

Burnout and work engagement of South African blue-collar workers: The development of a new scale

L. Brand-Labuschagne, K. Mostert, S. Rothmann
Jnr & J.C. Rothmann

ABSTRACT

Research in South Africa on work-related well-being (specifically burnout and work engagement) has focused mainly on white-collar workers. Although blue-collar workers form a major part of the South African work force, no valid and reliable instruments exist to measure burnout and work engagement of blue-collar workers. The objectives of this study were (1) to develop a work-related well-being scale that measures burnout and work engagement of blue-collar workers; (2) to test the newly developed items using Rasch analysis; and (3) to test the factorial validity and reliability of the new scale. A cross-sectional survey design was used in a convenience sample of blue-collar workers in different industries in South Africa ($N=2769$). Following scale development procedures, a scale was developed to measure burnout (exhaustion and cynicism) and work engagement (vigour and dedication). Using Rasch analyses, two items were eliminated, resulting in an 18-item instrument. Five items were retained to measure exhaustion, five items to measure cynicism, four items to measure vigour and four items to measure dedication. The results of the confirmatory factor analysis showed that a two-factor model with two higher-order factors – burnout (consisting of exhaustion and cynicism) and work engagement (consisting of vigour and dedication) – fit the data best. All the scales were reliable.

Dr L. Brand-Labuschagne is at the School of Human Resource Sciences, WorkWell Research Unit for Economic and Management Sciences, North-West University, Potchefstroom Campus and Afriforte, the commercial arm of the WorkWell Research Unit for Economic and Management Sciences, North-West University, Potchefstroom. Prof. K. Mostert is at the WorkWell Research Unit for Economic and Management Sciences, North-West University; and Mr S. Rothmann Jnr & Dr J.C. Rothmann are at Afriforte, the commercial arm of the WorkWell Research Unit for Economic and Management Sciences, North-West University. E-mail: Karina.Mostert@nwu.ac.za

Key words: work-related psychological well-being, burnout, work engagement, Rasch analysis, factorial validity, reliability, blue-collar workers

Introduction

World-wide, the well-being of individuals at work has become an important focus area for researchers as well as practitioners. Work-related well-being of employees is emphasised, because it is related to lower accident rates, turnover intention and absenteeism; positive attitudes; productivity; profitability and safety; commitment; and job satisfaction (Harter, Hayes & Schmidt 2002; Saks 2006; Schaufeli & Bakker 2004a; Siu, Phillips & Leung 2004; Sonnentag 2003; Sparks, Farager & Cooper 2001).

Over the last decade, the importance of work-related well-being (particularly burnout and work engagement) has been emphasised in South Africa, and the research has focused mainly on the well-being of white-collar employees. However, the blue-collar workforce is regarded as the backbone of the South African economy and forms a significant part of the total workforce. Blue-collar workers in South Africa are mainly unskilled or semi-skilled workers that perform work of a manual nature (Lee & Mohamed 2006) and are employed mostly in the mining, construction, electrical, transport and security industries. The literacy levels of blue-collar employees vary from semi-literate to illiterate, and the tasks that they perform are mostly of a manual nature (Lee & Mohamed 2006). Some evidence was found that when the psychological well-being of mineworkers is negatively influenced, it has an effect on safety behaviour and accident proneness (Paul & Maiti 2007). It was also confirmed that negative psychological well-being has an influence on the accident-proneness of general blue-collar workers (Kirschenbaum, Oigenblick & Goldberg 2000).

Blue-collar workers in different industries are constantly faced with harsh and dangerous working conditions. In the mining industry, for example, employees work with explosives; test geological formations; operate load-haul-dump machines, scraper winches and heavy-duty machines; and maintain mining machinery in conventional mines (Calitz 2004). Employees are also exposed to harsh working conditions that include mining underground at temperatures in excess of 28° Celsius, long working hours, sometimes unsafe working conditions, highly unionised environments and enormous pressure to perform. The consequences of high environmental heat loads can be expressed in terms of impaired work capacity, errors of judgement and the occurrence of heat disorders, especially heat stroke, which is often associated with severe and irreversible tissue damage and high mortality rates (Calitz 2004). With more than one hundred miners killed every year in the South African mining

industry, this industry has proved to have the highest rates of fatal occupational injuries (McGwin et al. 2002).

Psychological well-being was first introduced in South Africa in the Construction Regulations (RSA 2003), section 15(12)(a). This section was included in the Occupational Health and Safety Act (Act No. 85 of 1993). In this Act, psychological well-being is referred to as 'psychological fitness'. The purpose of this Act is to regulate a particularly hazardous industry and to create a legal framework to ensure higher levels of health and safety (Deacon & Kew 2006), especially for employees who function in extremely hazardous working conditions, such as working at heights, operating cranes, construction vehicles and mobile plants. However, research has indicated that other hazardous industries, such as the mining industry, could also benefit from this Act, because occupational hazards also threaten psychological well-being in this industry (although the Act is currently restricted only to employees in the construction industry). Evaluating the psychological work-related well-being of employees in this industry could ensure better safety behaviour (Paul & Maiti 2007).

It is important to measure the levels of work-related well-being of blue-collar employees in order to enhance productivity and safety at work and to adhere to legislative requirements. However, no valid and reliable instrument exists to measure work-related well-being, particularly burnout and work engagement, of blue-collar employees. Although the underlying constructs (burnout and work engagement) will probably be the same for white- and blue-collar workers, some of the items of existing questionnaires measuring these constructs could be difficult for blue-collar workers to understand. This could be due to the length and construction of these items, or the choice of wording, which could be difficult for certain language groups to understand. Lee and Mohamed (2006) confirmed that low literacy levels and language difficulties hinder the validity and reliability of information obtained from measuring blue-collar employees.

Research objectives

Based on the limitations mentioned, the objectives of this study were (1) to develop a work-related well-being scale that measures burnout and work engagement of South African blue-collar workers; (2) to test the newly developed items using Rasch analysis; and (3) to test the factorial validity and reliability of the new scale.

Literature review

Burnout and work engagement

Schaufeli and Bakker (2004a) argue that burnout and work engagement are the main indicators of employee psychological well-being and are the mediators in health-impairment and motivational processes. Schaufeli and Enzmann (1998: 36) define burnout as “a persistent, negative, work-related state of mind in ‘normal’ individuals that is primarily characterised by exhaustion, which is accompanied by distress, a sense of reduced effectiveness, decreased motivation, and the development of dysfunctional attitudes and behaviours at work”. Burnout consist of two core dimensions, namely *exhaustion* (referring to fatigue and individual stress that result from the depletion of emotional and psychical resources) and *cynicism* (reflecting an indifferent or negative attitude towards one’s work due to the inability to deal with job demands). Although a lack of professional efficacy (which covers the self-evaluation dimension and refers to feelings of incompetence and lack of achievement both socially and non-socially) was also identified as a burnout dimension, recent studies have excluded this dimension from the burnout measure for various reasons. Firstly, professional efficacy shows fewer significant relationships with other variables and is perceived as the weakest burnout dimension (Lee & Ashforth 1996; Schaufeli 2003). Secondly, research indicates that personal efficacy develops independently, whereas cynicism develops in response to exhaustion (Leiter 1993). Thirdly, some researchers regard personal efficacy as a personality trait rather than part of the burnout dimension or as a possible consequence of the core negative emotional experience of burnout (Cordes & Dougherty 1993; Shirom 1989). In South Africa, Rothmann and Pieterse (2007) also reported that a two-factor structure of burnout (consisting of exhaustion and cynicism) fitted their data best.

With the introduction of the positive psychology paradigm, burnout research started to include the positive aspect of work engagement. Schaufeli, Salanova, González-Romá and Bakker (2002) proposed that work engagement should be operationalised and measured in its own right. When exposed to work stressors, some individuals do not show signs of burnout, but instead find pleasure in dealing with these stressors. These individuals can be described as being engaged in their work. When individuals experience work as meaningful, it can lead to ‘eustress’ (a positive psychological response to a stressor as indicated by a positive psychological state), which in turn promotes engagement, even in demanding situations (Nelson & Simmons 2003).

Schaufeli et al. (2002a) defined work engagement as a positive, fulfilling, work-related state of mind that is characterised by vigour, dedication and absorption.

Work engagement is also characterised by the two core dimensions of vigour and dedication. *Vigour* refers to high levels of energy and mental resilience while working, the willingness to invest effort in the work, not becoming exhausted easily, and persistence even in the face of difficulties. *Dedication* refers to experiencing a sense of significance from the work, feeling enthusiastic and proud of the job done, and feeling inspired and challenged by it. Absorption was also regarded as a core dimension, and referred to a feeling of being totally and happily engrossed in the work and experiencing difficulty in detaching from it. The experiences of time passing quickly and forgetting everything around one are evidence of this dimension (Schaufeli & Bakker 2001). However, in recent development in engagement research, absorption is excluded as a latent dimension of engagement. It is argued that absorption is not part of the core concept of engagement (Schaufeli & Bakker 2001) and, although absorption is regarded as playing a relevant role in engagement, some researchers have indicated that it is less critical and have questioned the relevancy of including absorption in the definition and measurement of work engagement (González-Romá, Schaufeli, Bakker & Lloret 2006; Montgomery, Peeters, Schaufeli & Den Ouden 2003). In South Africa, moreover, the absorption scale was found not to be reliable (Naudé & Rothmann 2004a). Several South African studies have reported that engagement, as measured by the Utrecht Work Engagement Scale (UWES), can be conceptualised as a construct consisting of two factors, namely vigour and dedication (Coetzer & Rothmann 2007; Jackson, Rothmann & Van der Vijver 2006; Rothmann & Jorgensen 2007; Rothmann & Pieterse 2007).

Burnout and work engagement are not two opposite poles; they function relatively independently of each other, are moderately negatively related and share between one-quarter to one-third of their variance (Schaufeli & Salanova 2007; Schaufeli et al. 2002). Although these constructs are related, they are conceptually distinct, and is it possible to distinguish empirically between these two dimensions of well-being. Instead of a single undifferentiated, common employee well-being factor, previous studies have found a structure that differentiates work-related well-being into two components, namely burnout (consisting of exhaustion and cynicism) and work engagement (consisting of vigour and dedication) (Schaufeli, Taris & Van Rhenen 2008; Schaufeli et al. 2002a).

