British Journal of Psychology (1990), **81**, 315–333 *Printed in Great Britain*

The psychometric properties of the Bortner Type A Scale

Jeffrey R. Edwards* and **A. J. Baglioni Jr**
Colgate Darden Graduate School of Business Administration, University of Virginia, PO Box 6550, Charlottesville, Virginia 22906–6550, USA

Cary L. Cooper
University of Manchester Institute of Science and Technology

Recent evidence indicates that the relationship between Type A behaviour pattern (TABP) and coronary heart disease (CHD) is dependent upon the method of measuring TABP. This suggests that the psychometric properties of TABP measures should be carefully investigated. This article examines one widely used TABP measure, the Bortner Scale, using data from 1320 working adults divided into three random samples. The reliability of the Bortner Scale as an overall TABP index is unacceptably low. However, further analyses indicate that, rather than reflecting a single dimension, the Bortner Scale contains two independent dimensions, one reflecting speed and the other reflecting competitiveness. The speed dimension was negatively related to job satisfaction and, to a lesser extent, positively related to anxiety and somatic symptoms, whereas the competitiveness dimension was positively related to job satisfaction. Implications for the use of the Bortner Scale are discussed.

Over the past 30 years, research into the effects of Type A behaviour pattern (TABP) on psychological and physical symptoms has grown considerably. Friedman & Rosenman (1959) initially described TABP as a combination of a competitive need for achievement, a sense of time urgency, aggressiveness and hostility. Friedman & Rosenman noted that this constellation of behaviours was more prevalent among cardiac patients with more severe coronary heart disease (CHD). These initial observations stimulated several large-scale prospective studies, most of which found a modest but significant relationship between TABP and the prevalence and incidence of CHD (French–Belgian Collaborative Group, 1982; Haynes, Feinleib & Kannel, 1980; Rosenman, Brand, Jenkins, Friedman, Straus & Wurm, 1975). However, the conclusiveness of these findings has been challenged by more recent evidence, which suggests that the relationship between TABP and CHD is largely dependent on the method used to measure TABP, with inconsistent results associated primarily with self-report measures (Booth-Kewley & Friedman, 1987; Matthews, 1988).

* Requests for reprints.

The inconsistent relationship between self-report TABP measures and CHD suggests that the psychometric properties of these measures deserve serious consideration. Most studies using self-report TABP measures have relied on the Jenkins Activity Survey (JAS) (Jenkins, Rosenman & Friedman, 1967), the Framingham Scale (Haynes, Levine, Scotch, Feinleib & Kannel, 1978) or the Bortner Scale (Bortner, 1969). These measures have been preferred over other self-report measures primarily because each has been validated against the structured interview (SI: Rosenman, 1978) and, in one or more studies, prospectively related to CHD (French–Belgian Collaborative Group, 1982; Haynes *et al.*, 1980; Rosenman *et al.*, 1975). However, with the exception of the JAS, the psychometric properties of these measures have received very little attention.

The purpose of this study is to examine the psychometric properties of one self-report TABP measure, the Bortner Scale. This scale is particularly worthy of attention, given its widespread use in cross-sectional (Bass, 1984; Bass & Akhras, 1987; Bass & Wade, 1982; Cottier, Adler, Vorkauf, Gerber, Hefer & Hurny, 1983; Heller, 1979; Kornitzer, Magotteau, Degre, Kittel, Struyven & Van Theil, 1982) and prospective (French–Belgian Collaborative Group, 1982; Koskenvuo, Kaprio, Rose, Kesaniemik, Sarna, Heikkila & Langinvainia, 1988) studies of CHD. Despite its widespread use, the psychometric properties of the Bortner Scale have received very little attention. We investigated the reliability and dimensionality of the Bortner Scale and the relationship between these dimensions and self-reported psychological and physical symptoms. The results of this study provide a basis for evaluating the Bortner Scale as a measure of TABP and suggest specific dimensions within the Bortner Scale which may predict symptomatology.

History of the Bortner Scale

Development of the Bortner Scale

The Bortner Scale (hereafter, 'the Bortner') was developed using data from 76 male insurance and business executives (Bortner & Rosenman, 1967). The original version of the Bortner (Bortner, 1969) contained 14 items, each consisting of two phrases placed at opposite ends of a continuum ranging from extreme TABP to the absence of TABP (i.e. Type B behaviour pattern). Respondents were asked to place a vertical mark through a 1.5 inch line separating the two phrases to indicate their position between the two extremes. Later versions have adopted numerical response formats, such as the adaptation by Cooper consisting of an 11-point scale centred at zero and ascending to five in both directions (Cooper & Marshall, 1979). Bortner (1969) also presented a seven-item short scale, which was formed by selecting items that, after controlling for the remaining items, independently contributed to the prediction of SI A/B classifications. Bortner (1969) also developed a weighted version of the short scale, with weights derived from a multiple regression of SI A/B classifications on the seven items forming the short scale.

Reliability of the Bortner Scale

Only a handful of studies has examined the reliability of the Bortner. Bortner (1969) originally reported a reliability estimate of .68. More recently, Mayes, Sime & Ganster (1984) reported a lower reliability estimate of .60. Data from the seven-item Bortner have yielded similar results. For example, Ray & Bozek (1980) administered the seven-item Bortner to a sample of 119 Australian adults and reported a reliability estimate of .53. Koskenvuo, Kaprio, Langinvainia, Romo & Sarna (1981) also administered the seven-item Bortner to 11 364 Finnish men and women and reported correlation matrices of the items. Based on these matrices, we derived reliability estimates (Cronbach's alpha) of .15 for men and .17 for women. Thus, available evidence suggests that the reliability of the Bortner is marginal, with reliability estimates ranging from .53 to .68 for the full scale and from .15 to .53 for the seven-item scale.

Dimensionality of the Bortner Scale

Studies using the Bortner have typically summed its items to form an overall measure of TABP (Bortner, 1969; Cooper & Marshall, 1979; French–Belgian Collaborative Group, 1982). This reflects an implicit assumption that the Bortner is a unidimensional scale. However, this assumption has not been systematically evaluated, and the limited available evidence suggests that the Bortner may, in fact, contain several distinct dimensions. For example, Pichot, De Bonis, Somogyi, Degre-Coustry, Kittel-Bossuit, Rustin-Vandenhende, Dramaix & Bernet (1977) conducted a principal component analysis of the scale, using data from 715 adults. Four factors were identified, reflecting time urgency, hard-driving, competitiveness and expressiveness. In contrast, Johnston & Shaper (1983) attempted to factor analyse the Bortner using data from 135 British executives and reported they could not find a stable solution. Koskenvuo *et al.* (1981) conducted a principal component analysis of the seven-item Bortner and reported that they found only one eigenvalue greater than one. However, we reanalysed their data based on the reported correlation matrices, using principal component analysis with varimax rotation. For both men and women, three-factor solutions were found, explaining 56 per cent of the variance for men and 54 per cent for women. Based on items with high loadings, we labelled these factors time urgency, polyphasic activity and competitive/emphatic behaviour. Taken together, this evidence is inadequate to draw form conclusions regarding the dimensionality of the Bortner, though it appears that time urgency may be a central component.

