Journal of Applied Psychology
1990, Vol. 75, No. 4, 440–454

0021-9010/90/$00.75

Examining the Relationships Among Self-Report Measures of the Type A Behavior Pattern: The Effects of Dimensionality, Measurement Error, and Differences in Underlying Constructs

Jeffrey R. Edwards and A. J. Baglioni, Jr.
Colgate Darden Graduate School of Business Administration
University of Virginia

Cary L. Cooper
Manchester School of Management
University of Manchester Institute of Science and Technology
Manchester, England

The most widely used self-report measures of the Type A behavior pattern (TABP) are the Bortner scale, the Framingham scale, and the Jenkins Activity Survey (JAS). Though high scores on each of these measures have been linked to the development of coronary heart disease, their intercorrelations are rather low, suggesting that they may reflect different aspects of TABP. This study indicates that the low correlations among the Bortner scale, the Framingham scale, and the JAS are due not only to differences in underlying constructs but also to measurement error and multidimensionality. These results also identify several psychometric problems, which raise serious questions regarding the use of these measures in TABP research. Suggestions for the development of new measures of TABP are offered.

Extensive research has been conducted on the role of the Type A behavior pattern (TABP) in the development of coronary heart disease (CHD). TABP was initially identified by Friedman and Rosenman (1959), who noted that patients with more severe CHD tended to exhibit competitiveness, time urgency, aggressiveness, and hostility. Since this initial work, numerous studies have examined the relationship between TABP and the prevalence and incidence of CHD (e.g., French–Belgian Collaborative Group, 1982; Haynes, Feinleib, & Kannel, 1980; Rosenman et al., 1975). Though somewhat mixed, available evidence indicates a modest relationship between TABP and the development of CHD (Booth-Kewley & Friedman, 1987; Matthews, 1988).

Though there is general consensus regarding the importance of TABP in the development of CHD, there is considerable controversy regarding its exact meaning and measurement (Matthews, 1983). TABP was originally measured with the Structured Interview (SI; Rosenman, 1978). Though the SI is generally accepted as a valid measure of TABP (Rosenman et al., 1975; Rosenman, 1978), it is impractical for use in large-scale survey research. As a result, various self-report measures of TABP have been developed. Among these, the Bortner scale (Bortner, 1969), the Framingham scale (Haynes, Levine, Scotch, Feinleib, & Kannel, 1978), and the Jenkins Activity Survey (JAS; Jenkins, Rosenman, & Friedman, 1967) each compare favorably with the SI and, in at least one major study, have been linked to the development of CHD (Bortner, 1969; French–Belgian Collaborative Group, 1982; Haynes et al., 1980; Rosenman et al., 1975).

Though the Bortner scale, the Framingham scale, and the JAS each purport to measure TABP, their intercorrelations are rather low, rarely exceeding .60 (Byrne, Rosenman, Schiller, & Chesney, 1985; Haynes et al., 1980; Johnston & Shaper, 1983; Mayes, Sime, & Ganster, 1984; Price & Clarke, 1978). This low intercorrelation calls into question the use of these measures as alternative indicators of TABP, hindering the generalization and accumulation of findings across studies. Of the various possible reasons for the low correlations, three are most likely. First, these measures may simply tap different underlying constructs, perhaps reflecting different aspects of TABP (Matthews, 1982, 1983). Second, these measures may contain substantial measurement error, which would attenuate their correlations with one another. Third, these measures may contain multiple dimensions, some of which are shared across measures and others of which are unique to a given measure. Collapsing these unique dimensions with the shared dimensions would reduce the correlations among the overall measures. Unfortunately, there is little valid evidence concerning the dimensionality, measurement error, and underlying constructs associated with the Bortner scale, the Framingham scale, and the JAS. As a result, it is impossible to determine the degree to which these factors contribute to the low correlations among these mea-

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

We gratefully acknowledge the Darden School Executive Program staff for their cooperation and Carol Steele LeMay for her assistance in data base management.

Correspondence concerning this article should be addressed to Jeffrey R. Edwards, Colgate Darden Graduate School of Business Administration, University of Virginia, P. O. Box 6550, Charlottesville, Virginia 22906–6550.

sures, which hinders their interpretation and provides little basis for developing new measures with a higher degree of convergence.

The purpose of this article is to determine the degree to which the low correlations among the Bortner scale, the Framingham scale, and the JAS are attributable to differences in underlying constructs, measurement error, and multidimensionality. We first review and critique available information about these factors and then empirically examine their extent and effects on the correlations among the Bortner scale, the Framingham scale, and the JAS. We conclude with recommendations concerning the use of these measures and the development of new measures of TABP.

Interrelationships Among the Bortner Scale, the Framingham Scale, and the JAS: A Review of the Evidence

Differences in Underlying Constructs

As stated earlier, one explanation for the low correlations among the Bortner scale, the Framingham scale, and the JAS is that they simply reflect different underlying constructs (Matthews, 1982, 1983). The strongest evidence regarding the constructs underlying these measures derives from their content validity, or the degree to which their items constitute a representative sample from a clearly defined domain of content (Cronbach & Meehl, 1955; Nunnally, 1978).

Bortner (1969) defined TABP in terms of hard-driving, ambitious, and time-conscious behavior. The procedure used to generate the Bortner items is not described, however, making it difficult to determine whether a domain-sampling procedure was used.

The definition of TABP underlying the Framingham measure involved aggressiveness, ambition, competitive drive, impatience, chronic time urgency, and a high need for achievement (Haynes et al., 1978, 1980). To construct the scale, a pool of 300 items was assembled from various sources, and three experts selected items that measured "behavior patterns initially conceptualized in the design of the questionnaire" (Haynes et al., 1978, p. 364). Though this suggests a domain sampling procedure, the relative emphasis placed on different aspects of TABP is not described, making it difficult to interpret the overall scale score.

The definition of TABP underlying the JAS reflected multiple dimensions, including competitiveness, achievement striving, aggressiveness, haste, impatience, restlessness, hyperalertness, explosive speech, muscular tension, external time pressure, challenging responsibility, and job commitment (Jenkins, Zyzanski, & Roseman, 1979). The items constituting the JAS, however, were apparently not designed to represent this domain of content but instead to mimic the SI (Jenkins et al., 1967, 1979), which itself was designed to assess behaviors prevalent in coronary patients (Rosenman, 1978). Subsequent revisions of the JAS retained items that predicted either the SI (Jenkins, Zyzanski, & Rosenman, 1971) or earlier versions of the JAS itself (Jenkins, 1978) and added items developed through "clinical and psychometric experience with the Type A construct" (Jenkins et al., 1979, p. 6). Thus, the development of the Bortner scale, the Framingham scale, and the JAS suggests little emphasis on content validity and therefore provides minimal information regarding the constructs underlying these measures.

When evidence regarding content validity is lacking, item content is often examined post hoc and, on the basis of relevant theory, meaning is assigned to a measure. Though few studies have systematically examined the content of the Bortner scale, the Framingham scale, and the JAS, indirect evidence is provided by factor analyses, in which item content is used to interpret the obtained factor structure.[1] Factor analyses of the Bortner scale suggest that it primarily reflects speed and time urgency, with a secondary emphasis on competitiveness (Edwards, Baglioni, & Cooper, in press; Pichot et al., 1977). An unpublished factor analysis of the Framingham scale has been mentioned in the literature (Chesney, Black, Chadwick, & Rosenman, 1981; MacDougall, Dembroski, & Musante, 1979); it purportedly emphasizes competitiveness and time pressure.

The original factor analysis of the JAS (Zyzanski & Jenkins, 1970) identified three factors, labeled Speed and Impatience (S), Job Involvement (J), and Hard-Driving and Competitive (H). Weighted composites derived from these factors were subsequently used to construct the corresponding JAS scales, which have been added to the global TABP (AB) scale as standard features of the JAS scoring procedure (Jenkins et al., 1979). Despite their widespread use, the interpretation of these scales has been challenged by Begley and Boyd (1985), who factor analyzed the JAS and concluded that, contrary to their labels, the S, J, and H scales reflect hurry and anger, job demands, and comparisons with other workers, respectively. Begley and Boyd (1985) also factor analyzed the JAS–AB scale and concluded that it reflects hard-driving behavior and competitiveness and, to a lesser extent, hurrying. Unfortunately, interpretations of the JAS scales based on factor analysis are prone to error because they typically overlook the effects of the item weighting procedure, which heavily favors certain items and, in some cases, assigns opposite weights to items with similar content. As we will demonstrate, this procedure substantially affects the interpretation of the overall scale scores.

In sum, available evidence suggests several differences in content among the Bortner scale, the Framingham scale, and the JAS, thereby supporting the notion that they measure different underlying constructs. This evidence is drawn indirectly from factor analyses rather than from systematic examinations of the measures, however, and, in the case of the JAS, overlooks the effects of item weights on total scale scores.

