities (X_9) has on learning performance (Y_{GPA}). According to Table 4.1, a definite but small relationship ($r=0,233$) exist between crystallized abilities (X_9) and learning performance (Y_{GPA}). The probability for this low correlation under H_0 was significant ($p=0,000$). H_{011} can therefore not be rejected. Given the argument presented earlier (paragraph 2.7, p. 35) on the attenuating effect of prior learning and language proficiency on the correlation between crystallized abilities developed via academic learning and the crystallized abilities developed via the APIL learning task, the finding that crystallized abilities (X_9) does not drastically influence learning performance (Y_{GPA}) is not altogether surprising. A substantially stronger correlation would, however, be expected between a measure of the crystallized abilities developed via academic learning and learning performance (Y_{GPA}).

4.2.7 The relationship between prior learning and learning performance

Hypothesis 6 proposes that prior learning (X_{12}) has a positive directional effect on learning performance (Y_{GPA}). As seen in Table 4.1, the results indicate a substantial relationship and moderate correlation that was significant ($r=0,431$ and $p=0,000$). This result supports

Hypothesis 6, suggesting that prior learning (X_{12}) significantly influences learning performance (Y_{GPA}).

4.2.8 The extent to which speed, accuracy and flexibility of information processing each significantly explains unique variance in the steepness of the learning curve in automatization

Table 4.2. Regression of speed, accuracy and flexibility of information processing on the steepness of the learning curve in automatization

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	47815.724	3	15938.575	32.570	.000[a]
	Residual	92488.987	189	489.360		
	Total	140304.711	192			

a. Predictors: (Constant), X5, X4, X3
b. Dependent Variable: X6
R Squared = .341 (Adjusted R Squared = .330)

Model		Unstandardized Coefficients		Standardized Coefficients			Correlations		
		B	Std. Error	Beta	t	Sig.	Zero-order	Partial	Part
1	(Constant)	-43.448	7.663		5.669	.000			
	Speed (X3)	.303	.066	.368	4.561	.000	.534	.315	.269
	Accuracy (X4)	.002	.001	.183	2.655	.009	.392	.190	.157
	Flexibility (X5)	.321	.191	.147	1.675	.096	.493	.121	.099

a. Dependent Variable: X6

Table 4.2 indicates that speed, accuracy and flexibility of information processing accounts for only 34% of the variance in the gradient of the learning curve in automatization. In other words, only 34% of the variability in the steepness of the learning curve in automatization (X_6) can be accounted for by variability in the weighted linear composite of speed (X_3), accuracy (X_4) and flexibility (X_5) of information processing. However, Hypothesis 7 proposes that speed (X_3), accuracy (X_4) and flexibility (X_5) of information processing each significantly explain unique variance in the steepness of the learning curve in automatization (X_6). The independent contributions of each independent variable to the prediction of the steepness of the learning curve in automatization as depicted in Table 4.2 will be discussed next.

After controlling for the other two independent variables in the predictor and the dependent variable, the unique variance in the speed of information processing (X_3) explains about 10% ($0,315^2$) of the variance in the steepness of the learning curve in automatization (X_6). When the effect of accuracy (X_4) and flexibility (X_5) of information processing is removed only from the predictor variance, 7% ($0,269^2$) of the total variance in the steepness of the learning curve in automatization (X_6) can be attributed to the unique variance in the speed of information processing (X_3). Table 4.2 indicates that the speed of information processing main effect does significantly (p=0,000) explain variance in the steepness of the learning curve in automatization (X_6) when included in a model already containing accuracy (X_4) and flexibility (X_5) of information processing. H_{013} can therefore be rejected.

When the accuracy of information processing (X_4) is correlated with the steepness of the learning curve in automatization (X_6), after controlling for the other two independent variables (in the predictor as well as the dependent variable), approximately 4% ($0,190^2$) of the variance in the steepness of the learning curve in automatization (X_6) can be attributed to the accuracy of information processing (X_4) above and beyond the effect of speed and flexibility of information processing. The semi-partial correlation between the accuracy of information processing (X_4) and the steepness of the learning curve in automatization (X_6) is $0,157^2$. This indicates that only approximately 2,5% of the variance in X_6 can be explained by the unique variance in $X_{4.}$ Accuracy of information processing main effect does significantly (p=0,009) explain variance in the steepness of the learning curve in automatization when included in a model already containing X_3 and X_5. H_{014} can therefore be rejected.

After controlling for speed of information processing (X_3) and accuracy of information processing (X_4) in the predictor as well as the dependent variable, the independent contribution of the flexibility of information processing (X_5) to the prediction of the steepness of the learning curve in automatization (X_6) is merely 1,5% ($0,121^2$). When controlling for X_3 and X_4 only in the predictor variable, the variance in X_6, which can be attributed to unique variance in X_5, is only 1% ($0,099^2$). In other words the flexibility of information processing (X5) does not really explain variability in the steepness of the learning curve in automatization (X6) above and beyond what can be explained by speed and accuracy of information processing. Furthermore, the probability of the partial regression coefficient sample estimate associated with X_5 under H_{015} is not significant (p=0,096). The non-significance of X_5 can be attributed to the fact the X_3 and X_5 are correlated (r=0,683), as are X_4 and X_5 (r=0,517). H_{015} can therefore not be rejected in favour of the alternative hypothesis.

The standardized regression coefficients, partial correlation coefficients and semi-partial correlation coefficients associated with the two significant effects in this model indicate that the speed of information processing (X_3) is the more important of the two significant predictors in explaining the steepness of the learning curve in automatization (X_6). The flexibility of information processing does not explain variance in the steepness of the learning curve in automatization that is not explained by speed and accuracy of information processing.

4.2.9 The extent to which speed, accuracy and flexibility of information processing each explain unique variance in the total amount of work done in automatization

Table 4.3. Regression of speed (X_3), accuracy (X_4) and flexibility (X_5) of information processing on the total amount of work done in automatization (X_7)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1586819.957	3	528939.986	92.490	.000[a]
	Residual	1080874.327	189	5718.912		
	Total	2667694.283	192			

a. Predictors: (Constant), X5, X4, X3
b. Dependent Variable: X7
R Squared = .595 (Adjusted R Squared = .588)

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.	Correlations Zero-order	Partial	Part
1	(Constant)	-56.307	26.198		-2.149	.033			
	Speed (X3)	1.864	.227	.518	8.199	.000	.725	.512	.380
	Accuracy (X4)	.007	.002	.167	3.077	.002	.465	.218	.142
	Flexibility (X5)	2.049	.654	.216	3.132	.002	.655	.222	.145

a. Dependent Variable: X7

Table 4.3 above shows that speed, accuracy and flexibility of information processing accounts for 59% of the variance in the total amount of work done in automatization. In other words, almost 60% of the variability in the total amount of work done in automatization (X_7) can be accounted for by variability in the weighted linear composite of speed (X_3), accuracy (X_4) and flexibility (X_5) of information processing. However, Hypothesis 8 proposes that speed (X_3), accuracy (X_4) and flexibility (X_5) of information processing each significantly explain unique variance in the total amount of work done in automatization (X_7).

When controlling for the other two independent variables in the predictor and the dependent variable, the unique variance in the speed of information processing (X_3) explains about 26% ($0{,}512^2$) of the variance in the total amount of work done in automatization (X_7). When the effect of accuracy (X_4) and flexibility (X_5) of information processing is removed only from the predictor variance, about 14% ($0{,}380^2$) of the total variance in the dependant variable (X_7) can be attributed to the unique variance in the speed of information processing (X_3). Table 4.3 indicates that the speed of information processing main effect does significantly (p=0,000) explain variance in the total amount of work done in automatization (X_7) when included in a model already containing accuracy (X_4) and flexibility (X_5) of information processing. H_{016} can therefore be rejected.

When the accuracy of information processing (X_4) is correlated with the total amount of work done in automatization (X_7), controlling for the other two independent variables (in the predictor as well as the dependent variable), approximately 5% ($0{,}218^2$) of the variance in the dependent variable (X_7) can be attributed to variance in the accuracy of information processing (X_4) above and beyond the effect of X_3 and X_5. The semi-partial correlation between the accuracy of information processing (X_4) and the total amount of work done in automatization (X_7) is 0,142. Therefore only approximately 2% of the variance in X_7 can be explained by the unique variance in $X_{4.}$ The accuracy of information processing main effect does significantly (p=0,002) explain variance in the steepness of the learning curve in automatization when included in a model already containing X_3 and X_5. H_{017} can therefore be rejected.

When controlling for X_3 and X_4 in the predictor and dependent variables, the independent contribution of the flexibility of information processing (X_5) to the prediction of the total amount of work done in automatization (X_7) is only about 5% ($0{,}222^2$). When controlling for X_3 and X_4 only in the predictor variable, the variance in X_7, which can be attributed to unique variance in X_5, is only 1% ($0{,}145^2$). In other words the flexibility of information processing does not explain a lot of variability in the total amount of work done in automatization above and beyond what can be explained by speed and accuracy of information processing. The partial regression coefficient estimate associated with X_5 is significant (p=0,002) when X_5 is included in a model already containing X_3 and X_4. However small the extent, X_5 nonetheless does significantly explain unique variance in the dependent variable. H_{018} can therefore be rejected. Again it is evident that the correlation between X_3 and X_5 as well as the correlation between X_4 and X_5 (r=0,683 and r=0,517 respectively) explain why X_5 explains only a small proportion of unique variance in X_7 despite the moderate zero-order correlation between X_7 and X_5 in isolation.

The standardized regression coefficients, partial correlation coefficients and semi-partial correlation coefficients associated with the three effects included in this model once again indicate that the speed of information processing (X_3) is the most important of the three predictors in explaining variance in the total amount of work done in automatization (X_7). Speed, accuracy and flexibility of information processing, however, all three significantly explain unique variance in the total amount of work done in automatization. Accuracy and flexibility of information processing are more or less of equal importance in the regression model in their effect on the total amount of work done in automatization

4.2.10 The extent to which the steepness of the learning curve in automatization and the total amount of work done in automatization each significantly explain unique variance in crystallized abilities.

Table 4.4 indicates that the steepness of the learning curve in automatization and the total amount of work done in automatization accounts for almost 47% of the variability in crystallized abilities. In other words, nearly 47% of the variability in crystallized abilities (X_9) can be accounted for by variability in the weighted linear composite of steepness of the learning curve in automatization (X_6) and the total amount of work done in automatization (X_7). However, Hypothesis 9 proposes that steepness of the learning curve in automatization (X_6) and the total amount of work done in automatization (X_7) each significantly explain unique variance in crystallized abilities (X_9).

