1995-7

# A revision of the dyadic adjustment scale for use with distressed and nondistressed couples: Construct hierarchy and multidimensional scales

Dean M. Busby
*Brigham Young University - Provo*, dean_busby@byu.edu

Clark Christensen
*Syracuse University*

D. Russell Crane
*Brigham Young University - Provo*

Jeffry H. Larson
*Brigham Young University - Provo*
Follow this and additional works at: https://scholarsarchive.byu.edu/facpub

Part of the Other Social and Behavioral Sciences Commons

# A revision of the dyadic adjustment scale for use with distressed and nondistressed couples: Construct hierarchy and multidimensional scales

Busby, Dean M; Christensen, Clark; Crane, D Russell; Larson, Jeffry H . Journal of Marital and Family Therapy ; Hoboken  Vol. 21, Iss. 3,  (Jul 1995): 289.

## ABSTRACT (ABSTRACT)

A study of the effectiveness of the Dyadic Adjustment Scale, commonly used as an instrument for measuring adjustment in relationships, is presented. A revised scale is proposed.

## FULL TEXT

INTRODUCTION

The Dyadic Adjustment Scale (DAS; Spanier, 1976) is a frequently used instrument for measuring adjustment in relationships (Spanier, 1985). The DAS has particular value for both researchers and clinicians since it is relatively short (32 items) yet is multidimensional in that it contains four subscales.

Most of the researchers who have investigated the psychometric properties of the DAS have utilized nonclinical samples and have focused on the applicability of the DAS for use in survey research (Sabatelli, 1988; Sharpley &Cross, 1982; Spanier &Thompson, 1982). However, recently researchers have paid more attention to the appropriateness of the DAS for use in clinical research (Crane, Allgood, Larson, &Griffin, 1990; Kazak, Jarmas, &Snitzer, 1988; Spanier, 1988).

The DAS has been shown to distinguish reliably between distressed and nondistressed samples (Crane et al., 1990), but at least two of the subscales (Dyadic Satisfaction subscale and Affectional Expression subscale) have been problematical (Crane, Busby, &Larson, 1991; Sharpley &Cross, 1982; Spanier &Thompson, 1982). This study was an attempt to understand the problems with the subscales of the DAS by evaluating their adherence to standards of construct hierarchy for multidimensional instrument development (Comrey, 1988). In addition, this study was an attempt to revise the DAS by following the standards more carefully so that the instrument would be more appropriate for clinical and research use.

REVIEW OF THE LITERATURE

In 1976 Spanier created an instrument to measure the separate components of marital adjustment and called it the Dyadic Adjustment Scale (DAS). The separate dimensions of marital adjustment were reported to be the following: (a) consensus on matters of importance to marital functioning, (b) dyadic satisfaction, (c) dyadic

cohesion, and (d) affectional expression.

After presenting the DAS, Spanier (1976) concluded that it could be used as an overall measure of marital adjustment, or the specific subscales could be used independently without losing confidence in the reliability or validity of the measure. Since the creation of the DAS, it has become one of the most widely used instruments in the family field (Crane et al., 1990; Sabatelli, 1988; Spanier, 1985, 1988).

An empirical follow-up on the DAS was done in 1982 by Spanier and Thompson with a sample of 50 separated persons and 155 divorced persons. This article was the first of what were to be five factor analyses on the DAS (Crane et al., 1991; Kazak et al., 1988; Sabourin, Lussier, Laplante, &Wright, 1990; Sharpley &Cross, 1982; Spanier &Thompson, 1982) An important result from their study was that the goodness of fit statistic implied that it was very likely that another structure, rather than the four factors designed by Spanier(1976), would represent the data better.

Spanier and Thompson (1982) analyzed the DAS in a variety of ways to find a more acceptable solution. Alternative structures were compared to the original factor structure and the results verified that a four-factor solution was still the most appropriate. The final results from their analyses indicated that the Consensus, Satisfaction, and Cohesion subscales were "replicated fairly well," although the negative and positive items of the Satisfaction subscale did not group together as originally hypothesized. Finally, the Affectional Expression subscale was problematical since two of the four items had small loadings.