Theoretical frameworks for understanding work-related psychological fitness

The following theoretical frameworks can be used to explain the mechanisms of burnout and work engagement: the Job Demands-Resources (JD-R) model

(Demerouti, Bakker, Nachreiner & Schaufeli 2001), Effort-Recovery (E-R) theory (Meijman & Mulder 1998) and Conservation of Resources (COR) theory (Hobfoll 1998).

Job Demands-Resources (JD-R) model

Demerouti et al. (2001) developed the JD-R model and argued that everyday work has its own set of factors associated with work-related psychological well-being. These factors can be divided into two main categories, namely job demands and job resources, so that this model could be applied to various occupational settings. *Job demands* refer to physical, emotional and organisational aspects of work that require employees to invest physical and mental energy and in return are associated with physical and psychological costs (Demerouti et al. 2001). *Job resources* refers to those physical, emotional and organisational aspects of work that enable employees to achieve work goals, reduce the physical and psychological costs of high job demands and ensure the personal growth and development of employees (Demerouti et al. 2001). Although the main function of job resources is to curb the effect of job demands, they are also important in their own right (Bakker & Demerouti 2007). Job resources can be categorised in terms of job-specific resources (for example, variety of tasks, adequate job information), organisational resources (for example, the opportunity to advance in a career) and social resources (for example, support from colleagues and supervisors) (Demerouti et al. 2001).

One of the main assumptions of the JD-R model is that the relationship between job demands/job resources and psychological well-being (burnout and engagement) is driven by two very distinct processes. The first of these processes is hypothesised as a *health-impaired process* that links job demands with health problems via burnout. This implies that when high job demands are present, psychological and physical resources are overstretched, which may lead to negative job strain (burnout), turning into health problems (i.e. cardiovascular problems) and negative organisational outcomes (i.e. turnover) (Demerouti et al. 2001). The second is the *motivational process*, where the lack of resources prevents employees from effectively dealing with high job demands, which fosters mental withdrawal and disengagement, causing the depletion of performance (Bakker, Demerouti, De Boer & Schaufeli 2003; Bakker, Demerouti & Schaufeli 2003; Schaufeli & Bakker 2004a). Conversely, high job resources are associated with higher engagement levels, which in turn contribute to higher commitment levels (Xanthopoulou, Bakker, Demerouti & Schaufeli 2007).

Effort-Recovery (E-R) theory

Meijman and Mulder (1998) developed the E-R theory and suggested that recovery and unwinding play a central role in ensuring work-related psychological well-being and health. According to the E-R theory, the exposure to job demands requires effort, which is associated with the development of short-term psycho-physiological reactions (for example, increased heart rate and mood changes) or load reactions (for example, fatigue). These psycho-physiological reactions are adaptive and reversible, implying that when individuals are no longer confronted by demands, the psycho-physiological systems previously affected by these demands return to the pre-demand level, and recovery can occur (Meijman & Mulder 1998). As a result, fatigue and other effects of the stressful situation are reduced.

The exposure to high job demands will not necessarily have negative consequences as long as sufficient recovery takes place. If, however, employees are continuously exposed to high work demands and have insufficient recovery opportunities, they do not recover from demands. Still in a sub-optimal state, employees therefore need to invest additional effort to cope with high job demands, resulting in increasing intensity of the negative load reaction and placing higher demands on the recovery process (Geurts et al. 2005). In the long run, insufficient recovery might seriously jeopardise the health (for example, by causing psychosomatic health problems or sleep loss) and well-being of employees, particularly in terms of higher exhaustion levels (linked to burnout) and lower levels of work engagement (Åkerstedt 2006; Sluiter, Frings-Dresen, Van der Beek & Meijman 2001; Van Hooff et al. 2005). However, Sonnetag (2003) confirmed that when sufficient effort recovery takes place, higher levels of work engagement are evident. This implies that sufficient recovery will ensure that individuals will be better enabled to confront stressful situations, will have enough resources to become involved in their work and will be able to concentrate fully on the job and ignore irrelevant cues (Sonnetag 2003).

Conservation of Resources (COR) theory

The COR theory implies that an adverse work situation threatens or harms a person's resources, such as work-related psychological well-being, health and even functioning in other domains of life. When employees work long hours, vigour decreases and fatigue and tension increase. In order to restore his/her resources, the individual needs to invest additional resources, such as social support and additional time (Hobfoll 1998). The main assumption of the COR theory is that people strive to obtain, retain and protect their resources (Hobfoll 1998), assuming that stress will occur when resources are threatened or lost, or when the investment of resources exceeds the gain

of resources (Sonnentag 2001). In the light of this theory, resources are referred to as “objects, personal characteristics, and energies that are either themselves valued for survival, directly or indirectly, or that serve as a means to achieving these resources” (Hobfoll 1998: 45).

Negative outcomes – such as burnout, turnover intentions and health complaints – occur when valued resources are lost or threatened, or are insufficient to meet the demands (Taris, Schreurs & Van Iersel-Van Silfhout 2001). In research that applied the COR theory to burnout, it was found that job demands and a lack of job resources are potential sources of job stress and convert into physical and emotional exhaustion. Conversely, available resources overcome the need for defensive coping, enhance self-efficacy and thus counteract burnout (Hobfoll & Freedy 1993). On a more positive note, a recently conducted longitudinal study indicated that higher job resources contribute to employees feeling more engaged in their work, but also that engaged workers tend to recognise, activate or create resources more easily (Xanthopoulou et al. 2007).

Measuring burnout and work engagement

Maslach Burnout Inventory-General Survey (MBI-GS)

The Maslach Burnout Inventory (MBI) is the most widely used instrument to assess burnout, and is used in approximately 93% of the studies measuring this construct (Schaufeli & Enzmann 1998; Schaufeli, Leiter & Maslach 2009). Internationally, researchers that administered the MBI-GS in different industries and among different occupations report a three-factor model of burnout consisting of exhaustion, cynicism and professional efficacy (Bakker, Demerouti & Schaufeli 2002; Kitaoka-Higashiguchi et al. 2004; Leiter & Schaufeli 1996; Richardsen & Martinussen 2004; Schutte, Toppinen, Kalimo & Schaufeli 2000; Taris, Schreurs & Schaufeli 1999). Similarly, the majority of South African studies also report on a three-factor burnout structure (Coetzer & Rothmann 2007; Campbell & Rothmann 2005; Naudé & Rothmann 2004b; Storm & Rothmann 2003a). During the past few decades, numerous studies have confirmed a two-factor model of burnout, consisting of the exhaustion and cynicism factors. The theoretical ground for this two-factor structure is based on the fact that exhaustion and cynicism constitute the core of the burnout syndrome, and that professional efficacy is only loosely related to the burnout phenomenon (Lee & Ashforth 1996; Rothmann & Pieterse 2007; Schaufeli & Bakker 2004a; Schaufeli et al. 2002a).

Schaufeli, Leiter, Maslach and Jackson (1996) found satisfactory internal consistency ($\alpha > 0.70$) (Nunnally & Bernstein 1994) for the exhaustion subscale of the MBI-GS, with Cronbach's alphas ranging between 0.87 and 0.89, alphas for the cynicism subscale ranging between 0.76, and 0.84 and for the professional efficacy subscale ranging between 0.76 and 0.84. In South Africa, acceptable internal consistency was found for the original three-factor structure of burnout consisting of exhaustion (alphas ranging between 0.79 and 0.92), cynicism (alphas ranging between 0.70 and 0.78) and professional efficacy (alphas ranging between 0.66 and 0.81) (Buys & Rothmann 2010; Coetzee & Rothmann 2004; Jackson & Rothmann 2005; Rothmann & Barkhuizen 2008; Storm & Rothmann 2003a). Schaufeli and Bakker's (2004) two-factor model of burnout had acceptable internal consistencies for both the exhaustion subscale (alphas ranging between 0.82 and 0.90) and the cynicism subscale (alphas ranging between 0.72 and 0.80). Similarly, a South African study by Rothmann and Pieterse (2007) revealed good internal consistencies for the exhaustion ($\alpha = 0.76$) and cynicism ($\alpha = 0.75$) factors.

Utrecht Work Engagement Scale (UWES)

The authors of the UWES originally hypothesised that work engagement should be regarded as a three-factor structure consisting of vigour, dedication and absorption (Schaufeli et al. 2002a). Indeed, a three-factor structure was found in various studies across different countries (Schaufeli et al. 2002a; Schaufeli, Bakker & Salanova 2006; Schaufeli, Martinez, Pinto, Salanova & Bakker 2002). Although the three-factor structure was also confirmed in South Africa among academic staff of higher education institutions (Barkhuizen & Rothmann 2006), some studies reported that the internal consistency of the absorption subscale was not acceptable (Naudé & Rothmann 2004a). This finding is consistent with arguments that the core of engagement consists of vigour and dedication (see Schaufeli & Bakker 2001). Consequently, some studies excluded absorption as a dimension of work engagement (González-Romá et al. 2006; Montgomery et al. 2003). Based on the findings of engagement research in South Africa, a two-factor structure consisting of vigour and dedication was hypothesised and confirmed in various studies (Coetzer & Rothmann 2007; Jackson et al. 2006; Rothmann & Jorgensen 2007; Rothmann & Pieterse 2007).

The internal consistency of the three subscales of the UWES has been reported in various international studies with acceptable Cronbach's alpha coefficients for vigour ranging between 0.68 and 0.88, dedication ranging between 0.71 and 0.96, and absorption ranging between 0.73 and 0.98 (Demerouti, Bakker, De Jonge, Janssen & Schaufeli 2001; González-Romá et al. 2006; Hakanen, Bakker & Schaufeli

2006; Langelaan, Bakker, Van Doornen & Schaufeli 2006; Salanova & Schaufeli 2008; Schaufeli et al. 2002b; Schaufeli et al. 2006). In South Africa, acceptable internal consistency was found for the vigour subscale (alphas ranging between 0.70 and 0.78) and the dedication subscale (alphas ranging between 0.79 and 0.89) (Jackson et al. 2006; Pienaar & Willemse 2008; Rothmann & Joubert 2007; Storm & Rothmann 2003b). Although Storm and Rothmann (2003b) found evidence of internal consistency for the absorption subscale (0.78), some South African studies reported insufficient internal consistency (alphas ranging from 0.55 to 0.69) (Naudé & Rothmann 2004a; Pienaar & Willemse 2008).

