

# A CRITICAL-REFLECTIVE EVALUATION OF THE USEFULNESS OF THE MOTIVATIONAL STYLES QUESTIONNAIRE

DERRY-LEE LAWRENCE  
DERRYL@HKLM.CO.ZA  
CONRAD SCHMIDT

*Department of Human Resource Management  
University of Johannesburg*

## ABSTRACT

The purpose of this study was to examine the usefulness of Motivational Styles Questionnaire (MSQ) from a critical-reflective perspective. The MSQ was administered to a random sample of managerial and consulting employees within the banking and information technology sector (N = 221). In order to examine the underlying factor structure and the corresponding reliabilities, a principal factor analysis was conducted, which yielded promising results. The findings of the critical analysis further revealed a number of gaps in the recommendations regarding the use of this instrument.

**Key words:** Graduate internship programmes, differentiated career success orientations, ethnographic analysis, card sorting techniques, Interviews

Psychological testing has always been controversial in psychology, particularly as the results from these tests can change a career path or otherwise alter a person's life course (Gregory, 2000). As a result a large amount of responsibility is placed on the part of the users, as well as the developers and publishers of psychometric tests. For this reason, it is imperative to evaluate these instruments methodically in order to determine whether they actually serve the purpose for which they are used.

The functionalist perspective explicitly maintains that a test validator has an obligation to determine whether a practice has constructive consequences for individuals and institutions, and especially to safeguard against adverse outcomes (Messick, 1980). It is evident then that there needs to be clear evidence for the usefulness of a psychological test for particular applications. The onus for this resides within the scientific community and those who use psychological tests as a basis for making important decisions about people and their behaviour. Gregory (2000, p. 115) openly states that "Psychological assessment is not a neutral endeavour, it is an applied science that occurs in a social and political context". Hence an ethical obligation exists to ensure that these tests are not only appropriate, but also that they are being used appropriately with due consideration for test takers.

The Motivational Styles Questionnaire (MSQ) is a partially-ipsative and partially-normative questionnaire. It is a relatively new personality measure developed by Roland Tarleton (1997) and previously distributed by The Psychological Corporation, which is recognised as one of the oldest test distributors in the world. The more recent distributors, Get Feedback, state in electronic format, that the MSQ "conforms to British Psychological Society and American Psychological Association guidelines on personality test construction and use" (<http://www.getfeedback.co.uk>). Although The Psychological Corporation no longer distributes the MSQ, it serves the needs of those existing clients who have purchased the instrument in the past.

The MSQ aims to assess the fundamental dimensions of behaviour "beyond those aspects covered by personality measures and ability tests" (Tarleton, MSQ manual, 1997, p. 1). According to the promotional claims of the test, it is typically used for placing new recruits; identifying high-flyer potential; relocating staff after a restructuring; distinguishing between good and bad performers; choosing between equally competent individuals in a selection situation; understanding the reasons why employees are not performing as expected; analysing team interaction issues from a motivational perspective; adding the motivational element to succession planning; gaining the commitment of talented individuals and retaining highly valued staff ([www.innovact.co.za](http://www.innovact.co.za)). According to the manual, the MSQ aims to

define the kind of working environment in which motivators are going to work through recognising the importance of "being where you want to be" (Tarleton, MSQ manual, 1997, p.1).

The ipsative section of the MSQ measures what people want out of a work situation, dealing specifically with areas of work that are deemed to satisfy particular needs. This section is called the Work Styles Preferences (WSP) questionnaire. Here the respondent is asked to indicate the extent to which a work situation appeals to them, in relation to an alternative work situation. In this way the MSQ utilises a type of forced-choice format, seeing that respondents are encouraged to endorse responses that most or least characterise their preferences. It is important, however, to note that the MSQ utilises the forced-choice format in a somewhat different way to that typically encountered in ipsative instruments. Respondents are asked to indicate their *relative* strength of preference on a scale with five points, where the centre point indicates no preference. This format may raise the question as to whether the nature of the WSP questionnaire is in fact a forced-choice design. However, despite endorsing strength of preferences that might create the impression of a normative scale, the instrument remains inherently ipsative given that the individual is encouraged to select a relative position between the two options. Furthermore, the answer sheet of the MSQ states that the descriptives at each end of the scale must not be construed as bipolar opposites, which is often the case in forced-choice design.

The following is an example of the WSP item format used in the answer sheet of the MSQ:

I generally see myself as...

Happy-go-lucky      Domineering

The normative section of the MSQ determines how individuals prefer to operate in their careers generally, as well as the rate at which they will typically progress towards achieving their goals. This section is referred to as the Work and Life Attitudes (WLA) questionnaire. Here the respondent is asked to indicate the extent of agreement or disagreement with a number of statements on a five point scale.

The purpose of this present study was to subject the MSQ, as well as the advice regarding its use, to critical analysis. From an epistemological point of view, the study could be classified as falling within the critical-reflective paradigm as it not only addresses an analysis of psychometric properties of the MSQ, but also the consequences related to the use of this instrument

(Habermas, 1972; Kemmis, 2001). Given that any action has normative implications, this paradigm encourages practitioners to adopt the stance of becoming *reflective* practitioners who not only reflect critically on the technical aspects of their methodology, but also on the ethical and social consequences of their actions in the process of implementing the technology (Argyris, 1990; Schön, 1983).

The basic strategy of applying a critical-reflective perspective to the evaluation of a psychological instrument proceeds from the premise that human beings form purposes in pursuit of their interests and that all action (including that of scientists and practitioners) is informed by some form of knowledge or theory about how to accomplish these purposes. The repertoire of knowledge in service of action, however, does not only include instrumental knowledge – that is, knowledge about how to do something – but it also includes a second-order kind of knowledge through which human beings are able to monitor their own actions. These may be seen as the “epistemic principles” (Geuss, 1981, p. 61), or alternatively the governing values (Argyris & Schön, 1974), which guide action. Critical reflection occurs when actors are made aware of the degree to which their actions are consistent with the espoused governing values (Argyris, Putnam & Smith, 1985). This enables researchers to reflect on their ways of thinking and acting, with particular reference to their own epistemology.

Existing practices can therefore be critiqued in light of the epistemic principles that a researcher holds. When people act in contradiction to their epistemic principles, their actions work against the realisation of their own objectives. However, a lack of reflection insulates the actor from becoming aware of such inconsistencies. According to Geuss (1981, p. 81) it is necessary to present agents “with evidence drawn from their own behaviour”, drawing on the notion of internal criticism.

The principle of internal criticism is founded on the requirements of consistency (where one is self-critical of one’s actions) and justice (where one does not hold others to claims which they themselves do not advocate). Internal criticism acts as a self-regulatory mechanism which directs scientists towards greater effectiveness from the perspective of their own values, serving productive rather than unproductive purposes. The intention is that once people are made aware of unintended inconsistencies, they will seek to redesign their actions accordingly.

Having regard to the format of the MSQ, much literature exists concerning the use of ipsative measures and scales. A critical-reflective stance proceeds from the premise that an ipsative instrument should be used consistently with its design. To this end, psychometric perspectives regarding inherent nature of ipsative scales are briefly reviewed.

### The inherent nature of ipsative scales

According to Murphy and Davidshofer (1994) there are two major schools of thought regarding the development of psychological tests. One school advocates the *normative* format where a test score reflects how a person has performed in relation to other people – the norm or comparison group – who have taken the same test. Proponents of the other school focus on the individual test taker and the interrelation of various psychological traits within the person. In this case the comparison is within the individual and not between the individual and other people. Interpretive differences exist between these two test construction strategies, as they provide different types of information. It is generally accepted that the normative format provides information regarding the strength of preference for a particular alternative as expressed on an intensity or frequency scale. The ipsative format, on the other hand, provides information about the relative strength of preference in relation to other possible preferences.

From a technical point of view, all data is considered ipsative when the raw scale scores sum to the same constant for different

individuals. There are different types of ipsative scales, one of which elicits forced-choice responses. Meade (2004, p. 531) identified “forced-choice ipsative data (FCID)” as a unique scale type, which is most commonly used within the organisational context of employee selection. Meade (2004) holds that it is common practice for researchers to treat all ipsative data as if they are identical. However, he further argues that this approach presents a problem because the psychometric properties of each design differ considerably from one another. Hence it is necessary to distinguish amongst the different types of ipsative data with which one is working. In spite of this, all data are described as ipsative when a given set of responses always sum to the same constant. This predominant property of ipsative measures introduces a fundamental problem in the interpretation of such data and in this way, challenges the principles of traditional test theory (Baron, 1996; Johnson, Wood & Blinkhorn, 1988; Meade, 2004).

The inherent nature of ipsative scales was initially articulated by Clemans (1966). He provides mathematically sound evidence to support the premise that the confinement of the overall sum of scale scores, introduces dependence amongst the scale scores. A number of authors cite his findings as a common point of departure in the analysis of ipsative inventories (Baron, 1996; Cornwell & Dunlap, 1994; Johnson et al., 1988; Meade, 2004). Clemans (1966) deduced the following mathematical findings:

- The sums of the columns, or rows, of an ipsative covariance matrix must equal zero;
- The sums of the columns, or rows, of an ipsative intercorrelation matrix will equal zero if the ipsative variances are equal;
- The average intercorrelations of ipsative variables has  $-1/(m-1)$  as a limiting value where  $m$  is the number of scales
- The sum of the covariances obtained between a criterion and a set of ipsative scores equals zero;
- The sum of ipsative validity coefficients will equal zero if the ipsative variances are equal

The mathematical derivations of Clemans (1966) provide more specific predictions of the technical consequences of using ipsative measures as if they were normative (Johnson et al., 1988). For an in-depth discussion of the statistical attributes of ipsative scales the reader is referred to the article by Clemans (1966). Johnson et al. (1988) further provide practical evidence for the mathematical results of ipsative data in relation to a number of different tests. However for the purposes of this present study, counterarguments concerning the reliability and validity findings of ipsative measures in more recent studies will be examined.

