APPLICATION OF A NOMOLOGICAL MODEL OF SELECTION VALIDITY

By

Tanya Justine Bartmann

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

Supervisor: Dr A Duvenage

March 2001

Declaration

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

Signature:

Date: 13 February 2001

Note. This thesis was submitted to meet the format requirements of journal articles for the relevant masters degree.

ABSTRACT

The purpose of this study was to determine the applicability of a nomological model for the assessment of psychological measurement validity with reference to personnel selection. The model, which can be traced to Thorndike (1949), Campbell and Fiske (1959), Nunnally (1967, cited in Duvenage, 1990), Guion (1974, cited in Duvenage, 1990), Binning and Barrett (1989) and Duvenage (1990), provides the evidential bases for two empirically founded validation strategies, namely criterion- and construct-related validity. Theoretically, the Nomological Model for Psychological Measurement Validity Assessment proved to be applicable for validating that a measure is able to predict job performance.

An easily understandable empirical investigation was required to verify the scientific utility of the model for selection validation. The model was used to validate the procedure for selecting machine operators in a large food manufacturer. The investigation illustrated that by utilising Route 1 of the nomological model, evidence of criterion-related validity, as described by legal and professional standards, is generated. Furthermore, the study indicated that by utilising Route 2 of the model, evidential requirements of construct-related validity as set by legal and professional standards are met. The results of the study proved that the Nomological Model for Psychological Measurement Validity Assessment is theoretically, professionally, legally and practically applicable for the validation of personnel selection procedures.

OPSOMMING

Die doel van hierdie studie was om die toepasbaarheid van 'n nomologiese model vir die bepaling van die geldigheid van sielkundige meting ten opsigte van personeelkeuring vas te stel. Die model, wat sy oorsprong in die werk van Thorndike (1949), Campbell en Fiske (1959), Nunnally (1967, aangehaal uit Duvenage, 1990), Guion (1974, aangehaal uit Duvenage, 1990), Binning en Barrett (1989) en Duvenage (1990) het, verskaf getuienis basisse vir twee empiries gefundeerde valideringsstrategieë, naamlik kriterium- en konstruk-verwante geldigheid. Daar is teoretiese bewyse gevind vir die toepasbaarheid van die Nomologiese Model vir Sielkundige Meting Geldigheids Raming vir die validering van metingsvoorspelbaarheid van werksprestasie.

'n Maklik verstaanbare empiriese ondersoek was nodig om die model se wetenskaplike nut vir keuringsgeldigheid te bepaal. Die model was gebruik vir die validering van keuringsprosedure-geldigheid van masjienoperateurs in 'n groot voedselverwerkingsmaatskappy. Die ondersoek het geïllustreer dat deur die toepassing van Roete 1 van die nomologiese model, bewyse gegenereer is vir kriterium-verwante geldigheid, soos voorgeskryf deur relevante wetgewing en professionele standaarde. Verder, deur die toepassing van Roete 2 van hierdie model, word voldoende bewyse vir konstruk-verwante geldigheid, soos voorgeskryf deur wetlike en professionele standaarde, verkry. Die resultate van die studie bewys dat die Nomologiese Model vir Sielkundige Meting Geldigheids Raming teoreties, professioneel, wettig en prakties toepasbaar is vir die validering van personeelkeuringsprosedures.

APPLICATION OF A NOMOLOGICAL MODEL OF SELECTION VALIDITY

The central concern of the present study is the validity of human resource testing and selection procedures. The ultimate research purpose being to empirically apply and evaluate a complete model of measurement validity that is suitable for personnel selection validation.

In the milieu of human resource selection, measuring instruments are used with the intention of making inferences about an individual's future job performance. These deductions assist decision-makers in distinguishing the most suitable person for the job. To predicate these conclusions, one needs to prove them valid. Without proof of validity, an instrument that is thought to be useful for the identification of potentially successful workers has no evidence to support the claim. This may result in erroneous inferences being made, and ultimately the incorrect person being selected (Binning & Barrett, 1989; Gatewood & Field, 1994; Nunnally, 1978). Furthermore, the recent Labour Relations Act (LRA) (1995), the Basic Conditions of Employment Act (BCEA) (1997) and the Employment Equity Act (EEA) (1998) form the legislative trilogy which monitors employment practices in South Africa. Democracy and equality are the order of the day.

The EEA, said to be the potentially most far-reaching of all labour measures thus far enacted, prohibits unfair discrimination in selection practices. Specific attention is paid to human resource assessment procedures. Paragraph 8(a) of this Act prohibits psychological testing and other similar assessments of an employee unless it has been scientifically proven that the measure is valid and reliable. If a case of unfair discrimination is alleged, according to paragraph 11 of this Act, the burden of proof falls on the employer. Subsequently, if an organisation does not prove its selection procedure to be valid, it will not be able to legally defend itself should the decision be challenged in a court of law (de Bruin, 1998; Grogan, 1998a, 1998b; Kriek, 2000; Levy & Associates, 1999).

Despite the exigency to conduct validation studies on assessment measures, a dearth exists. This could be attributed to the diverse conceptions of the term validity. The issue is also one of extreme sensitivity. Industrial Psychologists are expected to develop selection procedures in a non-racial society, so job performance of multi-cultural groups can be predicted. The

final deterrent may well be the gross shortage of basic validity research available in South Africa to facilitate this complex and technical validation process (Binning & Barrett, 1989; Saville & Holdsworth (SHL), 2000a; Society for Industrial Psychology (SIP), 1998; Tustin, 1992; United States of America (USA), 1978).