Validity of the Bortner Scale

Studies of the validity of the Bortner have focused primarily on concurrent and predictive validity, with less emphasis on content and construct validity (Cronbach & Meehl, 1955). Concurrent validity has been assessed by examining the relationship between the Bortner and the SI. In the initial report, Bortner (1969) analysed data from 76 men and found concordance between the Bortner and the SI of only 64 per cent. However, Rustin, Dramaix, Kittel, Degre, Kornitzer, Thilly & DeBacker (1976) found 75 per cent agreement between the Bortner and the SI. More recently, Pichot *et al.* (1977) reported agreement between the Bortner and SI of 71.5 per cent for a sample of 130 adults and 74.9 per cent for a sample of 267 adults. Mayes *et al.* (1984) collected data from 63 women and found a correlation of .45 between the SI and the Bortner, though this relationship was higher for content ($r = .57$) than for stylistics ($r = .19$).

The predictive validity of the Bortner has been assessed in prospective studies of the relationship between the scale and symptoms of ill health, particularly CHD. The most conclusive evidence to date was presented by the French–Belgian Collaborative Group (1982), who followed 2811 factory and civil service workers for approximately six years. After controlling for standard risk factors, the Bortner was positively related to the incidence of both myocardial infarction and angina pectoris. In contrast, Koskenvuo *et al.* (1988) followed 3679 Finnish men and women for approximately three years and found no relationship between the seven-item Bortner and the incidence of ischaemic heart disease. However, because the seven-item Bortner excludes several items which describe core TABP dimensions (e.g. impatience, hard-driving, ambitious), its interpretation differs from the full 14-item Bortner. As a result, it is impossible to determine whether the null result found by Koskenvuo *et al.* (1988) indicates problems associated specifically with the seven-item Bortner or suggests an inconsistent relationship between the Bortner and CHD, similar to that found for the Framingham Scale and JAS (Booth-Kewley & Friedman, 1987; Eaker, Abbott & Kannel, 1989; Matthews, 1988). Unfortunately, no other prospective studies using the full 14-item Bortner are available to clarify this issue. Thus, available evidence provides moderate support for the concurrent and predictive validity of the Bortner, but this evidence is based on a very limited number of studies.

Summary

Available evidence raises several questions regarding the psychometric properties of the Bortner Scale. The reliability of the Bortner is apparently rather low, indicating that it reflects poorly the presumed underlying general TABP construct. This may indicate problems with the items comprising the Bortner, the presence of multiple underlying TABP dimensions concealed within the overall Bortner score, or both. The limited evidence regarding the dimensionality of the Bortner suggests that multiple dimensions may exist, though this evidence is far from conclusive. Despite these difficulties, the Bortner apparently contains some degree of concurrent and predictive validity, but available evidence is scant and somewhat mixed. From this, we may conclude that the Bortner Scale taps certain dimensions of TABP, but it remains unclear how many dimensions it taps, what these dimensions are and the degree to which these dimensions are reflected in the overall Bortner score. This study presents evidence to help clarify these issues.

Method

Sample

Data were collected from 1320 working adults from three occupations, including accountants, dentists and nurses, as described below:

1. Accountants. Baglioni, Haskins & Cooper (1988) administered a survey to 338 senior accountants in one of the top eight firms in the US and public accountants at all levels in small and medium-sized firms throughout Virginia. A total of 271 completed surveys were returned (80.1 per cent response rate). The accountants averaged 30.3 years of age and 5.7 years of experience, though the range of experience was from less than one to nearly 40 years. About 63 per cent were male and 56 per cent were married.

2. Dentists. Cooper, Watts, Baglioni & Kelly (1988) administered a survey to a random sample of 587 UK dentists. A total of 484 completed surveys were returned (82.5 per cent response rate). The median age of the dentists was about 35 years (no information on years of experience was available). Approximately 82 per cent were male and 74 per cent were married.

3. Nurses. Hingley & Cooper (1986) administered a survey to 594 female nurse managers in one UK National Health Service Authority. A total of 475 completed surveys was returned (80 per cent response rate). The nurses had a mean age of 43.7 and a median of 20 years' experience in nursing. About 60 per cent were married.

Measures

The full 14-item Bortner Scale was administered to all respondents. The version of the Bortner used in this study differed slightly from the original (Bortner, 1969) in two respects. First, to facilitate scoring, we used a numerical response format centred at zero and ascending to five in both directions (Cooper & Marshall, 1979). Secondly, we reworded one anchor on item 14, such that the opposite of 'ambitious' became 'unambitious' rather than the original 'satisfied with job'. This modification was based on our experience in earlier studies, in which participants objected to the presumed bipolarity of the original anchors (Cooper & Marshall, 1979). Respondents also completed the eight-item Anxiety, Depression and Somatic Symptom subscales from the Crown–Crisp Experiential Index (Crown & Crisp, 1966). These subscales have been validated against clinical diagnoses (Crisp, Ralph, McGuinness & Harris, 1978) and have demonstrated adequate reliability (Alderman, Mackay, Lucas, Spry & Bell, 1983). Finally, respondents completed two measures of job satisfaction. These measures were included for two reasons. First, though job satisfaction is not itself an index of mental or physical health, it has demonstrated a consistent negative relationship with indices of mental and physical symptoms, including CHD (Edwards & Cooper, 1988). Secondly, the relationship between TABP and job satisfaction itself has been a focus of several major studies of occupational stress (e.g. Cooper & Marshall, 1979). One job-satisfaction measure was a list of facet satisfaction items which differed for each occupation.* To achieve comparable indices across occupations, facet satisfaction items were averaged to create an index of overall job satisfaction and then converted to the same numeric scale. A second measure was a single general job-satisfaction item, which was the same for each occupation. Though this measure contained only a single item, it was consistent across occupations; therefore, we felt its inclusion was worth while.