In accordance with classical test theory, an observed score consists of two components, namely a true score and an error score. Kerlinger & Lee (2000, p. 648) states that “reliability is defined, so to speak, through error: the more error, the greater the unreliability; the less error, the greater the reliability”. Johnson et al. (1988) further declare that the purpose of estimating the reliability of a test is to quantify this error. Yet ipsative measures have no error component by definition (Cornwell & Dunlap, 1994; Hicks, 1970). Therefore logic indicates that classical test theory cannot yield valuable estimates of reliability coefficients in respect of ipsative measures, since there is no error score and the score on any scale can be perfectly predicted from the other  $k - 1$  scales (Johnson et al. 1988; Meade, 2004). This deduction supports the widespread belief that standard statistics cannot be applied; as ipsative procedures systematically violate the basic assumptions upon which such statistics depend.

Many research studies demonstrate spurious reliability coefficients as a result of the fundamental scale interdependence within ipsative measures (Baron, 1996; Bartram, 1996; Johnson et al., 1988; Meade, 2004). For example, Gregory (2000, p. 517) reports on the reliability findings of the Edwards Personal Preference Schedule (EPPS), which he describes as “not particularly exciting”. The test-retest reliability of the ipsative,

forced-choice format of the EPPS reflects reliabilities ranging from 0.55 to 0.87. However, the results found within these studies are regarded as questionable for a number of reasons. In the first place, the stability within one scale creates consistency within some, or even most of the other scales (Johnson et al., 1988). Secondly, artifactual internal consistencies will be produced because the internal consistencies of the scales are essentially inter-dependent, which in turn affects construct interpretations (Tenopyr, 1988). The reliability coefficients therefore tend to be exaggerated, and so there is a possibility that the 'true' reliabilities are even lower than expected (Bartram, 1996). Furthermore, high reliability is not necessarily a guarantee of good scientific results. Yet, despite these deductions, numerous authors maintain that the reliability findings of ipsative scales may in fact be usefully interpreted, provided a number of prerequisites are met (Baron, 1996; Bartram, 1996; Saville & Willson, 1991).

The normative version of the Occupational Personality Questionnaire (OPQ) factor model reported substantially higher internal consistency reliabilities than that of the forced-choice format (Baron, 1996). This was substantiated by the work of Bartram (1996) whose findings indicate that ipsative reliabilities tend to be lower than normative reliabilities. However, both these studies demonstrate that an increase in the number of scales (30+) will inevitably increase the reliability of the measure, provided the constructs are not highly positively correlated. An additional study by Saville and Willson (1991) found that ipsative data based on a relatively large number of scales can be used legitimately in estimating reliabilities. Cornwell and Dunlap's study (1994), however, contests this view. These authors demonstrate that not all correlations are necessarily reliability estimates and that the apparent lack of information within ipsative scores prevents them from being used as substitutes for normative scores.

It is apparent that there is no agreement with regard to the reliability of ipsative measures. Proponents maintain that the reliability coefficients of ipsative scores are interpretable as they fall slightly below those found within normative scores for a reasonably large set of scales, given that the constructs are not highly positively correlated. On the other hand opponents argue that ipsative reliability coefficients are inflated and at all times, inadequate for meaningful interpretation. A further area of contention amongst these authors is whether the use of factor analytic procedures of ipsative data is a legitimate means of demonstrating theoretical validity.

Factor analysis is perceived as a construct validity tool because it pays particular attention to common factor variance (Kerlinger & Lee, 2000). However, if construct validity is defined as common factor variance, then these authors hold that factor analysis is, by definition, firmly tied to measurement theory. Factor analysis is in this way regarded as a constitutive meaning method which seeks to investigate the constitutive meanings of constructs. However, Johnson et al. (1988) state that factor analyses of ipsative data produce results which contradict both statistical and test theory, and in so doing yield misleading results. This is based upon the apparent psychometric constraints present within scale intercorrelations. Every scale score is perfectly predictable from the remaining scores because the subject totals sum to a constant, which results in primarily negative correlations amongst the scores. Since the scales are proven to be mathematically dependent and the average intercorrelations have negative values, ipsative scale correlations of any sort are not statistically interpretable (Closs, 1996; Cornwell & Dunlap, 1994; Johnson et al., 1988).

This view is supported by a number of studies which analysed the statistical properties of ipsative measures against equivalent normative ones. For example, a study conducted by Piedmont, McCrae, and Costa (1992) confirmed that the ipsative, forced-choice format of the EPPS apparently lowered validity coefficients

and decreased convergent and discriminant validity (Gregory, 2000). The results demonstrated that the relationships were strongest and most theory-confirming when the EPPS was scored in a normative fashion. A similar study was used to determine the validity of the DISCUS, an ipsative personality test (Martinussen, Richardsen & Varum, 2001). Both an ipsative and normative version of the DISCUS were administered to a sample of undergraduate students (N=103). The outcome of the study revealed that the normative and ipsative versions of the test were not equivalent and that caution is required when using the ipsative version for psychometric evaluations, as in validation studies.

Conflicting evidence, however, is presented in a study by Saville and Willson (1991), which involved an analysis of the OPQ data, based on computer simulated data sets and real data. They demonstrated that ipsative data may in fact be factor analysed as validly as normative data, where individuals can be compared on a scale by scale basis. Saville and Willson (1991) concluded that ipsative data, based on a relatively large number of scales (30+), is no less valid than normative data.

Baron (1996) supports the observation that ipsative results become more like normative ones with a notable increase in the number of scales. However she maintains that ipsative reliabilities tend to hold more promising results than factor analyses, and that there is little chance that ipsative data will be useful in determining underlying constructs. The results of Cornwell and Dunlap's study (1994) further refute the notion that there is a high level of correspondence between factor analyses of ipsative and normative inventories. In addition to this, Meade (2004, p. 544) remarks that there is no legitimate procedure for conducting factor analyses with particular reference to the current design of FCID measures, since it requires "many more parameters to be estimated than are possible with the observed variance/covariance matrix". Despite the dispute regarding factor analytic research findings of ipsative tests, a number of authors retain the view that, in some circumstances, ipsativity may actually increase validity (Bartram, 1996; Kerlinger & Lee, 2000; Saville & Willson, 1991; Toner, 1987).

Saville and Willson (1991) argue that ipsative measures have specific advantages over normative ones, in that they provide a certain degree of control over various forms of response bias. Forced-choice items control response styles and "item desirability" as they allow the respondent to choose among alternatives which on the surface, 'appear' to be equally favourable or unfavourable (Kerlinger & Lee, 2000, p. 719). For this reason the forced-choice format is accepted as a means to control for certain forms of response bias, of which ipsativity is an unwanted side-effect (Bartram, 1996).

Toner's (1987) study found that forced-choice responses offer a viable solution to acquiescent response bias, a distorting influence in the measurement of attitudes. However, Ray (1990) refutes Toner's conclusions based on a number of limitations inherent within forced-choice scales. Firstly, he mentions that these scales are based upon unmet assumptions where there is no way of checking whether opposite-seeming items are actually perceived as opposites by respondents on different occasions. Secondly, if the items are not equal in social desirability, there is a large possibility that the choice between alternatives will be determined by social desirability considerations. Finally, as shown, forced-choice scales cannot be meaningfully intercorrelated (Johnson et al., 1988). Therefore Ray (1990) concludes that forced-choice scales are not a viable option for response bias.

Anastasi and Urbina (1997) supports the contention that the forced-choice technique does not prove to be as effective in controlling faking or social desirability as anticipated. Findings based on the Edwards Personal Preference Schedule (EPPS) showed that items presented in free-choice format correlated highly with

those of the forced-choice format (Anastasi & Urbina, 1997). In addition to this, the social desirability of particular items is not constant for all purposes and may differ for different occupations (Anastasi & Urbina, 1997; Ray, 1990). Meade (2004) further indicates that, to date, no study has adequately demonstrated that FCID reduces response bias accurately, and when FCID is recognised as a decision-making process, it is clear that faking may still occur. Therefore it is not strange when authors caution test developers to avoid using ipsative scales, unless distortions within the normative formats are shown to be a significant threat to validity (Bartram, 1996; Hicks, 1970).

The existing research based on the reliability and validity findings of ipsative measures, highlights the opposing views concerning the utilisation of standard analytic procedures to interpret ipsative data. This in turn questions the appropriateness of using ipsative scores as a basis for inter-personal comparisons. Evidence has shown that the use of ipsative personality tests is common within organisations for making pertinent decisions about people, which results in a comparison amongst people (Johnson et al., 1988; Martinussen, Richardsen & Varum, 2001; Meade, 2004; Pickworth & Schoeman, 2000; Saville & Willson, 1991). Hence it is clear that there is still some confusion with regard to the use of ipsative instruments as tools for comparing people against one another.

Inter-individual comparisons require the conversion of raw scores to standard scores using a standardisation sample. This, however, is regarded as nonsensical in terms of the mathematical findings of ipsative measures, given that all scale scores sum to the same constant. Since there is not a total scale score across the scales, ipsative scores are generally accepted as relative, as opposed to absolute measures of individual preference. This was originally demonstrated as far back as 1951 in a study by Allport, Vernon and Lindzey (1951, p. 6), who stated "The test measures only the relative strength of the six evaluative attitudes". Hence ipsative tests measure the relative preference among a number of choices the respondent is faced with. Consequently each respondent's scores distribute about the mean of that particular individual, and not the population mean (Clemans, 1966; Johnson et al., 1988).

Closs (1996) supports this contention based on the nature of ipsative scale inter-dependencies, which cause a high score on one scale at the expense of a lower score on the remaining scales. According to him normative interpretations should never be applied to ipsative data as they can grossly distort an individual's true position on the factors being measured. Closs' (1996) study demonstrates that peoples' interests may be grossly overestimated or underestimated, which in turn may produce devastating effects when used within the context of personnel psychology. Closs (1996) therefore expands on Johnson et al's (1988) contention that the application of norms to ipsative scores is not only an act of futility and a waste of time, but it is also harmful to the individual. Hence, despite the findings which show that ipsative measures may approximate their normative counterparts, it appears to be widely accepted that ipsative tests cannot be used normatively and should be avoided altogether as a basis for making inter-personal comparisons.