Measurement Error

A second potential reason for the low correlations among the Bortner scale, the Framingham scale, and the JAS is the presence of measurement error. Most techniques for estimating measurement error rely on estimates of true score variance, drawing from the classical measurement theory assumption that an observed score is determined by an underlying true

[1] A complete review of factor analyses of the Bortner, Framingham, and JAS measures is provided later in this article, where we discuss the dimensionality of these measures.

score and measurement error (Lord & Novick, 1968). It follows that by estimating the proportion of true score variance, one may indirectly obtain an estimate of error variance. Studies of the Bortner scale, the Framingham scale, and the JAS have typically estimated true score variance with Cronbach's alpha (Cronbach, 1951). Overall, these studies suggest that the measures contain substantial measurement error. Reliability estimates have ranged from .50 to .68 for the Bortner scale (Bortner, 1969; Edwards et al., in press; Mayes et al., 1984) and from .54 to .71 for the Framingham scale (Haynes et al., 1978; Lee, King, & King, 1987). Reliability estimates ranging from .73 to .85 have been reported for the JAS scales by Jenkins et al. (1979), but Mayes et al. (1984) noted that, rather than using alpha, Jenkins et al. (1979) used a technique based on test–retest and item–total correlations, which tends to overestimate reliability. Using alpha, Mayes et al. (1984) reported reliability estimates for the JAS scales ranging from .24 to .56. Similarly, Shipper, Kreitner, Reif, and Lewis (1986) reported an alpha of .56 for the JAS–AB but noted that this estimate could be raised to .60 by deleting three items with negative item–total correlations.

Overall, the evidence cited above suggests that measurement error is moderate for the Framingham and Bortner measures and substantial for the JAS. This evidence should be considered tentative for two reasons. First, it is drawn from a small number of studies, some of which relied on very small samples (Bortner, 1969; Mayes et al., 1984). Second, most studies used alpha to estimate reliability. Despite its widespread use, alpha rests on several assumptions that are often overlooked in practice, two of which are particularly relevant here. First, alpha assumes that test items are tau equivalent, meaning that they load equally on the common underlying construct (Novick & Lewis, 1967). When this assumption is violated, alpha underestimates reliability. Second, alpha assumes that the items constituting a measure are congeneric, meaning that they reflect a single underlying construct (Jöreskog, 1971; Nunnally, 1978). When this assumption is violated, alpha cannot be readily interpreted as an index of true score variance (Campbell, 1976). Unfortunately, none of the studies reviewed here reported whether these assumptions had been met. As we will show, these assumptions are violated by the Bortner scale, the Framingham scale, and the JAS, often to a considerable extent.

Dimensionality

A third potential explanation for the low correlations among the Bortner scale, the Framingham scale, and the JAS is that these measures actually contain multiple dimensions, some of which are common across measures and others of which are unique to a given measure. Though few studies have examined the dimensionality of the Bortner scale, the Framingham scale, and the JAS, available evidence suggests that they each contain multiple dimensions. Most of this evidence is based on exploratory factor analysis.

For example, Pichot et al. (1977) identified four factors in the Bortner scale reflecting time urgency, hard-driving behavior, competitiveness, and expressiveness. In contrast, Edwards et al. (in press) identified two factors representing speed and competitiveness. Johnston and Shaper (1983) attempted to factor analyze the Bortner scale but reported that they could not find a stable solution. The Framingham scale has apparently been factor analyzed (Chesney et al., 1981; MacDougall et al., 1979), but the dimensionality of the obtained solution has not been described.

As indicated earlier, Zyzanski and Jenkins (1970) identified three factors in the JAS, which were used to construct the S, J, and H scales currently in use. A similar factor structure was found by Waldron, Zyzanski, Shekelle, Jenkins, and Tannanbaum (1977), but only for White respondents. Shipper et al. (1986) factor analyzed the JAS and found 19 factors with eigenvalues greater than 1, but these factors were apparently not rotated or interpreted. Begley and Boyd (1985) found five factors, three reflecting the original S, J, and H factors, one concerning comparisons with the average worker, and another reflecting eating too fast. Begley and Boyd also summarized the results of previous factor analyses and concluded that the S, J, and H factors each contained about 10 items consistently loading over .30, which is far fewer than the 20 to 24 items contained in the corresponding scales. The JAS–AB scale has been factor analyzed in two studies. Shipper et al. found eight factors with eigenvalues greater than 1, but they apparently did not rotate or interpret this solution. Begley and Boyd found two factors, one reflecting hard-driving behavior and competitiveness and another reflecting eating too fast. To our knowledge, the dimensionality of the S, J, and H scales has not been examined.

This evidence suggests that the full 52-item JAS contains at least three dimensions, whereas the dimensionality of the Bortner scale, the Framingham scale, and the JAS subscales remains unclear. Because this evidence is based almost exclusively on exploratory factor analyses, however, it should be considered tentative. Though widely used, exploratory factor analysis has several inherent weaknesses when used to assess dimensionality. First, exploratory factor analysis typically "underfactors" data by putting items that reflect correlated but conceptually distinct constructs on the same factor (Hunter & Gerbing, 1982). Second, the eigenvalues obtained from exploratory factor analysis, upon which decisions regarding dimensionality are typically based, are influenced by factors other than dimensionality, such as the number of items analyzed, the magnitude of item loadings, and the correlations among the factors. Third, decision rules based on eigenvalues (e.g., Cattell, 1966; Kaiser, 1960) provide only general guidelines, leaving the ultimate decision regarding dimensionality to the researcher. Finally, though exploratory factor analysis is useful in the absence of sufficiently detailed theory, confirmatory factor analysis should be used when items can be logically grouped according to relevant theory (Gerbing & Anderson, 1988; Hunter & Gerbing, 1982). Given the abundant theoretical literature regarding TABP (e.g., Dembroski, Weiss, Shields, Haynes, & Feinleib, 1978; Friedman & Rosenman, 1974; Glass, 1977; Matthews, 1982), sufficient information is available to classify items constituting TABP measures on conceptual grounds. In any case, factor structures obtained through exploratory factor analysis should always be evaluated and further refined using confirmatory techniques (Gerbing & Anderson, 1988). With the exception of Edwards et al. (in press), none of the studies reviewed here used confirmatory techniques. In the present study, we used a combination of exploratory and confirmatory techniques to examine the Bortner scale, the Framingham scale and the JAS, thereby providing firmer evidence regarding their dimensionality.

Summary

Available evidence regarding differences in underlying constructs, measurement error, and dimensionality of the Bortner scale, the Framingham scale, and the JAS is scant and inconclusive. The constructs underlying these measures are not readily apparent, primarily because these measures were not designed to reflect a clearly defined domain of content. Reliability estimates for these measures suggest that they contain considerable measurement error, though the procedures used to derive these estimates are in question. The dimensionality of these measures has not been examined extensively, and available evidence is generally inconclusive. In the present study, we examined these three aspects of the Bortner scale, the Framingham scale, and the JAS while attempting to overcome problems with earlier investigations, thereby elucidating the reasons for the low correlations among these measures.

Method

Sample

Surveys were mailed to 651 executives who attended 6-week summer executive programs at a major graduate business school in the eastern United States between 1983 and 1988. A total of 352 executives (54%) returned the survey, a moderately high response rate for mailed surveys (Kerlinger, 1973). These executives represented 11 principal industries, including finance, electronics, petroleum, automotive, banking, consumer goods, food, insurance, telecommunications, textiles, and utilities. Nearly one third held general management positions, and the rest were distributed approximately evenly among administration, financial, sales, production, and technical positions. Most of the executives were male (94%) and Caucasian (96%). The average age was 46, and the average total job experience was 20 years. Respondents did not differ significantly from nonrespondents in gender, race, age, or years of job experience.

The analyses for this study involved both exploratory and confirmatory procedures. To reduce capitalization on chance variation, we randomly divided respondents into two samples; primary analyses were conducted on the first sample and cross-validated on the second sample (Anderson & Gerbing, 1988; Campbell, 1976; Cudeck & Browne, 1983). After cases with missing data were deleted, the final sample sizes were 162 and 168. Previous research suggests that these sample sizes were adequate for the relevant analyses (Anderson & Gerbing, 1984; Boomsma, 1982).

Measures

Data were obtained with the Bortner scale, the Framingham scale and the JAS as part of a larger study of stress, coping, and well-being among executives. The version of the Bortner scale we used differed slightly from the original (Bortner, 1969) in two ways. First, to facilitate scoring, we used an 11-point numerical response format centered at 0 and ascending to 5 in both directions (cf. Cooper & Marshall, 1979). Second, one anchor on Item 14 was reworded. The opposite of "ambitious" became "unambitious" rather than the original "satisfied with job." This modification was based on our experience in earlier studies, where respondents objected to the presumed bipolarity of the original anchors (Cooper & Marshall, 1979). Following Bortner's (1969) original recommendations, we reversed the scoring of Items 1, 3, 4, 6, 8, 9, and 10. We used the original 10-item Framingham scale presented by Haynes et al. (1978). The first five items were divided by the appropriate constant to equate the response scales for each item. We used Form C of the JAS (Jenkins et al., 1979). Following the recommended scoring procedure, we assigned weights to responses to the original 52 items to obtain the 21-item JAS–AB scale, the 21-item JAS–S scale, the 24-item JAS–J scale, and the 20-item JAS–H scale.