The Society for Industrial Psychology (SIP) (1998) emulated the United States Equal Employment Opportunity Commission (USA, 1978) and the American Psychological Association (APA, 1985) and published local Guidelines for Validation and Use of Assessment Procedures for the Workplace. The main aim of the booklet was to provide individuals and organisations with a set of guidelines for fair practice in the choice, development, evaluation and use of selection procedures (SIP, 1998). However, no model applicable to selection assessment validation accompanied these guidelines. In fact, no model of human resource selection validation is evident in South African literature. Given the confusion and sensitivity surrounding selection validity, an applicable model should be deemed indispensable. Akkerman (1989) emphasized the need to build a common theoretical framework that will increase understanding and the quality of assessment. The drive behind this research is thus to elucidate and operationalize a suitable model that delineates the framework, interconnection and complexity of the concept validity and validation in the personnel selection context.

Validity, when broadly used, refers to the scientific accuracy of a measuring instrument. If an instrument measures what it claims to measure, the instrument is valid. In the context of selection, if an instrument measures an aspect of future job performance, then the instrument is valid. However, an instrument cannot actually measure future job performance. Rather, it measures a certain ability or trait, the scores of which are then used to make inferences about future job performance. Analogous to the hypothesis that scores of a particular test can infer some aspect of an applicant's performance on the job. Hence, one validates the hypothesized inferences made from test information, and not the test itself. An inference is valid to the extent that there is firm evidence to support it. Validity is thus the degree to which inferences from scores on tests are justified by evidence (Binning & Barrett, 1989; Duvenage, 1990; Guion, 1978; Lawshe, 1985; Nunnally, 1978).

The method that is used to accumulate or generate evidence to support such inferences is referred to as a validation strategy. The product thereof constitutes the evidential base that determines or defends the strength of inferences made from test scores (the hypothesis) (Binning & Barrett, 1989; Duvenage, 1990; Lawshe, 1985; SIP, 1998). Evidence may be accumulated in many ways. For convenience purposes the validation strategies have traditionally been grouped into different categories. But the use of category labels does not imply distinct types of validity. Certain incorrect perceptions imply that a specific strategy is seen as a more appropriate method of validation depending on the purpose of the test or measure (Cronbach, 1970; Lawshe, 1985). Validity is a unitary concept, and rigorous distinctions between the categories are not possible (APA, 1985). Legislation and professional standards recognize two validation strategies and evidential bases namely, criterion-related validity and construct-related validity.

Much contention surrounds the recognition of a third validation strategy labeled content validity. The pervasive argument exists that content validity resembles construct-related validity, short of empirical evidence (Duvenage, 1990; Guion, 1978, 1987). To further aggravate the situation, no specific standards exist against which the qualitative evidence of content validity can be evaluated, leaving judges and courts free to their own interpretation (Kleiman & Faley, 1978). Consequently, criterion-related and construct-related validity are recognized as the two empirically based validation strategies. However, one cannot lose sight of the fact that the inference of ultimate importance in any selection testing is the prediction of job performance based on available test scores, and an ideal validation includes several types of evidence which span across both validation strategies (APA, 1985; Binning & Barrett, 1989; Guion, 1978, 1998; Landy 1991; SIP, 1998).

For a model of measurement validity to be considered suitable for selection procedure validation, three prerequisites must be satisfied. The model must address (a) the two empirically-based validation strategies, (b) the interrelation between the strategies, and (c) the nature associated with this inter-connection, i.e. that which occurs in the observable perceptual plane of science which is linked to that which occurs in the unobservable conceptual plane (Duvenage, 1990). Theoretically, the Nomological Model for Psychological Measurement Validity Assessment (see Figure 1) satisfies these prerequisites, as is established in the next main section of this article.