* The differences in the facet satisfaction items for each sample suggest these indices may not be comparable across occupations. However, for each occupation, the sum of the facet satisfaction items was highly correlated with the single global satisfaction item (all correlations greater than .70, $p < .0001$), indicating that they served as viable alternative indicators of the same underlying construct. Furthermore, the comparability of the facet satisfaction items for each occupation is not a central issue, because the three occupations were equally represented in each sample used in the study. Therefore, any differences in results between our samples cannot be attributed to differences in the facet satisfaction measures.

Analysis

Analyses consisted of reliability estimation and both exploratory and confirmatory factor analysis. For these analyses, the total sample was divided into three random samples, with occupation and gender represented equally in each sample. The first and second samples were used to analyse the dimensionality of the Bortner, following a double cross-validation procedure (Campbell, 1976; Cudeck & Browne, 1983). The second and third samples were used to analyse the relationships between the factors underlying the Bortner and the anxiety, depression, physical symptom and job satisfaction measures, again following a double cross-validation procedure. All three samples were used for reliability estimation. After deleting cases with missing data, the final sample sizes were 393 for the first sample, 403 for the second sample and 393 for the third sample. Recent Monte Carlo studies indicate that these sample sizes are more than adequate for the statistical techniques employed (Anderson & Gerbing, 1984; Boomsma, 1982; Gerbing & Anderson, 1985; Guadagnoli & Velicer, 1988).

Results

Reliability

Two techniques were used to estimate the reliability of the Bortner. First, Cronbach's alpha was calculated (Cronbach, 1951). Though Cronbach's alpha is undoubtedly the most widely used reliability estimator, it is based on the assumption that the items comprising a scale are tau equivalent, meaning that they load equally on the hypothesized underlying construct (Gerbing & Anderson, 1988; Lord & Novick, 1968; McDonald, 1985; Nunnally, 1978). To the extent this assumption is violated, Cronbach's alpha underestimates the actual reliability of a scale (Smith, 1974). Therefore, we also applied a second estimator, omega (Heise & Bohrnstedt, 1970), which adopts the more general Spearman model by relaxing the assumption of tau equivalence (Joreskog, 1971; McDonald, 1985). If the assumption of tau equivalence is met, then omega reduces to alpha (Heise & Bohrnstedt, 1970). However, if the scale items are not tau equivalent, omega provides a better lower limit for the actual reliability of the scale (McDonald, 1985; Smith, 1974).

The results of these analyses are presented in Table 1. Assuming tau equivalence, reliability estimates were rather low, ranging from .504 to .548. Further inspection revealed that these low values were partly attributable to item 12, which was negatively correlated with over half of the remaining items for each sample. Next, we used LISREL VI (Joreskog & Sorbom, 1986)* to test the assumption of tau equivalence by estimating a single factor measurement model with each item loading unconstrained (i.e. free to take on whatever value best fits the data), re-estimating the model with item loadings constrained to be equal to one another, and examining the deterioration in fit associated with this constraint. If the unconstrained model provides a significantly better fit, the assumption of tau equivalence is rejected. Results indicated that, for all three samples, the unconstrained model provided a much better fit than the constrained model ($p < .001$). Therefore, we re-estimated the reliability of the Bortner, this time using omega. As expected, these estimates were somewhat higher than those obtained using alpha, ranging from .562 to .604. Nonetheless, both sets of estimates indicate that the reliability of the Bortner is unacceptably low (Nunnally, 1978).

* For all LISREL analyses, covariance matrices were used as input and standardized solutions are reported.

Table 1. Reliability estimates for the Bortner

	Sample 1	Sample 2	Sample 3
Alpha	.548	.527	.504
Omega	.604	.594	.562

Exploratory analysis of dimensionality

To examine the dimensionality of the Bortner, exploratory factor analyses were conducted on data from the first and second samples, using maximum-likelihood factoring with oblique rotation. For both samples, a scree test (Cattell, 1966) suggested either a two- or three-factor solution. Inspection of factor patterns revealed that the primary difference between these solutions was that the three-factor solution separated one of the factors comprising the two-factor solution. However, little information was gained by this separation, as the resulting factors were conceptually similar and several items loaded highly on both factors. Solutions with greater dimensionality were also rejected, in that each contained factors that were poorly defined and not conceptually distinct. Therefore, the two-factor solution was selected as a parsimonious but reasonably comprehensive representation of the structure of the Bortner (see Table 2). For both samples, the first factor primarily reflected speed and the second factor represented competitiveness. The correlation between these two factors was $-.030$ for the first sample (n.s.) and .118 for the second sample ($p < .05$).

The factor structures obtained from the first and second samples were then analysed with LISREL VI (Joreskog & Sorbom, 1986). This provided statistical information required to test, respecify and cross-validate the factor structures obtained for the two samples. For both samples, initial measurement models were constructed by fixing factor loadings less than .15 to zero and estimating parameters corresponding to the remaining factor loadings, measurement error for each item and the correlation between the factors. Then, factor loadings which were not significant at the .05 level were sequentially eliminated. Finally, factor loadings were sequentially added, starting with those with the largest modification indices, and stopping when the change in chi-square was no longer significant at the .05 level (Anderson & Gerbing, 1982; Sorbom, 1975). For the first sample, this procedure resulted in adding item 5 to the speed factor and adding item 1 to the competitiveness factor. For the second sample, items 5 and 11 were added to the speed factor, and item 10 was added to the competitiveness factor.

The final measurement models for the first and second samples are presented in Table 3. For both samples, the respecified measurement model was very similar to the original exploratory factor structure, and the essential interpretation of each factor remained unchanged. Furthermore, the measurement models obtained for both samples are quite similar, with 15 out of 17 loadings exhibiting similar magnitude and direction. The sole difference between the two models is that item 13 (no outside interests) loaded negatively on the competitive factor only in the first

Table 2. Exploratory factor analysis of the Bortner

	Sample 1		Sample 2	
Item	Factor 1	Factor 2	Factor 1	Factor 2
1. Never late	.742	−.112	.737	.001
2. Competitive	−.051	.803	−.098	.815
3. Anticipate	−.050	.083	.048	.079
4. Rushed	.404	.003	.544	.099
5. Impatient	.146	.256	.133	.204
6. Goes all out	.766	−.041	.809	.116
7. Do lots at once	.446	.001	.521	.089
8. Forceful	.302	.007	.236	.058
9. Wants job recognized	.537	.094	.510	.162
10. Fast	.830	.031	.787	.017
11. Hard-driving	.190	.416	.130	.464
12. Hide feelings	−.257	.051	−.286	−.071
13. No outside interests	.186	−.213	.189	−.070
14. Ambitious	−.097	.749	−.132	.701
Eigenvalue	2.74	1.52	2.86	1.51
% variance explained	19.6	10.9	20.4	10.8

sample, whereas item 10 (fast) loaded negatively on the competitive factor only in the second sample. The correlation between the two factors was not significant in either sample.