Results

Table 1 presents Pearson product-moment correlations and reliability estimates for the total scores obtained from the Bortner, Framingham, and JAS measures. The Bortner, Framingham, and JAS–AB scales were moderately correlated, with correlations ranging from .55 to .62, whereas the S, J, and H scales of the JAS were essentially uncorrelated. The relationships between the global TABP measures (the JAS–AB, the Framingham scale, and the Bortner scale) and the JAS scales were fairly consistent, with each global measure demonstrating the highest correlation with the S scale, followed by the H and J scales. Overall, these findings are consistent with earlier research, indicating only a moderate degree of convergence among the Bortner scale, the Framingham scale, and the JAS (Byrne et al., 1985; Haynes et al., 1980; Johnston & Shaper, 1983; Mayes et al., 1984; Price & Clarke, 1978).

Differences in Underlying Constructs

The content of each measure was examined to identify the underlying constructs. First, we examined each instrument to identify TABP dimensions that were addressed by at least one measure. This procedure revealed six dimensions: Speed/Impatience, Hard-Driving/Competitiveness, Time Pressure, Job Involvement, Anger/Temper, and Job Responsibility. Taken together, these dimensions comprise the core components of TABP discussed in the literature (Friedman & Rosenman, 1974; Glass, 1977; Matthews, 1982). Next, we constructed operational definitions of these dimensions, based on discussions of TABP in the literature (see Appendix). Finally, we recruited five colleagues who had training in various areas of applied psychology but were naive to the purpose of the study, explained the operational definitions of the TABP dimensions, and asked them to classify each item into one of the six TABP categories (a seventh "don't know" category was provided for items that could not be readily classified). Average interrater agreement (Cohen's kappa) was modest but statistically significant ($p < .001$), ranging from .313 for the JAS–H to .571 for the Framingham scale.

The results of the item classification procedure were analyzed by summing the number of items placed in each category separately for each measure. For the JAS scales, items were weighted according to the recommended scoring procedure before summing, thereby reflecting the differential contribution of each item to the overall scale score (Jenkins et al., 1979). These sums were then divided by the sum of the item weights for each scale (for the Bortner and Framingham scales, this simply involved dividing by the total number of items in the scale). The resulting values reflected the relative contribution of

Table 1
Correlations and Reliability Estimates for the Bortner Scale, Framingham Scale, and JAS

Measure	1	2	3	4	5	6
1. Bortner scale	(.676/.692)					
2. Framingham scale	.610**	(.694/.703)				
3. JAS–AB	.622**	.547**	(.436/.494)			
4. JAS–S	.536**	.497**	.644**	(.459/.533)		
5. JAS–J	.140	.173*	.267**	.149	(.025/.091)	
6. JAS–H	.173*	.227**	.429**	.081	.012	(.115/.235)

Note. Reliability estimates (alpha/omega) are reported on the diagonal. JAS = Jenkins Activity Survey; AB = global Type A behavior scale; S = Speed and Impatience; J = Job Involvement; H = Hard-Driving and Competitive.
* $p < .05$ (two-tailed test). ** $p < .01$ (two-tailed test).

each TABP dimension to the overall scale score, thereby facilitating comparisons across measures.

These values were then analyzed using a two-way multivariate analysis of variance with judges and measures constituting the factors and scores for the seven classification categories comprising the dependent variables. Differences in content of the Bortner scale, the Framingham scale, and the JAS were expected to be reflected in a significant main effect for the measures factor. Multivariate results confirmed this hypothesis, $F(35, 61) = 848.16$, $p < .001$. Univariate results were significant for all six TABP content dimensions (the effect for the seventh category failed to reach significance).

Further inspection revealed that these differences were largely due to the JAS–S and JAS–J scales, which contained the largest proportions of Speed/Impatience and Job Responsibility content, respectively (see Figure 1). The JAS–S scale also differed from the remaining scales by reflecting very little Hard-Driving/Competitiveness content, and the JAS–J and JAS–H scales differed from the remaining scales by reflecting very little Speed/Impatience content. The JAS–S and JAS–J scales also differed from one another in terms of Anger/Temper content, which contributed positively to the JAS–S scale and negatively to the JAS–J scale. The multivariate effect for judges was not significant, $F(28, 51) = 1.67$, $p > .05$.

These results reveal substantial differences in content among the TABP measures. The differences were primarily associated with the S, J, and H scales of the JAS, which were expected to differ from one another and from the global TABP measures. It remained unclear whether the global TABP measures differed from one another. Therefore, a second analysis was conducted with only the Bortner, Framingham, and JAS–AB measures. Multivariate results again revealed a significant effect for measures, $F(14, 4) = 9.70$, $p < .025$. Univariate results were significant for Time Pressure and for Anger/Temper. Further inspection revealed that these effects reflected relatively greater emphasis on Time Pressure in the Framingham scale and on Anger/Temper in the JAS–AB scale. The multivariate effect for judges again failed to reach significance. Overall, these results support the contention that the low correlations among the Bortner scale, the Framingham scale, and the JAS are, at least in part, attributable to differences in underlying constructs. Incidentally, these results also indicate that a sizable number of the items comprising these measures could not be readily classified into *any* of the six TABP categories included in this procedure (see Figure 1).

Measurement Error

The proportion of measurement error in the Bortner scale, the Framingham scale, and the JAS was estimated by calculating reliability coefficients for each measure and subtracting the obtained value from 1 (Lord & Novick, 1968). Reliability was initially estimated with alpha to facilitate comparisons with earlier research. Results indicated that measurement error accounted for approximately one third of the variance of the Bortner and Framingham measures, slightly over half of the variance of the JAS–AB and JAS–S scales, and almost all of the variance of the JAS–J and JAS–H scales (see Table 1).

Next, we tested the assumption of tau equivalence by estimating single-factor measurement models for each measure with LISREL VII (Jöreskog & Sörbom, 1988) and examining the deterioration in fit caused by constraining item loadings to be equal (for these analyses, the Bortner and JAS items were treated as continuous variables, and the Framingham items were treated as ordinal variables). This constraint significantly worsened the fit of all six measurement models ($p < .001$), indicating that the measures were not tau equivalent. Omega was then used to estimate reliability (Heise & Bohrnstedt, 1970). Omega is insensitive to departures from tau equivalence and reduces to alpha when items are, in fact, tau equivalent. Therefore, omega will generally provide a better estimate of reliability than alpha and will never provide a worse estimate (Greene & Carmines, 1980; Smith, 1974). These analyses yielded slightly lower estimates of measurement error for the Bortner and Framingham measures and moderately lower estimates for the JAS, though the overall pattern of results was not substantially affected (see Table 1).

Further inspection of the single-factor measurement models revealed two important findings (see Table 2). First, none of these models fit the data well, as reflected in the highly significant chi-square for each model. Given the dependence of chi-square on sample size, however, other indices of fit should be routinely considered, such as the adjusted goodness-of-fit index (AGFI) provided by LISREL (Jöreskog & Sörbom, 1988), the normed fit index (NFI; Bentler & Bonett, 1980), and the Tucker–Lewis index (TLI; Tucker & Lewis, 1973). For these