The purpose of the present study was to critically evaluate the recommended use of the MSQ. The rationale behind a critical-reflective analysis is to ascertain whether the development and advice regarding its use is in line with established psychometric principles. Given the partially ipsative nature of the MSQ, particular emphasis is placed on the extent to which the inferences drawn from test scores, claims regarding its areas of application, as well as the advice regarding its use, is consistent with the design of ipsative instruments. Furthermore, to the best of the present author's knowledge, no data has been reported in respect of the psychometric properties of the MSQ by independent researchers. Therefore, it was considered necessary to evaluate the psychometric properties of the test, in addition to the test's format.

## RESEARCH DESIGN

### Participants/respondents

The research was conducted within a number of different organisations, primarily from banking and information technology sectors. A convenience sample of 226 managerial and consulting employees of diverse ages, education and work experience, participated in the study. The managers ranged from top to lower levels of management. The sample is regarded as representative of a managerial and consulting population within the economic sector, based on the fact that a broadband of managers and consultants were randomly selected from several organisations. By virtue of the confidential nature of the study, the precise number of individuals within each ethnic group cannot be reported on. However, the author can state with confidence that all four groups (African, White, Indian and Coloured South Africans) were represented. As far as gender is concerned 58% were male and 33,6% were female. Missing information accounted for 8,4%. The average age of the respondents was 34,3 years with a standard deviation of 7,875 - ranging from 19 years to 59 years with the majority of respondents falling within the 25 to 41 years age group.

### Measuring Instrument

As noted earlier, the MSQ is a self report measure in two parts, designed to assess the kind of work situation to which a person is best suited. The promoters of the test claim that the MSQ is not concerned with what individuals can or cannot do, but with what they actually will do in the reality of the work situation. Therefore, the MSQ is used to predict future performance by comparing the strengths of a person's motivational preferences (www.innovact.co.za). The development of the MSQ derives from a synthesis of various need and process approaches to explaining what motivation is fundamentally about.

The approach of the MSQ is founded upon the 'Valence-Expectancy Theory', which suggests that motivation is based on valuing the reward and expecting that exerting the effort will result in attaining it (Tarleton, MSQ manual, 1997). Since the first half deals with what people want from their jobs, it is about the valence in the sense of valuing the job. The second half deals with the way people like to operate and in so doing relates to expectancy, in the sense of feeling that they can do it. Hence the MSQ was developed by "adapting" some well established techniques in the work situation (Tarleton, MSQ manual, 1997, p. 5).

The MSQ is administered as a paper-and-pencil test. For the purposes of this study, the tests were hand-scored using the assigned scoring keys and the scoring sheets. The two sections will now be explained in terms of scales, dimensions and scoring procedures.

#### *The ipsative section of the MSQ (Work Style Preferences Questionnaire)*

The original version of this section consisted of 96 items and three scales labelled Achievement, Security and Power, which were considered the "Big Three" needs encompassing all others (Tarleton, MSQ manual, 1997, p. 49). However a factor analysis on a sample of 300 revealed a seven scale framework, namely Achievement (ACH); Independence (IND); Structure (STR), Affiliation (AFF); Systems power (SYS); People Power (PEO) and Personal Power (PER). The ACH scale is designed to assess a person's need for personal achievement. The IND scale measures an individual's need for independence and autonomy. The STR scale focuses on an individual's need for structure and clearly defined targets and objectives. The AFF scale measures a person's need for team support and their desire to seek other people's views. The SYS scale identifies an individual's need for power in organisational systems, whereas the PEO scale measures a need for power and the desire to organise people. The PER scale, on the other hand, measures an individual's need for personal power.

The final WSP questionnaire consists of 45 forced-choice items, where respondents are instructed to focus on their view of their career as a whole, rather than just on their present job. The questionnaire also instructs the respondent not to regard the descriptive words at the end of each scale as opposites, but rather that they should choose the answer along the scale which best applies to them. In terms of scoring, the scale scores (plus and minus scores) are added together, along with a designated scale constant listed in the scoring sheet. The converted scale scores sum to a "check total" of 270 (MSQ scoring sheet) and are transformed to standard scores or stens for the purpose of comparing an individual's profile with a broad sample of managers and professionals (N=1269). The sten scores are then used to plot an individual's profile.

#### *The normative section of the MSQ (Work Life Attitudes Questionnaire)*

The original version of this section consisted of 105 items and three scales, specifically, Short-term striving (SHO); Medium-term striving (MED) and General Orientation (GEN). The SHO scale is designed to assess global confidence and an individual's approach towards getting started on new tasks. Here one may be slow and careful to consider the nature of the task, or on the other hand, one may be confident and keen to have a go. The MED scale assesses internal-external control and the way in which individuals drive projects through to completion, where some may be cautious and keen to ensure that they are on the right track, as opposed to those who are goal focussed and less likely to consider alternatives. The GEN scale pertains to resultant achievement motivation where one may have an operational focus or a personal focus. An operational focus would include those people who achieve objectives quickly and efficiently (consolidators), whereas those with a personal focus will be inclined to seek out new challenges and do the best job possible.

The SHO scale was "slightly" reworded in terms of Rosenberg's (1965) Self-Esteem Scale (Tarleton, MSQ manual, 1997). Factor analysis indicated that all ten items should be retained, however it was doubled in length by writing new items that were approximately opposite to Rosenberg's, as the standard deviation was only 4.01. The developer was granted permission to use the expanded version based on Rosenberg's Scale. The MED scale was based on Rotter's (1966) Internal-External Control Scale. Factor analysis suggested that only 19 statements loading on the first factor should be included in the new version of the measure, and these included six full items and seven half items from Rotter's scale. A twentieth item was devised which is now question 32: "Achieving good results depends on having the right opportunities". Permission was granted to the developer, by Professor Rotter, to use items from his original scale. The GEN scale was derived from an existing measure of resultant achievement motivation, which was subsequently dropped. Factor analysis showed that the best items were those about striving for success or thinking about it, as well as avoiding failure or worrying about it. Therefore 20 new items were written and aimed at presenting a clear distinction between two ways of operating.

The WLA questionnaire consists of 60 items, where the respondent is required to indicate the extent of their agreement or disagreement with the statements on a five point scale. The scoring procedure is similar to that of the WSP scores; however in this case, there is no check on the grand total.

#### **Procedure**

The MSQ was distributed to pre-selected individuals from a number of organisations, mainly from within human resource departments, who assisted in the random selection of participants, in addition to the administration of the MSQ. The questionnaires were completed by the research participants in the various locations at different times. The completed questionnaires were returned to the author, and individually hand-scored in accordance to the instructions set out in the test manual. Individual profiles were created for each participant and returned, along with the standard MSQ report sheet explaining the results. Many of the

respondents preferred to remain anonymous and in so doing they utilised a coded title, such as AB1, instead of their real names. This may account for some of the missing biographical information evident in some of the returned questionnaires.

#### **Data Analysis**

The results of the study were analysed in two ways, by means of statistical analyses and critical-reflective methodology. The results of the WSP questionnaire were evaluated, followed by the outcomes of the WLA questionnaire.

The statistical analyses were conducted by the Statistical Consultation Service of the University of Johannesburg.

#### *The ipsative section of the MSQ (Work Style Preferences Questionnaire)*

The items of the WSP questionnaire were intercorrelated to produce an inter-item correlation matrix. The frequencies of paired comparisons amongst the scales were tabulated. Standard analyses were not applied to this portion of the MSQ. The critical method was used in order to evaluate the treatment of this scale as if it were normative as prescribed by the MSQ manual (Tarleton, 1997), in relation to ipsative test theory.

#### *The normative section of the MSQ (Work Life Attitudes Questionnaire)*

In order to examine the gradients of the correlations in an intercorrelation matrix of continuous variables, the means, standard deviations, coefficients of skewness and kurtosis were computed. To overcome the effects of differential skewness of the items of the WLA questionnaire, the procedure suggested by Schepers (2004) was followed in order to determine the underlying factor structure and the corresponding reliabilities of the scales. The critical method was used in order to analyse the degree to which the postulated factors relate to the original scales of the WLA questionnaire.

## RESULTS

#### *The ipsative section of the MSQ (Work Style Preferences Questionnaire)*

The 45 items of the WSP questionnaire were intercorrelated to form an inter-item correlation matrix in order to determine the quantity of negative values. Inspection of the table revealed numerous statistically significant negative correlations, along with a large number of insignificant correlations. As the intercorrelation matrix is of order 45x45, it is too large to be reproduced in this paper. The seven WSP scales were tabulated (Table 1) in order to evaluate the frequency of paired comparisons amongst the scales. From an inspection of Table 1 it is evident that each scale was not equally paired with every other scale.

**TABLE 1**  
**FREQUENCY OF COMPARISONS**

SCALES	ACH	IND	STR	AFF	SYS	PEO	PER	Row Totals
ACH	X	0	4	1	3	0	3	11
IND	0	X	1	1	0	3	0	5
STR	3	2	X	0	3	1	1	10
AFF	2	1	0	X	0	1	1	5
SYS	2	0	2	0	X	0	0	4
PEO	0	2	2	1	0	X	0	5
PER	2	0	1	2	0	0	X	5
Column Totals	9	5	10	5	6	5	5	45

A critical analysis of the WSP questionnaire revealed the following results:

- The WSP scale is “ipsative in format” (Tarleton, MSQ manual, 1997, p. 15).
- The total scale scores always sum to a total of 270 (MSQ scoring sheet).
- The “scores are normed” and converted to standard scores or stens (MSQ manual, 1997, p. 15).
- Ipsative raw scores are regarded as of “little value” (MSQ manual, 1997, p.15).