To provide summary evaluations of the measurement models derived from the first and second samples, several indices of fit were calculated (see Table 3). The chi-square for both samples was statistically significant, indicating that the reproduced covariance matrices deviated significantly from the original covariance matrices. However, chi-square is highly sensitive to sample size, such that the likelihood of obtaining a significant chi-square associated with a given model increases as sample size increases (Bentler & Bonett, 1980; Joreskog & Sorbom, 1979). To overcome this, other indices of fit that are less sensitive to sample size have been proposed. Though consensus regarding the relative merits of these indices has not been reached (Anderson & Gerbing, 1984; Marsh, Balla & McDonald, 1988; Mulaik, James, Van Alstine, Bennett, Lind & Stilwell, 1989; Wheaton, 1987), we selected five indices that are widely used and have received theoretical and empirical support in the currently available literature. These indices included the goodness-of-fit index (GFI) and adjusted goodness-of-fit index (AGFI) provided by LISREL (Joreskog & Sorbom, 1986), the normed fit index (NFI: Bentler & Bonett, 1980), the parsimonious normed fit index (PNFI: James, Mulaik & Brett, 1982; Mulaik *et al.*, 1989) and the Tucker–Lewis Index (TLI: Tucker & Lewis, 1973). Each of these indices reflects the relative amount of information in the observed variables explained by the model, but differs in its precise formula, correction for the number of parameters estimated and susceptibility to sample size fluctuations (for further details, see Anderson &

Table 3. Measurement models for the Bortner

	Sample 1		Sample 2	
Item	Factor 1	Factor 2	Factor 1	Factor 2
1. Never late	.729**	−.119**	.723**	−.122**
2. Competitive	—	.786**	—	.822**
3. Anticipate	—	—	—	—
4. Rushed	.407**	—	.544**	—
5. Impatient	.170**	.248**	.159**	.186**
6. Goes all out	.764**	—	.809**	—
7. Do lots at once	.449**	—	.520**	—
8. Forceful	.305**	—	.235**	—
9. Wants job recognized	.539**	—	.512**	—
10. Fast	.833**	—	.768**	−.115**
11. Hard-driving	.229**	.420**	.195**	.452**
12. Hide feelings	−.255**	—	−.288**	—
13. No outside interests	.166**	−.223**	.186**	—
14. Ambitious	—	.766**	—	.745**
Indices of fit				
Chi-square	300.76**		216.74**	
d.f.	73		73	
GFI	.889		.927	
AGFI	.841		.895	
NFI	.773		.836	
PNFI	.620		.670	
TLI	.770		.854	

* $p < .05$; ** $p < .01$.
Note. Paths fixed at zero are not shown. Abbreviations are explained in the text.

Gerbing, 1984; Marsh *et al.*, 1988; Mulaik *et al.*, 1989; Wheaton, 1987). Though critical values of these indices are difficult to justify (Bentler & Bonett, 1980; Marsh *et al.*, 1988; Wheaton, 1987), values of .90 or greater are often interpreted as representing adequate fit for GFI, AGFI, NFI and TLI (critical values for PNFI have not been established). Using these criteria, it appears that both models provided only marginal fit, though the model derived from the second sample provided consistently better fit than the model derived from the first sample.

Confirmatory analysis of dimensionality

Due to the lack of existing information regarding the psychometric properties of the Bortner, the procedure used to derive the measurement models described above was largely data-driven. Models resulting from such procedures are of little use unless they are successfully cross-validated (Anderson & Gerbing, 1982, 1988; Hayduk, 1987). Furthermore, the measurement models derived from the first and second samples differ somewhat and, therefore, do not suggest a single measurement model that can best describe the Bortner. To address these issues, the double cross-validation procedure outlined by Cudeck & Browne (1983) was applied. In essence, this procedure consists of estimating alternative models on one sample (the calibration sample), imposing the resulting estimates on another sample (the validation sample) and then repeating the procedure with the roles of the calibration and validation samples reversed. The model that results in the least deterioration in fit when imposed on each validation sample is then selected as superior (Cudeck & Browne, 1983). Our analyses focused on four models, including those derived separately from the first and second samples, a common model that included only those loadings that were significant in both samples (thereby dropping items 10 and 13 from the competitiveness factor), and a simple structure model formed by locating items in the common model that loaded significantly on both factors and fixing the lower loading to zero (thereby dropping items 5 and 11 from the speed factor and item 1 from the competitiveness factor). A saturated (14-factor oblique) model and a baseline (uncorrelated variables) model were also included, thereby establishing upper and lower limits for the cross-validation indices (Cudeck & Browne, 1983). Results indicated that the common model and the model derived from the first sample performed equally well, and both were superior to the simple structure model and the model derived from the second sample. Given this equivocal result, the common model was deemed superior because it contained only those item loadings that successfully cross-validated (i.e. were statistically significant) in both samples.

Exploratory analysis of concurrent validity

The concurrent validity of the Bortner was examined by analysing the relationships between its underlying constructs and four criteria, covering anxiety, depression, somatic symptoms and job satisfaction. For these analyses, the common measurement model for the Bortner was used. Measurement models for the criterion measures were also used, thereby correcting the structural coefficients for attenuation. These models were then estimated using data from the second and third samples to confirm that all item loadings were significant and in the expected direction (see Table 4). Next, item loadings and measurement errors obtained from the measurement models were incorporated as fixed parameters in the structural model. This prevented the respecification of the structural model from changing the parameter estimates in the measurement models and, thus, the interpretation of the corresponding underlying constructs (Burt, 1976). To obtain comparable results across both samples, a multigroup analysis was conducted to obtain a common set of parameters for the measurement models. This analysis indicated that the common set of parameters did not significantly differ from those obtained separately for the second and third samples, thereby suggesting that a common set of parameters adequately described the measurement models for both samples. Structural models were then estimated for both samples, starting with all eight structural parameters free, and then sequentially eliminating those that were not statistically significant.