The following hypotheses are proposed for this study:

- Hypothesis 1: A two-factor higher order model, consisting of burnout (exhaustion and cynicism) and engagement (vigour and dedication) will fit the data closely.
- Hypothesis 2: The measurement of the four well-being dimensions (exhaustion, cynicism, vigour and dedication) is internally consistent.

Research design

Phase 1: Questionnaire development

The procedure of scale development suggested by DeVellis (2003) was followed and included the conceptualisation of the construct, item generation and evaluation, item development and item refinement.

Conceptualisation of the constructs

Burnout is defined as a persistent, negative, work-related state of mind in 'normal' individuals that is primarily characterised by exhaustion, which is accompanied by distress, decreased motivation, and the development of dysfunctional attitudes and behaviours at work (Schaufeli & Enzmann 1998). Burnout consists of the following two dimensions:

- **Exhaustion:** fatigue and individual stress resulting from the depletion of emotional and psychical resources
- **Cynicism:** an indifferent or negative attitude towards one's work due to the inability to deal with job demands.

Work engagement has been defined as a positive, fulfilling, work-related state of mind that is characterised by vigour and dedication (Schaufeli et al. 2002a). Work engagement is characterised by the following two core dimensions:

- Vigour: high levels of energy and mental resilience while working, the willingness to invest effort in the work, not becoming exhausted easily, and persistence even in the face of difficulties.
- Dedication: experiencing a sense of significance from the work, by feeling enthusiastic and proud of the job done, and feeling inspired and challenged by it.

Item generation

During the next stage, items from a preliminary item pool were identified (60 items), tapping into existing research scales measuring burnout and work engagement (Demerouti, Bakker, Verdakou & Kantas 2002; Maslach & Jackson 1986; Pines & Aronson 1988; Schaufeli & Bakker 2004a; Schaufeli & Bakker 2004b; Schaufeli et al. 2002a; Schaufeli et al. 1996; Shirom 2003). With the target population in mind, the following criteria were used to evaluate each item (DeVellis 2003): (1) each item should reflect the definition of the construct/dimension it intends to measure; (2) each item should be clear and concise in terms of problematic wording; and (3) the appropriate grammatical structure and word choice for each item are important. During the item evaluation process, a panel of subject experts (i.e. researchers in the area of work-related psychological well-being) rated each item in the categories high, moderate or low, using the evaluation criteria mentioned. Where moderate to low ratings were indicated, these judges (subject experts) were asked to propose an alternative to the specific item. The ratings of these judges were then discussed, items were adapted where necessary, and items were identified for the process to be followed. During this process, 33 items were identified for use in the new scale based on the evaluation criteria.

Item development

During the item development phase, items were carefully scrutinised and further developed and adapted where necessary to ensure that they fitted the proposed definitions of the constructs/dimensions, and were written in English that was accessible, reader-friendly and understandable to lower-level workers. With respect to the scale development process, it has been proposed that five- to seven-point frequency-based category scales are the ideal number of categories to be used (Green

& Frantom 2002). Items that did not fit the particular seven-point frequency-based format scale were adapted. Responses vary from never (0), to almost never (1), infrequently (2), sometimes (3), quite frequently (4), regularly (5) and always (6). Items were categorised in four dimensions: exhaustion (nine items), cynicism (eight items), vigour (eight items) and dedication (eight items).

Item evaluation and refinement

Following the item development process, a panel of five researchers (all of whom were industrial psychologists) was asked to judge the items. The judges were provided with the definitions of each of the constructs, and the items had to be categorised according to the definitions of the different dimensions. The judges were also asked to evaluate the item clarity and to point out any uncertainties. The judges reported that all the items were correctly categorised, and made some suggestions regarding item wording. Based on these inputs, changes were made to refine the items. All the items were translated into the eleven official languages of South Africa by five accredited language experts following a multistage translation process (Shanahan, Anderson & Mkize 2001). The questionnaires in the ten languages other than English were then translated back to English by five other language experts. The translators were clearly informed of the type of person that would participate in the survey and that the translations had to be clear and reader-friendly. The ten back-translated questionnaires were then compared with the original English questionnaire. Where the back-translated version did not correspond well with the original version, the problems were discussed and the questionnaires were adapted until the best fit was reached. The questionnaire was then finalised in all eleven official languages of South Africa.

After the item development had been completed, the items were administered to a convenience pilot sample of mine workers ($n=265$) in South Africa. The majority of the sample comprised males (70.31%), and almost 60% (59.76%) of the respondents indicated an African language as their preferred language. Most of the respondents were African (61.72%); 25.39% were between the ages of 20 and 29, and only 5.08% were older than 50.

During this process of identifying poorly performing items, specific guidelines were followed from existing literature (DeVellis 2003; Curbow, Spratt, Ungaretti, McDonnell & Breckler 2006; Foxcroft & Roodt 2005). These guidelines included the evaluation of item mean scores, standard deviation, item variance, inter-item correlations and item-total correlations. In order to eliminate poor performing items, the following criteria were followed: (1) mean item scores near the centre of the range

(value of four) were more desirable; (2) standard deviations lower than 1.00 were required; (3) item variances with relatively high values were more desirable; (4) inter-item correlation was more desirable if items had a strong correlation with all the other items and a minimum significance level of 0.05 (items not correlating sufficiently were removed from the scale); (5) positive item-total correlations indicated that the item measured the construct it intended to measure; and (6) values near zero did not discriminate between high and low scores.

The total sample was analysed simultaneously in order to identify poor items. Items that did not meet these guidelines were eliminated from the final questionnaire. After completing the elimination process, 20 items were retained to measure the following dimensions: exhaustion (five items), cynicism (five items), vigour (five items) and dedication (five items). The main objective of the item evaluation study was to evaluate the performance of the remaining items of the newly developed scale, so that the items that performed poorly could be eliminated from the scale, and the items that performed more desirably could be retained for further validation. The elimination of items was based on the Rasch model analysis (Rasch 1960) and is discussed in the results section.

Phase 2: Factorial validity and reliability

Research approach

A cross-sectional survey design was utilised (Struwig & Stead 2001). With a cross-sectional design, data can be collected on more than one incident at a single point in time, and analysis can be done to detect patterns of relationships (Bryman & Bell 2003).

Research participants

This study was conducted on a convenience sample ($N=2769$) from different industries in South Africa. The participants consisted of mine workers, construction workers (including height workers, operators of mobile plants and construction vehicles, and crane operators), electrical supply height workers, security guards, train drivers and general municipal workers. The sample included male (82.90%) and female (17.10%) participants. Most of the participants were single (51.80%); 41.60% were married. The racial groups included in the study were mainly African (74.32%), but the sample also comprised White (18.78%), Coloured (4.41%) and Indian (2.35%) groups. The majority of the participants were isiZulu speaking (19.80%), but the study

also included the other official language groups in South Africa: English (17.70%), Afrikaans (17.40%), Sepedi (10.00%), Sesotho (9.20%), Setswana (9.00%), isiXhosa (8.70%), Siswati (2.60%), Xitsonga (2.30%), Tshivenda (1.50%) and isiNdebele (1.40%). Most of the participants were aged 20–29 (33.41%) and 30–39 (32.32%). A total of 63.95% of the respondents had some form of high-school education, but only 33.33% had a grade 12 qualification. The majority of the respondents worked as electrical height workers in the electrical industry (39.76%), and construction workers (height workers, operators of mobile plants and construction vehicles, crane operators) in the construction (28.17%) industry.

Research procedure

Permission was granted by the management of each organisation for conducting the research and using the data anonymously for research purposes. During a one-day training session, fieldworkers were trained in administering the instrument by means of facilitation. This method of administering a questionnaire requires that the fieldworker ask the question in one of the eleven languages and wait for the participant to respond. During the training session, fieldworkers were equipped with the necessary skills and research tools to successfully administer the questionnaire. Careful attention was given to ensuring that the fieldworkers were trained regarding the basic concepts of the questionnaire, how to make use of the provided assessment tools (including flashcards to serve as aids in answering the questionnaire), and not to lead the participants in answering the questions.

The flashcards consisted of the rating scale categories in the form of a volume indicator, which went from very small to very large, in order to assist the participants in understanding the frequency scale. Flashcards were available for all the language groups. The fieldworkers were provided with examples of questions that they could use to explain to the participants what would be expected from them when answering the questions. The fieldworkers utilised the flashcards by asking the questions, and participants then indicated on the flashcard which option best described themselves.

Before the questionnaire was administered, the workforce was informed of the purpose of the research, and gave their informed consent. Participants with lower levels of literacy were assisted in completing the questionnaire by means of facilitation. However, the participants with higher literacy levels completed the questionnaire in their mother tongue on their own and returned the completed questionnaire to the fieldworker. The participants had between ten and twenty minutes to complete the questionnaire.

Measuring instrument

Psychological fitness

The adapted instrument (after two items had been removed during the Rasch analysis, see Table 3) was utilised in this study. Exhaustion was measured using five items (for example, “I feel tired before I arrive at work”). Cynicism was measured using five items (for example, “I am uncertain whether my work is important”). Vigour was measured with four items (for example, “I feel energetic in my job”). Work devotion was also measured using four items (for example, “I am passionate about my job”). All items were measured on a seven-point frequency rating scale ranging from zero (‘never’) to six (‘always’). Higher scores on exhaustion and cynicism are an indication of burnout, while higher scores on vigour and work devotion are indicators of work engagement.