*The normative section of the MSQ (Work Life Attitudes Questionnaire)*

The normative section of the MSQ was subjected to a factor analysis in order to evaluate the factor structure of the WLA questionnaire. Initially, however, analyses were performed to determine whether the sample was adequate and if there was justification to perform a factor analysis on this scale. The Kaiser-Meyer-Olkin Measure of Sampling Adequacy (MSA) indicated  $r = 0,733$  and this was considered adequate for factor analysis. The result of Bartlett's test of Sphericity was  $\chi^2 = 4319.953$ ;  $p < 0.0001$ . The significance of this test means that the variables are substantially mutually correlated and therefore provide the necessary justification to continue with the factor analysis.

As a first step, the means, standard deviations, coefficients of skewness and kurtosis of the 60 items of the WLA questionnaire were computed. The skewness and kurtosis of the items were evaluated, and are given in Table 2. The coefficients of skewness vary from -2,616 to 2,410. Table 2 indicates that the items are negatively skewed with a mean of -0,106 and standard deviation of 1,217. The coefficients of kurtosis vary from -1,296 to 12,011. A coefficient of 7 signifies a leptokurtic distribution and indicates a very low reliability of the measure (West, Finch & Curran, 1995). Items 9, 29 and 56 yielded values greater than 7.

Next, the 60 items were intercorrelated and subjected to a principal factor analysis: To determine the dimensionality of the vector space of items, the items were subjected to a principal factor analysis. As a first step the eigenvalues of the unreduced intercorrelation matrix were calculated. It was found that 19 eigenvalues were greater than unity and in accordance with Kaiser's (1961) criterion 19 factors were extracted. The obtained communalities were considered reasonable in respect of the 60 items, and varied from 0,306 to 0,746. The factor matrix was rotated to simple structure by means of Varimax Rotation with Kaiser Normalisation. The MSQ2 intercorrelation matrix (60x60), eigenvalues and factor matrix are not reproduced here due to limited space.

To counteract the effect of differential skewness of items, subscores were computed in respect of each of the factors by adding all the items with high loadings on a factor, together (Schepers, 2004). A second factor analysis was then performed based on the subscores. The descriptive statistics of the second round of factor analysis revealed an MSA of 0,74 and Bartlett's test of Sphericity gave  $\chi^2 = 0,743$ ;  $p < 0.0001$ . Since the MSA was not optimal, an anti-image factor analysis was performed. This resulted in subscore 19 being discarded as it represented a single item that did not make a meaningful contribution. The subscores were then intercorrelated and yielded the matrix of intercorrelations given in Table 3.

The eigenvalues of the subscores were calculated and given in Table 4. Five of the eigenvalues were greater than unity, suggesting five factors (Kaiser, 1961).

Accordingly five factors were extracted and rotated to simple structure by means of a direct Direct Oblimin rotation. The Rotated Factor Matrix is given in Table 5 including the communalities ( $h^2_j$ ) which range from 0,228 to 0,689. From an

inspection of Table 5 it is evident that four of the five factors are well determined, with three or more high loadings. Factors 1 and 3 are well determined, with several high loadings. Similarly, Factors 2 and 4 are well determined with three substantial loadings. However, Factor 5 is a doublet with two moderate loadings.

**TABLE 2**  
**COEFFICIENTS OF SKEWNESS AND KURTOSIS**

	Skewness		Kurtosis	
	Statistic	Std. Error	Statistic	Std. Error
Item 1	0.356	0.164	-1.235	0.327
Item 2	-1.083	0.164	0.459	0.326
Item 3	1.061	0.164	0.476	0.326
Item 4	0.977	0.164	-0.336	0.326
Item 5	-0.088	0.164	-1.227	0.326
Item 6	-1.035	0.164	0.284	0.326
Item 7	0.675	0.164	-0.594	0.326
Item 8	-0.234	0.164	-1.224	0.326
Item 9	-2.616	0.164	10.293	0.326
Item 10	-0.961	0.164	-0.007	0.327
Item 11	-1.749	0.164	4.304	0.326
Item 12	-0.728	0.164	-0.282	0.326
Item 13	2.41	0.164	5.887	0.326
Item 14	1.239	0.164	0.811	0.326
Item 15	-0.086	0.164	-1.192	0.327
Item 16	1.841	0.164	2.757	0.326
Item 17	-0.014	0.164	-1.151	0.326
Item 18	1.827	0.164	3.271	0.326
Item 19	-1.132	0.164	0.343	0.326
Item 20	-1.901	0.164	4.849	0.326
Item 21	0.701	0.164	-0.77	0.326
Item 22	1.314	0.164	1.01	0.326
Item 23	-0.046	0.164	-1.05	0.327
Item 24	-1.295	0.164	1.386	0.326
Item 25	0.32	0.164	-0.992	0.326
Item 26	-0.591	0.164	-0.836	0.327
Item 27	1.127	0.164	0.818	0.326
Item 28	-0.074	0.164	-1.198	0.326
Item 29	-2.304	0.164	9.059	0.326
Item 30	-0.593	0.164	-0.836	0.326
Item 31	-0.359	0.164	-1.074	0.326
Item 32	-0.348	0.164	-1.086	0.326
Item 33	0.056	0.164	-1.054	0.326
Item 34	-0.118	0.164	-1.092	0.326
Item 35	-0.217	0.164	-1.111	0.326
Item 36	0.747	0.164	-0.49	0.326
Item 37	0.317	0.164	-1.066	0.326
Item 38	-2.257	0.164	6.064	0.326
Item 39	0.097	0.164	-1.348	0.326
Item 40	-1.351	0.164	2.645	0.326
Item 41	-0.405	0.164	-0.875	0.326
Item 42	-0.359	0.164	-0.998	0.326
Item 43	0.039	0.164	-1.256	0.326
Item 44	0.367	0.164	-1.019	0.326
Item 45	1.486	0.164	1.732	0.327
Item 46	1.033	0.164	0.519	0.326
Item 47	1.052	0.164	0.061	0.327
Item 48	0.683	0.164	-0.442	0.326
Item 49	-2.102	0.164	6.096	0.327
Item 50	0.924	0.164	0.306	0.326
Item 51	0.214	0.164	-1.145	0.327
Item 52	-0.452	0.164	-1.087	0.327
Item 53	-0.357	0.164	-1.168	0.326
Item 54	1.147	0.164	0.175	0.326
Item 55	0.033	0.164	-1.296	0.326
Item 56	3.137	0.164	12.011	0.326
Item 57	-1.953	0.164	5.054	0.326
Item 58	-2.037	0.164	6.054	0.326
Item 59	-1.288	0.164	1.596	0.326
Item 60	-1.378	0.164	1.944	0.326

**TABLE 3**  
**UNREDUCED INTERCORRELATION MATRIX OF SUBSCORES**

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18
F1	1	0.579	0.236	-0.375	0.140	0.149	0.165	0.052	0.041	0.163	0.327	-0.103	-0.287	-0.222	0.203	0.230	-0.253	0.075
F2	0.579	1	0.369	-0.342	0.205	0.284	0.262	-0.032	0.105	0.245	0.266	-0.096	-0.319	-0.204	0.133	0.244	-0.244	0.309
F3	0.236	0.369	1	-0.157	0.062	0.180	0.196	-0.082	0.236	0.110	0.201	0.116	-0.184	0.087	0.073	0.239	-0.321	0.401
F4	-0.375	-0.342	-0.157	1	-0.199	-0.191	-0.332	0.110	0.119	-0.034	-0.350	0.172	0.329	0.201	-0.121	-0.103	0.280	-0.010
F5	0.140	0.205	0.062	-0.199	1	0.274	0.161	0.016	0.036	0.267	0.149	0.012	-0.110	-0.075	0.229	0.132	0.079	-0.066
F6	0.149	0.284	0.180	-0.191	0.274	1	0.326	-0.151	0.194	0.131	0.247	-0.293	-0.193	0.003	0.215	0.055	-0.024	0.099
F7	0.165	0.262	0.196	-0.332	0.161	0.326	1	-0.197	0.118	0.026	0.345	-0.178	-0.142	-0.006	0.240	0.186	-0.014	0.096
F8	0.052	-0.032	-0.082	0.110	0.016	-0.151	-0.197	1	0.062	0.106	-0.066	0.226	0.107	0.081	-0.016	0.030	0.029	-0.028
F9	0.041	0.105	0.236	0.119	0.036	0.194	0.118	0.062	1	0.146	0.045	0.103	0.055	0.264	0.095	-0.014	-0.010	0.216
F10	0.163	0.245	0.110	-0.034	0.267	0.131	0.026	0.106	0.146	1	0.129	-0.004	-0.024	0.019	0.105	0.219	0.008	0.085
F11	0.327	0.266	0.201	-0.350	0.149	0.247	0.345	-0.066	0.045	0.129	1	-0.075	-0.172	-0.063	0.174	0.246	-0.086	0.051
F12	-0.103	-0.096	0.116	0.172	0.012	-0.293	-0.178	0.226	0.103	-0.004	-0.075	1	0.190	0.126	-0.042	0.101	-0.015	0.155
F13	-0.287	-0.319	-0.184	0.329	-0.110	-0.193	-0.142	0.107	0.055	-0.024	-0.172	0.190	1	0.131	-0.061	-0.101	0.206	-0.130
F14	-0.222	-0.204	0.087	0.201	-0.075	0.003	-0.006	0.081	0.264	0.019	-0.063	0.126	0.131	1	0.122	-0.018	0.047	0.046
F15	0.203	0.133	0.073	-0.121	0.229	0.215	0.240	-0.016	0.095	0.105	0.174	-0.042	-0.061	0.122	1	0.157	-0.005	-0.074
F16	0.230	0.244	0.239	-0.103	0.132	0.055	0.186	0.030	-0.014	0.219	0.246	0.101	-0.101	-0.018	0.157	1	-0.091	0.037
F17	-0.253	-0.244	-0.321	0.280	0.079	-0.024	-0.014	0.029	-0.010	0.008	-0.086	-0.015	0.206	0.047	-0.005	-0.091	1	-0.097
F18	0.075	0.309	0.401	-0.010	-0.066	0.099	0.096	-0.028	0.216	0.085	0.051	0.155	-0.130	0.046	-0.074	0.037	-0.097	1