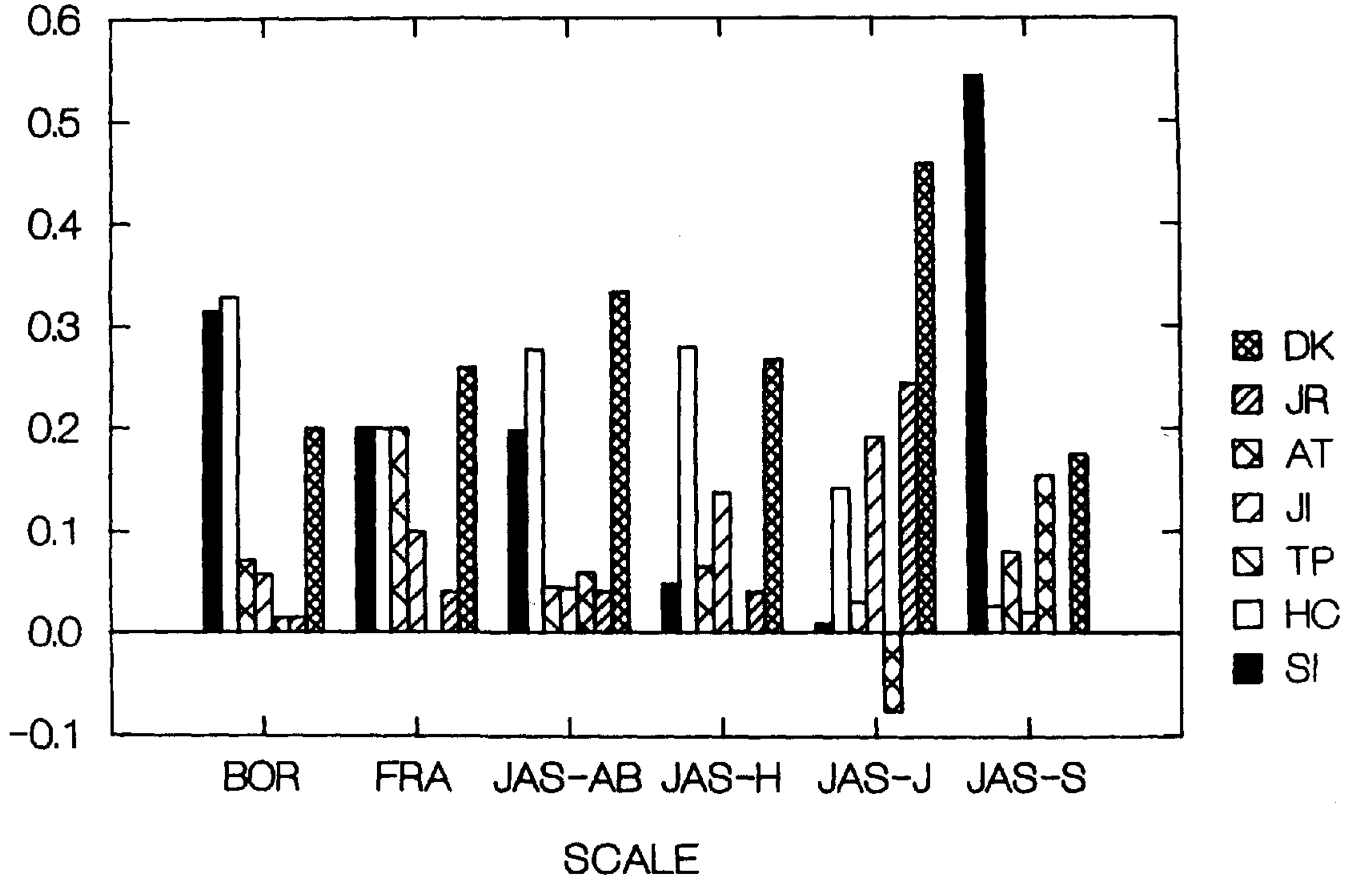

Figure 1. Proportional content of the Bortner scale (BOR), the Framingham scale (FRA), and the Jenkins Activity Survey (JAS). (JAS scales: AB = global Type A Behavior scale, H = Hard-Driving and Competitive, J = Job Involvement, S = Speed and Impatience. Type A behavior pattern dimensions: SI = Speed/Impatience, HC = Hard-Driving/Competitiveness, TP = Time Pressure, JI = Job Involvement, AT = Anger/Temper, JR = Job Responsibility, DK = Don't Know).

indices, values greater than .9 indicate acceptable fit. This criterion was not met by any of the models. Second, with the exception of the Framingham scale, each measure contained items that did not load significantly on the common underlying factor. This was particularly severe for the JAS–J, for which 17 of 24 items were not significant. Third, all four JAS scales contained items with loadings that were significant but negative, providing one reason why the corresponding reliability estimates were so low. Further examination revealed that, in each case, items with negative loadings were similar in content to items with positive loadings, but the recommended weighting procedure scored them in the opposite direction. Reversing items with negative loadings substantially increased the reliability estimates obtained for the JAS scales, though the proportion of measurement error was still moderate for the JAS–AB and JAS–H scales and substantial for the JAS–J scale (see Table 3).

The preceding results suggest that both measurement error and negative item loadings lower the correlations among the Bortner scale, the Framingham scale, and the JAS. The impact of these factors is readily observed by reversing items with negative loadings and then correcting these correlations for attenuation (see Table 3).[2] Reversing items with negative loadings substantially increased the correlations among the affected scales, particularly the JAS–S and JAS–H. Correcting for attenuation further increased the correlations, though the correlation between the JAS-AB and JAS–S scales was apparently overcorrected. This is probably attributable to specific item variance, which deflates reliability estimates and thus produces overcorrection (Lord & Novick, 1968). The correction procedure is further suspect in that reliability estimation assumes congeneric (i.e., unidimensional) measurement (Jöreskog, 1971; Nunnally, 1978). As we will show, this assumption is untenable for the Bortner, Framingham, and JAS measures.

[2] We did not correct the original correlations for attenuation because this procedure assumes that the obtained reliability estimates are less than unity only because of random measurement error. Reliability estimates may be lowered by other factors, however, such as specific item variance (Lord & Novick, 1968) and systematic scoring errors (as in the present case). When these deflated reliability estimates are used to correct for attenuation, the resulting "corrected" correlations are biased upward and, in some cases, may exceed 1. This line of reasoning was confirmed by the corrected correlations between the JAS–AB scale and the remaining JAS scales, which exceed 1.25.

Table 2
Single-Factor Measurement Models for the Bortner Scale, Framingham Scale, and JAS

Item	Standardized loading
Bortner scale[a]	
1. Never late	.107
2. Competitive	.363**
3. Anticipates others	.352**
4. Always rushed	.459**
5. Impatient	.388**
6. Goes all out	.637**
7. Does many things at once	.323**
8. Emphatic speech	.518**
9. Wants job recognition	.118
10. Fast (walking, eating)	.605**
11. Hard-driving	.695**
12. Expresses feelings	.205*
13. Few outside interests	−.027
14. Ambitious	.356**
Framingham scale[b]	
1. Hard-driving/competitive	.558**
2. Pressed for time	.725**
3. Bossy or dominating	.544**
4. Strong need to excel	.440**
5. Eat too quickly	.508**
6. After work, pressed for time	.413**
7. After work, think about work	.176*
8. After work, felt stretched	.395**
9. After work, felt dissatisfied	.272**
10. Upset while waiting	.295**
JAS–AB[c]	
1. Problems/challenges in life	.147
2. Act immediately under pressure	.316**
3. First one finished eating	.410**
4. Spouse says eat too fast[g]	−.297**
5. Frequently feel like hurrying others	.373**
6. Put words in others' mouths	.381**
7. Never late	.002
8. Hard-driving/competitive when younger	.382**
9. Hard-driving/competitive now[g]	−.713**
10. Spouse says hard-driving/competitive	.712**
11. Spouse says too active	.534**
12. More energy than most people	.399**
13. Fiery temper when younger	.305**
14. Daily job deadlines	.313**
15. Set quotas for self	.151
16. Keep two jobs moving forward[g]	−.302**
17. Take less than allotted vacation	.156
18. Bring work home	.143
19. Chosen as group leader	.312**
20. More responsible than others	.197*
21. More serious than others	.293**
JAS–S[d]	
1. Trouble finding time for haircut	.391**
2. Job stirs into action[g]	−.292**
3. First one finished eating	.502**
4. Spouse says eat too fast	.428**
5. Do two things at once	.343**
6. Frequently feel like hurrying others	.495**
7. Put words in others' mouths	.471**
8. Always hurry[g]	−.483**
9. Do things while waiting	.110
JAS–S[d] (*continued*)	
9. Do things while waiting	.110
10. Impatient while waiting	.287**
11. Hard-driving/competitive now[g]	−.543**
12. Spouse says hard-driving/competitive	.566**
13. Easily irritated	.577**
14. Do most things in a hurry	.674**
15. Fiery temper when younger	.271**
16. Fiery temper now	.447**
17. Set quotas for self	.069
18. Keep two jobs moving forward	.260**
19. Take less than allotted vacation	.141
20. Keep working when tired	.179*
21. Hurry more than others[g]	−.437**
JAS–J[e]	
1. Problems/challenges in life	.384**
2. Faced with annoyances/frustrations	.004
3. Put words in others' mouths	.028
4. Never late[g]	.129
5. Hard-driving/competitive now[g]	−.264**
6. Less energy than most people[g]	.095
7. Enjoy contests/try to win	.058
8. Fiery temper when younger[g]	−.119
9. Set quotas for self	−.219*
10. Work better against deadlines	−.146
11. Keep two jobs moving forward	−.125
12. Desire for promotion[g]	.095
13. Prefer promotion to raise	−.128
14. Increase in income	.551**
15. Bring work home	.147
16. Work nights/weekends	.010
17. Make lists	−.111
18. Hurry more than others[g]	−.322**
19. Increased hours worked on job	.175
20. Increased job responsibility	.389**
21. Higher job level	.776**
22. Change in job title	.175
23. Education level	−.134
24. Group officer in school	.051
JAS–H[f]	
1. Job stirs into action	.381**
2. Problems/challenges in life	.160
3. Spouse says eat too fast	−.020
4. Never let children win games	−.031
5. Hard-driving/competitive when younger	.395**
6. Hard-driving/competitive now	.748**
7. Spouse says hard-driving/competitive[g]	−.782**
8. Spouse says too active[g]	−.553**
9. Take work too seriously	.549**
10. Easily irritated[g]	−.430**
11. Enjoy contests/try to win	.333**
12. Fiery temper now	.292**
13. Irritated when interrupted	.184*
14. Job deadlines carry pressure	.122
15. Work better against deadlines	−.020
16. More effort than others	.306**
17. More responsible than others	.253**
18. More precise than others	.149
19. More serious than others	.326**
20. Education level[g]	−.066