The results of these analyses are presented in Table 5. For both samples, job satisfaction was negatively related to the speed factor and positively related to the competitiveness factor. Depression was negatively related to speed and competitiveness only for the first sample, whereas anxiety and somatic symptoms were positively related to speed only for the second sample. For both samples, the chi-square was statistically significant, and none of the remaining indices of fit reached

Table 4. Measurement model for criterion measures

	Sample 2				Sample 3			
	Anxiety	Depression	Somatic symptoms	Job satisfaction	Anxiety	Depression	Somatic symptoms	Job satisfaction
1. Upset for no reason	.482**	—	—	—	.543**	—	—	—
2. Might faint	.271**	—	—	—	.373**	—	—	—
3. Uneasy and restless	.530**	—	—	—	.589**	—	—	—
4. Panicky	.498**	—	—	—	.499**	—	—	—
5. Worrying	.587**	—	—	—	.600**	—	—	—
6. Strung up inside	.629**	—	—	—	.560**	—	—	—
7. Going to pieces	.588**	—	—	—	.584**	—	—	—
8. Bad dreams	.409**	—	—	—	.272**	—	—	—
9. Think slower than usual	—	.260**	—	—	—	.292**	—	—
10. Life is too much effort	—	.529**	—	—	—	.537**	—	—
11. Regret past behaviour	—	.337**	—	—	—	.232**	—	—
12. Wake up unusually early	—	.297**	—	—	—	.320**	—	—
13. Sadness	—	.622**	—	—	—	.694**	—	—
14. Extra effort to face crises	—	.394**	—	—	—	.388**	—	—
15. Need to cry	—	.274**	—	—	—	.533**	—	—
16. No sympathy for others	—	.263**	—	—	—	.151**	—	—
17. Dizzy or short of breath	—	—	.510**	—	—	—	.447**	—
18. Indigestion	—	—	.441**	—	—	—	.396**	—
19. Tingling sensations	—	—	.566**	—	—	—	.416**	—
20. Loss of appetite	—	—	.199**	—	—	—	.225**	—
21. Unduly tired	—	—	.472**	—	—	—	.627**	—
22. Sleep difficulties	—	—	.154**	—	—	—	.168**	—
23. Sweating or fluttering heart	—	—	.465**	—	—	—	.426**	—
24. Loss of sexual interest	—	—	.246**	—	—	—	.272**	—
25. Overall job satisfaction	—	—	—	.864**	—	—	—	.897**
26. General job satisfaction	—	—	—	.864**	—	—	—	.897**
Indices of fit								
Chi-square	—	—	737.42**	—	—	—	559.31**	—
d.f.	—	—	294	—	—	—	294	—
GFI	—	—	.864	—	—	—	.898	—
AGFI	—	—	.838	—	—	—	.878	—
NFI	—	—	.674	—	—	—	.753	—
PNFI	—	—	.610	—	—	—	.682	—
TLI	—	—	.747	—	—	—	.849	—

* $p < .05$; ** $p < .01$.

Note. To achieve identification, loadings on the job satisfaction factor were set equal to one another. For both samples, the anxiety, depression and somatic symptom factors were positively correlated with one another ($p < .01$). For the second sample, job satisfaction was negatively related to anxiety and somatic symptoms ($p < .05$).

Table 5. Structural models for the Bortner

Criterion	Sample 1		Sample 2	
	Speed	Competitiveness	Speed	Competitiveness
1. Job satisfaction	−.550**	.324**	−.553**	.256**
2. Anxiety	—	—	.112**	—
3. Somatic symptoms	—	—	.123*	—
4. Depression	−.147**	−.222**	—	—
Indices of fit				
Chi-square	1727.50**		1580.00	
d.f.	721		721	
GFI	.812		.828	
AGFI	.786		.804	
NFI	.619		.629	
PNFI	.572		.581	
TLI	.710		.733	
RNFI	.998		.965	
All structural parameters	225.36**		137.42**	

* $p < .05$; ** $p < .01$.
Note. Paths fixed at zero are now shown.

the .90 threshold, indicating that the models did not fit the data well. However, these indices reflect the combined fit of the structural and measurement models and, therefore, do not provide a specific assessment of the structural models under consideration. To overcome this, the relative normed fit index (RNFI: Mulaik *et al.*, 1989) was calculated. RNFI reflects the relative fit of the structural model independently of the measurement model. For both samples, RNFI exceeded .96, indicating that the derived structural models accounted for nearly all of the covariance among the latent variables. Furthermore, the simultaneous test of all structural parameters was highly significant for both samples ($p < .001$). Taken together, these results provide strong support for the structural models derived from the second and third samples.

Confirmatory analysis of validity

As with the measurement models derived from the first and second samples, the procedure used to derive the structural models for the second and third samples was largely data-driven, and several differences emerged between the resulting models. Therefore, the Cudeck & Browne (1983) double cross-validation procedure was again applied. Our analyses focused on three models, including those derived from the second and third samples, and a common model that included only those parameters that were significant in both samples (thereby only including structural parameters from speed and competitiveness to job satisfaction). The saturated model included all eight structural parameters, and the baseline model eliminated all eight parameters. Results indicated that the model derived from the third sample was slightly superior to the common model and the model derived from the second sample, though the differences among these models were not great.

Discussion

Summary of results

Our results regarding the reliability, dimensionality and concurrent validity of the Bortner Scale may be summarized as follows. First, assuming equal item weightings, the reliability of the Bortner is approximately .53, though assigning differential weightings to reflect the relative contribution of each item to the overall scale increases the reliability to approximately .59. In either case, the reliability of the Bortner falls below typical standards for self-report measures (Nunnally, 1978). Secondly, the Bortner apparently contains two distinct, independent dimensions, one representing speed and the other representing competitiveness. Though item loadings differed somewhat across samples, the essential nature of these dimensions remained consistent. Comparing these results with earlier studies reveals that the speed factor in our data subsumes the time urgency and polyphasic activity factors in the Koskenvuo *et al.* (1981) data and the time urgency and expressiveness factors found by Pichot *et al.* (1977). Similarly, the hard-driving and competitiveness factors found by Pichot *et al.* (1977) are contained within a single competitiveness factor in our data. Thus, the two-factor solution we obtained provides a somewhat more parsimonious representation of the Bortner than those obtained in earlier studies. Thirdly, the speed and competitiveness dimensions comprising the Bortner were significantly related to validity criteria, though these relationships differed across dimensions and across samples. In both samples, the speed dimension was negatively related to job satisfaction, whereas the competitiveness dimension was positively related to job satisfaction. However, the speed dimension was positively related to anxiety and somatic symptoms only in the second sample, but negatively related to depression in the third sample. Furthermore, the competitiveness dimension was unrelated to anxiety, somatic symptoms and depression in the second sample, but negatively related to depression in the third sample. Cross-validation suggested that the parameters linking the speed dimension to anxiety and somatic symptoms should probably be retained, but the magnitude of these parameters was rather small. The instability of the results involving the anxiety, somatic symptoms and depression measures was probably influenced by the general absence of severe clinical symptoms in our samples, which produced skewed distributions for these measures. When this occurs, the resulting parameter estimates and standard errors must be interpreted very cautiously (Joreskog & Sorbom, 1986).