Table 4.4. Regression of the steepness of the learning curve in automatization (X_6) and the total amount of work done in automatization (X_7) on crystallized abilities (X_9)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2364.616	2	1182.308	84.870	.000[a]
	Residual	2702.582	194	13.931		
	Total	5067.198	196			

a. Predictors: (Constant), X7, X6

b. Dependent Variable: X9

R Squared = .467 (Adjusted R Squared = .461)

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Correlations Zero-order	Correlations Partial	Correlations Part
1	(Constant)	10.655	1.000		10.657	.000			
	Steepness of the learning curve (X6)	.038	.018	.207	2.138	.034	.627	.152	.112
	Total amount of work done (X7)	.022	.004	.500	5.154	.000	.674	.347	.270

a. Dependent Variable: X9

After controlling for total amount of work done in automatization (X_7) in the predictor and the dependent variable, the unique variance in the steepness of the learning curve in automatization (X_6) explains about 2% ($0{,}152^2$) of the variance in crystallized abilities (X_9). When (X_7) is controlled for only in the predictor variance, approximately 1% ($0{,}112^2$) of the total variance in the crystallized abilities can be attributed to the unique variance in the steepness of the learning curve in automatization. Table 4.4 indicates that the effect of X_6 is significant (p=0,018) when the steepness of the learning curve in automatization is included in a model already containing the total amount of work done in automatization. H_{019} can therefore be rejected.

When the steepness of the learning curve in automatization (X_6) is correlated with the crystallized abilities (X_9), controlling for the steepness of the learning curve in automatization (X_6) in the predictor as well as the dependent variable, 12% ($0{,}347^2$) of the variance in crystallized abilities can be attributed to the unique variance in the total amount of work done in automatization. The semi-partial correlation between (X_7) and (X_9) is 0,270. This indicates that about 7% of the variance in X_9 can be explained by the unique variance in X_7. The total amount of work done in automatization does significantly (p=0,000) explain variance in crystallized abilities when included in a model already containing the steepness of the learning curve in automatization. H_{020} can therefore be rejected.

The standardized regression coefficients, partial correlation coefficients and semi-partial correlation coefficients associated with the two independent variables included in this model indicate that the total amount of work done in automatization (X_7) is the most important of the two predictors in accounting for differences in crystallized abilities (X_9).

4.3 THE PREDICTIVE VALIDITY OF THE INDIVIDUAL PREDICTORS OF THE SELECTION BATTERY

4.3.1 The extent to which the each of the predictors significantly explains variance in the composite criterion, first semester first year weighted grade point average

A correlational analysis (see Table 4.1 Matrix of zero-order Pearson correlation coefficients and the corresponding conditional probabilities) was used to test Hypothesis 10. All the predictors significantly ($p<0,05$) explained variance in the criterion except X_4 ($p=0,346$). H_{021} – H_{031} with the exception of H_{023} can therefore be rejected. Accuracy of information processing does not significantly explain variance in first year grade point average. However, all of the significant correlations were low, indicating a definite but small relationship. The significant correlations of these predictors with the criterion ranged from 0,141 to 0,431. Prior learning (X_{12}) explained the most variance in the criterion ($r=0,431^2$). There are therefore eleven predictors in the selection battery under investigation that provide relevant information about first year first semester grade point average. Whether all eleven these predictors explain unique variance in Y_{GPA} that is not explained by the other predictors in the battery will depend on the inter-correlation amongst the predictors.

4.3.2 The extent to which Global Learning potential explains variance in the composite criterion, first semester first year weighted grade point average.

From Table 4.5 below (p. 79) it is evident that a definite but small ($r=0,20$) significant ($p<0,05$) relationship exists between global learning potential (X_1) and first year weighted grade point average (Y_{GPA}). It can thus be concluded that X_1 significantly explains a small portion of variance in the composite criterion. H_{032} can therefore be rejected. The APIL global score is calculated from the subscale scores of the APIL. The various subscale measures contributing to the calculation of the global score will therefore as a mathematical necessity correlate with the global score. There is therefore no point in trying to combine both the individual APIL measures and the global score in a regression model.

Table 4.5. Zero-order Pearson correlation coefficients and the corresponding conditional probabilities for (X_1,Y_{GPA})

		X1	Y
X1	Pearson Correlation	1.000	
	Sig. (1-tailed)		
	N	192	
Y	Pearson Correlation	.200	1.000
	Sig. (1-tailed)	.003	
	N	192	197.000

4.3.3 The extent to which the each of the predictors significantly explains unique variance in the composite criterion, not explained by the other variables.

The criterion Y_{GPA} was subsequently regressed on the array of predictors X_2 – X_{12} including accuracy of information processing (X_4) that was earlier found not to correlate significantly with the criterion. Table 4.6a contains the results of the standard multiple regression analysis used to test Hypothesis 12 (H_{033} – H_{043}). Analysing the results depicted in Table 4.6a clearly indicates that memory and understanding (X_9) and prior learning (X_{12}) are the only two predictors that significantly (p=0,031 and p=0,000 respectively) explain unique variance in the composite criterion, first semester, first year weighted grade point average (Y_{GPA}), that is not explained by the other variables included in the prediction model. H_{040} and H_{043} can therefore be rejected, indicating that X_9 and X_{12} uncover relevant and unique information about first year first semester grade point average not conveyed by the remaining predictors in the model. The standardized partial regression coefficients for these two predictors are however small (X_9 (β=0,209) and X_{12} (β=0,384) respectively) indicating that the criterion is not very responsive to increases in the predictors.

Table 4.6a. Regression of conceptual reasoning ability, speed of information processing, accuracy of information processing, flexibility of information processing, steepness of learning curve in automatization, total amount of work done in automatization, transfer, memory and understanding, English Vocabulary, English Reading Comprehension and prior learning on first semester, first year weighted grade point average

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6262.415	11	569.310	5.032	.000[a]
	Residual	19348.059	171	113.147		
	Total	25610.474	182			

a. Predictors: (Constant), X12, X4, X10, X9, X3, X6, X2, X5, X8, X11, X7
b. Dependent Variable: Y
R Squared = .245 (Adjusted R Squared = .196)

Model		Unstandardized Coefficients		Standardized Coefficients			Correlations		
		B	Std. Error	Beta	t	Sig.	Zero-order	Partial	Part
1	(Constant)	28.130	5.931		4.743	.000			
	Conceptual reasoning ability (X2)	.033	.227	.014	.145	.885	.201	.011	.010
	Speed of information processing (X3)	.021	.041	.057	.510	.610	.158	.039	.034
	Accuracy of information processing (X4)	.000	.000	-.077	-.919	.360	.010	-.070	-.061
	Flexibility of information processing (X5)	-.011	.103	-.011	-.106	.916	.118	-.008	-.007
	Steepness of learning curve (X6)	.022	.057	.051	.389	.698	.140	.030	.026
	Total amount of work done (X7)	-.023	.018	-.224	-1.283	.201	.146	-.098	-.085
	Transfer (X8)	-.188	.175	-.126	-1.079	.282	.108	-.082	-.072
	Memory and understanding (X9)	.496	.228	**.209**	2.180	**.031**	.244	.164	.145
	English vocabulary (X10)	.355	.190	.185	1.872	.063	.291	.142	.124
	English reading comprehension (X11)	.028	.238	.014	.118	.907	.272	.009	.008
	Prior learning (X12)	.394	.092	**.384**	4.303	**.000**	.430	.313	.286

a. Dependent Variable: Y

Figure 3.3 hypothesizes X_9 and X_{12} to directly influence Y_{GPA}. The significant partial regression coefficients for these two predictors support the argument depicted in Figure 3.3. The insignificant partial regression coefficients for X_{10} and X_{11}, however, fail to corroborate the argument underlying Figure 3.3.

H_{032}, H_{033} to H_{039} as well as H_{041} cannot be rejected. Per implication the remaining predictors in the selection battery can be considered redundant because they provide no new information not already conveyed by X_9 and X_{12}. This suggests that only X_9 and X_{12} reveal information about determinants of performance on the criterion that is not provided by the other predictors in the model.

However it could be argued that accuracy of information processing (X_4)should never have been included in the foregoing regression model because the prior correlation analysis (Table 4.1; paragraph 4.3.1) indicated that X_4 does not significantly (p>0,05) explain variance in Y_{GPA} whereas all the other predictors do individually significantly (p<0,05) explain variance in Y_{GPA}. Table 4.6b depicts the results when regressing the criterion YGPA on the array of predictors X_2 – X_{12} excluding accuracy of information processing (X_4) because of the earlier finding that X_4 did not to correlate significantly (p>0,05) with the criterion.

Table 4.6b indicates that the removal of X_4 from the array of predictors affects the significance of the partial regression coefficients of X_9 and X_{10}. Removing X_4 from the variable set on which the criterion is regressed causes the unique contribution of X_9 to become insignificant (p>0,05) and the unique contribution of X_{10} to become significant (p<0,05).

Table 4.6b. Regression of conceptual reasoning ability, speed of information processing, flexibility of information processing, steepness of learning curve in automatization, total amount of work done in automatization, transfer, memory and understanding, English Vocabulary, English Reading Comprehension and prior learning on first semester, first year weighted grade point average

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.490[a]	.240	.196	10.72134
a. Predictors: (Constant), X12, X8, X10, X9, X2, X6, X5, X3, X11, X7				

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6375.838	10	637.584	5.547	.000[a]
	Residual	20230.706	176	114.947		
	Total	26606.544	186			
a. Predictors: (Constant), X12, X8, X10, X9, X2, X6, X5, X3, X11, X7						
b. Dependent Variable: Y						

Table 4.6b. Regression of conceptual reasoning ability, speed of information processing, flexibility of information processing, steepness of learning curve in automatization, total amount of work done in automatization, transfer, memory and understanding, English Vocabulary, English Reading Comprehension and prior learning on first semester, first year weighted grade point average (continued)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	27.031	5.642		4.791	.000			
	X2	-.043	.223	-.018	-.192	.848	.205	-.014	-.013
	X3	.015	.039	.041	.378	.706	.168	.028	.025
	X5	-.027	.100	-.028	-.271	.786	.132	-.020	-.018
	X6	.013	.055	.030	.239	.812	.155	.018	.016
	X7	-.021	.017	-.204	-1.202	.231	.160	-.090	-.079
	X8	-.146	.172	-.096	-.849	.397	.122	-.064	-.056
	X9	.441	.228	.186	1.937	.054	.241	.144	.127
	X10	.388	.189	.202	2.046	**.042**	.305	.152	.134
	X11	-.038	.238	-.019	-.159	.874	.277	-.012	-.010
	X12	.422	.091	.409	4.628	**.000**	.437	.329	.304

a. Dependent Variable: Y

When, based on the findings derived from Table 4.6a, Y_{GPA} is regressed on the weighted combination of X_9 and X_{12} Table 4.7a indicates that only X_{12} significantly ($p<0,05$) explains unique variance in Y_{GPA} when included in a regression model already containing the other predictor.