Statistical analysis

Rasch analysis

Rasch analysis was conducted utilising the WINSTEPS program (Linacre 2005). The Rasch model assumes that the relationship between a person’s ability and item response can be modelled as a probabilistic function, implying that if the ability level of a person increases on a specific latent trait, the probability of scoring higher on each item increases as well (Fox & Jones 1998). This implies that the whole continuum of the latent trait is evaluated through the items, and the item parameter estimates are invariant across groups differing in terms of their ability on the latent trait (Hagquist 2007).

Rasch analysis was performed on all the items to evaluate the validity and reliability of each dimension and to evaluate the rating scale categories of each dimension. The reliability of the rating scale is determined by both the item (item separation index and item reliability index) as well as the person (person separation index and person reliability index). Person and item separations refer to the distribution of the items or people over the continuum of the measured latent trait. Both person and item separation indexes should be at least 2.00 for an instrument to be regarded as useful (Fox & Jones 1998). Higher values on separation indicate greater distribution of items or people along the measured latent trait (Green & Frantom 2002).

Person separation reliability is comparable to the traditional internal consistency reliability in terms of Cronbach’s alpha coefficients, which estimate the true person variance. Rasch’s reliability estimates have the advantage that neither the sample

size nor sample specifics influence the reliability. This implies that these estimates measure a person's ability according to the responses on the specific test regardless of the sample to which they belong (Boone & Rogan 2005). Fox and Jones (1998) propose that the person separation index should be used as an alternative to person reliability, as it is bounded by 0 and 1 and is more useful for comparing reliability across different analyses of the same data.

Item reliability indicates how well difficulty levels of the item are distributed along the measured latent variable. This is calculated by dividing the true variance by the observed variance when utilising WINSTEPS (Fox & Jones 1998). The values for item separation reliability vary between 0 and 1 (Cervellione, Lee & Bonanno 2009). The higher the item reliability index, the better the chance of replicating the item placement in other samples (Bond & Fox 2007). The reliability of a measure is negatively influenced when the distributions of the items are too narrow and the standard errors are too big (Boone & Rogan 2005).

Fit statistics are utilised to evaluate the validity of each dimension through identifying individuals and items that function differently than expected, identifying problematic items as well as the persons whose responses are idiosyncratic (Boone & Rogan 2005). Chi-square values are used to determine how well the data fit the prescribed model. These chi-square values are provided in infit and outfit mean square statistics divided by their degree of freedom, with an expected value of +1 and ranges from 0 to positive infinity (Bond & Fox 2007). In order to evaluate the unidimensionality of the scale, item fit mean-square statistics (MNSQ) are utilised. These statistics indicate how well the item measures the intended underlying construct, and the ideal value is 1. Infit and outfit are used to measure the fit of the data (Cervellione et al. 2009). Infit statistics are less sensitive than outfit statistics when extreme responses are evident (Green & Frantom 2002).

Item fit refers to whether the items provide logical and useful information for all the participants, and thus whether the item would provoke the same answers for participants in another setting. The reasons for item misfit include the complexity of the item, confusion from participants, or the item not measuring the construct it is supposed to measure (Green & Frantom 2002). Person fit refers to the responses of individuals to items in a consistent manner. Misfit might be evident due to a variety of reasons such as being bored with the task, confusion occurring or the item evoking a different answer from the individual than was expected (Green & Frantom 2002). When items or persons underfit, it means that they cause noise or eradication performance and are not sufficiently predictable to make the Rasch model useful. This is detected when the fit statistics are higher than the cut-off point of 1.30. Overfit occurs when the items are not independent and provide the same information and no

new information on the measured variable. This is detected when the fit statistics are lower than the cut-off point of 0.70 (Bond & Fox 2007).

Evaluating the rating scale categories guides the researcher in deciding whether the categories are sufficient or whether some categories should be collapsed. A very basic way of examining the rating scale is by examining the category frequencies indicating how many respondents chose a particular rating category. In addition, average measures can give useful information about the rating categories. Average measures refer to the average ability estimate for all the respondents in the sample who chose that particular response category. These average measures are expected to increase in size as the measured variable increases (Bond & Fox 2007). Another way to investigate the rating scale categories is through the fit statistics. Linacre (1999) suggests that outfit statistics higher than 2.00 indicate more misinformation than information. These categories might need to be collapsed with broader categories.

Confirmatory factor analysis and reliability

Confirmatory factor analysis (CFA), as implemented by means of Mplus 6.1 (Muthén & Muthén 2007), was used to analyse the data. The input type was the covariance matrix. The robust maximum likelihood estimator was used to accommodate the lack of multivariate normality in the item distribution (Muthén & Muthén 2007). A seven-point frequency scale was used, and the variables were thus analysed as continuous variables. Individual items were used as indicator variables. Goodness-of-fit was evaluated using the traditional χ^2 statistic, comparative fit index (CFI), Tucker–Lewis index (TLI), root mean-square error of approximation (RMSEA) and standardised root mean square residual (SRMR). Although there is little consensus on cut-off values for adequate fit (Lance, Butts & Michels 2006), conventional guidelines were followed whereby fit is considered adequate if CFI and TLI values are larger than 0.90 (Byrne 2010), RMSEA is smaller than 0.05 (MacCallum, Browne & Sugawara 1996) and SRMR is smaller than 0.05 (Hu & Bentler 1999). The Akaike information criterion (AIC) and sample adjusted Bayesian information criterion (BIC) were used to compare the fit of competing models. The reliability of the four subscales was evaluated using Cronbach's alpha coefficients.

In addition, descriptive statistics (means and standard deviations) were used to describe the data, and product-moment correlations were used to determine relationships between the variables.

Results

Rasch analysis

Table 1 indicates the internal consistency of the measurement in terms of the item separation index and reliability, person separation index and reliability, person reliability in terms of Cronbach's alpha coefficients, the average measure of each dimension per person and item, as well as the infit and outfit statistics for each dimension in terms of the person and item infit and outfit statistics for each dimension.

Table 1: Person and item summary statistics

Dimension	Average measure (SD)	Infit (SD)	Outfit (SD)	Separation	Reliability	α
Exhaustion						
Person	-1.05 (1.03)	0.98 (0.76)	0.98 (0.85)	1.95	0.79	0.81
Item	0.00 (0.51)	1.00 (0.07)	0.98 (0.06)	24.74	1.00	-
Cynicism						
Person	-0.50 (0.57)	0.99 (0.80)	1.00 (0.80)	1.01	0.50	0.77
Item	0.00 (0.13)	1.00 (0.14)	1.00 (0.12)	9.69	0.99	-
Vigour						
Person	1.26 (0.95)	0.99 (0.81)	1.00 (0.97)	1.49	0.69	0.71
Item	0.00 (0.43)	1.08 (0.30)	1.01 (0.23)	20.14	1.00	-
Dedication						
Person	1.09 (0.83)	1.02 (0.79)	0.97 (0.75)	1.18	0.58	0.81
Item	0.00 (0.29)	1.01 (0.23)	0.97 (0.21)	14.63	1.00	-

Table 1 shows acceptable item reliability for all four dimensions, indicating that these items differentiated well among the measured variable (equal to or greater than 0.80). The item separations for all the dimensions were sufficient compared to the guideline of at least 2.00, as indicated by Bond and Fox (2007). The person separation indexes for all the dimensions were somewhat lower than the proposed guideline (>2.00). Cronbach's alpha coefficients for all dimensions were acceptable. The mean item fit and person fit were acceptable. It is evident that, on average, the responses do not underfit or overfit.

Table 2: Rating scale categories for all four dimensions

Dimension	Rating category	Category frequency	Average measure	Infit	Outfit
Exhaustion	0	3025	-2.15	1.00	1.00
	1	1871	-1.61	0.89	0.87
	2	1367	-0.92	0.89	0.83
	3	4177	-0.54	1.04	1.04
	4	648	0.25	0.84	0.84
	5	423	0.75	1.05	0.97
	6	204	1.27	1.29	1.39
Cynicism	0	4748	-0.83	1.05	1.03
	1	2550	-0.69	0.66	0.95
	2	1244	-0.34	0.83	0.73
	3	2277	-0.17	0.84	0.78
	4	596	-0.01	0.71	0.62
	5	580	0.09	0.85	0.84
	6	980	0.02	1.28	1.42
Vigour	0	191	-0.57	1.65	2.06
	1	281	-0.31	1.17	1.23
	2	504	0.01	0.95	0.96
	3	2124	0.49	1.04	1.04
	4	1425	0.81	0.76	0.62
	5	2940	1.42	0.87	0.86
	6	4050	2.05	1.04	1.03
Dedication	0	304	-0.28	1.55	1.55
	1	306	-0.24	1.01	1.01
	2	427	0.04	0.89	0.82
	3	1548	0.40	0.98	0.95
	4	1333	0.64	0.83	0.65
	5	2920	1.20	0.86	0.95
	6	4792	1.63	1.09	1.04

Table 2 shows the functionality of the rating scale categories of all four dimensions. As can be seen in Table 2, the seven-point frequency-based scale functioned satisfactorily for all four dimensions. It is some concern that certain categories were underutilised. This is evident for the last three categories (4, 5 and 6) of exhaustion and the first three categories (0, 1 and 2) of vigour and dedication. None of the categories, in any of the dimensions, showed outfit statistics higher than 2.00.

Table 3 indicates the item fit statistics for all four dimensions in terms of the measurement intensity of each item and the infit mean square and outfit mean square for each item.