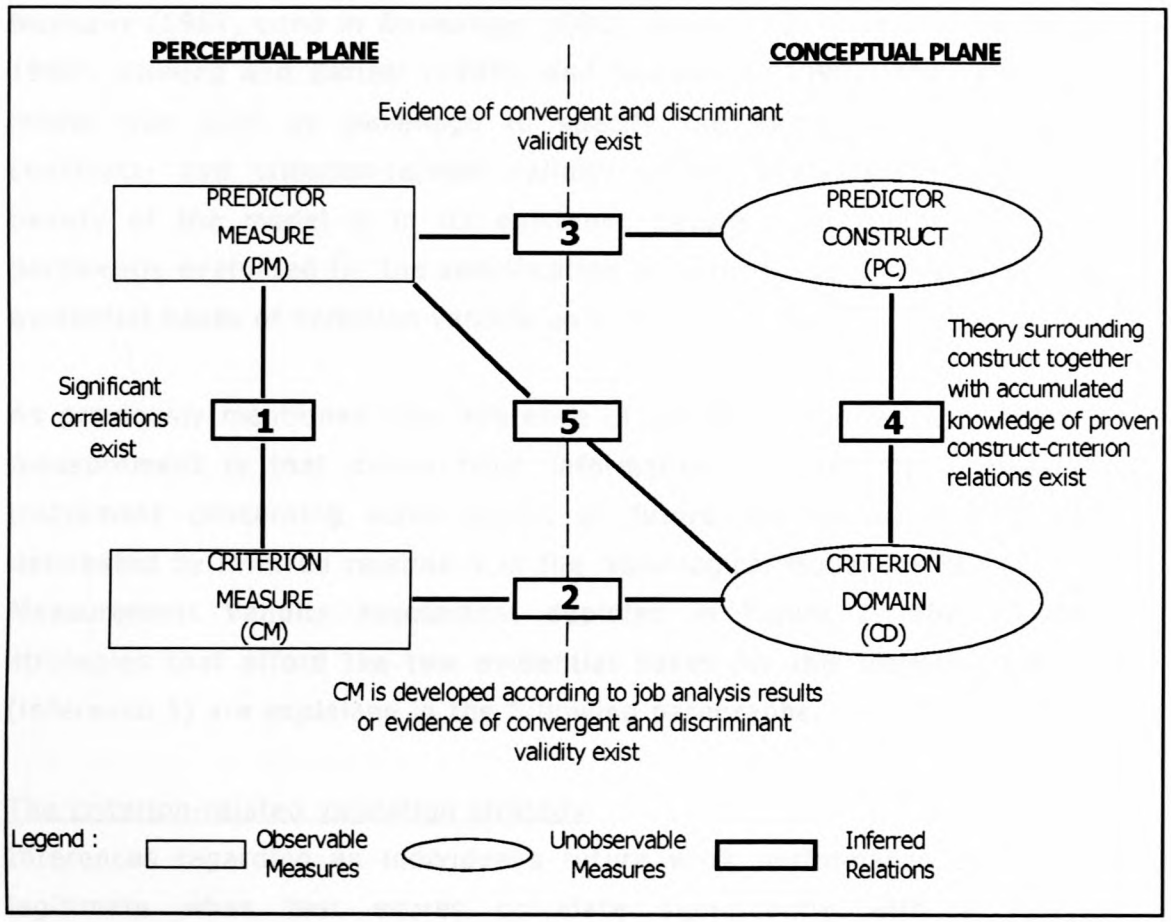

Figure 1 Nomological Model for Psychological Measurement Validity Assessment.

(Duvenage, 1990).

Nomo is Greek for law and *logos,* Greek for reason, argument or comprehension. A nomological network is:

> A system of interlocking explanatory and predictive laws in the form of links, connections or correlations between observables and observables on an empirical or perceptual plane (P-plane), unobservables and unobservables on a theoretical or conceptual plane (C-plane), and between observables on the P-plane and unobservables on the C-plane (Duvenage, 1990, p. 145).

Duvenage completed a survey on theories and models that combined nomologics and psychological measurement. He asserted that early pioneers such as Stevens (1935 & 1939), Hull (1937 & 1943), MacCorquodale and Meehl (1948), and Cronbach and Meehl (1955) (cited in Duvenage, 1990) provided the foundations upon which theorists could build. The actual conception of the Nomological Model for Psychological Measurement Validity

Assessment can be traced to Thorndike (1949), Campbell and Fiske (1959), Nunnally (1967, cited in Duvenage, 1990), Guion (1974, cited in Duvenage, 1990), Binning and Barrett (1989), and Duvenage (1990). The nomological model was used by Duvenage to specify the evidential bases for the construct- and criterion-related validity of job evaluation systems. The beauty of the model is in its epistemic nature, thus enabling it to be pertinently evaluated for the specification of criterion- and construct-related evidential bases of selection validity as well.

As previously mentioned, the inference of ultimate importance in selection measurement is that drawn from information obtained by a measuring instrument concerning some aspect of future job performance. This is delineated by inferred relation 5 in the Nomological Model for Psychological Measurement Validity Assessment depicted in Figure 1. The validation strategies that afford the two evidential bases for this ultimate inference (inference 5) are explained in the following paragraphs.

<u>The criterion-related validation strategy</u>

Inferences regarding an individual's future work performance are deemed legitimate when test scores correlate significantly with a relevant operational measure of successful job performance. This evidence is generated via the criterion-related validation strategy and is represented in the nomological model by inferences 1 and 2. Inference 1 depicts the relationship between the test and the measure of job performance, both of which are of an observable nature. The test is referred to as a predictor measure because it is used with the intention of predicting performance on the job. The nature of the predictor measure is irrelevant when proving criterion-related validity; it can be a measure of personality, hair colour or shoe size. Provided that a significant correlation exists between the predictor measure and the job performance measure, inference 1 is valid (Binning & Barrett, 1989; Duvenage, 1990; Nunnally, 1978).

The measure of job performance is referred to as a criterion measure. Ideally, one would like to predict overall successful job performance. However, this ultimate criterion is multiple and complex, and is rarely, if ever, available in measurable terms. So, a criterion measure that is quantifiable and resembles a sample of the ultimate criterion is used as a substitute. It is an operationalisation of the ultimate criterion (Akkerman,

1989; Thorndike, 1949). Different categories of criterion measures exist, namely supervisory ratings, measures of productivity, ancillary measures (e.g. absenteeism and turnover), and knowledge tests. Landy and Rastegary (1989) reviewed the recent work on various measures and identified a new class of criterion measure referred to as hands-on measures. Irrespective of the measure chosen, the tired old – yet pertinent - "criterion problem" will always raise its head. Issues of relevance, reliability and contamination make for matters of concern when proving criterion-related validity (Guion, 1978, 1998; Thorndike, 1949). The prime essential of a criterion measure is its relevance to the ultimate criterion. Typically, scientific concerns in personnel selection do not go beyond computing validity coefficients to generating evidence of criterion measure relevance. This results in a truncated evidential base (Binning & Barrett, 1989; Borman, Rosse & Abrahams, 1980; Landy & Rastegary, 1989; Nunnally, 1978; Thorndike, 1949).

To have full confidence of criterion-related validity, evidence that the criterion measure is relevant to the ultimate criterion must exist. This is delineated by inference 2 of the nomological model. The ultimate criterion – successful job performance - is termed the criterion domain and occurs in the conceptual plane. Before any testing or validation commences, the practitioner needs to understand what clusters of job behaviours and outcomes constitute the criterion domain. This calls for a job analysis to be conducted. In addition, a job analysis is essential for enabling selection decisions that are founded on inherent requirements of the job, as obligated by item 6(2)(a) of the Employment Equity Act (1998)(Levy & Associates, 1999).

Numerous methods of job analysis are available, ranging from task-orientated to worker-orientated (Dunnette, 1966; Gatewood & Field, 1994; Goldstein, 1993; McCormick, 1979; Spencer, McClelland & Spencer, 1992; Thorndike, 1949). From the identified job tasks, and subsequent knowledge, skills and abilities required, the practitioner is able to delineate the criterion domain. This in turn allows for a criterion measure's relevance to be evaluated. If a criterion measure is developed according to the findings of the job analysis, one can conclude that it is relevant. However, this is not always the case. Subsequently, evidence of a measure's relevance would be generated via empirically based relationships that are convergent and

discriminant in nature. Convergent relations imply that (a) criterion measure scores will relate to other measurement scores of the same criterion, (b) criterion measure scores will vary according to the way in which different people perform on the job, and (c) criterion scores will correlate with the scores of other criteria that are expected to be related. Discriminant relations on the other hand imply that criterion measure scores cannot be linked to scores of other criterion measures that are not related (Binning & Barrett, 1989; Cambell & Fiske, 1959; Duvenage, 1990; SIP, 1998).