It is interesting to note that the speed and competitiveness dimensions consistently exhibited opposite relationships with job satisfaction. One possible explanation for this pattern of results is that the hurried, rushed behaviours associated with the speed dimension may actually hinder the achievement of job-related goals, such as monetary rewards, status and promotion, thus decreasing job satisfaction. In contrast, the ambitious, hard-driving behaviours associated with the competitiveness dimension may facilitate the achievement of job-related goals, thus enhancing job satisfaction. Unfortunately, because the data used in this study were cross-sectional, we cannot rule out explanations based on reverse causality. For example, individuals who are satisfied with their jobs may adopt a less hurried and frenetic approach to their work. Analogously, individuals who are satisfied may view their jobs as energizing and, therefore, engage in ambitious, competitive and hard-driving behaviours. Though investigations into these explanations would be illuminating, it should be emphasized that the relationships found in this study pertain to TABP *as measured by the Bortner*, and the generalizability of these relationships must be assessed using other TABP measures that tap similar dimensions. Furthermore, this study did not include hard measures of cardiovascular functioning, which constitute the most relevant criteria regarding the concurrent and predictive validity of TABP measures. Because of these shortcomings, additional research is needed to reveal more about the process by which specific TABP dimensions, including those measured by the Bortner, are related to psychological and physical symptoms.

Implications for the use of the Bortner Scale

Our results suggest several implications regarding the use of the Bortner Scale. First, as typically used, the Bortner should be viewed as primarily an index of speed and, to a lesser extent, competitiveness. These results indicate some degree of overlap between the Bortner and other TABP measures, particularly the S and H subscales of the JAS (Zyzanski & Jenkins, 1970). However, the Bortner apparently contains little specific information regarding certain TABP dimensions contained in other measures, such as hostility, aggressiveness, job involvement and negative affect (Dembroski, MacDougall, Shields, Petitto & Lushene, 1978; Jenkins, Rosenman & Friedman, 1966; Matthews, 1983; Matthews, Glass, Rosenman & Bortner, 1977; Zyzanski & Jenkins, 1970). This limited overlap may partially explain the modest correlations typically found between the Bortner and other TABP measures (Byrne, Rosenman, Schiller & Chesney, 1985; Johnston & Shaper, 1983; Mayes *et al.*, 1984; Price & Clarke, 1978). However, these modest correlations are also attributable to the low reliability of the Bortner, which attenuates its relationship with other variables, including TABP measures. In any case, these results imply that the Bortner should *not* be considered as interchangeable with other TABP measures, but instead should be viewed as a measure of specific TABP components.

Secondly, given the distinct dimensions contained within the Bortner and their differential relationships with outcomes, scores based on a simple summation of all 14 Bortner items should be abandoned in favour of subscale scores reflecting each underlying dimension. To maintain a simple structure and enhance interpretability, each item should be assigned to only one subscale (Burt, 1976; Gerbing & Anderson, 1988). For items with double loadings, content should be the primary criterion for assignment (Gerbing & Anderson, 1988). Based on this procedure, a Bortner Speed subscale may be formed by summing items 1, 4, 5, 6, 7, 8, 9, 10, 12 (after reversal) and 13. Similarly, a Bortner Hard-driving and Competitiveness subscale may be formed by summing items 2, 11 and 14. Using omega, we obtained reliability estimates of approximately .75 for the Speed subscale and .72 for the Competitiveness subscale, both of which are substantially higher than those obtained for the overall Bortner. Dropping items with loadings less than .4 from the Speed subscale (thereby retaining items 1, 4, 6, 7, 9 and 10) raised its reliability to approximately .81.

Thirdly, the response format of the Bortner raises certain issues. First, we used a modified response format which, unlike the unsegmented line originally used by Bortner (1969), contained a Likert scale centred at zero and ascending to five in both directions. Strictly speaking, our findings should be generalized only to versions of the Bortner using the same response format. The generalizability of our findings to versions of the Bortner using other response formats, including the original presented by Bortner (1969), awaits empirical verification. A more fundamental question regarding the response format of the Bortner is whether each item, in fact, represents opposite ends of a continuum, as implicitly assumed by Bortner (1969). We have three reasons to question this assumption. First, a handful of our respondents circled numbers on *both* sides of the centred zero, implying that they simultaneously agreed with descriptors at both ends of a single continuum. For example, item 1 is anchored by 'never late' and 'casual about appointments'. It is not difficult to imagine a person who calmly arrives early for meetings and is, therefore, *both* never late *and* casual about appointments. Secondly, our experience using the Bortner with executives has revealed objections regarding the presumed bipolarity of the continua. Thirdly, simply examining the content of the items raises doubts regarding bipolarity, and item 1 again stands out in this regard. The bipolarity of the items comprising the Bortner may be empirically evaluated by separating each item into two questions, with each question addressing the extent to which the characteristic or behaviour at one end of the continuum exists. If pairs of questions derived from the original items are negatively correlated, then bipolarity may be assumed; otherwise, the questions should remain separate. In any case, it is generally safer *not* to assume bipolarity, instead asking separate questions for the descriptors at each end of the present Bortner items and subsequently letting the data determine whether, in fact, the original pairs of descriptors are negatively correlated.

In sum, the Bortner Scale contains several serious limitations, including low reliability, the presence of multiple underlying dimensions and the unsubstantiated assumption of item bipolarity. Based on these limitations, we recommend that the Bortner Scale should no longer be used in its present form. At a minimum, only subscales based on the dimensions underlying the Bortner should be used, based on the specifications presented above. Though these subscales are likely to demonstrate improved reliability and unidimensionality over the original Bortner, the Competitiveness subscale contains only three items, and both subscales still rely on the unsubstantiated assumption of item bipolarity. For these reasons, it is difficult to recommend the use of the Bortner Scale in any form.

Implications for the measurement of Type A behaviour pattern

The results of the present study further support the notion that TABP is a multidimensional construct (Dembroski *et al.*, 1978; Jenkins, Zyzanski & Rosenman, 1978; Zyzanski & Jenkins, 1970). Furthermore, these dimensions may have differential relationships with criterion measures. For the Bortner, the speed dimension was negatively related to job satisfaction and, to a lesser extent, positively related to anxiety and somatic symptoms, whereas the competitiveness dimension was positively related to job satisfaction. Therefore, summary indices of TABP may inappropriately combine distinct dimensions having differential relationships with outcomes, thus concealing important information. For this reason, we recommend that summary measures of TABP should be abandoned in favour of measures specifically directed towards the hypothesized underlying components of TABP. By doing this, researchers will achieve more valid assessments of the dimensions of TABP and further clarify the relationships between these dimensions and other relevant variables.

