



Burnout and work engagement: Independent factors or opposite poles?

Vicente González-Romá ^{a,*}, Wilmar B. Schaufeli ^b,
Arnold B. Bakker ^b, Susana Lloret ^a

^a *University of Valencia, Department of Methodology of Behavioral Sciences,
Av. Blasco Ibanez, 21, 46010-Valencia, Spain*

^b *Utrecht University, Department of Social and Organizational Psychology, The Netherlands*

Received 14 October 2004
Available online 3 March 2005

Abstract

Burnout researchers have proposed that the conceptual opposites of emotional exhaustion and cynicism (the core dimensions of burnout) are vigor and dedication (the core dimensions of engagement), respectively (Maslach & Leiter, 1997; Schaufeli, Salanova, González-Romá, & Bakker, 2002). We tested this proposition by ascertaining whether two sets of items, exhaustion–vigor and cynicism–dedication, were scalable on two distinct underlying bipolar dimensions (i.e., energy and identification, respectively). The results obtained by means of the non-parametric Mokken scaling method in three different samples ($N_s = 477, 507,$ and 381) supported our proposal: the core burnout and engagement dimensions can be seen as opposites of each other along two distinct bipolar dimensions dubbed energy and identification.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Job burnout; Work engagement; Bipolar dimensions; Mokken Scaling

* Corresponding author. Fax: + 34 96 3864697.

E-mail address: Vicente.Glez-Roma@uv.es (V. González-Romá).

1. Introduction

Burnout is a reaction to chronic occupational stress characterized by emotional exhaustion (i.e., the draining of emotional resources), cynicism (i.e., a negative, callous, and cynical attitude towards one's job) and lack of professional efficacy (i.e., the tendency to evaluate one's work negatively). A major part of the research on burnout has focused on identifying its antecedents and outcomes (Lee & Ashforth, 1996; Schaufeli & Buunk, 2003). Only in recent years have burnout researchers started to pay attention to the conceptual opposite of burnout: *work engagement* (Maslach, Schaufeli, & Leiter, 2001; Schaufeli et al., 2002). Work engagement is a multidimensional construct defined as a positive, fulfilling, work-related state of mind that is characterized by vigor, dedication and absorption. *Vigor* is characterized by high levels of energy and mental resilience while working, the willingness to invest effort in one's work, and persistence even in the face of difficulties. *Dedication* is characterized by a sense of significance, enthusiasm, inspiration, pride and challenge. *Absorption* is characterized by fully concentrating on and being deeply engrossed in one's work, where time passes quickly and one has difficulty detaching oneself from work.

Emotional exhaustion and cynicism are considered the core burnout dimensions (Green, Walkey, & Taylor, 1991). Reduced efficacy was added as a constituting element of burnout after it emerged as a third factor from a factor-analysis of a preliminary version of the Maslach Burnout Inventory (MBI) (Maslach, 1993). Conceptually, professional efficacy has been criticized as reflecting a personality characteristic rather than a genuine burnout-component (Coders & Dougerthy, 1993; Shirom, 2003). Empirically speaking, this criticism is supported by the relatively low correlation of professional efficacy with both of the other burnout dimensions (Lee & Ashforth, 1996) and by the fact that cynicism seems to develop in response to exhaustion, whereas professional efficacy seems to develop independently and parallelly (Leiter, 1993).

Analogously, vigor and dedication are considered the core dimensions of engagement, and absorption was found to be a relevant aspect of engagement after some thirty in-depth interviews were carried out (Schaufeli & Bakker, 2001).

Theoretically, vigor is conceived as the opposite of emotional exhaustion, and dedication is conceived as the opposite of cynicism (Maslach & Leiter, 1997; Schaufeli et al., 2002). According to this view, vigor items and emotional exhaustion items measure a single underlying *bipolar* dimension, and, consequently, they should be scalable on a single underlying bipolar dimension. The same applies to the other pair of opposite constructs: dedication items and cynicism items should be scalable on a single underlying *bipolar* dimension as well. These underlying bipolar dimensions have been labeled energy and identification, respectively (Schaufeli & Bakker, 2001; Schaufeli et al., 2002).