Thus, once the criterion domain is delineated based on job analysis information, and evidence exists proving the criterion measure is an operational measure of the criterion domain (inference 2), the correlation between the criterion measure and the predictor measure can be computed (inference 1). If significant correlations are obtained, the researcher can postulate that the predictor measure is able to predict job performance (inference 5) on the basis of criterion-related validity. In summary, Route 1 of the nomological model entails generating evidence to support inference 1 and inference 2 which predicate inference 5 on the grounds of criterion-related validity. Hence, Route 1 of the nomological model represents the criterion-related validation route for personnel selection.

The construct-related validation strategy

An inference regarding job performance based on an applicant's possession of a trait or attribute is considered valid if evidence of construct-related validity is present. The trait or attribute is labeled a construct. It is a theoretical concept that underlies human behaviour. This concept is abstract since it cannot be observed directly but possesses a domain of co-varying behaviour that researchers - through scientific investigation - have demonstrated to be associated with the construct-label. When proving construct-related validity, one verifies that the construct is important for job performance and that the measuring instrument assesses the behaviour associated with the construct (Binning & Barrett, 1989; Guion, 1987; Nunnally, 1978; SHL, 2000a; SIP, 1998).

The construct-related validation strategy is depicted by inference 3 and 4 of the nomological model. Inference 4 represents the relationship between the criterion domain (successful job performance) and the underlying construct. Akin to the criterion-related validation strategy, it is necessary to conduct a

job analysis to determine what clusters of behaviours and outcomes constitute the criterion domain on the conceptual plane. Founded on this domain, a construct that is assumed to underlie the desired criterion is identified. As discussed in the previous paragraph, a construct is abstract. Hence, the criterion domain and the construct both occur in the conceptual plane and therefore the relationship cannot be of an empirical nature. Subsequently, evidence of this underlying connection (inference 4) takes the form of a well-articulated theory surrounding what the construct comprises, together with accumulated knowledge of proven relationships between the construct and identifiable demands of the job (Cronbach, 1970; Duvenage, 1990; SIP, 1998).

Once evidence has established that the identified construct is important for successful job performance (inference 4), focus is placed on the measuring instrument. If the researcher verifies that the test sufficiently measures the construct that it claims to, then it is deemed construct-valid. Inference 3 of the nomological model delineates the construct validity of a test. The label *predictor* precedes the terms *measure* and *construct* as both are used with the intention of inferring future performance on the job. Proof demonstrating the relationship between the predictor measure on the perceptual plane and the predictor construct on the conceptual plane takes the form of empirically based relationships that are convergent and discriminant in nature – similar to that described in the criterion-related validation strategy (Binning & Barrett, 1989; Campbell & Fiske, 1959; Duvenage, 1990; Guion, 1987; Nunnally, 1978).

Thus, once the criterion domain is delineated according to job analysis information, and evidence exists proving the importance of the predictor construct to the criterion domain (inference 4), the convergent and discriminant correlations between the predictor measure and the predictor construct can be computed (inference 3). If relevant correlations are obtained, the researcher can postulate that the predictor measure is able to predict job performance (inference 5) on the basis of construct-related validity. In summary, Route 2 of the nomological model entails generating evidence to support inference 3 and 4 which predicate inference 5 on the grounds of construct-related validity. Hence, Route 2 of the nomological model represents the construct-related validation route of selection.

Although each of the validation strategies has been individually discussed, the unitary nature of validity emphasized earlier in the discussion cannot be ignored. Take the case of an intelligence test being used as a predictor measure of future job performance. Criterion-related validity could prove that there is a relationship between those employees who score high on the intelligence test, and those who rate high on the job-related criterion measure (inference 1 and 2). However, intelligence is a construct and hence no evidence exists to ensure that the observables that are being assessed by the predictor measure are underlying variables of job performance. Thus, construct-related validity needs to be proven. If evidence is generated to prove that the test measures the observables that it claims to (inference 3), and one can motivate that the observables are underlying variables of job performance (inference 4), then one can confidently claim that the predictor measure is a valid predictor of future job success (inference 5).

In view of the above discussions, the conclusion may be drawn that, theoretically, the Nomological Model for Psychological Measurement Validity Assessment can be lucratively employed for human resource selection validation. However, an empirical investigation is required to verify the scientific applicability of the nomological model for selection validation since it has only ever been used for job evaluation validation purposes in RSA. Subsequently, the present research aimed to utilize the model to validate the procedure for selecting machine operators in a large food manufacturer. It was essential that an easily understandable study be conducted to prevent the obvious psychometric advantages of the nomological model from overshadowing the empirical aspect of this investigation, which is the main objective of this research. Hence, the option to employ the model for validating management-potential selection on the basis of an occupational personality questionnaire was rejected.

METHOD

Sample

Machinery upgrading in one of the Gauteng-based manufacturing units resulted in the reduction of machine operator positions. Employees who rejected the option of voluntary retrenchment were asked to apply for the "new" operator positions. As a pre-screening, the applicants were tested for English literacy and numeracy. The tests comprised of sections from different Adult Basic Education and Training level 3 Independent

Examination Board exams from 1998 and 1999. It was agreed with the union that a cut-off score of 40% would be set to determine which applicants would form the selection pool for future machine operators. Of the 107 tested, 78 applicants achieved 40% or higher. Of the 78 subjects, job performance questionnaires could not be obtained or were unusable for 11 of the applicants, reducing the sample size to 67. All (100%) of the subjects were Black, of which 82% were female and 18% male. The mean age of the sample was 42 years (SD = 7,10). The educational level ranged from standard one to standard ten, with 69% of the candidates falling within the standard five to standard eight bracket and 11% within the standard nine to ten bracket.