**TABLE 3**  
**UNREDUCED INTERCORRELATION MATRIX OF SUBSCORES**

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18
F1	1	0.579	0.236	-0.375	0.140	0.149	0.165	0.052	0.041	0.163	0.327	-0.103	-0.287	-0.222	0.203	0.230	-0.253	0.075
F2	0.579	1	0.369	-0.342	0.205	0.284	0.262	-0.032	0.105	0.245	0.266	-0.096	-0.319	-0.204	0.133	0.244	-0.244	0.309
F3	0.236	0.369	1	-0.157	0.062	0.180	0.196	-0.082	0.236	0.110	0.201	0.116	-0.184	0.087	0.073	0.239	-0.321	0.401
F4	-0.375	-0.342	-0.157	1	-0.199	-0.191	-0.332	0.110	0.119	-0.034	-0.350	0.172	0.329	0.201	-0.121	-0.103	0.280	-0.010
F5	0.140	0.205	0.062	-0.199	1	0.274	0.161	0.016	0.036	0.267	0.149	0.012	-0.110	-0.075	0.229	0.132	0.079	-0.066
F6	0.149	0.284	0.180	-0.191	0.274	1	0.326	-0.151	0.194	0.131	0.247	-0.293	-0.193	0.003	0.215	0.055	-0.024	0.099
F7	0.165	0.262	0.196	-0.332	0.161	0.326	1	-0.197	0.118	0.026	0.345	-0.178	-0.142	-0.006	0.240	0.186	-0.014	0.096
F8	0.052	-0.032	-0.082	0.110	0.016	-0.151	-0.197	1	0.062	0.106	-0.066	0.226	0.107	0.081	-0.016	0.030	0.029	-0.028
F9	0.041	0.105	0.236	0.119	0.036	0.194	0.118	0.062	1	0.146	0.045	0.103	0.055	0.264	0.095	-0.014	-0.010	0.216
F10	0.163	0.245	0.110	-0.034	0.267	0.131	0.026	0.106	0.146	1	0.129	-0.004	-0.024	0.019	0.105	0.219	0.008	0.085
F11	0.327	0.266	0.201	-0.350	0.149	0.247	0.345	-0.066	0.045	0.129	1	-0.075	-0.172	-0.063	0.174	0.246	-0.086	0.051
F12	-0.103	-0.096	0.116	0.172	0.012	-0.293	-0.178	0.226	0.103	-0.004	-0.075	1	0.190	0.126	-0.042	0.101	-0.015	0.155
F13	-0.287	-0.319	-0.184	0.329	-0.110	-0.193	-0.142	0.107	0.055	-0.024	-0.172	0.190	1	0.131	-0.061	-0.101	0.206	-0.130
F14	-0.222	-0.204	0.087	0.201	-0.075	0.003	-0.006	0.081	0.264	0.019	-0.063	0.126	0.131	1	0.122	-0.018	0.047	0.046
F15	0.203	0.133	0.073	-0.121	0.229	0.215	0.240	-0.016	0.095	0.105	0.174	-0.042	-0.061	0.122	1	0.157	-0.005	-0.074
F16	0.230	0.244	0.239	-0.103	0.132	0.055	0.186	0.030	-0.014	0.219	0.246	0.101	-0.101	-0.018	0.157	1	-0.091	0.037
F17	-0.253	-0.244	-0.321	0.280	0.079	-0.024	-0.014	0.029	-0.010	0.008	-0.086	-0.015	0.206	0.047	-0.005	-0.091	1	-0.097
F18	0.075	0.309	0.401	-0.010	-0.066	0.099	0.096	-0.028	0.216	0.085	0.051	0.155	-0.130	0.046	-0.074	0.037	-0.097	1

**TABLE 4**  
EIGENVALUES OF THE UNREDUCED INTERCORRELATION MATRIX

Root	Eigenvalue	% of Variance	Cumulative %
1	3.575	19.863	19.863
2	1.856	10.308	30.172
3	1.611	8.950	39.122
4	1.433	7.960	47.082
5	1.106	6.147	53.229
6	0.987	5.484	58.713
7	0.866	4.809	63.522
8	0.854	4.743	68.266
9	0.779	4.326	72.592
10	0.766	4.253	76.845
11	0.665	3.695	80.539
12	0.639	3.548	84.088
13	0.611	3.395	87.482
14	0.551	3.063	90.545
15	0.488	2.711	93.256
16	0.447	2.482	95.738
17	0.432	2.403	98.140
18	0.335	1.860	100.000

Table 6 gives the matrix of factor correlations for the five factors. This table shows relatively low correlations, indicating that these factors are essentially uncorrelated and independent of one another.

**TABLE 6**  
INTERCORRELATIONS OF FACTORS

FACTOR	1	2	3	4	5
1	1.000	0.086	0.242	-0.098	-0.056
2	0.086	1.000	0.262	0.015	0.081
3	0.242	0.262	1.000	-0.188	0.263
4	-0.098	0.015	-0.188	1.000	0.029
5	-0.056	0.081	0.263	0.029	1.000

Next, separate scales were formed, corresponding to each of the factors, and subjected to item analysis.

The item statistics in respect of Scale 1 are given in Table 7.

From Table 7 it is clear that the item-total correlations vary from 0,348 to 0,648. The mean of the scale is 39,568 and the variance is 105,657. The scale is therefore internally highly consistent. Three items were rejected, namely items 42, 19 and 30, because their item-total correlations were too low. The reliability of the scale according to Cronbach's coefficient alpha is 0,847. From an inspection of the test items, it would appear that Scale 1 can be identified as Self-Regard, for example: I take a positive attitude to myself (Item 58) versus I have a rather negative self-image (Item 18).

The item statistics in respect of Scale 2 are given in Table 8.

**TABLE 5**  
ROTATED FACTOR MATRIX (DIRECT OBLIMIN ROTATION)

	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	h <sup>2</sup> <sub>j</sub>
F2	0.545	0.426			0.338	0.689
F1	0.538	0.136	0.235		0.201	0.503
F14	-0.510	0.156	0.111			0.270
F13	-0.356	-0.154	-0.101	0.183		0.248
F17	-0.277	-0.273	-0.119	-0.103	0.219	0.243
F3		0.655	0.226		-0.100	0.548
F18		0.634	-0.153			0.375
F9	-0.321	0.405			0.210	0.310
F7			0.512	-0.308		0.411
F11	0.108		0.504			0.317
F4	-0.406		-0.432	0.126	0.143	0.471
F15	-0.119		0.420		0.166	0.229
F16			0.382	0.226		0.228
F12	-0.189	0.152		0.560		0.379
F6		0.175	0.163	-0.539	0.282	0.489
F8				0.363	0.212	0.189
F10					0.463	0.260
F5		-0.115	0.252		0.384	0.267

According to Table 8 the item-total correlations range from 0,310 to 0,497. The mean of the scale is 28,298 and the variance is 38,312. The scale is therefore internally highly consistent. Items 59, 1, 12 and 33 were rejected. The reliability coefficient of Scale 2 two was lower than Scale 1, displaying a Cronbach coefficient alpha of 0,686. This, however, may be expected because the number of items associated with this scale is substantially lower than that of the previous scale, with only 10 items present. Upon inspection of the contents of the items with high item-total correlations, it is apparent that Scale 2 may be interpreted as Risk Taking.

The item statistics in respect of Scale 3 are given in Table 9.

With reference to Table 9, the item-total correlations range from 0,306 to 0,478, with a scale mean of 34,527 and a variance of 54,278. Therefore the scale is internally highly consistent. Items 57, 32, 35, 41 and 48 were rejected. The reliability of the scale according to Cronbach's coefficient alpha is 0,702. Although this scale is considered reliable, again there are too few items. From an inspection of the remaining items, it would appear that Scale 3 can be identified as External Locus of Control.

The item statistics in respect of Scale 4 are given in Table 10.

From Table 10 it is evident that the item-total correlations vary from 0,232 to 0,415, with a scale mean of 27,3699 and a variance of 25,381. Therefore the scale is internally highly consistent. Items 8, 50, 52 and 6 were rejected due to low item-total correlations. The reliability of the scale according to Cronbach's coefficient alpha is 0,610. With reference to the items, Scale 4 can be interpreted as Internal Locus of Control.

The item statistics in respect of Scale 5 are given in Table 11.

Table 11 indicates that the item-total correlations range from 0,353 to 0,488, with a scale mean of 14,88 and a variance of

**TABLE 7**  
**ITEM STATISTICS IN RESPECT OF SCALE 1: SELF-REGARD**

DESCRIPTION OF ITEM	N	MEAN OF ITEM	STANDARD DEVIATION OF ITEM	CORRECTED ITEM-TOTAL CORRELATION	CRONBACH'S ALPHA IF ITEM DELETED
Q2 On the whole I am satisfied with myself	218	1.904	0.991	0.444	0.840
Q9 I feel that I have a number of good qualities	218	1.339	0.603	0.429	0.842
Q11 I am able to do things as well as most people	218	1.596	0.781	0.348	0.843
Q13 I feel that I do not have much to be proud of	218	1.390	0.785	0.544	0.837
Q18 I have a rather negative self-image	218	1.518	0.844	0.648	0.833
Q19 My priority at work is to get jobs out of the way as quickly and efficiently as possible	218	2.028	1.191	****	****
Q20 I am generally happy about who I am and what I can do	218	1.514	0.751	0.525	0.838
Q22 I sometimes feel that I am not a very capable person	218	1.784	1.013	0.461	0.839
Q27 I often feel that I am struggling with things I should be able to do easily	218	2.018	1.038	0.381	0.842
Q29 I believe I have a lot to offer	218	1.395	0.622	0.479	0.840
Q30 I am more attracted to tasks which will take a lot of sorting out than to straightforward situations in which I can clearly see my way forward	218	2.436	1.255	****	****
Q31 I rarely suffer from self-doubt	218	2.679	1.247	0.441	0.840
Q36 I often feel that other people are more capable than I am	218	2.179	1.140	0.481	0.838
Q40 I consider myself to be a successful person	218	1.743	0.791	0.492	0.839
Q42 In the case of a well-prepared candidate, there is rarely, if ever such a thing as an unfair test	218	2.541	1.199	****	****
Q45 I do not have a very high opinion of myself	218	1.803	1.031	0.369	0.843
Q47 I certainly feel useless at times	218	1.890	1.089	0.551	0.835
Q49 I feel that I am a person of worth, at least on an equal basis with others	218	1.495	0.720	0.411	0.842
Q56 All in all I am inclined to feel a failure	218	1.289	0.661	0.558	0.838
Q58 I take a positive attitude to myself	218	1.468	0.706	0.603	0.836