However, to date, the assumption that exhaustion–vigor and cynicism–dedication constitute two bipolar dimensions representing energy and identification, respectively, has not been *tested*. Instead, the factor structure underlying the relationships among *total scores* on the three burnout dimensions and the three engagement dimensions was assessed (Salanova, Schaufeli, Llorens, Peiró, & Grau, 2001; Schaufeli & Bakker, 2004; Schaufeli et al., 2002). Using confirmative factor analysis,

these studies found support in various Dutch and Spanish samples for a two-factor model composed of a core burnout factor (including emotional exhaustion and cynicism) and an extended engagement factor (which included the three engagement dimensions plus professional efficacy).

The aim of the present study is to ascertain whether items measuring opposite burnout and engagement dimensions (i.e., exhaustion–vigor and cynicism–dedication) are scalable on a single underlying bipolar dimension. Since there are strong parallels with similar studies on the dimensionality of positive and negative affect, we will discuss some lessons from affect research first.

1.1. Lessons from research on affect

Research on the structure of affect has debated whether positive and negative affect are two independent factors or whether they are two poles of a single underlying bipolar dimension (see Russell & Carroll, 1999; Watson & Tellegen, 1999). Some studies have shown that when ambiguous response formats are used—like the common Likert-type response scale—the relationship between pairs of opposite affect items (e.g., sad and happy) is not linear. Actually, the corresponding bivariate frequency distribution is not diagonal, but roughly triangular instead (see Lloret & González-Romá, 2000, 2003; Russell & Carroll, 1999). This means that subjects responding “1. Not at all” to the item “Are you happy?” may respond “1. Not at all,” “2. Moderately,” or “3. Very much” to the opposite item “Are you sad?” This important finding explains why, when a set of items that are indicators of opposite poles of a single underlying bipolar dimension is submitted to factor analysis, instead of obtaining one single factor, two factors appear: one factor including the positive items, and another that includes the negative items (see Lloret & González-Romá, 2000). So, obviously, negative items are not highly and linearly related to positive items in the sense that low scores on the negative items are exclusively associated with high scores on the positive ones. In a set of items of this type, negative items show high linear relationships with other negative items (the same applies to positive items, of course), so that two clusters of highly linearly related items emerge that are reflected by two independent factors. Various authors have pointed out that the reason we obtain two factors when we expect a bipolar one is that data gathered by means of balanced Likert-type scales composed of positive and negative items do not fit the linear factor analysis model (Coombs & Kao, 1960; Davison, 1977; Van Schuur & Kiers, 1994).

So the conclusion is clear: if relationships among items are expected to not be linear, then factor analysis is not an appropriate technique for investigating whether items are scalable on a single underlying bipolar dimension. Instead, a scaling method that does not assume linear relationships must be used, such as the Mokken scaling technique (see below).

1.2. The present study

We believe that the problem raised above also applies to burnout and work engagement scales. Consider, for instance, the opposite exhaustion and vigor

dimensions. An employee responding “1. Never” to the item “I feel burned out from my job” may respond “1. Never,” “2. Sometimes,” or “3. Always” to the opposite item “At my work, I feel bursting with energy,” depending on the meaning attributed to “never burned out”. If the meaning is “some degree of energy,” the respondent will probably respond “2. Sometimes,” or “3. Always” to the vigor item, depending on the level of vigor felt. However, if the employee responds “never burned out” because he/she is at a neutral point on the latent bipolar energy continuum,” then the respondent will also respond with “never” to the opposite item. The same also applies to cynicism and dedication items, such as “I doubt the significance of my work” and “I find the work that I do full of meaning and purpose,” respectively.

The current study addresses three issues: (1) it will be ascertained whether or not a linear relationship exists between exhaustion and vigor items, and between cynicism and dedication items; (2) a series of principal components analyses is carried out to demonstrate that when opposite items are analyzed (i.e., exhaustion–vigor and cynicism–dedication), two components emerge instead of a single bipolar one; and (3) a more adequate non-linear scaling model is applied in order to assess to what extent exhaustion and vigor items, and cynicism and dedication items, assess two underlying bipolar dimensions (i.e., energy and identification).