Measures and model variables

To apply the nomological model, four selection procedure variables require identification. Each of these variables is discussed in the following paragraphs.

The unobservable criterion domain. Rudimentary to any personnel selection validation study is the delineation of the ultimate criterion that a predictor measure attempts to predict, namely successful job performance. This necessitates a systematic job analysis. The machine operator job was analysed using the SHL Work Profiling System (WPS)(SHL, 1999a). Results of the job analysis explicate that a key aspect of job success depends on the operators ability to collect work-related information from written documentation, as well as colleagues and seniors, to make the correct decision based on the obtained information regarding the operation of the machine. This constitutes the unobservable criterion domain of the nomological model.

The observable predictor measure. The paper-and-pencil Potential Index Battery (PIB) (Erasmus & Minnaar, 1995) was used as the predictor measure. It is a South African developed battery of 24 tests. Each test, referred to as an index, was developed to measure a certain dimension ranging from cognitive through to social and emotional elements. The PIB Index 3 establishes a candidate's 'potential to read and to remember [information] correctly; to draw objective, sensible conclusions from the reading matter' (Erasmus & Minnaar, 1997, p. 23). It comprises a story excerpt and a separate 25-item questionnaire with a 3-point response

format: A indicates the statement is true, B indicates the statement is not true, and C uncertain. Subjects are given five minutes to read the excerpt, thereafter it is collected and the questionnaire handed out. Subjects are required to answer whether the item statement is true, false or uncertain according to the information (or lack thereof) in the preceding document, within an eight minute time limit.

Studies of the battery's reliability and validity in financial and academic institutions can be located in the Uni-PIB 2000 manual (Erasmus & Minnaar, 1997).

The observable criterion measure. A relevant measure of job performance that is observable and quantifiable is required as a substitute for the unobservable criterion domain. It was decided that a supervisory rating measure would be used. A four-item Job Performance Evaluation Questionnaire was used to measure the subject's performance on the job. The criterion measure items were based on four of the key job tasks that were identified in the job analysis results. The items namely (a) entering numbers or facts or information into a computer, (b) checking completed work is according to a set standard, (c) listening to verbal instructions in English from superiors, and (d) asking questions in English to find out job-related information, were considered to be critical to successful job performance. For each item, a 5-point scale response ranging from one (unable) to five (excellent) existed. Seniors of the subjects were required to rate the employee's performance by marking the scale that was most representative or descriptive of the employee's task performance.

The unobservable predictor construct. Some factors that were considered determinants of successful machine operator job performance were (a) English vocabulary of a moderate level, (b) ability to read, (c) language usage, and (d) reading comprehension. These variables relate to an individual's ability to process, encode and decode verbal and non-verbal materials, referred to as verbal reasoning (Echols, Stanovich, West & Zehr, 1996; Langdon, Rosenblatt & Mellanby, 1998; Oosterveld & Vorst, 1996; Polk & Newell, 1995). Accordingly, the construct that was used in this present study to predict performance or outcomes on the ultimate criterion domain was verbal reasoning.

Procedure

The PIB Index 3 measure was administered as part of a more comprehensive battery of psychometric measures comprising of PIB Index 2 (Spatial Reasoning), PIB Index 4 (Calculations) and PIB Index 5 (Mental Alertness). A trade union official together with a psychometrist administered the measures to groups of 10 to 15 subjects over a two-week period. Standard instructions were provided in Zulu to the applicants.

The Job Performance Evaluation Questionnaires were sent out to each of the production managers, after authorisation had been given by the unit manager, with the understanding that the information would be used for research purposes only. The production manager either completed the questionnaire himself, or asked the senior controller of the job incumbent to complete the questionnaire.

RESULTS

The aim of the study is to empirically demonstrate and evaluate the nomological model's applicability to selection procedure validation. The research purpose is therefore not of a statistical nature. Correlations were computed between the PIB Index 3 and the PIB Spatial Reasoning measure (Index 2), the pre-screening literacy test, and the job performance measure. Each of the job task items was correlated individually and then the composite job performance score was correlated with the PIB Index 3. The results are represented in Table 1 below.

Table 1 Correlations between PIB Index 3 and tests that measure a similar construct, a different construct, and scores from a job performance measure.

Measure	PIB Index 3
Literacy	0,59**
Spatial reasoning	0,13
Enter into computer	0,26*
Check work to a set standard	0,28**
Listen to instructions	0,29**
Ask questions	0,31**
Composite job performance	0,37**

* $p < 0,05$

** $p < 0,01$

The question of ultimate importance in any selection procedure validation study is whether a predictor measure successfully predicts outcomes on a criterion domain. The critical question pertaining to the example in this study is whether scores obtained from the PIB Index 3 successfully predict machine operator job performance. Harnessed upon the Nomological Model for Psychological Measurement Validity Assessment, this question of ultimate importance is delineated by inference 5 depicted in Figure 2. In the passages that follow, this ultimate inference is proven valid on the basis of the two evidential validation strategies, as illustrated in the nomological model.