Note: Cronbach's Alpha = 0,847  
Number of items = 22

**TABLE 8**  
**ITEM STATISTICS IN RESPECT OF SCALE 2: RISK TAKING**

DESCRIPTION OF ITEM	N	MEAN OF ITEM	STANDARD DEVIATION OF ITEM	CORRECTED ITEM-TOTAL CORRELATION	CRONBACH'S ALPHA IF ITEM DELETED
Q1 When tackling problems at work, I tend to go for the simplest solution to achieve the objective rather than spend time on looking for the best possible answer	218	2.615	1.340	****	****
Q7 I am more attracted to jobs I can definitely cope with than ones which will take some getting used to	218	2.408	1.212	0.428	0.648
Q12 I tend to be more concerned about doing a good job than about my own personal ambition	218	3.638	1.153	****	****
Q15 I prefer situations in which I can make a quick decision and move on to the next job to ones which involve having to struggle with things I find hard to do	218	3.014	1.272	0.315	0.669
Q33 I am more stimulated by achieving operational targets than by my own personal successes along the way	218	3.051	1.173	****	****
Q37 I am more interested in making the most of a good system than in striving to come up with something new	218	2.748	1.224	0.426	0.648
Q43 I tend to stick to tasks I can definitely do well	218	2.959	1.212	0.497	0.635
Q51 I feel most comfortable in low risk situations	218	2.697	1.226	0.421	0.649
Q53 My main priority at work is to ensure that things do not go wrong	218	3.362	1.327	0.310	0.671

Note: Cronbach's Alpha = 0,686  
Number of items = 10

**TABLE 9**  
**ITEM STATISTICS IN RESPECT OF SCALE 3: EXTERNAL LOCUS OF CONTROL**

DESCRIPTION OF ITEM	N	MEAN OF ITEM	STANDARD DEVIATION OF ITEM	CORRECTED ITEM-TOTAL CORRELATION	CRONBACH'S ALPHA IF ITEM DELETED
Q3 Many unfortunate events in people's lives are due to bad luck	220	1.896	0.995	0.309	0.688
Q4 At times I think I am no good at all	220	2.073	1.269	0.324	0.686
Q10 Becoming a success is a matter of hard work, luck has little or nothing to do with it	220	2.109	1.134	0.306	0.688
Q14 It is not always wise to plan too far ahead because things may turn out to be a matter of good or bad fortune	220	1.959	1.128	0.454	0.670
Q16 Many times we might just well decide what to do by tossing a coin	220	1.568	0.956	0.341	0.685
Q21 Becoming the boss depends on being lucky enough to be in the right place at the right time	220	2.227	1.265	0.440	0.670
Q28 It is impossible for me to believe that chance or luck plays an important role in my life	220	2.823	1.286	0.357	0.681
Q32 Achieving good results depends on having the right opportunities	220	3.332	1.255	****	****
Q35 I tend to focus on longer term issues more than on immediate implications of my actions	220	2.850	1.178	****	****
Q41 I generally aim to come up with better solutions to problems than other people have, no matter how long it takes	220	3.482	1.129	****	****
Q44 Getting a good job depends mainly on being in the right place at the right time	220	2.573	1.220	0.391	0.677
Q46 People's lives are controlled to a large extent by accidental happenings	220	1.923	0.960	0.478	0.670
Q48 I tend to be more concerned about my enjoyment of the task than the outcome of my efforts	220	2.355	1.123	****	****
Q57 Getting people to do the right thing depends on leadership ability, luck has little to do with it	220	1.550	0.784	****	****
Q60 In my case, getting what I want has little or nothing to do with luck	220	1.809	0.922	0.381	0.681

Note: Cronbach's Alpha = 0,702  
 Number of items = 15

**TABLE 10**  
**ITEM STATISTICS IN RESPECT OF SCALE 4: INTERNAL LOCUS OF CONTROL**

DESCRIPTION OF ITEM	N	MEAN OF ITEM	STANDARD DEVIATION OF ITEM	CORRECTED ITEM-TOTAL CORRELATION	CRONBACH'S ALPHA IF ITEM DELETED
Q 6 In the long run, people get the respect they deserve in the world	219	3.854	1.148	****	****
Q 8 Performance ratings are frequently influenced by incidental events	219	2.753	1.279	****	****
Q 24 What happens to me is my own doing	219	4.059	1.014	0.410	0.555
Q 26 There is a direct connection between how hard I work and the credit I get	219	3.498	1.279	0.354	0.564
Q 34 People's misfortunes result from mistakes they make	219	3.073	1.217	0.315	0.576
Q 39 It is unfortunately often the case that people's real worth goes unrecognised no matter how hard they try	219	3.050	1.321	0.415	0.543
Q 50 Many times I seem to have little influence over the things that happen to me	219	3.772	1.076	****	****
Q 52 Most misfortunes are the result of lack of ability, ignorance, laziness or all three of these	219	3.311	1.349	****	****

Note: Cronbach's Alpha = 0,610  
 Number of items = 8

**TABLE 11**  
**ITEM STATISTICS IN RESPECT OF SCALE 5: CONSCIOUSNESS**

	DESCRIPTION OF ITEM	N	MEAN OF ITEM	STANDARD DEVIATION OF ITEM	CORRECTED ITEM-TOTAL CORRELATION	CRONBACH'S ALPHA IF ITEM DELETED
Q5	I am often diverted from the task at hand by the prospect of new challenges elsewhere	220	3.1	1.233	****	****
Q17	I am more inclined to seek out new challenges in less familiar areas than to concentrate my efforts on achieving immediate objectives	220	2.99	1.205	0.353	0.527
Q23	I attach more importance to doing a very thorough job than meeting current deadlines	220	3.1	1.186	****	****
Q25	I often find myself spending time trying to get one thing exactly right rather than getting on with the job	220	2.7	1.175	0.365	0.521
Q55	I tend to delve deeper into issues than is necessary to complete the job	220	2.99	1.310	0.488	0.443

Note: Cronbach's Alpha = 0,587  
 Number of items = 5

14,090. Two of the five items were rejected, namely items 5 and 23. The reliability of the scale according to Cronbach's coefficient alpha is 0,587. Upon inspection of Table 11, it is evident that the item-total correlations in respect of items 17, 25 and 55 are quite high. Hence Scale 5 may be interpreted as Conscientiousness.

A critical analysis of the WLA questionnaire involved an in-depth inspection of the development of the three scales in relation to the underlying theories. The SHO scale is based upon Rosenberg's (1965) Self-Esteem Scale, whereas the MED scale relates to Rotter's (1966) Scale of Locus of Control. The GEN scale as it stands is not based on any particular theory, but is related to the concept of resultant achievement motivation. The manual provides convergent validity studies which appear to validate the use of the scales in terms of their underlying theories. Two studies utilised self and colleague ratings as a basis for evaluating the correlation between each of the three scales and an adjectives and phrases checklist. The results demonstrated that many of the adjectives and phrases correlated highly with each of the scales. These studies are based on small samples with less than 100 managers and professionals.

A second study demonstrated the correlation between the MSQ and the Sixteen Personality Factor Questionnaire (16PF), for a sample of 64 managers and professionals. The findings showed that the SHO scale was strongly associated with Factors C (Affected by feelings - Emotionally stable), L (Trusting - Suspicious) and O (Unperturbed - Apprehensive). The MED scale was moderately associated with Factor Q1 (Conservative - Experimenting), whereas the GEN scale was moderately associated with Factors I (Tough-minded - Tender-minded) and N (Forthright - Shrewd).

A similar study was conducted between the MSQ and the OPQ Concept 5.2, with a sample of 66 managers and professionals. The GEN and SHO scale correlated on a number of similar factors, whereas all three of the scales correlated with Factors T6 (Change Orientation) and F9 (Achieving). It is of interest, however, to note that the manual reports on correlations between the scales, based on a sample of 726 managers and professionals. The findings of this study demonstrated that there was "very little correlation between the three WLA scales" (Tarleton, MSQ manual, 1997, p. 43). Yet the results of the correlation between the MSQ and the OPQ indicated that the scales measure similar dimensions and for that reason one would expect a correlation between the WLA scales.

The convergent validity studies indicate that the three scales are related to specific dimensions, some of which are shared such as Change Orientation and Achievement. The scales and the associated items clearly measure aspects of the underlying

theories in terms of Global Confidence, Locus of Control and Resultant Achievement Motivation. The rationale behind using these particular theories is explained in the manual. However, the use of the three scales SHO, MED and GEN in relation to these theories is not made clear. There is very little evidence in the manual which shows exactly why the SHO scale is related to belief in self, or why the MED scale is related to belief in control over situations. Nor is there a clear indication of the reasoning why the GEN scale should be interpreted as striving for personal success or avoiding operational failure.

## DISCUSSION

The principal objective of this study was to explore the recommendations regarding the use of the MSQ in its current form from a critical-reflective stance, which included a comment in terms of its psychometric properties. Both sections of the MSQ were evaluated. The ipsative section of the test, however, could not be analysed using standard analytic procedures based on the inherent nature of the items. The normative section was subjected to a principal factor analysis which yielded promising results. Both sections were subjected to a critical analysis.