Table 4.7a. Regression of memory and understanding (X_9) and prior learning (X_{12}) on first semester, first year weighted grade point average

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.433[a]	.187	.179	10.86534

a. Predictors: (Constant), X9, X12

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5145.395	2	2572.697	21.792	.000[a]
	Residual	22312.518	189	118.056		
	Total	27457.913	191			

a. Predictors: (Constant), X9, X12
b. Dependent Variable: Y

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	29.572	4.322		6.843	.000			
	X12	.426	.076	.410	5.602	.000	.431	.377	.367
	X9	.109	.173	.046	.628	.531	.229	.046	.041

a. Dependent Variable: Y

A similar conclusion emerges when, based on the findings derived from Table 4.6b, Y_{GPA} is regressed on the weighted combination of X_{10} and X_{12}. Only X_{12} significantly ($p<0,05$) explains unique variance in Y_{GPA} when included in a regression model already containing the other predictor (X_9).

Table 4.7b. Regression of English Vocabulary and prior learning on first semester, first year weighted grade point average

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.453[a]	.206	.197	10.68987

a. Predictors: (Constant), X12, X10

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5472.497	2	2736.249	23.945	.000[a]
	Residual	21140.552	185	114.273		
	Total	26613.049	187			

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	29.254	4.195		6.974	.000			
	X10	.260	.140	.136	1.852	.066	.305	.135	.121
	X12	.387	.076	.376	5.121	.000	.437	.352	.336

a. Predictors: (Constant), X12, X10

b. Dependent Variable: Y

These conclusions were verified by regressing the eleven significant predictors on the criterion by means of a stepwise multiple regression analysis. Only X_{12} was selected for inclusion in the multiple regression model.

Table 4.8. Regression of the APIL global score and prior learning on first semester, first year weighted grade point average

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.428[a]	.183	.174	10.83772

a. Predictors: (Constant), X1, X12

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4841.261	2	2420.631	20.609	.000[a]
	Residual	21611.944	184	117.456		
	Total	26453.205	186			

Table 4.8. Regression of the APIL global score and prior learning on first semester, first year weighted grade point average (continued)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	28.039	5.053		5.549	.000			
	X12	.479	.082	.463	5.807	.000	.424	.394	.387
	X1	-.160	.176	-.072	-.905	.366	.182	-.067	-.060

a. Predictors: (Constant), X1, X12

b. Dependent Variable: Y

The question prompted by Figure 3.3, however remains whether the addition of the global learning potential score of the APIL (X_1) to a battery already containing prior learning (X_{12}) would not significantly explain variance in Y_{GPA} that is not explained by X_{12}? Table 4.8 indicates that this is not the case. Global learning potential (X_1) does not significantly explain variance in the criterion that is not explained by prior learning (X_{12}). There is therefore no justification for including X_1 in the selection battery along with X_{12}.

4.4 ACTUARIAL DERIVATION OF A WEIGHTED LINEAR PREDICTION MODEL FROM A SET OF PREDICTOR AND CRITERION DATA

A correlational analysis was done to determine which of the predictors should be included in the weighted linear prediction model based on the magnitude and significance of the correlation of the predictors with the criterion (Y_{GPA}). All the predictors significantly ($p<0,05$) explained variance in the composite criterion except X_4 ($p=0,346$). However, all of these correlations were low, indicating a definite but small relationship. The correlations of these predictors with the criterion ranged from $r=0,141$ to 0,431. Moreover the predictors tend to correlate low to moderate and significantly ($p<0,05$) with each other. The result was that only prior learning (X_{12}) and English vocabulary (X_9) significantly explained unique variance in the criterion when included in a multiple regression model containing all twelve predictors. X_4, however, does not significantly ($p>0,05$) explain variance in Y_{GPA} whereas all the other predictors do individually significantly ($p<0,05$) explain variance in Y_{GPA}. When X_4 is removed from the variable set on which the criterion is regressed causes the unique contribution of X_9 to become insignificant ($p>0,05$) and the unique contribution of X_{10} to become significant ($p<0,05$). More importantly when Y_{GPA} is regressed on the weighted combination of X_9 and X_{12} only X_{12} significantly ($p<0,05$) explains unique variance in Y_{GPA}. A similar conclusion emerges when Y_{GPA} is regressed on the weighted combination of X_{10} and X_{12}. Only X_{12} significantly ($p<0,05$) explains unique variance in Y_{GPA}. Consequently there was no need to create a combined weighted linear predictor composite (X_{comp}) which would form the basis of the actuarial mechanical decision rule that would guide selection decisions. Prior learning is the only predictor that warrants inclusion in the actuarial mechanical prediction rule. The actuarial mechanical prediction rule is shown in Table 4.9 and expressed as Equation 2.

Table 4.9. Regression of prior learning on first semester, first year weighted grade point average

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.431[a]	.186	.181	10.84800

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5098.875	1	5098.875	43.329	.000[a]
	Residual	22359.038	190	117.679		
	Total	27457.913	191			

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	30.215	4.192		7.208	.000			
	X12	.447	.068	.431	6.582	.000	.431	.431	.431

a. Predictors: (Constant), X12

b. Dependent Variable: Y

$$E[Y_{GPA}|X_{12}] = a + b_1X_{12}$$
$$= 30{,}215 + 0{,}447X_{12} \quad 2$$

4.5 THE VALIDITY OF THE INFERENCES DERIVED FROM THE PREDICTION MODEL

Validity refers to the extent to which the inferences made from test scores are warranted; to the extent to which the interpretation (i.e. meaning) assigned to test scores is justified (Guion, 1991; 1998). Strictly speaking, what is being validated is therefore not the measuring instrument, nor the measures obtained from the instrument, but rather the inferences made from the measures. In the case of personnel selection the question a validation study needs to answer is therefore whether the clinical or mechanical inferences on the criterion derived from the scores obtained on the predictors are permissible. The regression equation depicted in Equation 1 (see Table 4.9) is the actuarial prediction rule that will form the basis of the selection decision rule. The expected criterion performance of all applicants ($E[Y|X_{12}]$) will be estimated by inserting the measures obtained during selection of prior learning into the regression equation depicted in Table 4.9. The resultant estimated criterion scores will be rank-ordered from high to low, the position of Y_k will be determined in the rank-ordered estimated scores and all applicants with $E[Y|X_{12}] > Y_k$ will be selected. This procedure could be regarded as permissible to the extent to which $E[Y|X_{12}]$ correlates significantly with Y_{GPA}. Table 4.9 indicates that $E[Y|X_{12}]$ correlates 0,431 and statistically significantly ($p<0,05$) with Y_{GPA} The predictions derived from Equation 1 are therefore valid.

Demonstrating that the derived actuarial prediction rule is valid is not sufficient to use the rule to control future admissions to the SA Military Academy. A critical question is whether the selection decision making based on the criterion estimates derived from Equation 1 will unfairly disadvantage any applicant groups?

4.6 AN EVALUATION OF THE FAIRNESS OF THE INFERENCES/ PREDICTIONS DERIVED FROM THE PREDICTION MODEL

The question whether the selection decision making based on the criterion estimates derived from Equation 1 will unfairly disadvantage members of any applicant groups is a difficult question to answer because of the elusive nature of the concept fairness. One man's fair is another man's foul. The term "fairness" is becoming more and more difficult to define. Part of the reason is the political nature of the concept. The concept "fairness" has an emotive connotation, and the judgement of "fairness" is therefore tinted by the glasses of the observer. However, Hunter and Schmidt (1976) have identified three fundamentally different ethical views on selection fairness. Of these three fundamental ethical positions technical guidelines on personnel selection procedures (Equal Employment Opportunity Commission, 1978; Society for Industrial and Organizational Psychology, 2003; Society for Industrial Psychology, 1998) seem to favour unqualified individualism as the basic ethical point of departure. The basic premise is therefore that applicants with an equal probability of succeeding on the job should have an equal probability of obtaining the job, irrespective of group membership (Guion, 1966; 1991; Huysamen, 2002).

More specifically, the technical guidelines on personnel selection procedures (Society for Industrial and Organizational Psychology, 2003; Society for Industrial Psychology, 1998) seem to favour the regression-based models of selection fairness (Cleary, 1968; Einhorn & Bass, 1971; Huysamen, 2002). Fairness, according to Cleary's model of selection fairness, is the absence of differences in regression slopes/or intercepts across the subgroups comprising the applicant population (Cleary, 1968). The Cleary model thus argues that selection decision-making, based on expected criterion performance, can be considered unfair or discriminatory if the position of members of specific groups in the rank-ordered predicted criterion performance is either systematically too low or systematically too high for members of a particular group. This would happen if group membership explains variance in the (unbiased) criterion, either as a main effect or in interaction with the predictors, which is not explained by the predictors, and the selection strategy fails to take group membership into account. Under these conditions the criterion inferences derived from selection instrument scores, could be said to exhibit predictive bias (Guion, 1991; 1998).

The presence of predictive bias in the use of Equation 1 will subsequently be evaluated by testing Hypothesis 13 to 16 (H_{044} – H_{048}). The criterion scores were firstly plotted against X_{12} with group membership as a plot symbol. The scatter plot is depicted in Figure 4.1. The regression of Y_{GPA} on X_{12} appears to differ between black and white students in terms of intercept as well as slope. The above scatter plot suggests that the single, undifferentiated prediction rule [Equation 1] will systematically underestimate the criterion performance of black students in the lower region of the X_{12} axis but that the converse will happen in the upper region of the X_{12} scale. This preliminary finding suggests that the single, undifferentiated prediction rule will make systematic group-related prediction errors when estimating the criterion performance of students. The question to be answered is whether the differences in slope and intercept are significant.

Figure 4.1 moreover seems to indicate that the White group tends to perform higher on average on the predictor than the Black group. A corresponding group difference is not evident on the criterion. This would suggest that the regression of the criterion on the predictor should differ in terms of intercept with the White group having a lower intercept. Again, however, the question is whether the difference in mean predictor performance across the two groups is statistically significant.