2. Method

2.1. Samples and procedure

The participants worked for one of three different organizations, namely a telecom company, a pension fund company or an insurance company. After meeting with the floor managers and the human resources department, it was agreed that all employees working in the call center of the telecom company would have the opportunity to fill out an electronic questionnaire (published on a secure website) during work time in a silent separate room. A newsletter and an email from management announced to all employees that the questionnaire could be filled out. Employees on sick leave received a paper-and-pencil questionnaire by surface mail at their home address. The employees from both other organizations received paper-and-pencil questionnaires and return envelopes at their work. The questionnaires were accompanied by a letter, in which the goal of the study was briefly introduced and the confidentiality and anonymity of the answers were emphasized. The employees were asked to fill out the questionnaire in private and post it afterwards in a special box in their departments.

Sample 1 ($N = 477$) included customer services employees working in the only call center in a large Dutch telecom company (response rate 88%). In all, 467 employees filled out the questionnaire online, and 10 sick employees filled out the paper-and-pencil version at home. This sample includes 205 males (43%) and 272 females (57%). Their mean age is 30 years ($SD = 8.80$), and the mean organizational tenure is one year ($SD = .83$). The call center is a built-in center in a telecom company that

takes care of incoming customer calls (inbound services). Thirteen per cent of the participants had a supervisory position.

Sample 2 ($N = 507$) consisted of employees in a pension fund company (response rate 83%). The sample included 173 males (34%) and 334 females (66%), and their mean age was 35 years ($SD = 9.5$). Mean organizational tenure was 6.5 years ($SD = 7.5$). The main activities of the employees in this organization were to collect premiums and administrate and pay out monthly allowances and pensions. Sixteen per cent of the participants had a supervisory position.

Sample 3 ($N = 381$) consisted of employees in an insurance company (response rate 61%). The sample included 261 males (69%) and 120 females (31%). The mean age was 40 years ($SD = 10.4$). Mean organizational tenure was 12.5 years ($SD = 10.5$ years). The services of this company include life and funeral insurance, disability insurance, pension funds, and home loans. Twelve per cent of the sample had a managerial position.

2.2. Measures

Exhaustion and *Cynicism* were assessed with the Dutch version (Schaufeli & Van Dierendonck, 2000) of the Maslach Burnout Inventory-General Survey (MBI-GS; Schaufeli, Leiter, Maslach, & Jackson, 1996). The exhaustion scale includes five items (e.g., “I feel emotionally drained from my work”), and the cynicism scale includes four items: e.g., “I doubt the significance of my work.” Internal consistencies (Cronbach’s α) range across the three samples from .85 to .89 for exhaustion, and from .78 to .81 for cynicism.

Vigor and *dedication* were assessed with the Utrecht Work Engagement Scale (UWES; Schaufeli et al., 2002). The vigor scale includes five items (e.g., “At my work, I feel bursting with energy”), and the dedication scale also includes five items (e.g., “I find the work that I do full of meaning and purpose”). Internal consistencies (Cronbach’s α) range across the three samples from .80 to .85 for vigor and from .88 to .91 for dedication. All burnout and engagement items are scored on a seven-point frequency rating scale, ranging from 0 (“never”) to 6 (“always”).

2.3. Analysis

First, we examined the bivariate frequency distributions for pairs of exhaustion–vigor and cynicism–dedication items and performed a series of principal components analyses with subsequent varimax rotation. Next, in order to answer our main research question, we applied the Mokken scaling method (Mokken, 1971) to each set of items. The Mokken model is a non-parametric Item Response Theory (IRT) model for cumulative items. A set of items is said to be cumulative if they can be ordered along an underlying unidimensional continuum in terms of the trait level required by each item for endorsement response. The Mokken model tests whether two or more items can be ordered along a unidimensional continuum. This model posits that subject j will probably endorse item i if the trait level of subject j (θ_j) equals or exceeds the trait level required by item i (β_i) for endorsement response.

The generalization of Mokken scale analysis to polytomous items was proposed by Molenaar (1982, 1991).