Table 3
Correlations Among the Bortner Scale, Framingham Scale, and JAS (Reversing Items With Negative Loadings)

Measure	1	2	3	4	5	6
1. Bortner scale	(.681/.697)	.853**	.893**	.881**	.481**	.606**
2. Framingham scale	.595**	(.694/.703)	.808**	.816**	.375**	.653**
3. JAS–AB	.630**	.575**	(.710/.720)	1.133	.394**	.955**
4. JAS–S	.649**	.606**	.852**	(.777/.785)	.300**	.736**
5. JAS–J	.280**	.220**	.234**	.186*	(.472/.490)	.200*
6. JAS–H	.417**	.453**	.670**	.539**	.116	(.666/.684)

Note. Reliability estimates (alpha/omega) are reported on the diagonal. Correlations above the diagonal are corrected for measurement error. JAS = Jenkins Activity Survey; AB = global Type A behavior scale, S = Speed and Impatience, J = Job Involvement, H = Hard-Driving and Competitive.
* $p < .05$ (two-tailed test). ** $p < .01$ (two-tailed test).

Dimensionality

Three procedures were used to determine whether the Bortner scale, the Framingham scale, and the JAS were unidimensional. First, item content was examined to determine whether the items in each measure shared a common meaning (Hunter & Gerbing, 1982). The information required for this procedure was provided by the item classification results, which indicated that the Bortner scale, the Framingham scale, and the JAS not only tap different underlying constructs but also reflect multiple TABP dimensions. Second, results from the unidimensional measurement models for each measure were reexamined to determine whether the product rule for internal consistency had been met. This rule states that, for unidimensional measures, the observed correlation between a pair of items should equal the correlation produced by multiplying the loadings of the items on the common underlying factor (Hunter & Gerbing, 1982). Results revealed several significant residuals for each model ($p < .01$), indicating that this rule had not been met. Third, results for the unidimensional measurement models were again examined to determine whether the interitem measurement errors were correlated. According to classical measurement theory, if a measure is unidimensional, then these correlations should not differ from zero (Hunter & Gerbing, 1982; Lord & Novick, 1968). Results revealed several significant correlations for each measure ($p < .01$), indicating that this criterion had not been met.

These results strongly suggest that the Bortner, Framingham, and JAS scales are not unidimensional but instead reflect multiple dimensions. To confirm this, we constructed a series of multidimensional measurement models and examined them for internal consistency, external consistency, significance of item loadings, and overall fit. To construct these models, our primary criterion was item content (Hunter & Gerbing, 1982), supplemented by the pattern of residuals and modification indices provided by the single-factor measurement models. Specifically, an item was assigned to a dimension if it was consistently classified by a simple majority (three out of five) of the judges and if it shared the dimension with at least one other item (thereby allowing estimation of item loadings). The remaining items were initially included with loadings fixed at zero, thereby providing statistical information (i.e., residuals, modification indices) to guide final decisions regarding item inclusion. These models were then estimated using LISREL VII, and items with nonsignificant loadings and modification indices were deleted. Items with significant modification indices were added to the indicated factor only if their content was clearly consistent with the other items on that factor. The resulting models were then cross-validated with data from the second sample, and items with nonsignificant loadings were deleted. The final models were then reestimated with data from the first sample to ensure that the overall integrity of the models was preserved.

The final multidimensional measurement models for the Bortner scale, the Framingham scale, and the JAS are presented

Note to Table 2 (opposite). The JAS–AB, JAS–S, JAS–J, and JAS–H item descriptions are from the Jenkins Activity Survey (JAS). Abbreviated Framingham scale items are taken from "The relationship of psychosocial factors to coronary heart disease in the Framingham study: I. Methods and risk factors," by S. G. Haynes, S. Levine, N. Scotch, M. Feinleib, and W. B. Kannel, 1978, *American Journal of Epidemiology, 107,* pp. 362–383. Items from the Bortner scale are reprinted with permission from the *Journal of Chronic Diseases,* Vol. 22, R. W. Bortner, "A short rating scale as a potential measure of pattern A behavior," 1969, Pergamon Press plc. JAS scales: AB = global Type A Behavior scale, S = Speed and Impatience, J = Job Involvement, H = Hard-Driving and Competitive.
[a] $\chi^2(77, N = 162) = 149.29$, $p < .01$; adjusted goodness of fit index (AGFI) = .830; Tucker-Lewis Index (TLI) = .699; normed fit index (NFI) = .602. [b] $\chi^2(35, N = 162) = 118.84$, $p < .01$; AGFI = .783; TLI = .574; NFI = .602. [c] $\chi^2(189, N = 162) = 410.70$, $p < .01$; AGFI = .752; TLI = .488; NFI = .406. [d] $\chi^2(189, N = 162) = 407.22$, $p < .01$; AGFI = .763; TLI = .561; NFI = .466. [e] $\chi^2(252, N = 162) = 456.17$, $p < .01$; AGFI = .750; TLI = .251; NFI = .206. [f] $\chi^2(170, N = 162) = 307.35$, $p < .01$; AGFI = .800; TLI = .627; NFI = .490. [g] The recommended recoding procedure reverses the scoring of this item.
* $p < .05$. ** $p < .01$.

in Table 4. These models indicate that the Bortner, Framingham, JAS–AB, and JAS–H scales each contain two dimensions, that the JAS–J scale contains three dimensions, and that the JAS–S scale contains four dimensions. The models for the Bortner and JAS–AB scales are most similar, reflecting an approximately equal balance between Speed/Impatience and Hard-Driving/Competitiveness. In contrast, the Framingham model reflects a combination of the Hard-Driving/Competitiveness and Time Pressure dimensions. The JAS–S model was dominated by the Speed/Impatience dimension, with a secondary emphasis on the Hard-Driving/Competitiveness, Time Pressure, and Anger/Temper dimensions. The JAS–J contained two dimensions reflecting job-related content (Job Involvement, Job Responsibility) and a third dimension reflecting Hard-Driving/Competitiveness. Finally, the JAS–H model was dominated by the Hard-Driving/Competitiveness dimension, with a secondary emphasis on Anger/Temper. Summary statistics indicated that each model fit the data rather well, though the Bortner and JAS–S models did not fare as well as the other models, primarily because fairly large residuals were associated with pairs of items that shared highly specific content (e.g., eating fast).

These results indicate that the Bortner scale, the Framingham scale, and the JAS contain multiple dimensions. To determine whether the correlations among these measures were attributable to a subset of shared dimensions, we constructed subscales corresponding to the measurement models and calculated correlations among them, first using the original scoring and then reversing items with negative loadings (see Table 5).[3] With the original scoring, significantly higher correlations associated with subscales measuring the same dimension emerged for the Bortner, Framingham, JAS–AB, and JAS–H scales. After items with negative loadings were reversed, this pattern was evident for all six measures. Overall, out of 63 possible comparisons, 33 reached significance with the original scoring, and 61 reached significance after items with negative loadings were reversed. These results provide moderate support for the contention that the low correlations among the Bortner scale, the Framingham scale, and the JAS, as originally scored, are due to combining shared dimensions with unique dimensions within each of these measures. When items with negative loadings are reversed, these results are far more pronounced.

Combined Measurement Model

The preceding analyses indicate that the Bortner scale, the Framingham scale, and the JAS contain multiple dimensions, some of which are shared across measures. Furthermore, the large residuals associated with items that shared highly specific content suggested the presence of a second-order factor structure, with these specific items defining a factor that itself loaded on a more general factor (Gerbing & Anderson, 1984; Jöreskog & Sörbom, 1988). To explore the overall structure of the Bortner scale, the Framingham scale, and the JAS more fully, we combined items from these measures into a single analysis. Because the JAS scales share multiple items, they contain redundant information. Therefore, the original 52 JAS items, rather than the 86 scale items, were used. Because of their multiple-choice format, JAS items were treated as ordinal variables for these analyses, with options more indicative of TABP receiving higher scores.

An initial model was constructed by combining items from different subscales that measured the same dimension. Items that had been excluded from the subscale models because they were the sole indicator of a given dimension were also added. This resulted in a 46-item, six-factor measurement model, which was then estimated using LISREL VII. As with the subscale models, the remaining 30 items were initially included with loadings fixed at zero, thereby supplementing item content with statistical information to guide final decisions regarding item inclusion.

Results suggested three major respecifications. First, conceptually homogeneous subsets of items assigned to the Speed/Impatience and Hard-Driving/Competitiveness factors had fairly large residuals, suggesting a second-order factor structure. Second, 5 of the 30 items not assigned to any of the six TABP factors appeared to describe a seventh factor. Third, most of the remaining 25 items had nonsignificant modification indices and were therefore dropped. The model was respecified and reestimated. Items with nonsignificant loadings were deleted, and the resulting model was cross-validated with data from the second sample. Items that were not significant in the second sample were deleted, and the resulting model was reestimated with data from the first sample to ensure that the overall integrity of the models was preserved.