### *The ipsative section of the MSQ (Work Style Preferences Questionnaire)*

The inter-item correlation matrix of the WSP questionnaire revealed numerous statistically significant negative correlations (25%), along with a large number of insignificant correlations (47%). This supports the findings of Clemans (1966) who indicated that at least two thirds of these correlations will be negative or very low. Based on these findings, it is evident that the majority of correlations are close to zero, which signifies that the items do not correlate. This, however, should not be the case since a number of items measure the same constructs. From inspection of Table 1, it is clear, for example, that the ACH scale is paired three times with the SYS and PER scales, and four times with the STR scale. Furthermore, the manual indicates that the seven scales are grouped according to three broad dimensions, namely, Achievement (ACH and IND), Structure (STR and AFF) and Power (SYS, PEO and PER). Therefore logic would indicate that there would be a number of significant positive correlations between the items. However, this is not the case with regard to the findings of the intercorrelation matrix of the WSP questionnaire. Hence the notion that correlations of ipsative data are spurious is supported (Closs, 1996; Cornwell & Dunlap, 1994; Johnson et al., 1988). In other words, the inter-item correlation matrix of ipsative data cannot be meaningfully interpreted.

Ipsative measures are modelled differently to normative measures in that an ipsative test is effectively a variant of the paired-comparisons technique (Johnson et al., 1988), in which

case Thurstone's Law of Comparative Judgement (1927) could be applied. The Law of Comparative Judgement would result in scaling a point estimate for each person on some notional scale. However, in the case of the MSQ, Table 1 illustrates that each scale was not equally paired with every other scale and so the Law of Comparative Judgement could not be applied appropriately (Torgerson, 1958). This, however, may have been expected, since the manual states that the seven scales are grouped according to three broad dimensions, those being Achievement, Structure and Power. As a result, those scales relating to one of the three dimensions would not be paired with each another, which is evident from Table 1. For example, the ACH scale is not paired with the IND scale, and the STR scale is not paired with the AFF scale. However, the concern here is the methodology used to categorise the seven scales according to the three broad dimensions.

The manual reports on the way in which the WSP questionnaire was developed and in this way serve as a basis for confirming the theoretical validity of the WSP scales. The results of a small sample (N=300) indicated that there were more than three factors to begin with. Consequently, a seven scale framework was postulated. However, there is no clarification of the exact format of the original 96 item questionnaire. Therefore one is left to question whether it was an ipsative or normative design, since it would be unreasonable to assume either. Consistent evidence has proven that ipsative data cannot be validly factor analysed due to inherent scale interdependence (Closs, 1996; Cornwell & Dunlap, 1994; Johnson et al., 1988). However, if the original version was in fact ipsative, like its successor, the seven scale structure of the WSP questionnaire is likely to be an inaccurate reflection of the underlying factors.

Inspection of Table 1 further indicates that certain scales are not paired with others, for any apparent reason. For example, the ACH scale is not paired with the PEO scale, and the IND scale is not paired with the SYS scale. This is unexpected, since these scales do not relate similarly to the three original dimensions. The MSQ manual provides no justification for why these scales were not paired with one another. The manual does, however, provide examples of why certain scales were paired with others. For instance, preference for high risk (SYS scale) is presented in the questionnaire opposite a preference for low risk (STR scale) or moderate risk (ACH scale). Yet, this statement implies that these scales are in some way related, as they all measure the underlying dimension of risk. Bearing in mind that correlations amongst ipsative scales are uninterpretable, it is important, however, to consider the findings presented in the manual with reference to the scale inter-correlation matrix (Tarleton, 1997, p. 43). The manual reports on some of the correlations in this table. For example: "Independence is strongly negatively associated with both Affiliation (-0.53) and People Power (-0.58)" (Tarleton, MSQ manual, 1997, p. 43). Yet, the manual does not report on the insignificant correlation between SYS and ACH, which is the second poorest (-0.03) in the table. This, in essence, would contradict the original statement that these scales are correlated and both measure a similar dimension.

The manual gives very little justification for the specific grouping of scales as indicated in Table 1. The explanation provided for pairing the SYS, STR and ACH scales is based on a common dimension, which further implies that these scales were paired as bipolar opposites. However on the answer sheet respondents are told that they are not, and upon inspection of Table 1 along with the items, it is evident that the scales are not paired as opposites. This makes the interpretation of the paired comparisons especially confusing for the user, and hence one would question why certain scales are paired with some and not with others. As a result there is a lack of *consistency* within the manual regarding the interpretation of the groupings of certain scales. Furthermore there is lack of clarification concerning the development of the original WSP questionnaire, from a three scale framework to seven scale framework.

Based on the results of a critical analysis of the WSP questionnaire there is evidence that the developer of the MSQ treats scores as if they are normative despite acknowledging their ipsative nature. For example, the manual states that the WSP questionnaire is ipsative because it is concerned with what is most important to individuals in the work situation and in this way, identifies the area of work in which they are most likely to feel that they are where they want to be. The developer argues that it would be futile to compare raw scores because a low raw score on one scale does not necessarily mean that individuals do not prefer something more than something else, but rather they prefer it *relative* to something else. The only way to evaluate whether an individual prefers something in relation to something else requires comparing the individual's preferences with a standardised sample (Tarleton, MSQ manual, 1997).

This explanation implies that the developer acknowledges the inherent qualities of the ipsative scales. However, the comparison made with a standardised sample, amounts to treating these measures normatively, which contradicts the very nature of ipsative test theory. These conversions will not reflect relative preferences because as soon as one transforms a raw score to a standardised score, the assumption is made that the original score was an absolute measure. Consequently, the ipsative format of the WSP questionnaire appears to be less obvious. Furthermore, the nature of the forced-choice format of the WSP items creates the impression that there is a continuum, which is typical of normative scales. By combining two dissimilar methodologies in the WSP questionnaire, the developer makes interpretation difficult and confusing, and as a result the user may interpret this section as a normative design.

The findings of this research strongly suggest that the format of the WSP questionnaire should be changed. A possible recommendation would be to change the ipsative item format to a normative format. For instance, if a scale of work preference was created, it would be more sensible to include separate scales for each system, rather than measuring both on an artificial continuum. Beyond the psychometric justification for measuring functional preferences independently, there would be practical benefits to this approach. It would yield more specific information about respondents' preferences for each work situation. In this case, the instrument may be used as a basis for making inter-personal comparisons, such as in the context of selection. On the other hand, if the test was to be used ipsatively, it must be utilised appropriately and in accordance with the principles of ipsative test design. In other words, the raw scores of the WSP questionnaire should not be converted to sten scores. In this format, the test may, for example, be used as a career development tool which is restricted to intra-individual comparisons, that is comparisons made within a person and not between individuals.

#### *The normative section of the MSQ (Work Life Attitudes Questionnaire)*

The principal factor analysis of the WLA questionnaire yielded five factors namely, Self-regard (Scale 1), Risk Taking (Scale 2), External Locus of Control (Scale 3), Internal Locus of Control (Scale 4) and Conscientiousness (Scale 5). The first three scales yielded acceptable reliabilities. Scale 1 is highly reliable with a coefficient alpha of 0,847. Scales 2 and 3 reflected rather depressed reliabilities compared to Scale 1, due to the number of items within the scales. It must, however, be noted that the reliabilities are reasonable for so few items, and this indicates that the items are of a high standard. On the other hand, the remaining scales display the lowest reliabilities, along with the fewest items. These scales are regarded as poor scales, and are not worth using. However, the remaining scales are shown to be reliable measures of Self-Regard, Risk Taking and External Locus of Control. In order to improve the reliabilities of Scales 2 and 3 for future use, the number of items associated with each scale need to be increased.

In terms of the critical analysis of the WLA questionnaire, the postulated items were evaluated in terms of the three predominant scales. In accordance with the WLA classifications set out in the manual, it would appear that Self-Regard is related to the SHO scale, which is designed to assess global confidence. On the other hand, Risk Taking would be linked to GEN scale in relation to operational focus. External Locus of Control is associated with the MED scale which is defined in terms of Rotter's (1966) scale. However, it is also possible that these factors relate to all of the WSP scales, and are not necessarily independent of one another as the findings in the manual specifies. For instance, Risk Taking can be related to Short-Term Striving in that people who prefer low risk situations may be "slow starting and careful to consider the nature of the task", whereas those who are inclined to take risks are "confident and keen to 'have a go'" (Tarleton, MSQ manual, 1997, p. 9). Similarly, Risk Taking can relate to the MED scale, where those who prefer less risky situations are "cautious" and keen to ensure they are on the right track as opposed to those who take risks and can be regarded as goal-focused and "less-likely to consider alternatives" (Tarleton, MSQ manual, 1997, p. 9). It is evident then that these factors can be interpreted differently and may not necessarily be independent of each of the WLA scales.

The WLA scales may therefore be regarded as 'inferred categories', in that the manual does not provide a theoretical rationale for linking these specific scales to the existing theories. In the same way, the present author can assume inferred categories for Scales 1, 2 and 3. Critical theory seeks to understand the connection between inference and results, based on the Ladder of Inference articulated by Argyris (1990). The premise is that we are so skilled in our thinking that we jump up the ladder from selecting data (inference) to drawing conclusions (results), without realising it. As a result, our conclusions feel so obvious that we see no need to retrace our steps. Thus the basic contention is that there is too much distance between inference and results, and accordingly the reasoning behind an inference needs to be made more explicit. The present author does not oppose the inferences made in terms of the WLA scales and their theoretical underpinnings; however, a critical approach would seek to understand how meanings were given to these scales in light of these theories.

The present study has certain limitations. The sample size was relatively small, and as a result it is not possible to generalise findings of the factor analysis to the economic sector. Furthermore, the sample was biased towards the banking and information technology sectors. It is recognised that factor analytic studies vary from country to country and so similar studies in different countries may reflect different findings. Recommendations for future research may include a similar study on a broader, more diverse sample of individuals across different sectors. Included in this may be the development of a parallel normative version of the MSQ WSP questionnaire. This, however, will depend upon the structure of the original version of the WSP questionnaire, in order to ascertain whether in fact the seven scale framework is a valid reflection of the original factors.