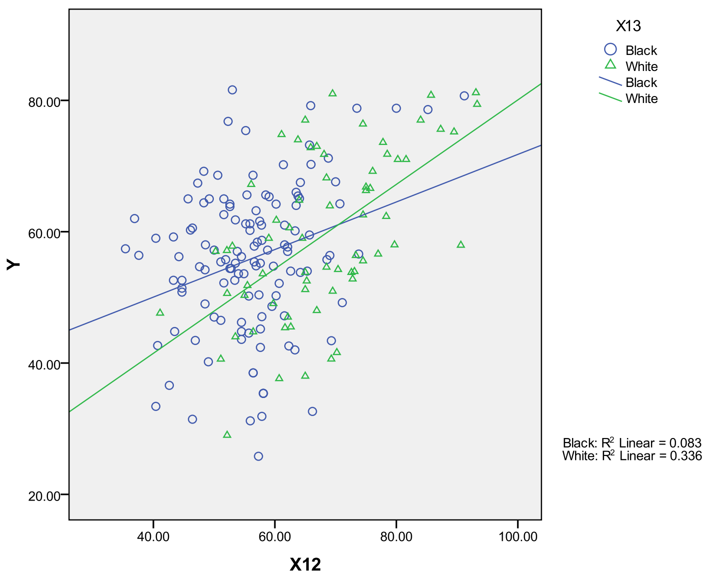

Figure 4.1. Scatter plot of the criterion against the predictor with group-specific regression lines fitted

The standardized residuals resulting from the use of Equation 1 to predict Y_{GPA} is plotted in Figure 4.2 against X_{12} with group membership a plot symbol.

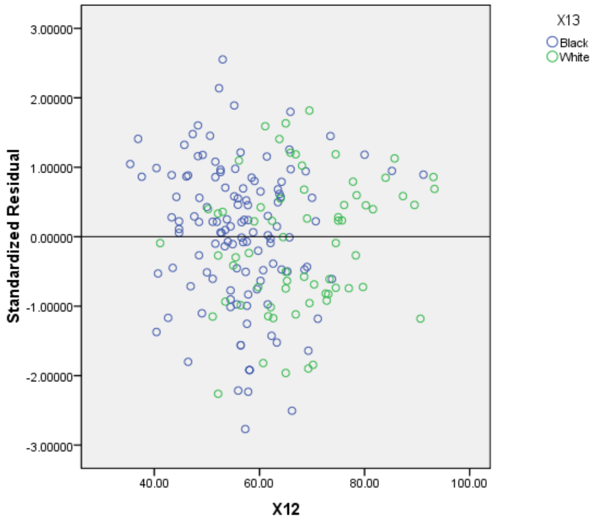

Figure 4.2. Scatter plot of the residuals with groups represented by different plot symbols

The residuals Y-E[Y|X] seems to be comparatively equally distributed. Projecting the residuals from the two groups on the Y-axis does not suggest a systematic group-related difference in the mean residuals. This would suggest that the use of the single, undifferentiated prediction rule (Equation 1) would lead to fair selection decisions. This inference is corroborated by the finding shown in Table 4.10 that the mean standardized residuals do not differ significantly across the two race groups.

Table 4.10. Independent sample T-test of the significance of the difference in the mean residuals obtained from predicting Y_{GPA} from the regression of Y_{GPA} on X_{12}

	X13	N	Mean	Std. Deviation	Std. Error Mean
Standardized Residual	Black	124	.0509462	1.02235154	.09180988
	White	67	-.1007025	.95547949	.11673040

			Standardized Residual	
			Equal variances assumed	Equal variances not assumed
Levene's Test for Equality of Variances		F	.000	
		Sig.	.988	
t-test for Equality of Means		t	1.001	1.021
		df	189	143.455
		Sig. (2-tailed)	.318	.309
		Mean Difference	.15164875	.15164875
		Std. Error Difference	.15154956	.14850939
	95% Confidence Interval of the Difference	Lower	-.14729716	-.14190066
		Upper	.45059467	.44519816

The foregoing findings do not provide a conclusive verdict as to whether the use of Equation 1 would result in unfair selection decision-making. The possibility exists that because of the differences in slope and the fact that the regression equations intersect approximately in the middle of the predictor distribution that the degree of systematic group-related over and under estimation of the criterion could cancel each other out. To obtain more conclusive evidence H_{044} –H_{048} were consequently tested.

The test statistics used to test H_{045} to H_{048} assume equal error variances across the two race groups. The square of the standard error of estimates of the regression of Y_{GPA} on X_{12} in the two race groups separately, required to test H_{044}, are shown in Table 4.11.

Table 4.11. Simple linear regression of the criterion on the predictor for black and white students separately

X13	Model		Sum of Squares	df	Mean Square	F	Sig.
Black	1	Regression	1368.642	1	1368.642	11.092	.001[a]
		Residual	**15053.455**	**122**	**123.389**		
		Total	16422.097	123			
White	1	Regression	3421.453	1	3421.453	32.834	.000[a]
		Residual	**6773.211**	**65**	**104.203**		
		Total	10194.664	66			

a. Predictors: (Constant), X12

b. Dependent Variable: Y

Hypothesis 13 (H_{044}) represents the assumption of equal error variances across black and white students. H_{044} was tested by calculating the following test statistic (Berenson, Levine & Goldstein, (1983) utilizing the data from Table 4.11:

$$F_b = S^2[Y|X;\pi 1]/S^2[Y|X;\pi 2]$$
$$= 123{,}389/104{,}203$$
$$= 1{,}184$$

The critical F-value is given by $F_k(n_B\text{-}2;n_W\text{-}2)$. The critical $F_k(122;65)$ value of 1,45 exceeds the calculated F value. Because $F_b < F_k$, H_{044} can therefore not be rejected ($p>0{,}05$) and equal error variances can be assumed.

Table 4.12. Univariate analysis of variance tests of between-subjects effects: saturated model

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	**5377.811**[a]	**3**	1792.604	15.358	.000
Intercept	4051.410	1	4051.410	34.710	.000
X12	1368.642	1	1368.642	11.726	.001
Race	440.500	1	440.500	3.774	.054
INT	364.018	1	364.018	3.119	.079
Error	21826.666	**187**	**116.720**		
Total	652971.618	191			
Corrected Total	27204.476	190			

Dependent Variable:Y
a. R Squared = .198 (Adjusted R Squared = .185)

Table 4.13. Univariate analysis of variance tests of between-subjects effects: reduced model containing only the predictor main effect

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	**5098.875**[a]	1	5098.875	43.329	.000
Intercept	6114.034	1	6114.034	51.955	.000
X12	5098.875	1	5098.875	43.329	.000
Error	22359.038	190	117.679		
Total	658329.858	192			
Corrected Total	27457.913	191			

Dependent Variable:Y
a. R Squared = .186 (Adjusted R Squared = .181)

Hypothesis 14 (H_{045}) was tested to determine whether the regression of the criterion on the predictor coincides for black and white students. H_{045} is tested by calculating the following test statistic (Berenson, Levine & Goldstein, 1983) utilizing the data from Table 4.12 and Table 4.13:

$$F_b=\{(SSR[b1,b2,b3]-SSR[b1])/[p-1]\}/MSE[b1,b2,b3]$$
$$= \{(5377{,}811-5098{,}875)/2\}/116{,}72$$
$$= 1{,}195$$

The critical F-value is given by $F_k(p-1;n-p-1)$. The critical $F_k(2,187)$ of 3,00 exceeds the calculated F-value. Because the calculated F_b (1,19) does not exceed F_k (3,00) H_{045} cannot be rejected ($p>0{,}05$). This result suggests that the slope and/or intercepts of the regression are not significantly different. Stated differently, according to the results, black and white students were sampled from the same population. Per implication, the use of the combined equation to calculate expected criterion performance will lead to fair selection decisions.

In the light of the above findings there is no need to test Hypothesis 15 (H_{046}) and 16 (H_{047}) to determine whether an interaction term should be added to a model already containing the group main effect.

4.7 AN EVALUATION OF THE UTILITY OF THE FAIR PREDICTION MODEL OVER RANDOM SELECTION

The utility of a selection procedure can be determined by means of several existing utility models, of which Taylor-Russell (1939), Naylor-Shine (1965), Brogden (1946) and Cronbach Gleser (1965) are the best known (Twigge, Theron, Steel & Meiring, 2005). To answer the question whether the selection procedure under investigation is adding any value to the organization, utility analysis was done based on the Taylor-Russell utility model as well as the Naylor-Shine interpretation of selection utility.

Taylor and Russell (1939) introduced the concepts **base rate** - the percentage of successful persons in the population of applicants, **selection ratio** - the percentage of applicants to be selected, **success ratio** - the proportion of selected applicants who will succeed, and **total utility** - the difference between the success ratio given a specific combination of validity, base rate and selection ratio minus the success ratio which results without knowledge of the test result (Holling, 1998). The success ratio that would be expected under random selection would be equal to the base rate. The Taylor-Russell utility model assumes a linear homoscedastic regression of a normally distributed standardized criterion on a normally distributed standardized predictor (Theron, 2001). This model of utility is important because it describes the usefulness of a selection practice in terms of the percentage successful selectees it will yield (Theron, 2001).

Table 4.14. Taylor-Russell utility estimates for the fair actuarial use of X_{12} as a predictor of first year first semester academic performance at the SA Military Academy

		BR								
		0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
	0,1	.283536	.462424	.599032	.707491	.794529	.86403	.918417	.959163	.986862
		.183536	.262424	.299032	.307491	.294529	.264030	.218417	.159163	.086862
	0,2	.231212	.397473	.533169	.646849	.742653	.822919	.888968	.941336	.979582
		.131212	.197473	.233169	.246849	.242653	.222919	.188968	.141336	.079582
	0,3	.199677	.355446	.488258	.603576	.704012	.790938	.864962	.925979	.972806
		.099677	.155446	.188258	.203576	.204012	.190938	.164962	.125979	.072806
	0,4	.176873	.323424	.452682	.568114	.671304	.762967	.843204	.911459	.966007
		.076873	.123424	.152682	.168114	.171304	.162967	.143204	.111459	.066007
SR	0,5	.158906	.297061	.422407	.537043	.641839	.737043	.822407	.897061	.958906
		.058906	.097061	.122407	.137043	.141839	.137043	.122407	.097061	.058906
	0,6	.144005	.274306	.395469	.508644	.614203	.712076	.801788	.882283	.951248
		.044005	.074306	.095469	.108644	.114203	.112076	.101788	.082283	.051248
	0,7	.131202	.253991	.370698	.481831	.587434	.687247	.780682	.866620	.942719
		.031202	.053991	.070698	.081831	.087434	.087247	.080682	.066620	.042719
	0,8	.119895	.235334	.347242	.455730	.560663	.661712	.758292	.849368	,932803
		.019895	.035334	.047242	.055730	.060663	.061712	.058292	.049368	.032803
	0,9	.109651	.217685	.324269	.429337	.532725	.634166	.733226	.829158	.920393
		.009651	.017685	.024269	.029337	.032725	.034166	.033226	.029158	.020393

Table 4.14 displays the success ratio values[8] [top] and total utility values [bottom] that would result from the use of Equation 1 for strict top-down selection for various possible selection ratios and base rates. Table 4.14 reveals rather modest but nonetheless not negligible gains in the proportion of selectees that would be successful if selection decisions would be based on criterion estimates derived actuarially from Equation 1 rather than random selection. Table 4.14 more specifically indicates that Taylor-Russell selection utility will be optimal for small selection ratios and a base rate approaching 0,40. Taylor-

[8] SPSS was used to generate the success ratio and utility values by using the probability density function of the standardized bivariate normal distribution.