Table 3: Item fit statistics for the four dimensions

Dimension	Item	Measure (θ)	Infit mean square	Outfit mean square
Exhaustion	EX1	-0.06	0.99	0.99
	EX2	-0.26	0.97	0.95
	EX3	0.91	1.13	1.10
	EX4	-0.65	0.98	0.96
	EX5	0.06	0.93	0.91
Cynicism	CY1	-0.03	0.88	0.85
	CY2	-0.14	0.85	0.86
	CY3	-0.11	1.19	1.14
	CY4	0.07	1.14	1.04
	CY5	0.21	0.95	1.10
Vigour	VI1	0.36	0.88	0.88
	VI2	0.03	0.72	0.69
	VI3	0.11	1.00	0.92
	VI4	0.32	1.22	1.22
	VI5	-0.82	1.57	1.31
Dedication	DE1	0.14	0.94	1.00
	DE2	-0.25	0.89	0.83
	DE3	-0.21	0.87	0.82
	DE4	-0.18	0.88	0.83
	DE5	0.50	1.48	1.37

The results of the exhaustion items indicate that EX3 has the highest measurement intensity ($\theta=0.91$), and EX4 has the lowest measurement intensity ($\theta=-0.65$). The infit and outfit statistics for all five exhaustion items are satisfactory. CY5 shows the highest measurement intensity ($\theta=0.21$) and CY3 the lowest intensity ($\theta=-0.11$). The infit and outfit statistics for the cynicism items are also acceptable. With respect to the infit and outfit statistics, all the dimensions proved to be acceptable.

The vigour item with the highest measurement intensity was VI1 ($\theta=0.36$), and the item with the lowest measurement intensity was VI5 ($\theta=-0.82$). With respect to the infit and outfit statistics, all the items showed good fit except for VI5 (infit=1.57, outfit=1.31). This item was regarded as an under-fitting item and did not provide

information consistent with the other vigour items. The dedication item with the highest measurement intensity was DE5 ($\theta=0.50$), and the item with the lowest measurement intensity was DE2 ($\theta=-0.25$). The infit and outfit of the dedication items were satisfactory for all the items except for DE5 (infit=1.48, outfit=1.37). This item also underfits and does not provide information consistent with the other dedication items.

Because of the weaker fit for VI5 and DE5 respectively, it was decided to remove these items from the scale because they do not demonstrate homogeneity. Table 4 shows the renewed person and item summary statistics for the dimensions of vigour and dedication when these items are removed.

Table 4: Person and item summary statistic with removed items

Dimension	Average measure (SD)	Infit (SD)	Outfit (SD)	Separation	Reliability	α
Vigour						
Person	1.14 (1.03)	0.98 (0.79)	0.97 (0.78)	1.49	0.69	0.72
Item	0.00 (0.16)	1.00 (0.20)	0.97 (0.20)	7.44	0.98	
Dedication						
Person	1.35 (1.03)	0.98 (0.91)	0.98 (0.91)	1.25	0.61	0.83
Item	0.00 (0.21)	1.03 (0.03)	0.98 (0.05)	8.94	0.99	-

When the two weaker items (VI5 and DE5) are removed, the item separation and reliability remain acceptable. The person separation index remained unchanged for vigour and improved for dedication. The Cronbach's alphas of both dimensions improved significantly ($\alpha>0.70$).

Table 5 illustrates that when VI5 and DE5 are removed, the infit and outfit statistics of the remaining items improved.

Confirmatory factor analysis

Next, the factorial validity was investigated, using the remaining 18 items. First, the hypothesised model (Model 1) was tested. This model consisted of two higher-order factors, namely burnout and work engagement. Two first-order factors loaded on each of these higher order factors – exhaustion and cynicism loaded on burnout, and vigour and dedication loaded on work engagement. Five items loaded on exhaustion

Table 5: Item fit after removal of items

Dimension	Item	Measure (θ)	Infit mean square	Outfit mean square
Vigour	VI1	0.18	0.90	0.90
	VI2	0.20	0.75	0.73
	VI3	0.11	1.07	0.98
	VI4	0.13	1.29	1.28
Dedication	DE1	0.36	1.07	1.06
	DE2	-0.17	1.00	0.93
	DE3	-0.17	1.00	0.94
	DE4	-0.08	1.05	0.98

and five items on cynicism. Four items loaded on vigour and four items on dedication. As can be seen in Table 6, the CFI and TLI values are larger than 0.90, RMSEA is 0.05 and the SRMR is smaller than 0.05. Model 1 therefore shows good model fit and can be considered a plausible explanation for the observed inter-item covariance matrix.

In order to test whether alternative models do not provide more plausible explanations for the observed inter-item covariance matrix, competing models were tested. Because several South African studies found superior fit for a one-factor engagement model (Naudé & Rothmann 2004a; Olivier & Rothmann 2008), Model 2 proposed two higher-order factors, burnout (consisting of exhaustion and cynicism) and work engagement (where all the work engagement items load on this factor). Model 3 consist of a two-factor model (with burnout as one factor and exhaustion and cynicism items loading on this factor, and engagement as the other factor and vigour and dedication items loading on this factor). Finally, Model 4 was tested as a one-factor well-being construct, with exhaustion, cynicism, vigour and dedication items loading on the factor. Table 6 shows the results of these analyses.

The hypothesised model (M1) fitted the data significantly better than Model 2 ($\Delta\chi^2=150.02$; $\Delta df=2$; $p < 0.05$), Model 3 ($\Delta\chi^2=825.09$; $\Delta df=4$; $p < 0.05$) or Model 4 ($\Delta\chi^2=2669.16$; $\Delta df=5$; $p < 0.05$). In accordance, the AIC and BIC values for Model 1 are the lowest. This provides support for Hypothesis 1.

Next, completely standardised factor loadings, significance, variance explained and completely standardised residual variances of the model tested are reported in Table 7.

Table 7 shows that not all communalities were sufficiently high (> 0.50). However, removing them will result in an inadequate number of items to measure each

Table 6: Fit statistics for the hypothesised and comparison models

Model	χ^2	df	p	CFI	TLI	RMSEA	SRMR	AIC	BIC
Model 1	759.52	114	0.00	0.95	0.94	0.05	0.04	157907.23	158061.17
Model 2	909.54	116	0.00	0.93	0.92	0.05	0.04	158114.65	158263.09
Model 3	1584.61	118	0.00	0.88	0.86	0.07	0.05	159028.54	159336.71
Model 4	3428.68	119	0.00	0.72	0.68	0.10	0.08	161625.48	161927.72

χ^2 =chi-square; df=degrees of freedom; p=statistical significance; CFI=comparative fit index; TLI=Tucker-Lewis index; RMSEA=root mean square error of approximation; SRMR=standardised root mean square residual; AIC=Akaike information criterion; BIC=Bayesian information criterion

Table 7: Factor loadings, significance, variance explained and residual variances

Model	Standardised loading	Loading significance	R ²	Residual variances	Residual variances significance
Exhaustion					
Item 1	0.68	0.00	0.46	0.54	0.00
Item 2	0.65	0.00	0.42	0.58	0.00
Item 3	0.63	0.00	0.39	0.61	0.00
Item 4	0.66	0.00	0.43	0.57	0.00
Item 5	0.74	0.00	0.55	0.45	0.00
Cynicism					
Item 1	0.73	0.00	0.54	0.47	0.00
Item 2	0.75	0.00	0.57	0.43	0.00
Item 3	0.56	0.00	0.31	0.69	0.00
Item 4	0.58	0.00	0.33	0.67	0.00
Item 5	0.56	0.00	0.32	0.69	0.00
Vigour					
Item 1	0.64	0.00	0.41	0.59	0.00
Item 2	0.74	0.00	0.55	0.45	0.00
Item 3	0.61	0.00	0.38	0.63	0.00
Dedication					
Item 1	0.72	0.00	0.52	0.48	0.00
Item 2	0.78	0.00	0.61	0.39	0.00
Item 3	0.74	0.00	0.54	0.46	0.00
Item 4	0.73	0.00	0.53	0.47	0.00
Burnout					
Exhaustion	0.78	0.00	0.61	0.39	0.00
Cynicism	0.89	0.00	0.79	0.21	0.00
Engagement					
Vigour	0.99	0.00	0.98	0.03	0.54
Dedication	0.87	0.00	0.75	0.25	0.00

dimension and will impede the validity of the constructs. All the residual variances are statistically significant, except for vigour, which in accordance has a substantially large percentage of variance explained ($R^2=0.99$). This provides evidence of a problem regarding the two-factor model for engagement.

Reliability and correlations between the latent variables

Next, the phi matrix is reported, displaying correlations between the latent variables.

Table 8: Descriptive statistics and correlations between the well-being dimensions

Dimension	M	SD	Exhaustion	Cynicism	Vigour	Dedication	Burnout
Exhaustion	0.00	1.07					
Cynicism	0.00	1.27	0.70				
Vigour	0.00	0.98	-0.55	-0.62			
Dedication	0.00	1.06	-0.48	-0.55	0.86		
Burnout	0.00	0.84	0.78	0.89	-0.70	-0.62	
Engagement	0.00	0.97	-0.55	-0.63	0.98	0.87	-0.71

Table 8 indicates that the dimensions are moderately to highly correlated. All correlations are statistically significant ($p<0.01$). Sufficient Cronbach’s alpha coefficients ($\alpha>0.70$) (Nunnally & Bernstein 1994) were obtained for all the dimensions (exhaustion=0.81, cynicism=0.76, vigour=0.70 and dedication=0.83), providing support for Hypothesis 2. The two burnout dimensions (exhaustion and cynicism) were statistically significantly associated with vigour and dedication.

Discussion

Work-related well-being has become a major focus in most organisations around the world. Research shows that work-related well-being of employees can influence accidents and injuries at work (Siu et al. 2004). Although different studies focus on the work-related well-being of white-collar workers, no research could be found on the measurement of work-related well-being of blue-collar workers in South Africa. The Construction Regulations of South Africa (RSA 2003) include psychological well-being as an important attribute to ensure the safety of construction workers

whose work is particularly hazardous by nature. Unfortunately no valid and reliable measuring instrument is available to measure the work-related well-being of blue-collar workers. The objectives of this research were therefore to (1) develop a work-related well-being scale that measures burnout and work engagement of blue-collar workers; (2) to test the newly developed items using Rasch analysis; and (3) to test the factorial validity and reliability of the new scale.

With respect to the first objective, and based on instruments measuring burnout and work engagement, items were identified from various questionnaires that measure the relevant dimensions. Close attention was paid to the wording of items during the development phase, because the work experiences of blue-collar workers, their work conditions and level of education differ considerably from those of white-collar workers (DeVellis 2003). The scale development procedure was strictly adhered to in the conceptualisation of the different constructs, item generation and evaluation, item development, item refinement and the correct translation process to include all eleven official languages of South Africa.