The final combined measurement model for the Bortner scale, the Framingham scale, and the JAS is presented in Table 6. As can be seen, the model contains 56 of the original 76 items. Of these, 21 were associated either with the General Speed factor or one of its five specific factors—Doing Many Things at Once, Punctuality, Eating Fast, Putting Words in the Mouths of Others, or Impatience. Of the remaining items, 14 described either the general Hard-Driving/Competitiveness factor or one of its three specific factors—Ambitious, Competitive, or Hard-Driving. The remaining items were distributed among five factors describing Time Pressure, Job Involvement, Anger/Temper, Job Responsibility, and the Desire to Acquire Titles and Credentials. Summary statistics indicated a fairly poor fit, with a highly significant chi-square and AGFI, NFI, and TLI values below .9. However, only 97 of the 1,540 residuals were significant at the .05 level (77 would be expected by chance), and none of the suggested respecifications could be justified on conceptual grounds. Thus, the poor fit of the model was apparently indicative of the highly restrictive structure imposed on the data.

Discussion

The results of this study indicate that differences in underlying constructs, measurement error, and multidimensionality each contribute to the low correlations among the Bortner scale, the Framingham scale and the JAS. This confirms earlier

[3] Correlations corrected for measurement error are not reported because several subscales contained fewer than four items, thereby rendering reliability estimates suspect (Anderson & Gerbing, 1982).

Table 4

Multidimensional Measurement Models for the Bortner Scale, Framingham Scale, and JAS

Item	Standardized loading					
	SI	HC	AT	TP	JI	JR
Bortner scale[a]						
2. Competitive		.466				
3. Anticipates others	.382					
4. Always rushed	.386					
5. Impatient	.412					
6. Goes all out		.671				
7. Does many things at once	.688					
10. Fast (walking, eating)	.569					
11. Hard-driving		.720				
14. Ambitious		.430				
Framingham scale[b]						
1. Hard-driving/competitive		.822				
2. Pressed for time				.974		
3. Bossy or dominating		.514				
4. Strong need to excel		.532				
6. After work, pressed for time				.484		
8. After work, felt stretched				.418		
JAS–AB[c]						
2. Act immediately under pressure	.432					
3. First one finished eating	.782					
4. Spouse says eat too fast[g]	−.662					
5. Frequently feel like hurrying others	.503					
6. Put words in others' mouths	.515					
8. Hard-driving/competitive when younger	.404					
9. Hard-driving/competitive now[g]		−.719				
10. Spouse says hard-driving/competitive		.875				
11. Spouse says too active		.531				
12. More energy than most people		.362				
JAS–S[d]						
1. Trouble finding time for haircut				.516		
3. First one finished eating	.589					
4. Spouse says eat too fast	.523					
JAS–S[d] *(continued)*						
5. Do two things at once				.456		
6. Frequently feel like hurrying others	.548					
7. Put words in others' mouths	.496					
8. Always hurry[g]	−.545					
10. Impatient while waiting	.230					
11. Hard-driving/competitive now[g]		−.755				
12. Spouse says hard-driving/competitive		.861				
13. Easily irritated			.793			
14. Do most things in a hurry	.663					
15. Fiery temper when younger			.451			
16. Fiery temper now			.665			
21. Hurry more than others[g]	−.432					
JAS–J[e]						
5. Hard-driving/competitive now[g]		−.733				
6. More energy than most people		.332				
7. Enjoy contests/try to win		.280				
9. Set quotas for self					.483	
14. Increase in income						.524
15. Bring work home					.381	
16. Work nights/weekends					.710	
20. Increased job responsibility						.410
21. Higher job level						.810
JAS–H[f]						
5. Hard-driving/competitive when younger		.389				
6. Hard-driving/competitive now		.760				
7. Spouse says hard-driving/competitive[g]		−.874				
8. Spouse says too active[g]		−.528				
9. Take work too seriously		.484				
10. Easily irritated[g]			−.864			
11. Enjoy contests/try to win		.336				
12. Fiery temper now			.511			
16. More effort than others		.255				

Note. All loadings are significant at the .05 level. The JAS-AB, JAS-S, JAS-J, and JAS-H item descriptions are adapted from the Jenkins Activity Survey (JAS). Copyright© 1965, 1966, 1969, 1979 by the Psychological Corporation. Reprinted by permission. All rights reserved. Abbreviated Framingham scale items are taken from "The relationship of psychosocial factors to coronary heart disease in the Framingham study: I. Methods and risk factors," by S. G. Haynes, S. Levine, N. Scotch, M. Feinleib, and W. B. Kannel, 1978, *American Journal of Epidemiology, 107,* pp. 362–383. Copyright by the Johns Hopkins University, School of Hygiene and Public Health. Reprinted by permission. Items from the Bortner scales are reprinted with permission from the *Journal of Chronic Diseases,* Vol. 22, R. W. Bortner, "A short rating scale as a potential measure of pattern A Behavior," 1969, Pergamon Press, plc. JAS scales: AB = global Type A Behavior scale, S = Speed and Impatience, J = Job Involvement, H = Hard-Driving and Competitive. Type A behavior pattern dimensions: SI = Speed/Impatience, HC = Hard-Driving/Competitiveness; TP = Time Pressure; AT = Anger/Temper, JI = Job Involvement, JR = Job Responsibility.

[a] $\chi^2(26, N = 162) = 57.18$, $p < .01$; adjusted goodness-of-fit index (AGFI) = .872; Tucker–Lewis index (TLI) = .806; normed fit index (NFI) = .779. [b] $\chi^2(8, N = 162) = 16.47$, $p < .05$; AGFI = .914; TLI = .906; NFI = .910; [c] $\chi^2(34, N = 162) = 54.93$, $p < .05$; AGFI = .896; TLI = .916; NFI = .853. [d] $\chi^2(84, N = 162) = 174.48$, $p < .01$; AGFI = .833; TLI = .785; NFI = .724. [e] $\chi^2(24, N = 162) = 23.07$, *ns*; AGFI = .946; TLI = 1.013; NFI = .835. [f] $\chi^2(26, N = 162) = 40.39$, $p < .05$; AGFI = .905; TLI = .928; NFI = .871. [g] The recommended recoding procedure reverses the scoring of this item.

Table 5
Correlations Among Subscales Created From the Bortner Scale, Framingham Scale, and JAS

Scale	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Bortner scale															
1. SI	—	.371**	.449**	.463**	**.528$_a$****	.376**	**.552$_a$****	.388**	.289**	.279**	.235**	.282**	.084	.264**	.305**
2. HC	.371**	—	**.649$_a$****	.289**	.242**	**.486****	.293**	**.574$_a$****	.369**	.209**	**.485$_a$****	.272**	.171*	**.487$_a$****	.336**
Framingham scale															
3. HC	.449**	**.649$_a$****	—	.332**	.327**	**.610$_a$****	.358**	**.681$_a$****	.407**	.226**	**.520$_a$****	.201**	.108	**.631$_a$****	.373**
4. TP	.463**	.289**	.332**	—	.215**	.358**	.245**	.386**	.029	**.375****	.207**	.355**	.006	.290**	.079
JAS–AB															
5. SI	**.529$_a$****	.199*	.373**	.198*	—	.262**	**.944$_a$****	.299**	.268**	.307**	.222**	.112	−.037	.198*	.255**
6. HC	.362**	**.441****	**.562$_a$****	.332**	.241**	—	.309**	**.837$_a$****	.360**	.213**	**.590$_a$****	.156*	.241**	**.710$_a$****	.289**
JAS–S															
7. SI	**.501$_a$****	.248**	.314**	.182*	**.731$_a$****	.229**	—	.350**	.292**	.320**	.211**	.099	−.005	.258**	.323**
8. HC	.126	**−.080**	**−.047**	−.012	.011	**.447****	−.017	—	.349**	.177*	**.631$_a$****	.187*	.252**	**.622$_a$****	.296**
9. AT	.289**	.369**	.407**	.029	.245**	.345**	.281**	.071	—	.162*	.205**	−.014	.167*	.325**	**.816$_a$****
10. TP	.279**	.209**	.226**	**.375****	.184*	.208**	.265**	.012	.162*	—	.169*	.161*	−.003	.141	.110
JAS–J															
11. HC	.006	**.072**	**.017**	−.131	−.012	**.086**	−.051	**.221$_a$****	−.063	.024	—	.236**	.179*	**.466$_a$****	.100
12. JI	.282**	.272**	.201**	.355**	.124	.134	.070	−.110	−.014	.161*	.108	—	.072	.217**	.006
13. JR	.084	.171*	.108	.006	−.001	.226**	−.040	.033	.167*	−.003	−.029	.072	—	.083	.111
JAS–H															
14. HC	.182*	**.410$_a$****	**.552$_a$****	.207**	.161*	**.569$_a$***	.183*	**−.078**	.275**	.076	**.005**	.151*	.038	—	.322**
15. AT	−.176*	−.129	−.102	−.127	−.202**	−.179*	−.244**	−.119	**−.135**	−.096	−.036	−.095	−.045	−.091	—

Note. Correlations below the diagonal correspond to original scale scoring; correlations above the diagonal correspond to scales reversing items with negative loadings. Correlations among subscales measuring the same dimension are in bold-faced type. Within each block, subscripted correlations are significantly higher than correlations in the same row and column ($p < .05$). Jenkins Activity Survey (JAS) scales: AB = global Type A Behavior scale, S = Speed and Impatience, J = Job Involvement, H = Hard-Driving and Competitive. Type A Behavior pattern dimensions: SI = Speed/Impatience, HC = Hard-driving/Competitive, TP = Time Pressure, AT = Anger/Temper, JI = Job Involvement, JR = Job Responsibility.
* $p < .05$. ** $p < .01$.