In Mokken scaling, the assumptions of unidimensionality and cumulateness of the scale are evaluated by means of a coefficient of scalability, H , developed by Mokken (1971) and based on Loevinger's scalability coefficient (Loevinger, 1948). This coefficient reflects the proportion of error patterns in respondents' answers. In the simple case of two dichotomous items, the error pattern occurs when the most difficult or extreme item is passed or endorsed while the easiest or least extreme item is failed or not endorsed. As in the Guttman scalogram, this response pattern is erroneous because it violates the ordering of the two items along the continuum. The scalability coefficient compares the observed frequency of this error pattern (F_{ij}) with the expected frequency of this pattern (E_{ij}), assuming that responses to items are independent across subjects and that the marginals of the 2×2 cross table of bivariate frequencies of items i and j are fixed. For two items, the scalability coefficient H_{ij} is defined as

$$H_{ij} = 1 - (F_{ij}/E_{ij}).$$

In a similar way, scalability coefficients for a given scale and a given item are also available, as well as the generalization of these coefficients to the polytomous case (Molenaar, 1991). The reference points to assess scalability coefficients were given by Mokken (1971). A set of items constitutes a weak unidimensional cumulative scale if $0.3 \leq H < 0.4$. Medium scalability is obtained if $0.4 \leq H < 0.5$, and a scale is considered to be a strong unidimensional cumulative scale if $0.5 \leq H < 1$. A set of items with $H < 0.3$ is considered unscalable. In this case, the set of items may be multidimensional. H can be interpreted as an index for the degree to which subjects can be accurately ordered by means of a set of items along the underlying continuum (Molenaar, Debets, Sijtsma, & Hemker, 1994).

The Mokken model is preferable to parametric IRT models when data under analysis are based on only a small number of items (e.g., less than 20) and small samples, which is the case in the present study. Mokken analyses were carried out using the MSP program (Molenaar et al., 1994), after reversing the exhaustion and cynicism items.

3. Results

3.1. Bivariate distributions of responses

As anticipated, the bivariate distributions of responses for pairs of exhaustion–vigor and cynicism–dedication items *did not* show the diagonal form characteristic of a linear relationship. In most cases, the distributions obtained approached a roughly triangular shape instead. For instance, the bivariate distribution for the item pair EX3 (“I feel burned out from my job”)—VII (“At my work, I feel bursting with energy”) in Sample 2 showed that one-third of the respondents with low scores on item EX3 (that is, with scores of 0, 1 or 2) had low scores on item VII, 28% had

medium scores on item VII (that is, a score of 3), and only 38% showed high scores (that is, scores of 4, 5 or 6) as expected from a linear relationship model. The bivariate distribution for the item pair CY1 (“I doubt the significance of my work”)—DE1 (“I find my work full of meaning and purpose”) in the same sample showed that 19% of the people with low scores on item CY1 also had low scores on item DE1, 28% had medium scores on item DE1, and 53% showed high scores on item DE1. Hence, these two examples illustrate that the assumption of linearity that underlies statistical techniques like factor analysis is violated.

3.2. Principal components analysis

What happens when under these conditions burnout and engagement items are nonetheless submitted to a principal components analysis? As outlined before, we expected to obtain two components for each of the exhaustion–vigor and cynicism–dedication item sets. To test this hypothesis, we subjected each set of items in each of the three samples to a principal components analysis with varimax rotation. As expected, in each sample, the final solution revealed two components with eigenvalues greater than 1 (see Table 1). Factor loadings showed that exhaustion and vigor were the two components underlying the exhaustion–vigor items, and that cynicism and dedication were the two components underlying the cynicism–dedication items.

3.3. Mokken scaling

Finally, to test whether items measuring opposite burnout and engagement dimensions (i.e., emotional exhaustion–vigor and cynicism–dedication) are scalable on a single underlying bipolar dimension, we applied the Mokken scaling method to each of the two sets of items in every sample. The scalability coefficients obtained were greater than .3 in every case, pointing out that each set of items was scalable on a single underlying dimension in every sample. The scalability coefficients (H) obtained for the exhaustion–vigor items were .36, .39, and .44, for samples 1, 2, and 3, respectively. The scalability coefficients obtained for the cynicism–dedication items equaled .50 in all three samples. That means that the exhaustion–vigor items

Table 1
Results from principal component analysis

Items	Sample 1		Sample 2		Sample 3	
	Eigenvalues	% S ²	Eigenvalues	% S ²	Eigenvalues	% S ²
Exhaustion–vigor	4.2	32.0	4.1	32.2	4.4	35.0
	1.9	29.1	1.9	26.7	2.2	31.3
Cynicism–dedication	4.8	38.8	4.8	39.7	4.4	37.6
	1.3	29.1	1.3	28.0	1.6	29.5

Note. % S², percentage of variance explained by the rotated components.

constitute a weak to moderate energy scale, and that the cynicism–dedication items constitute a strong identification scale.