After items had been developed and refined during the pilot study, the Rasch model was used to meet objective 2 (in order to evaluate items and eliminate poorly functioning items). The Rasch measuring model, as a function of the item response theory, was used in various research settings, but has become very popular in the psychometric evaluation of outcome scales (Tennant & Conaghan 2007). Furthermore, the Rasch measurement model was utilised because it assumes that all items are part of a unidimensional scale (Hagquist 2007; Rasch 1960). By utilising the Rasch analysis, the optimal number of rating scale categories for each construct was examined. All the rating scales seemed to function sufficiently. However, some rating scale categories for the exhaustion, vigour and dedication were under-utilised. On the exhaustion dimension, the categories indicating high exhaustion levels ('quite frequently', 'regularly' and 'always') were under-utilised. In the case of vigour and dedication, the three rating categories indicating vigour and dedication ('never', 'almost never' and 'infrequently') were under-utilised. This could imply that the seven-point rating scale might be too complex for low literacy employees and that a four-point scale ('never', 'almost never', 'sometimes', 'always') might have been more sufficient. Another reason might be that the respondents misunderstood the items and that they were reluctant to answer with the relevant intensity.

In order to evaluate the reliability of the measures obtained on the subscales of items, both the reliability (as indicated by Cronbach's alpha coefficients) and the person separation index were calculated by using Rasch analysis. The Rasch person separation index is thought to be similar to the Cronbach's alpha coefficient and provides the means to differentiate persons on the same construct or to indicate

whether the placement of people on other items measuring the same construct will be the same (Fox & Jones 1998). No problems were identified concerning item separation and reliability. This indicates that items were able to discriminate well across the investigated variables, and that items will probably be stable in another sample or research setting. Furthermore, based on the alpha coefficients, the reliability of the scales was sufficient. However, the person separation indexes for all the dimensions were somewhat less reliable, implying that the dimensions could have discriminated better among respondents with different abilities. Furthermore, it is possible that different items targeted the same ability excessively. Possible reasons for the lower person separation are that the respondents misunderstood some items or might have been reluctant to answer with the required intensity. Better person separation might thus be ensured if the wording and intensity of the items are explored and adapted.

One of the most important functions of utilising the Rasch analysis in instrument development is to identify problematic items by evaluating the item fit statistics (Boone & Rogan 2005). During this analysis, two problematic items were identified and eliminated based on the infit and outfit statistics of these items – item five of the vigour dimension and item five of the dedication dimension. Both items were causing an underfit of the model and were insufficient in predicting the specific dimension. Item five of vigour showed the lowest item measures, indicating the high endorsement of this item by participants. However, it seemed to underfit (based on infit statistics), indicating the unpredictability of this item. This could be because this item was longer than to the rest of the items. The wording could also have been problematic for the lower literacy employees, in that it was difficult for them to understand the exact meaning of this item. As a result, this item did not measure what it was supposed to measure (Green & Frantom 2002). Furthermore, item five of the dedication dimension showed the highest item measures (indicating low endorsement of items by participants), but seemed to underfit, meaning that this item did not provide information consistent with the other items measuring dedication. After these two problematic items were discarded from the scale, the fit of the remaining items as well as the person separation index and item reliability index improved. Consequently, the fit of the entire measure improved.

In order to meet objective 3, the factorial validity and reliability of the new questionnaire was tested (with the two poorly performing items identified by the Rasch analysis omitted). The hypothesised model was tested and compared with competing models. When competing models were tested, the results supported the hypothesised model, indicating that two higher order factors exist, namely burnout (consisting of exhaustion and cynicism) and work engagement (consisting of vigour and dedication), supporting previous findings (Coetzer & Rothmann 2007; Jackson

et al. 2006; Rothmann & Jorgensen 2007; Rothmann & Pieterse 2007). However, a problem is evident with the engagement construct. The residual variance for the vigour construct is not statistically significant and a substantially large percentage of variance is explained ($R^2=0.99$). This indicates that engagement might instead be a one-factor model, which also supports previous findings (Naudé & Rothmann 2004a; Olivier & Rothmann 2008). All the scales showed acceptable reliabilities. These findings confirm previous research on the reliability of these dimensions (Buys & Rothmann 2010; Coetzee & Rothmann 2004; Jackson & Rothmann 2005; Jackson et al. 2006; Pienaar & Willemse 2008; Rothmann & Joubert 2007; Storm & Rothmann 2003b).

In conclusion, the scale development procedure was followed, and a new questionnaire was developed to measure burnout and work engagement of blue-collar workers. Rasch analysis was utilised for the preliminary identification and elimination of problematic items. Evidence for factorial validity and reliability was also reported, as well as significant relationships with relevant outcome variables. However, the results showed some ambiguity regarding the engagement factor structure. The results of this study form the foundation for providing researchers and practitioners with a measuring instrument to evaluate the psychological well-being of blue-collar workers. Furthermore, the findings contribute to better understanding of the work experiences and work-related psychological well-being of blue-collar workers in South Africa.

Although this study is a valuable contribution to the measurement of work-related well-being, there are limitations to the study. The focus of this study was to develop a new instrument for psychological well-being and to test the factorial validity and reliability of this instrument. The questionnaire was translated from English into all the other official languages of South Africa. The view of Van de Vijver and Tanzer (2004) should, however, be taken into account, namely that from a psychological perspective, the quality of items might still be poor even though they have been grammatically correctly translated. South Africa is a multicultural society, and organisations employ individuals of diverse ethnic and cultural backgrounds. It can therefore not be taken for granted that scores obtained for one ethnic group can be compared across ethnic groups. Because of the inherent measurement issues in multicultural contexts, analyses should not only focus on internal consistency and factorial validity, but also on factorial invariance and item bias. Without a test of equivalence and bias, it is impossible to know the extent to which scores or constructs underlying an instrument can be compared across cultures (Van de Vijver & Leung 1997). It is therefore important to investigate the item bias and factorial invariance of

this instrument, because blue-collar employees in South Africa are very diverse and represent various cultural and language groups.

Another limitation was that the items did not differentiate well among the participants and that some rating scales were under-utilised, possibly because this study is the first to focus on blue-collar workers in South Africa. Limitations regarding items and scale categories were therefore anticipated. Another reason for this limitation may have been that the scale was translated into the eleven languages, but some languages do not have synonyms for the seven different options, as found in English. It is possible that by including more items and reducing the scale categories, the results would improve.

The results of this study were obtained using only self-report questionnaires. Method bias – also known as ‘common method variance’ or nuisance – might therefore be problematic when utilising this method. This implies that the investigated phenomenon becomes difficult to differentiate from measurement artifacts (Avolio & Bass 1991; Hufnagel & Conca 1994). Different researchers indicate that this type of bias can negatively influence the results of a study when using self-report questionnaires, although the severities thereof have not been confirmed (Crompton & Wagner 1994; Spector 1987). Few other methodologies are proposed as alternatives to self-report measures, since the employee is regarded as the most important source to offer information regarding his/her work (Frese & Zapf 1999). Subjective methods, including observers’ ratings, might be a good alternative, but they are not without problems of their own (for example, observer’s bias, halo and stereotyping effects).

Despite the limitations, recommendations can be made for future studies regarding the use of the newly developed instrument. It is recommended that future research should test for invariance of all language groups in South Africa. The cultural sensitivity of the new instrument should therefore be ensured for all cultural and language groups in South Africa. Consequently, in ensuring cultural sensitivity, researchers can ensure that all the items are understood in the same way and that the same construct is measured across different language and cultural groups. Future research should focus on the equivalence of this new instrument. According to Van de Vijver and Leung (1997), there are hierarchical distinctions of three types of equivalence, including construct equivalence (the extent to which the same construct is measured across all cultural groups), measurement unit equivalence (when two metric measures have the same measurement unit but different origins) and scalar equivalence (when two metric measures have the same measurement unit and the same origin). Because equivalence cannot be assumed, future studies should establish and report on this issue. If unacceptable construct equivalence is found, item bias should be computed (Van de Vijver & Leung 1997). Therefore, participants with an

equal standing on the theoretical construct underlying the instrument should have the same expected score on the item, irrespective of group membership. Item bias can be produced by sources such as incidental differences in appropriateness of the item content and inadequate item formulation. Because bias will lower the equivalence of a measuring instrument, this is an important aspect to be addressed in future research. Furthermore, cross-cultural Rasch analysis can be used in the future to ensure the cultural sensitivity of the instrument (Van de Vijver & Tanzer 2004), thus ensuring more reliable information in terms of psychological well-being for all blue-collar workers in South Africa.

As mentioned, measurement of the engagement construct was problematic, because the residual variance for the vigour construct was substantially large ($R^2=0.99$). More research is needed in order to confirm the factorial structure of work engagement. Based on the information regarding the items and rating scale categories of this study, it is recommended that additional items for all the different dimensions should be further explored. Considering the low literacy levels of these employees, careful attention should be given to the structuring of these items. It is evident that shorter items function better than items consisting of long sentences. Furthermore, research should also consider replacing the seven-point frequency scale with a four-point frequency scale. It is possible that the participants' abilities to discriminate meaningfully between the seven categories have been influenced. This may also be because of problematic wording in terms of different meanings in different language groups. DeVellis (2003) proposes that when respondents find it difficult to discriminate meaningfully between too many options, the number of options needs to be reduced. It is therefore more useful to have a shorter scale than to have a long scale in which some categories are under-utilised.

Author's note

This article was part of the first author's PhD thesis, completed at the North-West University, Potchefstroom Campus.

References

- Åkerstedt, T. 2006. 'Psychosocial stress and impaired sleep', *Scandinavian Journal of Work and Environmental Health*, 32(6): 493–501.
- Avolio, B.J. & Bass, B.M. 1991. 'Identifying common methods variance with data collected from a single source: An unresolved sticky issue', *Journal of Management*, 17: 571–587.