Acknowledgement

This research was supported in part by the Darden School Sponsors.

References

Alderman, K. J., Mackay, C. J., Lucas, E. G., Spry, W. B. & Bell, B. (1983). Factor analysis and reliability of the Crown–Crisp Experiental Index (CCEI). *British Journal of Medical Psychology*, **56**, 329–345.

Anderson, J. C. & Gerbing, D. W. (1982). Some methods for respecifying measurement models to obtain unidimensional construct measurement. *Journal of Marketing Research*, **19**, 453–460.

Anderson, J. C. & Gerbing, D. W. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika*, **49**, 155–173.

Anderson, J. C. & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, **103**, 411–423.

Baglioni, A. J., Jr, Haskins, M. & Cooper, C. L. (1988). A model of the sources, moderators, and psychological symptoms of stress among audit seniors. Unpublished manuscript, Colgate Darden Graduate School of Business Administration, University of Virginia.

Bass, C. (1984). Type A behaviour in patients with chest pain: Test–retest reliability and psychometric correlation of the Bortner scale. *Journal of Psychosomatic Research*, **28**, 289–300.

Bass, C. & Akhras, F. (1987). Physical and psychological correlates of severe heart disease in men. *Psychological Medicine*, **17**, 695–703.

Bass, C. & Wade, C. (1982). Type A behaviour: Not specifically pathogenic. *Lancet*, **10**, 1147–1149.

Bentler, P. M. & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, **88**, 588–606.

Boomsma, A. (1982). The robustness of LISREL against small sample sizes in factor analysis models. In K. G. Joreskog & H. Wold (Eds), *Systems under Indirect Observation: Causality, Structure, Prediction*, pp. 149–173. New York: North-Holland.

Booth-Kewley, S. & Friedman, H. S. (1987). Psychological predictors of heart disease: A quantitative review. *Psychological Bulletin*, **101**, 343–362.

Bortner, R. W. (1969). A short rating scale as a potential measure of pattern A behavior. *Journal of Chronic Diseases*, **22**, 87–91.

Bortner, R. W. & Rosenman, R. H. (1967). The measurement of pattern A behavior. *Journal of Chronic Diseases*, **20**, 525–533.

Burt, R. S. (1976). Interpretational confounding of unobserved variables in structural equation models. *Sociological Methods and Research*, **5**, 3–52.

Byrne, D. G., Rosenman, R. H., Schiller, E. & Chesney, J. A. (1985). Consistency and variation among instruments purporting to measure the Type A Behavior Pattern. *Psychosomatic Medicine*, **47**, 242–261.

Campbell, J. P. (1976). Psychometric theory. In M. Dunnette (Ed.), *Handbook of Industrial and Organizational Psychology*, pp. 185–222. Chicago: Rand McNally.

Cattell, R. B. (1966). The scree test for the number of factors. Multivariate Behavioral Research, **1**, 946–976.

Cooper, C. L. & Marshall, J. (1979). *Executives under Pressure: A Psychological Study*. New York: Praeger.

Cooper, C. L., Watts, J., Baglioni, A. J., Jr & Kelly, M. (1988). Occupational stress amongst general practice dentists. *Journal of Occupational Psychology*, **61**, 163–174.

Cottier, C., Adler, R., Vorkauf, H., Gerber, R., Hefer, T. & Hurny, C. (1983). Pressured pattern or Type A behavior in patients with peripheral arteriovascular disease: Controlled retrospective exploratory study. *Psychosomatic Medicine*, **45**, 187–193.

Crisp, A. H., Ralph, P. C., McGuinness, B. & Harris, G. (1978). Psychoneurotic profiles in the adult population. *British Journal of Medical Psychology*, **51**, 293–301.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, **16**, 297–334.

Cronbach, L. J. & Meehl, P. C. (1955). Construct validity in psychological tests. *Psychological Bulletin*, **52**, 281–302.

Crown, S. & Crisp, A. H. (1966). A short clinical diagnostic self-rating scale for psychoneurotic patients. *British Journal of Psychiatry*, **112**, 917–923.

Cudeck, R. & Browne, M. W. (1983). Cross-validation of covariance structures. *Multivariate Behavioral Research*, **18**, 147–167.

Dembroski, T. M., MacDougall, J. M., Shields, J. L., Petitto, J. & Lushene, R. (1978). Components of the Type A coronary-prone behavior pattern and cardiovascular responses to psychomotor performance challenge. *Journal of Behavioral Medicine*, **1**, 159–176.

Eaker, E. D., Abbott, R. D. & Kannel, W. B. (1989). Frequency of uncomplicated angina pectoris in Type A compared with Type B persons (the Framingham study). *American Journal of Cardiology*, **63**, 1042–1045.

Edwards, J. R. & Cooper, C. L. (1988). The impacts of positive psychological states on physical health: A review and theoretical framework. *Social Science and Medicine*, **27**, 1447–1459.

French–Belgian Collaborative Group (1982). Ischemic heart disease and psychological patterns. *Advances in Cardiology*, **29**, 25–31.

Friedman, M. & Rosenman, R. H. (1959). Association of specific overt behavior pattern with increases in blood cholesterol, blood clotting time, incidence of arcus senilis and clinical coronary artery disease. *Journal of the American Medical Association*, **169**, 1286–1296.

Gerbing, D. W. & Anderson J. C. (1985). The effects of sampling error and model characteristics on parameter estimation for maximum likelihood confirmatory factor analysis. *Multivariate Behavioral Research*, **20**, 255–272.

Gerbing, D. W. & Anderson, J. C. (1988). An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research*, **25**, 186–192.

Guadagnoli, E. & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, **103**, 265–275.

Hayduk, L. A. (1987). *Structural Equation Modeling with LISREL*. Baltimore, MD: Johns Hopkins University Press.

Haynes, S. G., Feinleib, M. & Kannel, W. B. (1980). The relationship of psychosocial factors to coronary heart disease in the Framingham study: III. Eight-year incidence of coronary heart disease. *American Journal of Epidemiology*, **111**, 37–58.

Haynes, S. G., Levine, S., Scotch, N., Feinleib, M. & Kannel, W. B. (1978). The relationship of psychosocial factors to coronary heart disease in the Framingham study: I. Methods and risk factors. *American Journal of Epidemiology*, **107**, 362–383.