Table 6
Combined Measurement Model for the Bortner Scale, Framingham Scale, and JAS

Item	Standardized loading
General Speed	
B4. Always rushed	.505
B8. Emphatic speech	.662
J5. Act immediately under pressure	.552
J12. Always hurry	.455
J23. Do most things in a hurry	.729
J44. Hurry more than others	.465
Doing Many Things at Once	
B7. Does many things at once	.514
J8. Do two things at once	.457
J32. Keep two jobs moving forward	.339
Eating Fast	
B10. Fast (eating, walking)	.695
F5. Eat too quickly	.897
J6. First one finished eating	.847
J7. Spouse says eat too fast	.675
Putting Words in the Mouths of Others	
B3. Anticipates others	.542
J9. Frequently feel like hurrying others	.707
J10. Put words in others' mouths	.669
Impatience	
B5. Impatient	.746
F10. Upset while waiting	.602
J14. Impatient while waiting	.490
Punctuality	
B1. Never late	.744
J11. Never late	.744
Time Pressure	
F2. Pressed for time	.797
F6. After work, pressed for time	.597
F8. After work, felt stretched	.508
J1. Trouble finding time for haircut	.442
J28. Daily job deadlines	.272
General Hard-Driving/Competitiveness	
F1. Hard-driving/competitive	.778
J16. Hard-driving/competitive when younger	.426
J17. Hard-driving/competitive now	.858
J18. Spouse says hard-driving/competitive	.787
Hard-Driving	
B6. Goes all out	.535
B11. Hard-driving	.771
J19. Spouse says too active	.558
Ambitious	
B14. Ambitious	.359
F4. Strong need to excel	.412
J20. Take work too seriously	.301
J42. More effort than others	.259
J46. More serious than others	.219
Competitive	
B2. Competitive	.634
J24. Enjoy contests/try to win	.634
Job Involvement	
F7. After work, think about work	.161
J2. Job stirs into action	.478
J30. Never late	.266
J37. Bring work home	.491
J38. Work nights/weekends	.338
Job Responsibility	
J36. Increase in income	.575
J48. Increased job responsibility	.336
J49. Higher job level	.754
Anger/Temper	
F3. Bossy or dominating	.600
J22. Easily irritated	.759
J25. Fiery temper when younger	.605
J26. Fiery temper now	.731
Titles and Credentials	
J33. Desire for promotion	.723
J34. Prefer promotion to raise	.512
J50. Change in job title	.339
J51. Education level	.371

Note. All loadings are significant at the .05 level. For identification, loadings for factors with only two items were constrained to be equal. For each item, the letter prefix indicates the source scale (B = Bortner scale, F = Framingham scale, J = JAS). $\chi^2(1457, N = 162) = 2207.37$, $p < .01$; adjusted goodness-of-fit index = .645; Tucker–Lewis index = .696; normed fit index = .468. JAS–AB, JAS–S, JAS–J, and JAS–H items were adapted from the Jenkins Activity Survey (JAS). Copyright© 1965, 1966, 1969, 1979 by the Psychological Corporation. Reproduced by permission. All rights reserved. Abbreviated Framingham scale items are taken from "The relationship of psychosocial factors to coronary heart disease in the Framingham study: I. Methods and risk factors," by S. G. Haynes, S. Levine, N. Scotch, M. Feinleib, and W. B. Kannel, 1978. *American Journal of Epidemiology, 107,* pp. 362–383. Copyright 1978 by the Johns Hopkins University, School of Hygiene and Public Health. Reprinted by permission. Items from the Bortner scale are reprinted with permission from the *Journal of Chronic Diseases,* Vol. 22, R. W. Bortner, "A short rating scale as a potential measure of pattern A behavior," 1969, Pergamon Press, plc. JAS scales: AB = global Type A behavior scale, S = Speed and Impatience, J = Job Involvement, H = Hard-Driving and Competitive.

speculations that these measures reflect different aspects of TABP (Matthews, 1982, 1983). These results also implicate measurement error and multidimensionality as important causes of the low correlations. Furthermore, the results indicate that, as commonly used, the Bortner scale, the Framingham scale and the JAS are hampered by several serious psychometric problems.

First, the proportion of true score variance is marginal for the Bortner and Framingham measures, unacceptably low for the JAS–AB and JAS–S scales, and virtually nil for the JAS–J and JAS–H scales. The primary reason for the poor performance of the JAS scales is that each contains pairs of items that, despite similar content, are scored in the opposite direction. This procedure makes little sense from a substantive viewpoint and, in fact, is inconsistent with the original factor analysis of the JAS used to derive these scales (Zyzanski & Jenkins, 1970). Reversing items with negative loadings dramatically improved reliability estimates for the JAS scales, though these estimates were still only marginal for the JAS–AB and JAS–H scales and unacceptable for the JAS–J scale.

Second, though the Bortner, Framingham, and JAS scales are often treated as unidimensional measures, they each contain multiple dimensions. This oversight has created several serious methodological and conceptual problems. For example, reliability estimates for these measures (including those presented in this study) are suspect because these estimates assume congeneric (i.e., unidimensional) measurement (Jöreskog, 1971; Nunnally, 1978). In addition, the overall interpretation of these measures is confounded because it is impossible to determine the relative contribution of each dimension to the overall score (Burt, 1976; Gerbing & Anderson, 1988; Wolins, 1982). Furthermore, relationships associated with specific dimensions within these measures cannot be identified. This is particularly problematic because of recent evidence that only a subset of TABP dimensions, particularly anger and hostility, are consistently related to CHD (Booth-Kewley & Friedman, 1987; Matthews, 1988). Finally, the assumption of congeneric measurement not only is fundamental to measurement theory but also underlies most techniques used to analyze structural relationships embedded in causal models (Anderson & Gerbing, 1982; Danes & Mann, 1984; Hattie, 1985; Hunter & Gerbing, 1982; Jöreskog, 1971).

Given these flaws, we suggest that the Bortner scale, the Framingham scale, and the JAS should no longer be used in their current form. Instead, items from these scales should be recombined, modified, and supplemented to construct unidimensional measures of specific TABP dimensions. If the Bortner scale, the Framingham scale, and the JAS are used separately, the multidimensional measurement models in Table 4 suggest one way to recombine their items. This procedure is unlikely to produce strictly unidimensional measures, however, given the second-order factor structure associated with the Speed/Impatience and Hard-Driving/Competitiveness dimensions. A superior strategy would be to recombine items according to the measurement model presented in Table 6.[4] Though this would certainly produce better measures than the existing measures, we must emphasize that the model in Table 6 is not flawless. It is simply the best that could be obtained from the available items. A far better strategy would be to precisely define the TABP dimensions of interest, develop items that convincingly represent the domain specified by these definitions, combine these items to form unidimensional measures, and test these measures for internal consistency, external consistency, and construct validity (cf. Churchill, 1979). We must also emphasize that, contrary to current practice with the JAS, items should be assigned to one and only one measure. Otherwise, the resulting measures will be conceptually confounded and, by construction, multidimensional (Gerbing & Anderson, 1988). This principle is violated by the JAS scoring procedure, which assigns 17 items to two scales, 7 items to three scales, and 1 item to all four scales. This pitfall should be avoided by constructing items that specifically reflect a single TABP dimension. If this procedure is followed, the resulting measures should exhibit greater content and construct validity, thereby providing a substantial improvement over current measures.

The discussion above explicitly urges the measurement of specific TABP dimensions rather than global TABP. This represents an obvious departure from traditional TABP research, which has focused primarily on global TABP. From its initial conceptualization, however, global TABP has been described as an inherently multidimensional construct (Friedman & Rosenman, 1974). Therefore, scales that comprehensively reflect global TABP are necessarily multidimensional and, therefore, are fraught with the methodological and conceptual problems described in this article. Furthermore, recent evidence challenges the predictive validity of global TABP, suggesting that certain TABP dimensions, particularly anger and hostility, are the most important predictors of CHD (Booth-Kewley & Friedman, 1987; Matthews, 1988). For these reasons, we recommend that global measures of TABP be abandoned in favor of measures of TABP dimensions that collectively reflect the constellation of behaviors constituting TABP.