## CONCLUSION

The factor analysis of the normative section of the instrument yielded promising results, seeing that three factors were extracted with reliabilities ranging from 0,686 to 0,847. However, some uncertainty exists in that there seems to be a lack of an apparent connection between the postulated factors within this study, and the original scales of the WLA questionnaire. The manual does not provide detailed reasoning for linking the three scales with particular underlying theories. Even though the postulations on the part of the developer could be theoretically justified, it would assist the user if these inferences were made explicit.

The findings of the ipsative version of the MSQ, however, appear to be somewhat contentious. The results have demonstrated that the ipsative section of the MSQ may be utilised in ways which are not consistent with ipsative test theory.

The promotional claims of the questionnaire imply that the instrument may be used to make both intra- and inter-individual comparisons. Inter-individual comparisons would be made if the instrument was used to distinguish between good and bad performers, or to choose between equally competent individuals in a selection situation. The partially-ipsative nature of the MSQ, however, disallows meaningful comparisons among people and according to established psychometric principles, should be avoided altogether. On the other hand, intra-individual comparisons would occur if the MSQ was used for career planning, placing of new recruits or the relocation of staff after a restructuring. These sorts of evaluations would be justified in the case of an appropriately designed ipsative measure. However, the conversion of the WSP raw scores to sten scores, referenced against a norm group, makes ipsative interpretation problematic as it is confounded by inter-individual comparisons. In this way, the information provided by the manual seems to suggest that the instrument may be used to make comparisons between people. However, this advice would be *inconsistent* given the underlying design of the instrument. Furthermore, the manual recognises the use of an ipsative-style format, yet in no way clarifies how to utilise the MSQ ipsatively, and not normatively. In this respect, gaps exist in the recommendations that will enable the practitioner to use the instrument consistently with its design.

There appears to be widespread agreement that ipsative measures should not be used in the same way as normative measures. The use of ipsative tests for making inter-individual comparisons is not only lacking in validity, but also "positively harmful and should be discontinued" (Closs, 1996, p. 46). The explanation provided in the MSQ manual for using ipsative scores normatively, tends to obscure the ipsativity of the WSP questionnaire. This may create the impression that the test is normative in nature and can be used accordingly. As a result the claims for using the MSQ to make inter-personal comparisons may appear to be justified to users who are not sensitive to the particular nature of ipsative instruments. However, this would amount to using the instrument inconsistently with its ipsative design. This statement does not suggest that the questionnaire is not useful, it does, however, imply that practitioners who intend to use the MSQ must be aware of the problems and implications associated with its current design.

There are a number of concerns evident in this study which needs to be taken into account by practitioners who wish to act consistently with psychometric principles. A lack of reflection prevents an actor from being aware of inconsistencies. Therefore, a corollary of this study was an attempt to promote critical-reflection amongst researchers and those who practice within the field of psychological testing. The intention behind such an approach is to make researchers aware of the extent to which their actions do or do not serve the principles to which they subscribe. Evaluating psychological instruments from a critical-reflective perspective provides a useful technique for determining whether there is consistency in the application of psychometric theory. This methodology enables internal reflection amongst researchers with the ultimate purpose of promoting self-knowledge. It proposed that such an approach could be followed in future evaluations of psychometric tests.

## ACKNOWLEDGEMENTS

I would like to dedicate this article to my lecturer, my mentor, my study leader, my colleague, but most of all, my dear friend, the late Dr Conrad Schmidt without whom I could not have embarked on such a challenging study. It was an honour and

privilege to work by his side, he taught me how to 'think', and think and then when all else fails... think some more... and for that I owe him a lifetime of appreciation and gratitude.

Special thanks must also be given to a truly brilliant man, Professor J.M. Schepers. Thank you for all your time, dedication and patience during the course of this study.

## REFERENCES

- Allport, G.W., Vernon, P.E., and Lindzey, G.A. (1951). *A study of values, revised manual of directions*. Boston: Houghton Mifflin. [In Clemans, W.V. (1966). An analytical and empirical examination of some properties of ipsative measures. *Psychometric Monographs*, 14, 1-54.]
- Anastasi, A & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, New Jersey: Prentice-Hall.
- Argyris, C. (1990). *Overcoming organizational defences: Facilitating organizational learning*. Needham, Massachusetts: Allyn & Bacon.
- Argyris, C. & Schön, D.A. (1974). *Theory in practice: Increasing professional effectiveness*. San Francisco: Jossey-Bass.
- Argyris, C., Putman, R. & McLain Smith, D. (1985). *Action science*. San Francisco: Jossey-Bass.
- Baron, H. (1996). Strengths and limitations of ipsative measurement. *Journal of Occupational and Organisational Psychology*, 69, 49-56.
- Bartram, D. (1996). The relationship between ipsatized and normative measures of personality. *Journal of Occupational and Organisational Psychology*, 69, 25-39.
- Clemans, W.V. (1966). An analytical and empirical examination of some properties of ipsative measures. *Psychometric Monographs*, 14, 1-54.
- Closs, S.J. (1996). On the factoring and interpretation of ipsative data. *Journal of Occupational Psychology*, 69, 1, p. 41-47.
- Cornwell, J.M. & Dunlap, W.P. (1994). On the questionable soundness of factoring ipsative data: a response to Saville & Willson (1991). *Journal of Occupational and Organisational Psychology*, 67, 89-100.
- Geuss, R. (1981). *The idea of critical theory*. Cambridge, England: Cambridge University Press.
- Gregory, R.J. (2000). *Psychological testing* (3rd ed.). Needham Heights, Massachusetts: Allyn & Bacon.
- Habermas, J. (1972). *Knowledge and human interests*. London: Heinemann.
- Hicks, L.E. (1970). "Some properties of ipsative, normative, and forced-choice normative measures". *Psychological Bulletin*, 74, 167-184.
- Johnson, C.E., Wood, R. & Blinkhorn, S.F. (1988). Spuriousness and spuriousness: the use of ipsative personality tests. *Journal of Occupational Psychology*, 61, 153-162.
- Kaiser, H.F. (1961). A note on Guttman's lower bound for the number of common factors. *British Journal of Statistical Psychology* 14(1), 1.
- Kerlinger, F. N., & Lee, H.B. (2000). *Foundations of behavioral research* (4th ed.). Australia, Wadsworth.
- Kemmis, S. (2001). Exploring the relevance of critical theory for action research: Emancipatory action research in the footsteps of Jurgen Habermas. [In P. Reason & H. Bradbury (Eds), *Handbook of action research: Participative inquiry and practice*. London: Sage.]
- Martinussen, M., Richardsen, A.M. & Varum, H.W. (2001). Validation of an ipsative personality measure (DISCUS). *Scandinavian Journal of Psychology*, 42, 411-416.
- Meade, A. W. (2004). An illustration of the psychometric properties of forced-choice ipsative data (FCID) as they pertain to a personnel selection context. *Journal of Occupational and Organisational Psychology*, 531-552.
- Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist*, 35, 1012-1027. [In Gregory, R.J. (2000). *Psychological testing* (3rd ed.). Needham Heights, Massachusetts: Allyn & Bacon.]
- Murphy, K.R. & Davidshofer, C.O. (1994). *Psychological testing: Principles and applications* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Pickworth, G.E. & Schoeman, W.J. (2000). The psychometric properties of the Learning Style Inventory and the Learning Style Questionnaire: two normative measures of learning styles. *South African Journal of Psychology*, 30, 44-52.
- Piedmont, R.L., McCrae, R., & Costa, P.T., Jr. (1992). An assessment of the Edwards Personal Preference Schedule from the perspective of the five-factor model. *Journal of Personality Assessment*, 58, 67-78. [In Gregory, R.J. (2000). *Psychological testing* (3rd ed.). Needham Heights, MA: Allyn & Bacon.]
- Ray, J.J. (1990). Acquiescence and problems with forced-choice scales. *Journal of Social Psychology*, 130, 397-399.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princetown, New Jersey: Princetown University Press. [In Tarleton, R. (1997). *The Motivational Styles Questionnaire Manual*. Psychological Corporation, UK: Harcourt Brace & Company.]
- Rotter, J.B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 80, No 1 (Whole No. 609). [In Tarleton, R. (1997). *The Motivational Styles Questionnaire Manual*. Psychological Corporation, UK: Harcourt Brace & Company.]
- Saville, P. & Willson, E. (1991). "The reliability and validity of normative and ipsative approaches in the measurement of personality". *Journal of Occupational Psychology*, 64, 219-239.
- Schepers, J.M. (2004). Overcoming the effects of differential skewness of test items in scale construction. *South African Journal of Industrial Psychology*, 30 (4), 27-43.
- Schön, D.A. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.
- Tarleton, R. (1997). *The Motivational Styles Questionnaire Manual*. Psychological Corporation, London: Harcourt Brace & Company.
- Tenopyr, M.L. (1988). "Artifactual reliability of forced-choice scales". *Journal of Applied Psychology*, 73, 749-751.
- Thurstone, L.L. (1927). A law of comparative judgement. *Psychological Review*, 34, 273-286. [In Torgerson, W.S. (1958). *Theory and methods of scaling*. New York: J. Wiley & Sons.]
- Toner, B. (1987). The impact of the agreement bias on the ranking of questionnaire response. *Journal of Social Psychology*, 127, 221-222. [In Ray, J.J. (1990). Acquiescence and problems with forced-choice scales. *Journal of Social Psychology*, 130, 397-399.]
- Torgerson, W.S. (1958). *Theory and methods of scaling*. New York: J. Wiley & Sons.
- West, S.G., Finch, J.F. & Curran, P.J. (1995). Structural equation models with nonnormal variables. Problems and remedies. [In R.H. Hoyle (Ed.). *Structural equation modelling: concepts, issues and applications* (pp. 56-75). Newbury Park, CA: Sage.]