Russell selection utility would, in addition, improve if the validity of the prediction rule could be improved.

The Naylor-Shine utility model likewise assumes a linear homoscedastic regression of a normally distributed standardized criterion on a normally distributed standardized predictor (Theron, 2001). The Naylor-Shine utility model interprets selection utility in terms of the improvement in the expected standardized criterion performance of the selected group of applicants affected by the selection procedure over standardized criterion performance that would be expected under random selection. For a standardized criterion and a standardized predictor, the regression of the standardized criterion on the standardized predictor can be written as (Theron, 2001):

$$E[Zy|Zx] = r[X,Y]Zx$$

The expected standardized criterion performance of the top-down selected applicants can then be written as:

$$E[Zy|E(Zx|Zx \geq Zxc)] = r[X,Y]E[Zx|Zx \geq Zxc]$$

The Naylor-Shine table is based on the fact that if normality of the predictor distribution is assumed, it can be shown that:

$$E[Zx|Zx \geq Zxc] = \lambda/\phi$$ [9]

It thus follows that (Theron, 2001):

$$\begin{aligned} E[Zy|E(Zx|Zx \geq Zxc)] &= r[X,Y]E[Zx|Zx \geq Zxc] \\ &= r[X,Y]\lambda/\phi \end{aligned}$$

Since the expected standardized criterion performance under random selection would be the mean of the standardized criterion distribution $E[Zy|E(Zx|Zx \geq Zxc)]$ simultaneously also can be interpreted as the improvement in criterion performance (expressed in standard deviation units) affected by the selection procedure over random selection.

Table 4.15 displays the expected standardized criterion performance of the top-down selected applicants at various selection ratios[10]. Table 4.15 indicates that the Naylor-Shine utility improves as the selection ratio decreases. Naylor-Shine utility will also increase if the validity of the selection procedure could be improved.

---

[9] The symbol λ denotes the height of an ordinate under the standardized normal distribution cutting of an upper proportion equal to $SR=\phi$

[10] The expected standardized criterion performance values were generated via SPSS.

Table 4.15. Naylor-Shine utility estimates for the fair actuarial use of X_{12} as a predictor of first year first semester academic performance at the SA Military Academy.

SR	$E[Zy \mid E(Zx \mid Zx \geq Zxc)]$[11]
0,1	.756405
0,2	.603400
0,3	.499529
0,4	.416238
0,5	.343852
0,6	.277492
0,7	.214084
0,8	.150850
0,9	.084045

4.8 THE DEVELOPMENT OF A CRITERION-REFERENCED NORM TABLE THAT EXPRESSES THE RISK OF FAILURE CONDITIONAL ON EXPECTED ACADEMIC PERFORMANCE

A criterion-referenced norm table that expresses the risk of failure conditional on expected academic performance was derived from the use of only Grade 12 results (X_{12}). This was the only predictor with a relatively strong significant relationship with First Year First Semester Academic Performance (Y_{GPA}) that warranted inclusion into the SA Military Academy selection battery. The table was developed by transforming the critical GPA of 50% in the conditional criterion distribution (conditional on X_{12}) to a standard normal score ($z_{critical}$-score) to determine the probability of obtaining a standard normal score of less than or equal to $z_{critical}$ at selected X_{12} values. The criterion-referenced norm table (see APPENDIX 1) was calculated using the regression coefficient and standard error of estimate sample estimates displayed in Table 4.9.

4.9 RESTRICTION OF RANGE

The data used in this study only reflects the first year first semester results of three different intake groups at the SA Military Academy. Students were selected based on their results obtained from the selection procedure under investigation. Although the

---

[11] The table values should be interpreted as the number of standard deviation units with which performance would increase if selection decisions would be based on actuarially derived criterion estimates from Equation 1 rather than on random predictions.

psychometric results of students who were rejected are available, no criterion data is available for these applicants. Per implication, the data has a restricted range and therefore the interpretation of the obtained results needs to take this into account.

Formulas to correct the validity coefficients for criterion unreliability and restriction of range have been derived from classical measurement theory. Case 2 [Case A] restriction of range will be applicable if the correlation to be corrected is between two variables X and Y, selection occurred directly on the variable X [or Y] through complete truncation on X at X_C [or on Y at Y_C] and both restricted and unrestricted variances are known only for the explicit selection variable X [or Y] (Theron, 1999).

Therefore, to estimate of the validity of this selection battery (used to select students for the SA Military Academy and for whom criterion scores are available), Case 2 [Case A] restriction of range is applicable. However, if Case 2 [Case A] selection takes place directly on the predictor X, then by assumption, neither the regression of Y on X nor the criterion variance conditional on X will be affected. For this reason no corrections to the parameters of the regression equation or the standard error of estimate is required. The regression of X on Y would be affected, but since it is of no real interest in selection validation research (Theron, 1999), the correction of validity coefficients for criterion unreliability and restriction of range will add no real value to the current study.

CHAPTER 5
CONCLUSIONS AND RECOMMENDATIONS

5.1 INTRODUCTION

The objective of this study was to determine if the psychometric evaluation procedure, used by the South African Military Academy to make selection decisions, can validly predict academic performance of first year learners. Furthermore it was important to determine whether this procedure can discriminate fairly between candidates, and whether the procedure is efficient. The sample used for this study consisted of three year groups enrolled at the SA Military Academy.

It was theorized that specific learning behaviours (learning competencies) are instrumental in achieving desired academic results. It was reasoned that differences in learning performance could be explained in terms of learning behaviours. Differences in learning behaviours in turn were attributed to differences in learning competency potential latent variables. To differentiate between candidates who have better or poorer training prospects in terms of a construct-orientated approach to selection, a performance hypothesis on the person-centred drivers of the learning competencies was developed. It was argued that the presence, or absence of the necessary cognitive processes/competencies that would assist in the understanding and interpretation of the learning material, the intellectual drivers of these learning competencies, proficiency in English and past academic performance should discriminate between better or poorer academic performance of learners attending the academic programmes at the SA Military Academy. The grade point average of the first year first semester academic results was used as a measure of the criterion construct, Learning Performance. Based on this proposition, specific learning competencies were hypothesized to be instrumental in attaining this desired performance outcome.

The following objectives were formulated for the study:

- To test the propositions made by the performance hypothesis depicted as a structural model in Figure 2.2;

- To determine the predictive validity of the individual predictors of the selection battery;
- To derive a weighted linear prediction model actuarially from a set of predictor and criterion data;
- To determine the validity of the inferences derived from prediction model;
- To evaluate the fairness of the inferences/predictions derived from the prediction model and adapt the model if necessary;
- To evaluate the utility of the fair prediction model over random selection; and
- To develop a criterion-referenced norm table that expresses the risk of failure conditional on expected academic performance.

The purpose of this chapter is to state the final conclusions and implications of this study. Recommendations for further research are put forward.

5.2 CONCLUSIONS

Almost all of the results obtained in this study support the theory and propositions made by the performance hypothesis. Only one variable, accuracy of information processing, did not perform as predicted. As anticipated, fluid intelligence has a positive directional effect on transfer. It was confirmed that all three components of information processing has an influence on the steepness of the learning curve in automatization, even though the accuracy with which information is processed only had a modest influence on the rate at which an individual becomes more skilled and efficient in performing a new task. It was further confirmed that the steepness of the learning curve in automatization and the total amount of work done in automatization both positively effects crystallized ability development. The positive directional relationship between transfer and crystallized abilities was confirmed as well. A disappointing small relationship was confirmed between crystallized abilities and learning performance. This finding, however, is not altogether surprising given the lack of alignment between the crystallized abilities developed on the APIL and the abilities required to succeed at the SA Military Academy. A positive directional effect of prior learning on learning performance was confirmed. Even though significant correlations were confirmed between the components of the performance hypothesis, these were only small relationships. Prior learning explained the most variance in the criterion ($r=0,431^2$).

The inter-correlation amongst the predictors was used to infer the proportion of unique variance each predictor would explain in the composite criterion. The nature of the inter-correlation amongst the predictors suggested that a rather disconcertingly large number of predictors that correlated significantly with the criterion ($p<0,05$) would become redundant in a selection battery. A regression of the composite criterion on the array of predictors $X_2 – X_{12}$ including accuracy of information processing (X_4) (earlier found not to correlate significantly with the composite criterion) revealed that only memory and understanding (X_9) and prior learning (X_{12}) significantly explained unique variance in the composite criterion, first semester, first year weighted grade point average, (Y_{GPA}) not explained by the other variables included in the prediction model. Stated differently, only X_9 and X_{12} uncovered relevant and unique information about determinants of performance on the criterion not conveyed by the remaining predictors in the model. The remaining predictors in the selection battery can, consequently, be considered redundant since they provide no new information not already conveyed by X_9 and X_{12}. The foregoing results do, however, not mean that when administering only X_9 and X_{12} each of these predictors would significantly ($p<0,05$) explain unique variance in the criterion. When Y_{GPA} is regressed on the weighted combination of X_9 and X_{12}, only X_{12} significantly ($p<0,05$) explains unique variance in Y_{GPA} when included in a regression model already containing X_9.

When regressing the criterion Y_{GPA} on the array of predictors $X_2 – X_{12}$ excluding accuracy of information processing (X_4) (because of the earlier finding that X4 does not significantly correlate with the composite criterion), the significance of the partial regression coefficients of memory and understanding (X_9) and English vocabulary (X_{10}) is affected. Removing X_4 from the variable set on which the criterion is regressed causes the unique contribution of X_9 to become insignificant ($p>0,05$) and the unique contribution of X_{10} to become significant ($p<0,05$). When, however, Y_{GPA} is regressed on the weighted combination of X_{10} and prior learning (X_{12}), again only X_{12} significantly ($p<0,05$) explains unique variance in Y_{GPA} when included in a regression model already containing the other predictor.