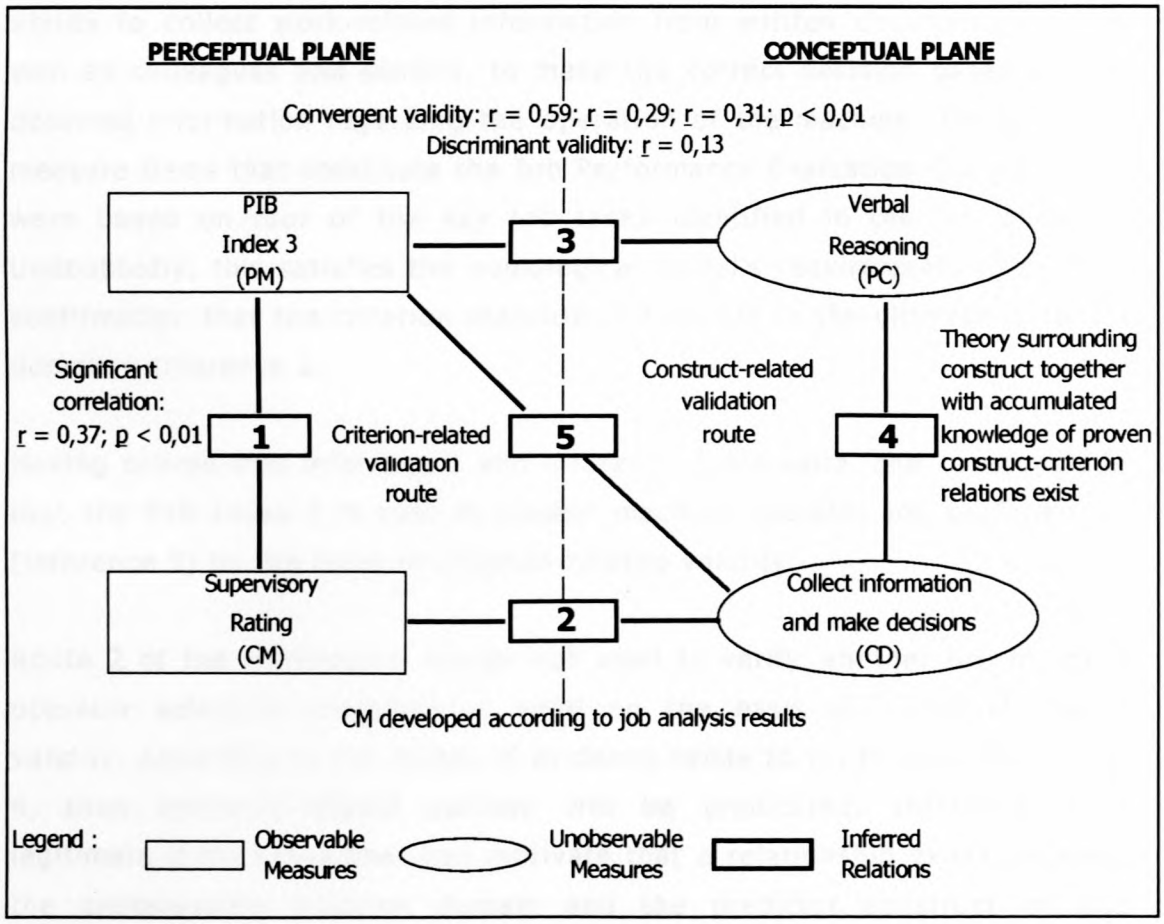

Figure 2 Operationalized Nomological Model for Psychological Measurement Validity Assessment for personnel selection.

Criterion-related validity of the machine operator selection procedure will be predicated if evidence is generated to support inferences 1 and 2 of the nomological model i.e. Route 1. As discussed in earlier paragraphs, inference 1 is valid provided that a significant correlation exists between the

observable predictor measure and the criterion measure. Validity of inference 2, according to the nomological model, requires that a job analysis is conducted to delineate the unobservable criterion domain, and confirmation is produced to verify the observable criterion measure is relevant to the ultimate criterion domain. Reference to Table 1 confirms that inference 1 can be made because a significant correlation does exist between the PIB Index 3 and the composite job performance scores (r = 0,37; p < 0,01).

The outcomes of the WPS job analysis were used to define the unobservable criterion domain of the machine operator job. As mentioned in the method section of this article, successful job performance depends on the operators ability to collect work-related information from written documentation, as well as colleagues and seniors, to make the correct decision based on the obtained information regarding the operation of the machine. The criterion measure items that constitute the Job Performance Evaluation Questionnaire were based on four of the key job tasks identified in the job analysis. Undoubtedly, this satisfies the nomological model's requirements to provide confirmation that the criterion measure is relevant to the ultimate criterion domain – inference 2.

Having proven that inference 1 and inference 2 are valid, one can conclude that the PIB Index 3 is able to predict machine operator job performance (inference 5) on the basis of criterion-related validity.

Route 2 of the nomological model was used to verify whether the machine operator selection procedure is valid on the basis of construct-related validity. According to the model, if evidence exists to verify inference 3 and 4, then construct-related validity will be predicated. Inference 4 is legitimate if the researcher can motivate that a relationship exists between the unobservable criterion domain and the predictor construct on the conceptual plane. To claim that verbal reasoning is a determinant of a machine operator's potential work performance (inference 4), theory surrounding what the construct comprises, together with accumulated knowledge of proven relationships between the construct and identifiable job demands must be provided. As discussed in the method section of this article, an individual's vocabulary, reading ability, language usage and reading comprehension are variables assumed to represent the theoretical

construct of verbal reasoning (Echols et al., 1996; Langdon et al., 1998; Oosterveld & Vorst, 1996; Polk & Newell, 1995). Local and international studies conducted on shop-floor production and technical staff in three different industries namely food, motor and telecommunications, reported significant correlations between predictors that measure aspects of verbal reasoning and job performance criteria that is considered relevant to the job of machine operator (SHL, 1988, 1989, 1999b). Criteria such as receiving information, working with information, assimilating information, communicating, working with machinery and overall job performance, correlated significantly with different predictor measures of aspects of verbal reasoning. In view of the construct verbal reasoning, and the findings of completed studies, one can motivate that a relationship exists between the criterion domain of the machine operator job and the unobservable predictor construct, thus verifying inference 4 of the nomological model.