- Bakker, A.B. & Demerouti, E. 2007. 'The Job Demands-Resources Model: State of the art', *Journal of Managerial Psychology*, 22(3): 309–328.
- Bakker, A.B., Demerouti, E. & Schaufeli, W.B. 2002. 'Validation of the Maslach Burnout Inventory–General Survey: An internet study', *Anxiety, Stress and Coping*, 15: 245 – 260.
- Bakker, A.B., Demerouti, E. & Schaufeli, W.B. 2003. 'Dual processes at work in a call centre: An application of the Job Demands-Resources Mode', *European Journal of Work Organizational Psychology*, 12: 393–417.
- Bakker, A.B., Demerouti, E., De Boer, E. & Schaufeli, W.B. 2003. 'Job demands and job resources as predictors of absence duration and frequency', *Journal of Vocational Behavior*, 62: 341–356.
- Barkhuizen, N. & Rothmann, S. 2006. 'Work engagement of academic staff in South African higher education institutions', *Management Dynamics*, 15(1): 38–46.
- Bond, T.G. & Fox, C.M. 2007. *Fundamental Measurement in the Human Sciences*, 2nd edition. Mahwah, NJ: Erlbaum Associates.
- Boone, W. & Rogan, J. 2005. 'Rigour in quantitative analysis: The promise of Rasch analysis techniques', *African Journal of Research in SMT Education*, 9(1): 25–38.
- Bryman, A. & Bell, E. 2003. *Business Research Methods*, 2nd edition. New York: Oxford University Press.
- Buyts, C. & Rothmann, S. 2010. 'Burnout and engagement of reformed church ministers', *South African Journal of Industrial Psychology*, 36(1): 1–11.
- Byrne, B.M. 2010. *Structural Equation Modeling with AMOS: Basic Concepts, Applications and Programming*, 2nd edition. Mahwah, NJ: Lawrence Erlbaum Associates.
- Calitz, P.L. 2004. The experience of women in the platinum mining industry. Unpublished master's dissertation. North-West University, Potchefstroom.
- Campbell, C. & Rothmann, S. 2005. 'A psychometric assessment of the Maslach Burnout Inventory (General Survey) in a customer-service environment', *Management Dynamics*, 14(2): 16–28.
- Cervellione, K.L., Lee, Y. & Bonnano, G. A. 2009. 'Rasch modeling of the self-deception scale of the balanced inventory of desirable responding', *Educational and Psychological Measurement*, 69(3): 438–458.
- Coetzee, S.E. & Rothmann, S. 2004. 'An adapted model of burnout for employees at a higher education institution in South Africa', *South African Journal of Industrial Psychology*, 30(3): 29–40.
- Coetzer, W.J. & Rothmann, S. 2007. 'A psychometric evaluation of measures of affective well-being in an insurance company', *South African Journal of Industrial Psychology*, 33(2): 7–15.
- Cordes, C.L. & Dougherty, T.W. 1993. 'A review and integration of research on job burnout', *Academy of Management Review*, 18: 621–656.
- Crampton, S.M. & Wagner III, J.A. 1994. 'Percept-percept inflation in microorganizational research: An investigation of prevalence and effect', *Journal of Applied Psychology*, 79: 67–76.

- Curbow, B., Spratt, K., Ungaretti, A., McDonell, K. & Breckler, S. 2006. 'Development of the child care worker job stress inventory', *Early Childhood Research Quarterly*, 18: 310–330.
- Deacon, C. & Kew, G. 2006. 'Psychological versus medical fitness: Health and safety', *Construction World*, Aug: 48–50.
- Demerouti, E., Bakker, A.B. Nachreiner, F. & Schaufeli, W.B. 2001. 'The Job Demands-Resources Model of burnout', *Journal of Applied Psychology*, 86: 499–512.
- Demerouti, E., Bakker, A.B., De Jonge, J., Janssen, P.P. & Schaufeli, W.B. 2001b. 'Burnout and engagement at work as a function of demands and control', *Scandinavian Journal of Work, Environment and Health*, 27: 279–286.
- Demerouti, E., Bakker, A.B., Vardakou, I. & Kantas, A. 2002. 'The convergent validity of two burnout instruments: A multitrait-multimethod analysis', *European Journal of Psychological Assessment*, 18(3): 296–307.
- DeVellis, R.F. 2003. *Scale Development: Theory and Application*, 2nd edition. Thousand Oaks, CA: Sage Publications.
- Fox, C.M. & Jones, J.A. 1998. 'Uses of the Rasch modeling in counselling psychology research', *Journal of Counseling Psychology*, 45(1): 30–45.
- Foxcroft, C. & Roodt, G. 2005. *An Introduction to Psychological Assessment in the South African Context*, 2nd edition. Cape Town: Oxford University Press Southern Africa.
- Frese, M. & Zapf, D. 1999. 'On the importance of the objective environment in stress and attribution theory', *Journal of Organizational Behavior*, 20: 761–766.
- Geurts, S.A., Taris, T.W., Kompier, M.A., Dikkers, J.S., Van Hooff, M.L. & Kinnunen, U.M. 2005. 'Work-home interaction from a work psychological perspective: Development and validation of a new questionnaire, the SWING', *Work and Stress*, 19: 319–339.
- González-Romá, V., Schaufeli, W.B., Bakker, A.B. & Lloret, S. 2006. 'Burnout and work engagement: Independent factors or opposite poles?' *Journal of Vocational Behavior*, 68: 165–174.
- Green, K.E. & Frantom, C.G. 2002. 'Survey development and validation with the Rasch model', Paper presented at the International Conference on Questionnaire Development, Evaluation, and Testing, Charleston, SC.
- Hagquist, C. 2007. 'The psychometric properties of the self-reporting SDQ: An analysis of Swedish data based on the Rasch model', *Personality and Individual Differences*, 43: 1289–1301.
- Hakanen, J.J., Bakker, A.B. & Schaufeli, W.B. 2006. 'Burnout and work engagement among teachers', *Journal of School Psychology*, 43: 495–513.
- Harter, J.K., Hayes, T.L. & Schmidt, F.L. 2002. 'Business-unit-level relationship between employee satisfaction, employee engagement and business outcomes: A meta-analysis', *Journal of Applied Psychology*, 87(2): 268–279.
- Hobfoll, S.E. 1998. *Stress, Culture, and Community: The Psychology and Physiology of Stress*. New York: Plenum.

- Hobfoll, S.E. & Freedy, J. 1993. 'Conservation of resources: A general stress theory applied to burnout', In Schaufeli, W.B., Maslach, C. & Marek T. (eds), *Professional Burnout: Recent Development in Theory and Research*. Washington, DC: Taylor & Francis.
- Hu, L.T. & Bentler, P.M. 1999. 'Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives', *Structural Equation Modeling*, 6: 1–55.
- Hufnagel, E.M. & Conca, C. 1994. 'User response data: The potential for errors and biases', *Information System Research*, 5: 48–73.
- Jackson, L.T.B. & Rothmann, S. 2005. 'An adapted model of burnout for educators in South Africa', *South African Journal of Education*, 25(2): 100–108.
- Jackson, L.T.B., Rothmann, S. & Van de Vijver, F.J.R. 2006. 'A model of work-related well-being for educators in South Africa', *Stress and Health*, 22: 263–274.
- Kirschenbaum, A., Oigenblick, L. & Goldberg, A.I. 2000. 'Well-being, work environment and work accidents', *Social Science and Medicine*, 50: 631–639.
- Kitaoka-Higashiguchi, K., Nakagawa, H., Morikawa, Y., Ishizaki, M., Miura, K., Naruse, Y., Kido, T. & Higashiyama, M. 2004. 'Construct validity of the Maslach Burnout Inventory-General Survey', *Stress and Health*, 20: 255–260.
- Lance, C.E., Butts, M.M. & Michels, L.C. 2006. 'The sources of four commonly reported cutoff criteria. What did they really say?' *Organizational Research Methods*, 9: 202–220.
- Langelaan, S., Bakker, A.B., Van Doornen, L.J.P. & Schaufeli, W.B. 2006. 'Burnout and work engagement: Do individual differences make a difference?' *Personality and Individual Differences*, 40: 521–532.
- Lee, G.J. & Mohamed, Y. 2006. 'Perceptions of workplace issues among blue-collar workers in South Africa's Gauteng Province', *South African Journal of Psychology*, 36(1): 45–62.
- Lee, R.T. & Ashforth, B.E. 1996. 'A meta-analytic examination of the correlates of three dimensions of burnout', *Journal of Applied Psychology*, 81: 123–133.
- Leiter, M.P. 1993. 'Burnout as a developmental process: Consideration of models', In Schaufeli, W.B., Maslach, C. & Marek T. (eds), *Professional Burnout: Recent Development in Theory and Research*. Washington, DC: Taylor & Francis.
- Leiter, M.P. & Schaufeli, W.B. 1996. 'Consistency of the burnout construct across occupations', *Anxiety, Stress and Coping*, 9: 229–243.
- Linacre, J.M. 1999. 'Investigating rating scale category utility', *Journal of Outcome Measurement*, 3(2): 103–122.
- Linacre, J.M. 2005. WINSTEPS Rasch measurement (computer program). Chicago: Winsteps.com.
- Maslach, C. & Jackson, S.E. 1986. *Maslach Burnout Inventory Manual*, 2nd edition. Palo Alto, CA: Consulting Psychologists Press.
- McGwin, G. Jnr, Valent, F., Taylor, A.J., Howard, H.J., Davis, G.G., Brissie, R.M. & Rue III, L.W. 2002. 'Epidemiology of fatal occupational injuries in Jefferson County, Alabama', *Southern Medical Journal*, 95(11): 1300–1312.