These conclusions were verified by regressing the eleven significant predictors on the criterion by means of a stepwise multiple regression analysis. Only prior learning (X_{12}) was selected for inclusion in the multiple regression model. Figure 3.3 hypothesises variables X_1, X_9, X_{10}, X_{11}, and X_{12} to directly influence Y_{GPA}. The significant partial regression coefficients for predictors X_9 and X_{12} mentioned earlier, support the argument depicted in

Figure 3.3. However the insignificant partial regression coefficients for X_{10} and X_{11}, fail to corroborate the argument underlying the proposed hypotheses. In addition, Global learning potential (X_1) also does not significantly explain variance in the criterion that is not explained by X_{12}. This validation study failed to uncover the required evidence to support the performance hypothesis stated earlier. No justification was found for the inclusion of predictors (X_1 – X_{11}) in the selection battery along with X_{12}.

In the light of the reported findings there was no need to create a combined weighted linear predictor composite (X_{comp}) which would form the basis of the actuarial mechanical decision rule that would guide selection decisions. Prior learning proved to be the only predictor that warrants inclusion in the actuarial mechanical prediction rule. The regression equation depicted in Equation 1 (see Table 4.9) is the derived actuarial prediction rule that will form the basis of selection decisions.

Expected criterion performance of all applicants ($E[Y|X_{12}]$) will be estimated by inserting the prior learning measures obtained during selection into Equation 1. The resultant estimated criterion scores will be rank-ordered from high to low, the position of the cut-off score (Y_k) will be determined in the rank-ordered estimated scores and all applicants with $E[Y|X_{12}] > Y_k$ will be selected. This procedure could be regarded as permissible since $E[Y|X_{12}]$ correlates 0,431 and statistically significantly ($p<0,05$) with Y_{GPA}. The predictions derived from Equation 1 are therefore valid.

Demonstrating that the derived actuarial prediction rule is valid is, however, not sufficient to use the rule to control future admissions to the SA Military Academy. It is critical to verify if the selection decision making based on the criterion estimates derived from the above mentioned Equation 1 will unfairly disadvantage any applicant groups. Since the concept “fairness” has an emotive connotation it is difficult to resolve “fairness” questions. Three fundamentally different ethical views on selection fairness have been identified (Hunter & Schmidt, 1976), of these three fundamental ethical positions technical guidelines on personnel selection procedures (Equal Employment Opportunity Commission, 1978; Society for Industrial and Organizational Psychology, 2003; Society for Industrial Psychology, 1998) seem to favour unqualified individualism. The basic premise is that applicants with an equal probability of succeeding on the job should have an equal probability of obtaining the job, irrespective of group membership (Guion, 1966; 1991;

Huysamen, 2002). The findings of this research suggest that black and white students were sampled from the same population and that the use of the single, undifferentiated prediction rule (Equation 1) would lead to fair selection decisions.

To answer the question whether the selection procedure under investigation is adding any value to the organization, utility analysis was done based on the Taylor-Russell utility model as well as the Naylor-Shine interpretation of selection utility:

- Taylor and Russell (1939) introduced the concepts **base rate** - the percentage of successful persons in the population of applicants, **selection ratio** - the percentage of applicants to be selected, **success ratio** - the proportion of selected applicants who will succeed, and **total utility** - the difference between the success ratio given a specific combination of validity, base rate and selection ratio minus the success ratio which results without knowledge of the test result (Holling, 1998). The success ratio that would be expected under random selection would be equal to the base rate. This model of utility is important because it describes the usefulness of a selection practice in terms of the percentage successful selectees it will yield (Theron, 2001). Taylor-Russell utility estimates for the fair actuarial use of X_{12} as a predictor of first year first semester academic performance at the SA Military Academy is displayed in Table 4.14. This table illustrates the success ratio values and total utility values that would result from the use of Equation 1 for strict top-down selection for various possible selection ratios and base rates. This research suggests modest but however not negligible gains in the proportion of successful selectees if selection decisions are based on criterion estimates derived actuarially from Equation 1 rather than random selection. Findings indicates that Taylor-Russell selection utility will be optimal for small selection ratios and a base rate approaching 0,40. It can be concluded that Taylor-Russell selection utility would improve if the validity of the prediction rule could be improved.

- The Naylor-Shine utility model interprets selection utility in terms of the improvement in the expected standardized criterion performance of the selected group of applicants affected by the selection procedure over standardized criterion performance that would be expected under random selection. Table 4.15 displays the expected standardized criterion performance of the top-down selected applicants at various selection ratios. This table indicates that the Naylor-Shine utility improves as the selection ratio decreases. Naylor-Shine utility will also increase if the validity of the selection procedure could be improved.

A criterion-referenced norm table that expresses the risk of failure conditional on expected academic performance was derived from the actuarial use of only X_{12}, since it is the only predictor with a relatively strong significant relationship with Y_{GPA} (see APPENDIX 1).

5.3 IMPLICATIONS

The results of this validation study confirm the proposition that success at learning in the past predicts success at learning in the future, although the reported correlation is not excessive. Insufficient support was found for the performance hypothesis underlying the selection procedure of the SA Military Academy. A large proportion of the variance in the criterion remains unexplained. Evidence suggests that there are specific learning behaviours (learning competencies) that are not accounted for in the proposed performance hypothesis. Per implication, we are left in the dark regarding the main drivers of academic success at the SA Military Academy.

The predictor construct prior learning (X_{12}) used in this validation study, was operationalized by a measure of actual performance of a student at school. The average of a student's matriculation examination results was used as an indication of his/her level of Prior Learning. Similarly, the criterion construct learning performance of a student was expressed as a Grade Point Average (Y_{GPA}), the average weighted score for all subjects the student has taken that semester. Not surprisingly, evidence suggests that performance at learning predict future performance at learning. Both X_{12} and Y_{GPA} is the average of a student's academic results used as an indication of his/her level of learning. The specific learning behaviours (learning competencies) instrumental in achieving desired academic results have not been directly assessed. It could be argued that the crystallized abilities developed through formal school education are transferred by fluid intelligence onto the novel educational problems presented by the SA Military Academy curriculum. If this transfer process could be simulated in assessment during selection into the Academy educational programme, such a learning competency measure, rather than the transfer process assessed by the APIL which assumes no prior learning, might well demonstrate a significant predictive relationship with the criterion. Essentially the same argument applied to automatization. Automatization needs to be assessed in terms of the learning content relevant to the criterion rather than in terms of learning content that is initially equally unfamiliar to everybody.

The same argument also applies with regards to the hypothesized language proficiency main effect on automatization and the hypothesized language proficiency x fluid intelligence interaction effects on transfer. The APIL purposefully uses geometric test stimuli with which all testees are equally unfamiliar, irrespective of the educational opportunities they might have had in life. In a world where problems to be solved are presented in English, English language proficiency can logically be expected to play a significant role in solving academic problems and automating those solutions. But the same is not true in the contrived and largely non-verbal reality created by the APIL-B.

The preceding argument gives rise to a critical question about changes to the South African school system. The change to an outcomes based education system may in future lead to a situation where Grade 12 results are only expressed in a “Competent/Not yet competent” statement. What would the impact of such a change be on admission requirements of tertiary institutions and prospective employers? The current research would suggest that such a development would seriously erode the predictive efficiency of the single best predictor of academic performance at the SA Military Academy.

Based on the results of this study it would seem as if the selection procedure used to select candidate officers into the academic programme of the SA Military Academy can be simplified. All of the dimensions of the psychometric evaluation procedure under investigation are redundant, because of its failure to successfully predict academic performance. The only successful predictor is obtained from matriculation results. Although the use of this procedure with only one predictor could be regarded as psychometrically permissible ($E[Y|X_{12}]$ correlates 0,431 and statistically significantly ($p<0,05$) with Y_{GPA}), it may be rather risky. The following recommendations should be pondered for further action.

5.4 A CAVEAT

This research study investigated the validity of the mechanical use of the dimension scores rendered by the APIL, the Academic Aptitude Test, the Concept Formation Test and matriculation results in predicting first year, first-semester academic achievement at the SA Military Academy. This study did not investigated the validity of the clinical use of the dimension scores rendered by the APIL, the Academic Aptitude Test, the Concept Formation Test and matriculation results in predicting first year, first semester academic achievement at the SA Military Academy. To have done so would have required clinical judges making explicit clinical criterion inferences for the subjects included in the validation sample based on the available predictor scores and correlating these clinically derived estimates $E_C[Y|X_i]$ with Y_{GPA}. Reviews of the accuracy of clinical prediction suggest that the clinical combination of predictor data very rarely exceed predictions made by actuarial prediction models and that statistical methods are in many cases more accurate than highly trained clinicians (Gatewood & Field, 1994, Grove & Meehl, 1996; Murphy & Davidshofer, 1988). In all likelihood, therefore $r(E_C[Y|X_i],Y_{GPA})$ will be lower than the $r(E[Y|X_{12}],Y_{GPA})$ obtained in this study.

The fairness and utility evidence generated in this study can also not be lead in defence of the clinical use of the SA Military Academy predictors. The current research evidence only reflects on the predictive bias and utility of the actuarial use of X_{12}.

5.5 RECOMMENDATIONS

5.5.1 Shortcomings

A shortcoming of this study is the inability at this stage to assess learning performance in terms of the ability to creatively utilize the newly derived knowledge in solving novel problems that could realistically be encountered in the work environment, and the fact that the researcher had to settle for the assessment of learning performance in terms of the consequences or outcomes of learning (i.e., crystallized knowledge) and competence during training. This study used first-year, first semester academic performance as an operational measure of academic performance. This measure should be regarded as deficient in as far as it constitutes a biased sample of the evaluations over the three year academic programme. The criterion measure, moreover, could be considered problematic in as far as it fails to provide separate assessments of the post programme standing on the latent abilities and the transfer competence in using these latent abilities in solving novel job relevant problems. The academic training offered by the SA Military Academy could be considered successful if the job-relevant crystallized abilities of students are affected by the programme and students are able to successfully transfer the newly developed abilities onto novel job-relevant problems.

The SANDF ultimately needs competent officers. A further critical question that arises is therefore whether good scholars become good officers? Is academic success at the SA Military academy instrumental in eventually achieving success as an officer in the SANDF. The foregoing argument and the argument in terms of which the grounding of the SA Military Academy has been motivated assumes that is the case. Whether academic performance actually significantly explains variance in officer success has not as yet been established. Another troublesome question flow from the first, namely what constitutes a good or successful officer?