Inference 3 of the model depicts the construct validity of a measuring instrument. Inference 3 is supported when correlations exist between the unobservable predictor construct and the observable predictor measure, that are convergent and discriminant in nature. To prove that the PIB Index 3 measures aspects of verbal reasoning, scores from Index 3 were hypothesized to correlate significantly with the pre-screening literacy test scores. Furthermore, two of the criterion measure items that were considered to relate to verbal reasoning, namely listening to verbal instructions in English from superiors and asking questions in English to find out job-related information, were hypothesized to correlate significantly with the PIB Index 3. To prove discriminant validity, a test that measures a construct that is different and unrelated to verbal reasoning was required. Accordingly, the PIB test of Spatial Reasoning (Index 2) was selected. It measures the 'ability to see [or] identify relationships between visual objects [and] to understand and reason two-and-three dimensionally' (Customer Focus, 2000, p. 1). The aim was to prove that no relationship exists between the PIB Index 3 and spatial reasoning. Research findings have however reported correlations between measures of verbal reasoning and measures of spatial reasoning. In the majority of studies, correlation coefficients ranged from 0,2 to 0,4 (Athanasou, 1999; Langdon et al., 1998; SHL, 1999b, 2000b). For the purposes of this study, it was reasoned that the correlations between measures of similar constructs would be significantly higher than those of measures of dissimilar constructs.

Significant correlations were obtained between the PIB Index 3 and the literacy measure (r = 0,59; p < 0,01), as well as the two construct-related criterion measure items (r = 0,29; p < 0,01) (r = 0,31; p < 0,01), thus proving evidence of convergent validity as required by the nomological model. An insignificant correlation (r = 0,13) was calculated between the PIB Index 3 and the measure of spatial reasoning, thus failing to reject the null hypothesis of no correlation and subsequently providing evidence of the measure's discriminant validity, as necessitated by the Nomological Model for Psychological Measurement Validity Assessment.

Therefore, the postulate that PIB Index 3 is able to predict machine operator job performance (inference 5) is supported because of proven inference 3 and proven inference 4.

DISCUSSION

Evaluation of the nomological model for validation of personnel selection procedures is the prime essential of this study. To establish the model's applicability, theoretical, professional, legal and practical determinants require deliberation.

Earlier discussions established that the nomological model addresses the two empirically based validation strategies, namely criterion- and construct-related validity. Furthermore, the model acknowledges the conceptual-perceptual interrelations between the two strategies (Duvenage, 1990). Consequently, the nomological model was deemed applicable for personnel selection validation on a theoretical basis.

According to the SIP guidelines (1998, p. 10), 'evidence for criterion-related validity typically consists of a demonstration of a useful relationship between the assessment procedure (predictor or predictors) and one or more measures of job-relevant behaviour (criteria or criterion)'. The Nomological Model for Psychological Measurement Validity Assessment affords this evidential base by substantiating inference 1 and 2, as was illustrated in the machine operator application. Inference 1 verified that a relationship exists between the predictor and the job performance measure on the basis of a significant correlation. Inference 2 predicated that the criterion measure was relevant to the job. It can thus be concluded that the nomological model is applicable for personnel selection validation because it

provides evidence of criterion-related validity, as described by professional standards.

Furthermore, according to the SIP guidelines (1998, p. 40), evidence of construct-related validity is two-fold:

> First is evidence that the construct is indeed important for job performance – that is, evidence must be grounded in a thorough knowledge of the job. Ordinarily, a job analysis can prove a part of the basis for identifying and defining psychological constructs which are important to job performance. Clarity of the articulation of the meaning and the nature of the construct, and well-informed expert judgement that logical relationships exist between the nature of the construct and identifiable demands of the job, is essential. The second is evidence that the instrument used as a assessment procedure is a valid measure of the construct and not of other constructs.

The nomological model presents this evidential base of construct-related validity in a systematic and somewhat less cumbersome manner – as was explicated in the machine operator investigation. Inference 4 corroborated that the construct was important for job performance. This was founded on the job analysis results and research findings of significant relationships between the construct and identified job demands. Inference 3, based on evidence of convergent and discriminant relations, established the construct validity of the measure. Consequently, the nomological model's utility for construct-related validation of human resource selection is sanctioned because it satisfies the evidential requirements as prescribed by professional standards.

The legal requirement, as stipulated by the EEA (1998), specifies that a measure used in psychological and similar assessments is prohibited unless it has been scientifically shown to be valid. A scientific investigation commands a planned, systematic, and empirical study (Kerlinger, 1986). The question must be asked as to whether a practitioner following the SIP guidelines would conduct such a planned and systematic study due to the rather cluttered nature of the guidelines. Conversely, the nomological model is an illustration of the study-components and the method required to ascertain scientific validity. Accordingly, the applicability of the nomological model for personnel selection validation is accentuated.

The final determinant of the model's applicability is a question of practicality. Is the model intelligible and lucid in character? Due to the diagrammatic nature of the model, and the consecutive inference numbering, the nomological model is able to elucidate the complexity of validation through depiction rather than copious written explanations. This further strengthens the argument of its utility.

In conclusion, the results of this study have proven that the Nomological Model for Psychological Measurement Validity Assessment is theoretically, professionally, legally and practically applicable to personnel selection validation procedures. This finding could be expected to have a significant influence on the SIP (1998) Guidelines for Validation and Use of Assessment Procedures for the Workplace, with the prospect that a third revision of the guidelines will result in the incorporation of such a model. Ultimately, the research results anticipate to impact on the understanding and manner in which practitioners conduct personnel selection validation studies in the future.

REFERENCES

Akkerman, A. (1989). Criteria and individual assessment. In M. Smith & I. Robertson (Eds), Advances in selection and assessment (pp. 67-79). Chichester: Wiley.

American Psychological Association (APA) (1985). Standards for Educational and Psychological Testing. Washington: APA.

Athanasou, J.A. (1999). Australian data for forms A, B, and C of the personnel selection tests. Psychological Reports, 84, 549-552.

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.