4. Discussion

The aim of the present study was to ascertain whether items measuring opposite burnout and engagement dimensions (i.e., exhaustion–vigor and cynicism–dedication) were scalable on a single underlying bipolar dimension. We showed that the bivariate distributions of responses for pairs of exhaustion–vigor and cynicism–dedication items *did not* show the diagonal form characteristic of a linear relationship. Instead, the observed distributions approached a roughly triangular shape. Thus, the linearity assumption of principal components analysis cannot be maintained, and, consequently, this method is inadequate for analyzing the underlying structure of sets of exhaustion–vigor and cynicism–dedication items. We showed that when this method is used with this purpose anyway, two components emerge, instead of a single, bipolar one. By using the Mokken scaling method (Mokken, 1971; Molenaar, 1982, 1991), a method that does not assume linear relationships among items, we showed that exhaustion and vigor items were scalable on a single underlying bipolar dimension labeled energy, whereas cynicism and dedication items were also scalable on a single, bipolar dimension labeled identification.

Our study makes a contribution to the theoretical development of the structure of employee well-being. A number of burnout researchers have proposed that vigor is the conceptual opposite of emotional exhaustion, and that dedication is the conceptual opposite of cynicism (Maslach & Leiter, 1997; Schaufeli et al., 2002). However, to our knowledge, no empirical study has addressed this proposition. Our study yields empirical evidence supporting this conceptualization of the core burnout and engagement dimensions as conceptual opposites that define two distinct underlying bipolar dimensions dubbed energy and identification.

An important practical implication of our study is that using the entire set of 19 items (the questionnaire for energy and identification, QUEENI), one can obtain subjects' scores on both bipolar dimensions of energy and identification. Energy scores can be obtained adding the five vigor items and the five recoded exhaustion items. A high score on energy indicates a high level of vigor, whereas a low score on energy indicates a high exhaustion level. Analogously, identification scores can be obtained by adding the five dedication items and the four recoded cynicism items. A high score on identification indicates a high level of dedication, whereas a low score on identification indicates a high level of cynicism.

A limitation of our study is that we did not provide empirical evidence on the validity of energy and identification scores. This was not the purpose of the present study, but it is a major aim for future research. A particular strength of our study is its multi-sample nature. This allowed us to observe that the two sets of items studied (i.e., exhaustion–vigor and cynicism–dedication) were scalable on two underlying bipolar dimensions (i.e., energy and identification, respectively) in the three

independent samples from different types of organizations. This result supports the validity of our findings.