The results of this research suggest that learning/academic performance is also shaped by a number of additional factors not taken into account by the existing selection procedure of the SA Military Academy and thus not reflected in the current performance hypothesis. To the extent that the current selection procedure fails to accurately reflect the manner in which important influential determinants of performance combine to affect learning performance it should be regarded as deficient.

A further shortcoming of the research is the fact that the actuarial prediction model was not cross-validated. The failure of the utility analysis to use the cross validated correlation in estimating the utility of the selection procedure therefore resulted in slightly over optimistic estimates of the efficiency of the selection procedure.

5.5.2 Recommendations

It would be premature to discard the learning potential latent variables examined in this study. Given the well documented superiority of actuarial prediction models (Gatewood & Field, 1994, Grove & Meehl, 1996; Murphy & Davidshofer, 1988) the challenge is to continue the search for a well fitting performance@learning structural model. The conviction remains that the learning competencies and the learning competency potential latent variables examined in this study has a fruitful role to play in an explanatory performance@learning structural model. The conceptualization of learning should, however, be broadened to include additional learning competencies over and above transfer and automatization. Behaviourally learning involves more than these two competencies. Possible behavioural learning performance dimensions that should be included in the performance@learning structural model over and above transfer and automatization would be time at task, self motivation and management of resources. Inclusion of these additional learning competencies would then open up the possibility of incorporating additional learning competency potential latent variables into the model like conscientiousness, tenacity, learning motivation and learning self efficacy. An attempt to identify possible extraneous variables that indirectly influence learning performance of students at the SA Military Academy may also prove fruitful. The conviction therefore is that the critical element that remains elusive is not so much the latent variables that should be included in a performance@learning structural model but rather the structural organization between the learning competencies, the learning competency potential latent variables and the learning outcomes.

The plea is therefore for a fresh structural equation modelling based approach to personnel selection. In terms of this approach latent scores would be estimated for the predictor constructs from the available indicator variable scores. The latent scores would then, subsequently, be estimated for the criterion construct(s) using the structural parameter estimates derived for the model.

In operationalizing the expanded performance@learning structural model the transfer and automatization latent variables will have to be measured in terms of learning content relevant to the criterion rather than in terms of learning content that is initially equally unfamiliar to everybody. The transfer and automatization measures need to reflect the ability to transfer Grade 12 crystallized abilities onto learning problems typically encountered in the SA Military Academy curriculum.

It moreover recommended that a generic performance@work structural model should be developed and eventually empirically tested in which officer competency potential latent variables are structurally mapped on generic officer competencies and in which the latter are again structurally mapped on a set of generic officer outcome variables for which military officers could be held accountable. To develop such a generic performance@work structural model, a systematic and thorough job analysis of the position of an officer in the SANDF should be performed. The job analysis should clearly and comprehensively specify the content of the job of an officer and the context in which the job of an officer is performed[12].

The performance@learning structural model should then be sequentially linked with the generic performance@work structural model. By assessing the work performance (in terms of the generic officer competencies and the generic latent outcome variables) of successful students after completion of their three year B Mil Degree in addition to the competency potential latent variables underlying officer performance and evaluating the fit of such a sequentially linked structural model, insight would be gained in the question raised earlier as to whether (and how) the Military Academy programmes serve officer competence.

If close fitting structural model should be found it becomes imperative to examine the cross validation of the model to another sample from the same applicant population.

[12] The down side of this suggestion is that the role of an officer in the SANDF may be too diverse to find a common set of competencies for all officers across the different functions (sharp-end and blunt-end personnel) within the different Arms of Service (Army, Air Force, Navy, Medical Services) relevant in times of peace as well as during war.

REFERENCES

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

Bartone, P.T., Snook, S.A. & Tremble, T.R. (2002). Cognitive and personality predictors of leader performance in West Point Cadets. *Military Psychology. 14*(4), 321-338.

Bartram, D. (2006). The Universal Competency Framework. *SHL White Paper*. SHL Group plc.

Behr, A.L. (1985). The Senior certificate examination as a predictor of University success. *South African Journal of Education*. *5*(3), 107-112.

Berenson, M.L., Levine, D.M. & Goldstein, M. (1983). *Intermediate statistical methods and applications: a computer package approach*. New York: Prentice-Hall, Inc.

Binning, J.F. & Barrett, G.V. (1989). Validity of Personnel decisions: a conceptual analysis of the inferential and evidential bases. *Journal of Applied Psychology*, 74(3), 478-494.

Bolton, B. (1988). *Multivariate Approaches to Human Learning*. In Handbook of multivariate experimental psychology: Perspectives on individual differeces (2nd ed.) Eds J. R. Nesselroade and R.B. Cattell. New York: Plenum Press.

Boudreau, J.W. (1991). Utility analysis for decisions in human resource management. In M.D. Dunnette & L.M. Hough (Eds.). *Handbook of industrial and organisational psychology* (Second edition, Volume 2, pp. 621-745). Palo Alto, California: Consulting Psychologists Press, Inc.

Bradley, J.P., Nicol, A.A.M., Charbonneau, D. & Meyer, J.P. (2002). Personality correlates of leadership development in Canadian Forces officer candidates. *Canadian Journal of Behavioural Science, 34*(2), 92-103.

Brogden, H.E. (1946). On the Interpretation of the Correlation Coefficient as a Measure of Predictive Efficiency. *Journal of Education and Psychology, 37*, 65-76.

Brogden, H.E. (1949a). When Testing Pays Off. *Personnel Psychology, 2*, 171-185.

Brogden, H.E. (1949b). A New Coefficient: Application to Biserial Correlation and to Estimation of Selective Efficiency. *Psychometrika, 14*, 169-182.

Cascio, W.F. (1991a). *Applied psychology in personnel management*. Englewood Cliffs, N.J.: Prentice-Hall.

Cascio, W.F. (1991b). *Costing human resources; the financial impact of behavior in organisations*. Boston: PWS-Kent Publishing Company.

Catrambone, R. & Holyoak, K.J. (1989). Overcoming contextual limitations on problem solving transfer. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *15*,1147-1156.

Cattell, R.B. (1971) *Abilities: Their structure, growth and action*. Boston: Houghton Mifflin.

Cleary, T.A. (1968). Test bias: prediction of grades of Negro and white students in integrated colleges. *Journal of Educational Measurement, 5,* 115-124.

Cronbach, L.J. & Gleser, G.C. (1965). *Psychological tests and personnel decisions* (2nd edition). Urbana, Illinois: University of Illinois Press.

Cronshaw, S.F. & Alexander, R.A. (1985). One answer to the demand for accountability: selection utility as an investment decision. *Organisational Behavior and Human Decision Processes*, 35, 102-118.

Defence Corporate Communication (2003). Military Skills Development System. *Department of Defence Bulletin*. No 02/03. 9 Jan.

De Goede, J. (2007). An investigation into the internal structure of the learning potential construct as measured by the APIL test battery. Unpublished master's thesis. University of Stellenbosch, Stellenbosch.

Diamantopoulos, A. & Siguaw, J.A. (2000). *Introducing LISREL*. London, New Delhi: Sage Publications.

Einhorn, H.J. & Bass, A.R. (1971). Methodological considerations relevant to discrimination in employment testing. *Psychological Bulletin*, *75*, 261-269.

Ellis, M.V. & Blustein, D.L. (1991). Developing and using educational and psychological tests and measures: the unificationist perspective. *Journal of Counseling and Development*, 69, 550-555.

Engelbrecht, M. (2000). Leerpotensiaal as voorspeller van akademiese sukses by eerstejaaruniversiteitlearnere. Ongepubliseerde PhD-verhandeling, Potchefstroomse Universiteit vir Christelike Hoër Onderwys, Potchefstroom.

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.

Erasmus, B.J. and Van Dyk, P.S. (1999). *Training Management in South Africa. (*Second edition) Halfway House: International Thomson.

Ferguson, G.A. (1956). On transfer and the abilities of man. *Canadian Journal of Psychology*, *10*, 121-131.

Gatewood, R.B. & Feild, H.S. (1994). *Human resource 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.

Glaser, R. (1981). The future of testing: A research agenda for cognitive psychology and psychometrics. *American Psychologist*, *36*(9), 923-936.

Grove, W.M. & Meehl, P.E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: the clinical- statistical controversy. *Psychology, Public Policy, and Law, 2,* 293-323.

Guion, R.M. (1966). Employment tests and discriminatory hiring. *Industrial Relations*, *5*, 20-37.

Guion, R.M. (1991). Personnel assessment, selection and placement. In M.D. Dunnette & L.M. Hough (Eds.). *Handbook of industrial and organisational psychology* (pp. 327-397)(Second edition, Volume 2, pp. 327-397). Palo Alto, California: Consulting Psychologists Press, Inc.

Guion, R.M. (1998). *Assessment, measurement and prediction for personnel decisions*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Hair, J.F., Anderson, R.E., Tatham, R.L. & Black, W.C. (1995). *Multivariate data analysis with readings* (Third edition). Englewood Cliffs, NJ: Prentice-Hall.

Harackiewicz, J.M., Barron, K.E., Tauer, J.M. & Elliot, A.J. (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of Educational Psychology*. *94*(3), 562-575.

Holling, H. (1998). Utility analysis of personnel selection: an overview and empirical study based on objective performance measures. *Methods of Research Online*, *3*, 1. Internet: http://www.pabst-publishers.de/mpr/

Howcroft, J.G. (1991). Self-esteem and academic achievement of Black and Coloured university learners. *Acta Academica*, *23*(3), 106-119.

Hunter, J.E. & Schmidt, F.L. (1976). Critical analysis of the statistical and ethical implications of various definitions of test bias. *Psycological Bulletin*, *83*, 1053-1071.

Huysamen, G.K. (2002). The relevance of the new APA standards for educational and psychological testing for employment testing in South Africa. *South African Journal of Psychology*, *32*(2), 26-33.

Jensen, A.R. (1986). *g*: Artifact or reality? *Journal of Vocational Behaviour*, *29*, 301-331.

Kenny, D. A., & Judd, C. M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological Bulletin*, *96*, 201-210.

Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of Behavioural Research. (Fourth edition).* Fort Worth: Harcourt College.

Kline, P. (1991). *Intelligence: The Psychometric View*. London: Routledge.

Kotze, M.E. (1993). Die verband tussen loopbaanvolwassenheid, rolbelangrikheid en akademiese prestasie: ‘n verkennende studie. *Tydskrif vir Bedryfsielkunde*, *19*(3), 13-17.