Borman, W.C., Rosse, R.L. & Abrahams, N.M. (1980). An empirical construct validity approach to studying predictor-job performance links. Journal of Applied Psychology, (65)(6), 662-671.

Campbell, D.T. & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 56(2), 81-105.

Cronbach, L.J. (1970). Essentials of psychological testing. (3rd ed.). New York: Harper & Row.

Customer Focus, (2000). PIB dimensions measuring trainability in a technical environment. Unpublished business proposal, Customer Focus, Gauteng.

De Bruin, C. (1998). Complying with the Employment Equity Act. Get your equity ducks in a row. People Dynamics, 16(11), 68-72.

Dunnette, M.D. (1966). Personnel selection and placement. California: Wadsworth.

Duvenage, A. (1990). Definition and application of a nomological validity model on job evaluation. Unpublished doctoral thesis, University of Stellenbosch.

Echols, L.D., Stanovich, K.E., West, R.F. & Zehr, K.S. (1996). Using children's literacy activities to predict growth in verbal cognitive skills: A longitudinal investigation. Journal of Educational Psychology, 88(2), 296-304.

Erasmus, P.F. & Minnaar, G.G. (1995). Realising human potential: The crux of the RDP. Human Resource Management, 11(1), 34-35.

Erasmus, P.F. & Minnaar, G.G. (1997). Uni-PIB 2000 manual. Unpublished manuscript, Uni-PIB, Pretoria.

Gatewood, R.D. & Field, H.S. (1994). Human resource selection. (3rd ed.). Fort Worth: Dryden Press.

Goldstein, I.L. (1993). Training in organizations. (3rd ed.). California: Brooks/Cole.

Grogan, J. (1998a). More equal than others. Employment Law, 14(2), 3.

Grogan, J. (1998b). Towards a rainbow workforce. The Employment Equity Bill. Employment Law, 14(2), 4-10.

Guion, R.M. (1978). 'Content validity' in moderation. Personnel Psychology, 31, 205-213.

Guion, R.M. (1987). Changing views for personnel selection research. Personnel Psychology, 40, 199-213.

Guion, R.M. (1998). Assessment, measurement, and prediction for personnel decisions. London: Lawrence Erlbaum.

Kerlinger, F.N. (1986). Foundations of behavioral research. (3rd ed.). USA: Harcourt.

Kleiman, L.S. & Faley, R.H. (1978). Assessing content validity: Standards set by the court. Personnel Psychology, 31, 701 – 713.

Kriek, H.J. (2000, June). Ensure fairness and best practice in selection. Paper presented at the SHL Employment Equity workshop, Cape Town.

Landy, F.J. (1991, June). Has the traditional selection model outlived its usefulness? If yes, what then? Paper presented at the Fourth Congress on Psychometrics for Psychologists and Personnel Practitioners, Johannesburg.

Landy, F.J. & Rastegary, H. (1989). Criteria for selection. In M. Smith & I. Robertson (Eds), Advances in selection and assessment (pp. 47-65). Chichester: Wiley.

Langdon, D.W., Rosenblatt, N. & Mellanby, J.H. (1998). Discrepantly poor verbal skills in poor readers: A failure of learning or ability? British Journal of Psychology, 89, 177-190.

Lawshe, C.H. (1985). Inferences from personnel tests and their validities. Journal of Applied Psychology, (70), 237-238.

Levy, A. & Associates. (1999, March). Non-discriminatory recruitment and selection. Paper presented at Labour Legislation workshop, Cape Town.

McCormick, E.J. (1979). Job analysis: methods and applications. New York: Amacom.

Nunnally, J.C. (1978). Psychometric theory. USA: McGraw-Hill.

Oosterveld, P. & Vorst, H.C.M. (1996). Note on reliability and validity of self-evaluation of ability scales. Perceptual and Motor Skills, 82, 991-994.

Polk, T.A. & Newell, A. (1995). Deduction as verbal reasoning. Psychological Review, 102(3), 533-566.

Republic of South Africa (RSA) (1995). Labour Relations Act (Act no. 66 of 1995). Government Gazette, 366(16861).

Republic of South Africa (RSA) (1997). Basic Conditions of Employment Act (Act no. 75 of 1997). Government Gazette, 390(18491).

Republic of South Africa (RSA) (1998). Employment Equity Act (Act no. 55 of 1998). Government Gazette, 400(19370).

Saville & Holdsworth (SHL) (1988). International validity study no. 36, shop floor staff (production) in motor industry. World Wide Web (http://www.shl.co.za).

Saville & Holdsworth (SHL) (1989). International validity study no. 38, process operatives in food production industry. World Wide Web (http://www.shl.co.za).

Saville & Holdsworth (SHL) (1999a, November). Job analysis and competency design. Paper presented at job analysis workshop, Cape Town.

Saville & Holdsworth (SHL) (1999b). National validity study no. V010, technical officers in telecommunications industry. World Wide Web (http://www.shl.co.za).

Saville & Holdsworth (SHL) (2000a). Guidelines for best practice in occupational assessment in South Africa. Pretoria: SHL Group.

Saville & Holdsworth (SHL) (2000b). National validity study no. V013, apprentices in transport industry. World Wide Web (http://www.shl.co.za).

Society for Industrial Psychology (SIP) (1998). Guidelines for the Validation and Use of Assessment Procedures for the Workplace. Gauteng: SIP.

Spencer, L.M., McClelland, D.C. & Spencer, S.M. (1992). Competency assessment methods. History and state of the art. Boston: Hay/McBer.

Thorndike, R.L. (1949). Personnel Selection. New York: Wiley.

Tustin, C.M. (1992, June). Guidelines for the validation and use of personnel selection procedures. Paper presented at the Congress on Psychometrics for Psychologists and Personnel Practitioners, Pretoria.

United States of America (USA) (1978). Uniform Guidelines on Employee Selection Procedures. Federal Register, 43(166).