Heise, D. R. & Bohrnstedt, G. W. (1970). Validity, invalidity, and reliability. In E. F. Borgatta & G. W. Bohrnstedt (Eds), *Sociological Methodology*, pp. 104–129. San Francisco: Jossey-Bass.

Heller, R. F. (1979). Type A behaviour and coronary heart disease. *British Medical Journal*, **2**, 368.

Hingley, P. & Cooper, C. L. (1986). *Stress and the Nurse Manager*. New York: Wiley.

James, L. R., Mulaik, S. A. & Brett, J. M. (1982). *Causal Analysis: Assumptions, Models and Data*. Beverly Hills, CA: Sage.

Jenkins, C. D., Rosenman, R. H. & Friedman, M. (1966). Components of the coronary-prone behavior pattern: Their relation to silent myocardial infarction and blood lipids. *Journal of Chronic Diseases*, **19**, 599–609.

Jenkins, C. D., Rosenman, R. H. & Friedman, M. (1967). Development of an objective psychological test for the determination of the coronary-prone patterns in employed men. *Journal of Chronic Diseases*, **20**, 371–379.

Jenkins, C. D., Zyzanski, S. J. & Rosenman, R. H. (1978). Coronary-prone behaviour: One pattern or several? *Psychosomatic Medicine*, **40**, 25–43.

Johnston, D. W. & Shaper, A. G. (1983). Type A behavior in British men: Reliability and intercorrelation of two measures. *Journal of Chronic Diseases*, **36**, 203–207.

Joreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika*, **36**, 109–133.

Joreskog, K. G. & Sorbom, D. (1979). *Advances in Factor Analysis and Structural Equation Models*. Cambridge, MA: ABT Books.

Joreskog, K. G. & Sorbom, D. (1986). *LISREL VI*. Chicago: National Educational Resources.

Kornitzer, M., Magotteau, V., Degre, C., Kittel, F., Struyven, J. & Van Thiel, E. (1982). Angiographic findings and the Type A behavior pattern assessed by means of the Bortner Scale. *Journal of Behavioral Medicine*, **5**, 313–320.

Koskenvuo, M., Kaprio, J., Langinvainia, H., Romo, M. & Sarna, S. (1981). Psychosocial and environmental correlates of coronary-prone behavior in Finland. *Journal of Chronic Diseases*, **34**, 331–340.

Koskenvuo, M., Kaprio, J., Rose, R. J., Kesaniemik, A., Sarna, S., Heikkila, K. & Langinvainia, H. (1988). Hostility as a risk factor for mortality and ischemic heart disease in men. *Psychosomatic Medicine*, **50**, 330–340.

Lord, F. M. & Novick, M. R. (1968). *Statistical Theories of Mental Test Scores*. Reading, MA: Addison-Wesley.

Marsh, H. W., Balla, J. R. & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*, **103**, 391–410.

Matthews, K. A. (1983). Assessment issues in coronary-prone behavior. In T. M. Dembroski, T. H. Schmidt & G. Blumchen (Eds), *Biobehavioral Bases of Coronary Heart Disease*, pp. 62–78. New York: Karger.

Matthews, K. A. (1988). Coronary heart disease and Type A behaviors: Update on an alternative to the Booth-Kewley and Friedman (1987) quantitative review. *Psychological Bulletin*, **104**, 373–380.

Matthews, K. A., Glass, D. C., Rosenman, R. H. & Bortner, R. W. (1977). Competitive drive pattern A, and coronary heart disease: A further analysis of some data from the Western Collaborative Group Study. *Journal of Chronic Diseases*, **30**, 489–498.

Mayes, B. T., Sime, W. E. & Ganster, D. C. (1984). Convergent validity of Type A behavior pattern scales and their ability to predict physiological responsiveness in a sample of female public employees. *Journal of Behavioral Medicine*, **7**, 83–108.

McDonald, R. P. (1985). *Factor Analysis and Related Methods*. Hillsdale, NJ: Erlbaum.

Mulaik, S. A., James, L. R., Van Alstine, J., Bennett, N., Lind, S. & Stilwell, C. D. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin*, **105**, 430–445.

Nunnally, J. C. (1978). *Psychometric Theory*. New York: McGraw-Hill.

Pichot, P., De Bonis, M., Somogyi, M., Degre-Coustry, C., Kittel-Bossuit, F., Rustin-Vandenhende, R. M., Dramaix, M. & Bernet, A. (1977). Étude metrologique d'une batterie de tests destinée à l'étude des facteurs psychologiques en epidemiologie cardio-vasculaire. *International Review of Applied Psychology*, **26**, 11–19.

Price, K. P. & Clarke, L. K. (1978). Behavioural and psychophysiological correlates of the coronary-prone personality: New data and an unanswered question. *Journal of Psychosomatic Research*, **22**, 409–417.

Ray, J. J. & Bozek, R. (1980). Dissecting the A–B personality type. *British Journal of Medical Psychology*, **53**, 181–186.

Rosenman, R. H. (1978). The interview method of assessment of the coronary-prone behavior pattern. In T. M. Dembroski, S. M. Weiss, J. L. Shields, S. G. Haynes & M. Feinleib (Eds), *Coronary-prone Behavior*, pp. 55–69. New York: Springer-Verlag.

Rosenman, R. H., Brand, R. J., Jenkins, C. D., Friedman, M., Straus, R. & Wurm, M. (1975). Coronary heart disease in the Western Collaborative Group Study: Final follow-up experience of $8\frac{1}{2}$ years. *Journal of the American Medical Association*, **233**, 872–877.

Rustin, R. M., Dramaix, M., Kittel, F., Degre, C., Kornitzer, M., Thilly, C. & DeBacker, G. (1976). Evaluation of techniques used to define Type A pattern in the Belgian Prevention Project of cardiovascular diseases. *Revue Epidemiologie et Santé Publique*, **24**, 497–507.

Smith, K. W. (1974). On estimating the reliability of composite indexes through factor analysis. *Sociological Methods and Research*, **2** (May), 485–510.
Sorbom, D. (1975). Detection of correlated errors in longitudinal data. *British Journal of Mathematical and Statistical Psychology*, **28**, 138–151.
Tucker, L. R. & Lewis, C. (1973). The reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, **38**, 1–10.
Wheaton, B. (1987). Assessment of fit in overidentified models with latent variables. *Sociological Methods and Research*, **16**, 118–154.
Zyzanski, S. J. & Jenkins, C. D. (1970). Basic dimensions within the coronary-prone behavior pattern. *Journal of Chronic Diseases*, **22**, 781–795.

Received 23 August 1989; revised version received 18 January 1990