Landy, F.J. (1986). Stamp collecting versus science; validation as hypothesis testing. *American Psychologist*, 41(11), 1183-1192.

Larson, G.E. & Saccuzzo, G.E. (1989). Cognitive correlates of general intelligence: Toward a process theory of *g*. *Intelligence*, *13*, 5-31.

Lavin, D.E. (1965). *The prediction of academic performance*. New York: Russel Sage Foundation.

Meehl, P. E. (1957). When shall we use our heads instead of the formula? *Journal of Counseling Psychology, 4,* 268-273.

Messick, S. (1989). *Validity*. In R.L. Linn (Ed.). *Educational measurement* (Third edition). New York: American Council on Education and McMillan Publishing Company.

Milkovich, G.T. & Boudreau, J.W. (1994). *Human resource management* (Seventh edition). Homewood, Illinois: Richard D. Irwin Inc.

Minutes of the Meeting on MSDS intake at the Military Academy held on 24 February 2003 at 12:00 in the Conference Room, Military Academy, Saldanha.

Mpofu, E. & Oakland, T. (2001). Predicting school achievement in Zimbabwean multiracial schools using Biggs’ Learning Process Questionnaire. *South African Journal of Psychology*, *31*(3), 20-28.

Murphy, K.R. and Davidshofer, C.O. (1988). *Psychological testing: principles and applications*. Englewood Cliffs, NJ: Prentice Hall.

Naylor, J.C. & Shine, L.C. (1965). A table for determining the increase in mean criterion score obtained by using a selection device. *Journal of Industrial Psychology, 3*, 33-42.

Nel, A. (1997). Die Voorspelling van Akademiese sukses binne konteks van 'n alternatiewe universitietstoelatingsbeleid. Ongepubliseerde MA-verhandeling, Randse Afrikaanse Universiteit, Johannesburg.

Nel, P.S., Gerber, P.D., Van Dyk, P.S., Haasbroek, G.D., Schultz, H.B., Sono, T. & Werner, A. (2001). *Human Resources Management* (Fifth edition). Cape Town: Oxford University Press Southern Africa.

Ormrod, J.E. (1990). *Human Learing: Theories, principles, and education*. New York: Merrill Publishing Company.

Owen, R. & De Beer, J.F. (1981). *Manual for the Academic Aptitude Test (University) (AAT)*. Pretoria: Human Science Research Council.

PEC, (2003). Guidelines on the submission of new academic programmes to the Programme Evaluation Committee. Unpublished manuscript. Stellenbosch University, Stellenbosch.

Pelser, S.K.S. (1992). Persoonlikheid en leerstrategieë van presterende en nie-presterende eerstejaarlearnere. Ongepubliseerde D Ed-verhandeling, Randse Afrikaanse Universiteit, Johannesburg.

Pienaar, J.L. (1991). Die voorspelling van akademiese sukses met behulp van die AAT (Universiteit). Ongepubliseerde MA-verhandeling, VISTA Universiteit, Pretoria.

Plug, Louw, Gouws & Meyer (1997). *Verklarende en vertalende sielkundewoordeboek*. Heinemann: Johannesburg.

Rademeyer, M.M. & Schepers, J.M. (1998). Voorspelling van die akademiese sukses van eerstejaarstudente. *Tydskrif vir Bedryfsielkunde*, *24*(1), 33-40.

Rosnow, R.L. & Rosental, R. (1998). *Beginning behavioural research: a conceptual primer*. (*Third edition*) Prentice-Hall: New Jersey.

Saville & Holdsworth. (2000). Competency design; towards an integrated human resource management system. *SHL Newsline*, March, 7-8.

Saville & Holdsworth. (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., Pearman, 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.

Schmitt, N. and Landy, F.J. (1992). The concept of validity. In N. Schmitt and W.C. Borman (Eds.), *Personnel selection* (pp. 275–309). San Francisco: Jossey-Bass.

Sherman, A., Bohlander, G. & Snell, S. (1998). *Managing Human Resources* (Eleventh edition). Cincinnati: South-Western College Publishing.

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

Schumacker, R.E. & Lomax, R.G. (1996). *A beginner's guide to structural equation modeling*. Mahaw, New Jersey: Lawrence Erlbaum Associates, Publishers.

Smith, G.J. (Geen datum aangedui). *Psigometrika: Aspekte van Toetsgebruik.* Pretoria: HUAM Opvoedkundige Uitgewers.

Spangenberg, H.H. (1990). Assessing managerial competence. Kenwyn: Juta.

Society for Industrial and Organizational Psychology. (2003). *Principles for the validation and use of personnel selection procedures*. Bowling Green: Author.

Society for Industrial Psychology (SIP). (1998). *Guidelines for the validation and use of assessment procedures for the workplace*. Johannesburg: SIP.

Sternberg, R.J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge University Press.

Sussmann, M. & Robertson, D.U. (1986). The validity of validity: an analysis of validation study designs. *Journal of Applied Psychology*, 71(3), 461-468.

Tabachnick, B.G. & Fidell, L.S. (1996). *Using Multivariate Statistics* (3rd ed.) New York: HarperCollins College Publishers.

Taylor, H.C. & Russel, J.T. (1939). The relationship of validity coefficients to the practical effectiveness of tests in selection: discussion and tables. *Journal of Applied Psychology, 23*, 565-578

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

Taylor, T.R. (2001). *Learning potential: theory and practical assessment*. Unpublished class notes (Research Methodology & Psychometrics 778/788, Set 5, The Assessment of learning potential), University of Stellenbosch.

Theron, C.C. (1999). Psychometric implications of corrections for attenuation and restriction of range for selection validation research. Unpublished doctoral thesis, University of Stellenbosch, Stellenbosch.

Theron, C.C. (2001). *Personnel Selection*. Unpublished class notes (Research Methodology & Psychometrics 778/788), University of Stellenbosch, Stellenbosch.

Theron, C.C. (2002). Unpublished class notes (Psychometrics and Research Methodology, 244), University of Stellenbosch.

Theron, C.C. (2007). Confessions, scapegoats and flying pigs: Psychometric testing and the law. *SA Journal of Industrial Psychology*, *33*(1), 102-117.

Thorndike, R.L. (1982). *Applied Psychometrics*. Boston: Houghton Mifflin Company.

Ting, S.M.R. & Robinson, T.L. (1998). First-year academic success: A prediction combining cognitive and psychosocial variables for Caucasian and African American learners. *Journal of College Learner Development*, *39*(6), 599-610.

Tredoux, C.G. & Durrheim, K.L. (2002). *Numbers, hypotheses and conclutions.* Cape Town: Juta.

Twigge, L., Theron, C., Steel, H. & Meiring, D. (2005). A psychometric investigation into the use of an adaptation of the Ghiselli Prediction Index in personnel selection. *SA Journal of Industrial Psychology*, *31*(1), 18-30.

Van Dyk, P.S., Nel, P.S., Loedolff, P. Van Z. & Haasbroek, G.D. (2001). *Training management: A multidisciplinary approach to human resources development in Southern Africa*. Cape Town: Oxford University Press.

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

Vernon, P.A. (1990). An overview of chronometric measures of intelligence. *School Psychology Review*, *19*, 399-410.

Vosloo, H.N. (1987). The relationship between aptitude, interest and adjustment and the academic achievement of first-year learners at college of education. Educare: *Tydskrif van die Fakulteit Opvoedkunde*, *16*(2), 126

Weinberg, R.A. (1989). Inteligence and IQ. Landmark issues and great debates. *American Psychologist*, *44*(2), 99-104.

Werther, Jr, W.B. & Davis, K. (1993). *Human resource and personnel management* (4th ed.). New York: McGraw-Hill.

APPENDIX 1 CRITERION REFERENCED NORM TABLE

X_{12}	PRED_Y	RISK
1	30.662	0.962677
2	31.109	0.959196
3	31.556	0.955455
4	32.003	0.951444
5	32.45	0.947148
6	32.897	0.942557
7	33.344	0.937657
8	33.791	0.932437
9	34.238	0.926886
10	34.685	0.920993
11	35.132	0.914746
12	35.579	0.908137
13	36.026	0.901156
14	36.473	0.893794
15	36.92	0.886043
16	37.367	0.877899
17	37.814	0.869354
18	38.261	0.860404
19	38.708	0.851046
20	39.155	0.841278
21	39.602	0.831099
22	40.049	0.82051
23	40.496	0.809514
24	40.943	0.798113
25	41.39	0.786313
26	41.837	0.774121
27	42.284	0.761546
28	42.731	0.748596
29	43.178	0.735283
30	43.625	0.721622
31	44.072	0.707625
32	44.519	0.69331
33	44.966	0.678694
34	45.413	0.663795
35	45.86	0.648635

36	46.307	0.633234
37	46.754	0.617616
38	47.201	0.601804
39	47.648	0.585823
40	48.095	0.569699
41	48.542	0.553458
42	48.989	0.537126
43	49.436	0.520732
44	49.883	0.504303
45	50.33	0.487866
46	50.777	0.47145
47	51.224	0.455082
48	51.671	0.43879
49	52.118	0.422601
50	52.565	0.406542
51	53.012	0.390639
52	53.459	0.374916
53	53.906	0.359399
54	54.353	0.34411
55	54.8	0.329072
56	55.247	0.314305
57	55.694	0.29983
58	56.141	0.285665
59	56.588	0.271826
60	57.035	0.258329
61	57.482	0.245188
62	57.929	0.232415
63	58.376	0.220021
64	58.823	0.208015
65	59.27	0.196404
66	59.717	0.185196
67	60.164	0.174393
68	60.611	0.163999
69	61.058	0.154016
70	61.505	0.144444
71	61.952	0.135281
72	62.399	0.126524
73	62.846	0.118171
74	63.293	0.110215
75	63.74	0.10265

76	64.187	0.095471
77	64.634	0.088668
78	65.081	0.082233
79	65.528	0.076156
80	65.975	0.070427
81	66.422	0.065035
82	66.869	0.059969
83	67.316	0.055218
84	67.763	0.050769
85	68.21	0.04661
86	68.657	0.042729
87	69.104	0.039114
88	69.551	0.035752
89	69.998	0.03263
90	70.445	0.029737
91	70.892	0.027059
92	71.339	0.024586
93	71.786	0.022306
94	72.233	0.020207
95	72.68	0.018277
96	73.127	0.016507
97	73.574	0.014886
98	74.021	0.013403
99	74.468	0.01205
100	74.915	0.010817