- Meijman, T.F. & Mulder, G. 1998. 'Psychological aspects of workload', In Drenth, P.J. & Thierry, H. (eds), *Handbook of Work and Organizational Psychology*, Volume 2. *Work Psychology*. Hove, England: Psychology Press.
- Montgomery, A.J., Peeters, M.C.W., Schaufeli, W.B. & Den Ouden, M. 2003. 'Work-home interference among newspaper managers: Its relationship with burnout and engagement', *Anxiety, Stress and Coping*, 16(2): 195–211.
- Muthén, L.K. & Muthén, B.O. 2007. *Mplus User's Guide*, 5th edition. Los Angeles, CA: Muthén & Muthén.
- Naudé, J.L.P. & Rothmann, S. 2004a. 'The validation of the Utrecht Work Engagement Scale for emergency medical technicians in Gauteng', *South African Journal of Economic and Management Sciences*, 7(3): 473–488.
- Naudé, J.L.P. & Rothmann, S. 2004b. 'The validation of the Maslach Burnout Inventory: Human services survey for emergency medical technicians in Gauteng', *South African Journal of Industrial Psychology*, 30(3): 21–28.
- Nelson, D.L. & Simmons, B.L. 2003. 'Health psychology and work stress: A more positive approach', In Quick, J.C. & Tetrick, L.E. (eds), *Handbook of Occupational Health Psychology*. Washington: DC: American Psychological Association.
- Olivier, A.L. & Rothmann, S. 2008. 'Antecedents of work engagement in a multinational oil company', *South African Journal of Industrial Psychology*, 33(3): 49–56.
- Paul, P.S. & Maiti, J. 2007. 'The role of behavioral factors on safety management in underground mines', *Safety Science*, 45: 449–471.
- Pienaar, J. & Willemsse, S.A. 2008. 'Burnout, engagement, coping and general health of service employees in the hospitality industry', *Tourism Management*, 29(6): 1053–1063.
- Pines, A. & Aronson, E. 1988. *Career Burnout: Causes and Cures*. New York: Free Press.
- Rasch, G. 1960. *Probabilistic Models for Some Intelligence and Attainment Tests*. Copenhagen: Danish Institute for Educational Research. (Expanded edition, Chicago: University of Chicago Press, 1980).
- Richardsen, A.M. & Martinussen, M. 2004. 'The Maslach Burnout Inventory: Factorial validity and consistency across occupational groups in Norway', *Journal of Occupational and Organizational Psychology*, 77: 377–384.
- Rothmann, S. & Barkhuizen, N. 2008. 'Burnout of academic staff in South African higher education institutions', *South African Journal of Higher Education*, 22(2): 439–456.
- Rothmann, S. & Jorgensen, L.I. 2007. 'A model of work-related well-being for police members in the North West province', *Acta Criminologica*, 20(4): 73–84.
- Rothmann, S. & Joubert, J.H.M. 2007. 'Job demands, job resources, burnout and work engagement of managers at a platinum mine in the North West Province', *South African Journal of Business Management*, 38(3): 49–61.
- Rothmann, S. & Pieterse, J. 2007. 'Predictors of work-related well-being in sector education training authorities', *South African Journal of Economic and Management Sciences*, 10(3): 298–312.

- RSA (Republic of South Africa). 2003. Construction Regulation R.7721 in terms of the Occupational Health and Safety Act (Act No. 85 of 1993). Pretoria: Government Printer.
- Saks, A.M. 2006. 'Antecedents and consequences of employee engagement', *Journal of Managerial Psychology*, 2(7): 600–619.
- Salanova, M. & Schaufeli, W.B. 2008. 'A cross-national study of work engagement as a mediator between job resources and proactive behavior', *International Journal of Human Resource Management*, 19(1): 116–131.
- Schaufeli, W.B. 2003. 'Past performance and future perspectives of burnout research', *South African Journal of Industrial Psychology*, 29(4): 1–15.
- Schaufeli, W.B. & Bakker, A.B. 2001. 'Werk en welbevinden: Naar een positieve benadering in de Arbeids- en Gezondheidspsychologie' [Work and well-being: Towards a positive approach in occupational and health psychology], *Gedrag en Organisatie*, 14: 229–253.
- Schaufeli, W.B. & Bakker, A.B. 2004a. 'Job demands, job resources and their relationship with burnout and engagement: A multi-sample study', *Journal of Organizational Behavior*, 25: 293–315.
- Schaufeli, W.B. & Bakker, A.B. 2004b. *Utrecht Work Engagement Scale, Preliminary Manual*. Occupational Health Psychology Unit, Utrecht University.
- Schaufeli, W.B. & Enzmann, D. 1998. *The Burnout Companion to Study and Practice: A Critical Analysis*. Philadelphia: Taylor & Francis.
- Schaufeli, W.B. & Salanova, M. 2007. 'Efficacy or inefficacy, that's the question: Burnout and work engagement and their relationship with efficacy beliefs', *Anxiety, Stress and Coping*, 20(2): 177–196.
- Schaufeli, W.B., Bakker, A.B. & Salanova, M. 2006. 'The measurement of work engagement with a short questionnaire: A cross-national study', *Educational and Psychological Measurements*, 66(4): 701–716.
- Schaufeli, W.B., Leiter, M.P. & Maslach, C. 2009. 'Burnout: 35 years of research and practice', *Career Development International*, 14(3): 204–220.
- Schaufeli, W.B., Leiter, M.P., Maslach, C. & Jackson, S.E. 1996. 'The Maslach Burnout Inventory: General Survey', In Maslach, C., Jackson S.E. & Leiter M.P. (eds), *MBI Manual*, 3rd edition. Palo Alto, CA: Consulting Psychologists Press.
- Schaufeli, W.B., Salanova, M., González-Romá, V. & Bakker, A.B. 2002a. 'The measurement of engagement and burnout: A two sample confirmatory factor analysis approach', *Journal of Happiness Studies*, 3: 71–92.
- Schaufeli, W.B., Martinez, I.M., Pinto, A.M., Salanova, M. & Bakker, A.B. 2002b. 'Burnout and engagement in university students: A cross-national study', *Journal of Cross-cultural Psychology*, 33(5): 464–481.
- Schaufeli, W.B., Taris, T.W. & Van Rhenen, W. 2008. 'Workaholism, burnout and engagement: Three of a kind or three different kinds of employee well-being', *Applied Psychology: An International Review*, 57: 173–203.

- Schutte, N., Toppinen, S., Kalimo, R. & Schaufeli, W. 2000. 'The factorial validation of the Maslach Burnout Inventory-General Survey (MBI-GS) across occupational groups and nations', *Journal of Occupational and Organizational Psychology*, 73: 53–66.
- Shanahan, S.F., Anderson, S.J. & Mkhize, N.J. 2001. 'Assessing psychological distress in Zulu-speakers: Preliminary findings from an adaption of the SCL-90-R', *South African Journal of Psychology*, 31(4): 1–12.
- Shirom, A. 1989. 'Burnout in working organizations', In Cooper, C.L. & Robertson, I.T. (eds), *International Review of Industrial and Organisational Psychology*. New York: Wiley.
- Shirom, A. 2003. 'Job-related burnout: A review', In Quick, J.C. & Tetrick, L.E. (eds), *International Review of Industrial and Organizational Psychology*. Chichester, UK: Wiley.
- Siu, O., Phillips, D.R. & Leung, T. 2004. 'Safety climate and safety performance among construction workers in Hong Kong: The role of psychological strains as mediators', *Accident Analysis and Prevention*, 36: 359–366.
- Sluiter, J.K., Frings-Dresen, M.H., Van der Beek, A.J. & Meijman, T.F. 2001. 'The relation between work-induced neuroendocrine reactivity and recovery, subjective need for recovery, and health status', *Journal of Psychosomatic Research*, 50: 29–37.
- Sonnentag, S. 2001. 'Work, recovery activities, and individual well-being: A diary study', *Journal of Occupational Health Psychology*, 6: 196–210.
- Sonnentag, S. 2003. 'Recovery, work engagement, and proactive behavior: A new look at the interface between nonwork and work', *Journal of Applied Psychology*, 88: 518–828.
- Sparks, K., Faragher, B. & Cooper, C.L. 2001. 'Well-being and occupational health in the 21st century workplace', *Journal of Occupational and Organizational Psychology*, 74: 489–509.
- Spector, P.E. 1987. 'Method variance as an artifact in self-reported affect and perceptions at work: Myth or significant problem?', *Journal of Applied Psychology*, 72: 438–443.
- Storm, K. & Rothmann, S. 2003a. 'A psychometric analysis of the Maslach Burnout Inventory-General Survey in the South African Police Service', *South African Journal of Psychology*, 33(4): 219–226.
- Storm, K. & Rothmann, S. 2003b. 'A psychometric analysis of the Maslach Burnout Inventory-General Survey in the South African Police Service', *South African Journal of Psychology*, 33(4): 219–226.
- Struwig, F.W. & Stead, G.B. 2001. *Planning, Designing and Reporting Research*. Cape Town, South Africa: Pearson Education.
- Taris, T.W., Schreurs, P.J.G. & Schaufeli, W.B. 1999. 'Construct validity of the Maslach Burnout Inventory-General Survey: A two-sample examination of its factor structure and correlates', *Work and Stress*, 13(3): 223–237.
- Taris, T.W., Schreurs, P.J. & Van Iersel-Van Silfhout, I.J. 2001. 'Job stress, job strain, and psychological withdrawal among Dutch university staff: Towards a dual-process model for the effects of occupational stress', *Work and Stress*, 15(4): 283–296.
- Tennant, A. & Conaghan, P.G. 2007. 'The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied and what should one look for in a Rasch paper?', *Arthritis and Rheumatism*, 15: 1358–1362.

- Van de Vijver, F. & Leung, K. 1997. *Methods and Data-analysis for Cross-cultural Research*. Thousand Oaks, CA: Sage.
- Van de Vijver, F. & Tanzer, N.K. 2004. 'Bias and equivalence in cross-cultural assessment: An overview', *European Review of Applied Psychology*, 54: 119–135.
- Van Hooff, M., Geurts, S.A., Taris, T.W., Kompier, M.A., Dikkers, J.S., Houtman, I.L. & Van den Heuvel, F. 2005. 'Disentangling the relationships between work–home interference and employee health: A longitudinal study among Dutch police officers', *Scandinavian Journal of Work, Environment and Health*, 31: 15–29.
- Xanthopoulou, D., Bakker, A.B., Demerouti, E. & Schaufeli, W.B. 2007. 'The role of personal resources in the Job Demands-Resources Model', *International Journal of Stress Management*, 14(2): 121–141.