The results obtained in this study are based on a predominantly white-collar, middle-aged, male, Caucasian sample. Though this sample is representative of those used in the design, development, and validation of the Bortner scale and the JAS (e.g., Bortner, 1969; French–Belgian Collaborative Group, 1982; Jenkins et al., 1967; Pichot et al., 1977; Rosenman et al., 1975; Zyzanski & Jenkins, 1970), studies of the Framingham scale have typically used more heterogeneous samples (Haynes et al., 1978; Lee et al., 1987). Though available evidence is limited, some studies suggest that the psychometric properties of TABP measures may vary somewhat with the race, gender, and occupational level of respondents (Haynes et al., 1980; Waldron et al., 1977). Hence, the results of this study need to be replicated with non-White, female, and blue-collar samples before the Bortner scale, the Framingham scale, and the JAS are entirely abandoned. Nevertheless, several problems identified here, such as apparent scoring reversals in the JAS and heterogeneous content in all six measures, transcend sample characteristics. Therefore, we suspect that the problems identified in this study are likely to emerge in future studies with different samples.

In sum, the low correlations among the Bortner scale, the Framingham scale, and the JAS are attributable to a combination of differences in underlying constructs, measurement error, and multidimensionality. These factors not only lower the correlations among the measures but also suggest serious flaws in their psychometric properties. Therefore, the measures need to be substantially modified or, preferably, replaced with measures of specific TABP dimensions, at least for investigations with samples similar to ours. By carefully designing TABP measures and examining their associations with indices of stress, coping, and health, researchers can make substantial advances in understanding the nature and consequences of the collection of behaviors that constitute TABP.

[4] Physically recombining JAS items with other items to construct new measures risks copyright infringement. To avoid this, the JAS should be administered intact and any item recombination performed after the data have been collected.

References

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.

Begley, T. M., & Boyd, D. P. (1985). The relationship of the Jenkins Activity Survey to Type A behavior among business executives. *Journal of Vocational Behavior, 27,* 316–328.

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. Jöreskog & 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.

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, M. 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.

Chesney, M. A., Black, G. W., Chadwick, J. H., & Rosenman, R. H. (1981). Psychological correlates of the Type A behavior pattern. *Journal of Behavioral Medicine, 4,* 217–229.

Churchill, G. A., Jr. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research, 16,* 64–73.

Cooper, C. L., & Marshall, J. (1979). *Executives under pressure: A psychological study.* New York: Praeger.

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.

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

Danes, J. E., & Mann, O. K. (1984). Unidimensional measurement and structural equation models with latent variables. *Journal of Business Research, 12,* 337–352.

Dembroski, T., Weiss, S., Shields, J., Haynes, S. G., & Feinleib, M. (Eds.). (1978). *Coronary prone behavior.* New York: Springer.

Edwards, J. R., Baglioni, A. J., & Cooper, C. L. (in press). The psychometric properties of the Bortner Type-A scale. *British Journal of Psychology.*

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. *JAMA: The Journal of the American Medical Association, 169,* 1286–1296.

Friedman, M., & Rosenman, R. H. (1974). *Type A behavior and your heart.* New York: Knopf.

Gerbing, D. W., & Anderson, J. C. (1984). On the meaning of within-factor correlated measurement errors. *Journal of Consumer Research, 11,* 572–580.

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.

Glass, D. C. (1977). *Behavior patterns, stress, and coronary disease.* Hillsdale, NJ: Erlbaum.

Greene, V. L., & Carmines, E. G. (1980). Assessing the reliability of linear composites. In K. F. Schuessler (Ed.), *Sociological methodology 1980* (pp. 160–175). San Francisco: Jossey-Bass.

Hattie, J. (1985). Methodology review: Assessing unidimensionality of tests and items. *Applied Psychological Measurement, 9,* 139–164.

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.

Hunter, J. E., & Gerbing, D. W. (1982). Unidimensional measurement, second order factor analysis, and causal models. In B. M. Staw & L. L. Cummings (Eds.), *Research in organizational behavior* (pp. 267–320). Greenwich, CT: JAI Press.

Jenkins, C. D. (1978). A comparative review of the interview and questionnaire methods in the 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. 71–88). New York: Springer-Verlag.

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 Disease, 20,* 371–379.

Jenkins, C. D., Zyzanski, S. J., & Rosenman, R. H. (1971). Progress toward validation of a computer-scored test for the Type A behavior pattern. *Psychosomatic Medicine, 33,* 193–202.

Jenkins, C. D., Zyzanski, S. J., & Rosenman, R. H. (1979). *Jenkins Activity Survey Manual.* New York: Psychological Corporation.

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.

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

Jöreskog, K. G., & Sörbom, D. (1988). *LISREL VII.* Chicago: SPSS Inc.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20,* 141–151.

Kerlinger, F. N. (1973). *Foundations of behavioral research* (2nd ed.). New York: Holt, Rinehart, & Winston.

Lee, D., King, D., & King, L. (1987). Measurement of the Type A behavior pattern by self-report questionnaire: Several perspectives on validity. *Educational and Psychological Measurement, 47,* 409–423.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores.* Reading, MA: Addison-Wesley.

MacDougall, J. M., Dembroski, T. M., & Musante, L. (1979). The structured interview and questionnaire methods of assessing coronary-prone behavior pattern in male and female college students. *Journal of Behavioral Medicine, 2,* 71–83.

Matthews, K. A. (1982). Psychological perspectives on the Type-A behavior pattern. *Psychological Bulletin, 91,* 293–333.

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 and alternative to the Booth-Kewley and Friedman (1987) quantitative review. *Psychological Bulletin, 104,* 373–380.

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.

Novick, M. R., & Lewis, C. (1967). Coefficient alpha and the reliability of composite measures. *Psychometrika, 32,* 1–13.

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). Etude métrologique d'une batterie de tests destinéee à l'étude des facteurs psychologiques en épidemiologie cardio-vasculaire [A methodological study of a battery of tests designed for the study of psychological factors in cardiovascular epidemiology]. *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.

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 1/2 years. *JAMA: The Journal of the American Medical Association, 233,* 872–877.

Shipper, F., Kreitner, R., Reif, W. E., & Lewis, K. E. (1986). A study of four psychometric properties of the Jenkins Activity Survey Type A scale with suggested modifications and validation. *Educational and Psychological Measurement, 46,* 551–564.

Smith, K. W. (1974). On estimating the reliability of composite indexes through factor analysis. *Sociological Methods and Research, 2,* 485–510.

Tucker, L. R., & Lewis, C. (1973). The reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38,* 1–10.

Waldron, I., Zyzanski, S., Shekelle, R. B., Jenkins, C. D., & Tannanbaum, S. (1977). The coronary-prone behavior pattern in employed men and women. *Journal of Human Stress, 3,* 2–18.

Wolins, L. (1982). *Research mistakes in the social and behavioral sciences.* Ames: Iowa State University Press.

Zyzanski, S. J., & Jenkins, C. D. (1970). Basic dimensions within the coronary-prone behaviour pattern. *Journal of Chronic Diseases, 22,* 781–795.

Appendix

Definitions of Type A Behavior Pattern Dimensions

The following definitions were provided to the judges prior to item classification:

Speed/Impatience: The tendency to do things fast, even when there is no external need to do so. This tendency is generated from within the person, not from the situation. This tendency is manifested in hurrying through both minor and major daily activities as well as encouraging others to do so.

Hard-Driving/Competitiveness: The tendency to constantly strive to improve oneself and/or one's performance, both relative to others and relative to one's own past performance. This tendency is manifested in excessive drive and ambition in all areas of life.

Time Pressure: The tendency to report many demands competing for one's time. This may involve a description of the sources of this pressure, or the response to this pressure (i.e., doing things fast or poorly). It should be emphasized that Speed/Impatience refers to doing things fast for internal reasons, whereas Time Pressure refers to doing things fast for external reasons.

Job Involvement: The tendency to identify one's sense of self-worth with one's job. This tendency is manifested in an excessive focus on job-related activities and an internal desire to spend inordinate amounts of time on these activities. Note that Job Involvement refers to spending excessive time on job-related activities for internal reasons, whereas Time Pressure may also occasionally refer to spending excessive time on job-related activities, but for external reasons.

Anger/Temper: The tendency to display anger, hostility, and a quick temper. This may be manifested either verbally or physically. This tendency originates primarily from within the person, such that these behaviors may be displayed without particular situational causes or cues.

Job Responsibility: This refers to having attained a high level of achievement in one's job. These persons have attained a high level in their organization, are accountable for large portions of work, are responsible for many subordinates, have high salaries, and in general have advanced to positions of power and status.

Don't Know: This includes items that were either designed poorly, do not fall into one of the above categories, or have no clear meaning at all.

Received August 24, 1989
Revision received March 1, 1990
Accepted March 1, 1990 ■