References

- Coombs, C. H., & Kao, R. C. (1960). On a connection between factor analysis and multidimensional unfolding. *Psychometrika*, *25*, 219–231.
- Corders, C. L., & Dougerthy, T. W. (1993). A review and integration of research on job burnout. *Academy of Management Review*, *18*, 621–656.
- Davison, M. L. (1977). On a metric, unidimensional unfolding model for attitudinal and developmental data. *Psychometrika*, *42*, 523–548.
- Green, D. E., Walkey, F. H., & Taylor, A. J. W. (1991). The three-factor structure of the Maslach Burnout Inventory. *Journal of Social Behavior and Personality*, *6*, 453–472.
- Lee, R. T., & Ashforth, B. E. (1996). A meta-analytic examination of the correlates of the three dimensions of job burnout. *Journal of Applied Psychology*, *81*, 123–133.
- Leiter, M. P. (1993). Burnout as a developmental process: Consideration of models. In W. B. Schaufeli, C. Maslach, & T. Marek (Eds.), *Professional burnout: Recent developments in theory and research* (pp. 237–250). Washington, DC: Taylor & Francis.
- Lloret, S., & González-Romá, V. (2000). Medición de constructos bipolares mediante escalas tipo Likert: por qué aparecen factores monopolares? [Measurement of bipolar constructs with Likert-type scales: Why do unipolar factors emerge?]. *Psicothema*, *12*, 471–479.
- Lloret, S., & González-Romá, V. (2003). How do respondents construe ambiguous response formats of affect items?. *Journal of Personality and Social Psychology*, *85*, 956–968.
- Loevinger, J. (1948). The technique of homogeneous tests compared with some aspects of scale analysis and factor analysis. *Psychological Bulletin*, *45*, 507–530.
- Maslach, C. (1993). Burnout: A multi-dimensional perspective. In W. B. Schaufeli, C. Maslach, & T. Marek (Eds.), *Professional burnout: Recent developments in theory and research* (pp. 19–32). New York: Taylor & Francis.
- Maslach, C., & Leiter, M. P. (1997). *The truth about burnout: How organizations cause personal stress and what to do about it*. San Francisco, CA: Jossey-Bass.
- Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology*, *52*, 397–422.
- Mokken, R. J. (1971). *A theory and procedure of scale analysis*. The Hague: Mouton.
- Molenaar, I. W. (1982). Mokken scaling revisited. *Kwantitatieve Methoden*, *8*, 145–164.
- Molenaar, I. W. (1991). A weighted Lowinger H-coefficient extending Mokken scaling to multicategory items. *Kwantitatieve Methoden*, *37*, 97–117.
- Molenaar, I. W., Debets, P., Sijtsma, K., & Hemker, B. T. (1994). *User's manual MSP, a program for Mokken Scale analysis for Polytomous Items* (version 3.0). Groningen, The Netherlands: ProGAMMA.
- Russell, J. A., & Carroll, J. M. (1999). On the bipolarity of positive and negative affect. *Psychological Bulletin*, *125*, 3–30.
- Salanova, M., Schaufeli, W. B., Llorens, S., Peiró, J. M., & Grau, R. (2001). Desde el 'burnout' al 'engagement': Una nueva perspectiva [From burnout to engagement: A new perspective]. *Revista de Psicología del Trabajo y Organizaciones*, *16*, 117–134.
- 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 Health Psychology]. *Gedrag & Organisatie*, *14*, 229–253.
- Schaufeli, W. B., & Bakker, A. B. (2004). 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., & Buunk, B. P. (2003). Burnout: An overview of 25 years of research and theorizing. In M. J. Schabracq, J. A. M. Winnubst, & C. L. Cooper (Eds.), *Handbook of work and health psychology* (pp. 383–425). Chichester: Wiley.

- Schaufeli, W. B., & Van Dierendonck, D. (2000). *UBOS—Utrechtse Burnout Schaal. Handleiding [UBOS—Test Manual]*. Lisse. The Netherlands: Swets Test Services.
- Schaufeli, W. B., Leiter, M. P., Maslach, C., & Jackson, S. E. (1996). Maslach burnout inventory—general survey. In C. Maslach, S. E. Jackson, & M. P. Leiter (Eds.), *The Maslach Burnout Inventory-test Manual* (3rd ed., pp. 22–26). Palo AltoCA: Consulting Psychologists Press.
- Schaufeli, W., Salanova, M., González-Romá, V., & Bakker, A. B. (2002). The measurement of engagement and burnout: A two sample confirmatory factor analytic approach. *Journal of Happiness Studies*, 3, 71–92.
- Shirom, A. (2003). Job-related burnout: A review. In J. C. Quick & L. Tetrick (Eds.), *Handbook of occupational health psychology* (pp. 245–264). Washington, DC: American Psychological Association.
- Van Schuur, W. H., & Kiers, H. A. L. (1994). Why factor analysis is often the wrong model for analysing bipolar concepts and what model to use instead. *Applied Psychological Measurement*, 18, 97–110.
- Watson, D., & Tellegen, A. (1999). Issues in the dimensional structure of affect effects of descriptors, measurement error, and response format: comment on Russell and Carroll (1999). *Psychological Bulletin*, 125, 601–610.