In conclusion, Spanier and Thompson (1982) reported that the DAS was originally designed as a global assessment device and as such it was a reliable and valid instrument. Second, they reported that the domains of the specific dimensions (subscales) of marital adjustment were not sampled and screened with the intent of developing distinct measures of the facets (subscales). Nevertheless, Spanier and Thompson (1982) reported that the four factors were robust and had distinct meaning, citing research in which the subscales were used separately (Thompson &Spanier, 1983). Finally, they suggested that future work should weed out items which did not mark the factors across samples.

The work by Spanier and Thompson (1982) initiated a debate over whether the DAS was a global unidimensional instrument or a multidimensional instrument. This debate has not been satisfactorily ended to date since some authors continue to present the instrument as unidimensional while others report that it is multidimensional (Sabourin et al., 1990; Spanier, 1988; Thompson, 1988).

There seems to be little practical use for 32 items if discriminating between distressed and nondistressed couples is the only valid way the DAS can be used; this can be done adequately by much shorter instruments like the Marital Adjustment Test (MAT; Locke &Wallace, 1959) or the Kansas Marital Satisfaction Scale (KMS; Schumm et al., 1986). This illustrates the primary distinction of the DAS in relation to other marital adjustment instruments; it has the potential of being multidimensional, thereby providing more information to clinicians and researchers. Virtually all of the studies exploring the psychometric strengths of the DAS have utilized factor analytic techniques.

Another factor analysis was performed on the DAS in 1982 by Sharpley and Cross. The sample in this study consisted of 95 married persons. They reported that although a four-factor solution was found to be appropriate for the data, the rotated factor solution did not resemble the subscales presented by Spanier (1976). A careful investigation of Sharpley and Cross's (1982) results shows that the Affectional Expression and Dyadic Satisfaction subscales were the most problematical subscales.

One of the few studies with a sample of distressed and nondistressed couples was that by Crane et al. (1991). They found results similar to those of Sharpley and Cross (1982) and Spanier and Thompson (1982) in that the Dyadic Satisfaction subscale did not fair well since half of the items did not produce high factor loadings on the appropriate subscale. In addition, when separate analyses were conducted with the distressed and nondistressed samples, the problems with the Dyadic Satisfaction subscale were amplified, and most of the subscales were not satisfactory with the nondistressed sample.

The next factor analysis of the DAS reviewed in this study was performed by Kazak et al. (1988). They chose to do a separate factor analysis for men and women, which makes it difficult to compare their results to those of previous studies. The samples for the study consisted of 219 women and 190 men. Kazak et al. summarized their results by saying that there was weak support for the presence of four subscales. Particular problems were apparent with the Consensus and Satisfaction subscales since items cross-loaded on both scales. The items of the Affectional Expression subscale, on the other hand, received better support than in previous studies. Kazak et al. concluded by proposing that the DAS only be used as a general assessment instrument and that the subscales not be used separately.

Spanier (1988) responded to Kazak et al.'s (1988) article, as did Thompson (1988). Spanier (1988) reiterated that the DAS worked best as a global summary measure and that he had not used subscale scores in any of his own research. He concluded by suggesting that future research be focused on specific refinements of the DAS and that another generation of measures should be developed without the problems of the current instruments.

Thompson (1988) responded by suggesting again that the DAS was best as a summary measure and should not be used for assessment of the separate dimensions of marital quality. Thompson complained that doubters continued to find fault with the DAS while believers continued to use it. A request was made to move beyond criticism of the scale to revision of it.

The final study considered in this review included the only treatment of the DAS in a hierarchical fashion (Sabourin et al., 1990). These researchers proposed that some of the problems with the DAS could have been a result of not taking into consideration the hierarchical nature of the instrument. Their solution to the unidimensional versus multidimensional argument was to consider a third option: that the DAS was hierarchical. They proposed that the DAS consisted of 4 first-level factors which combined to form a second-level global factor. Using a confirmatory factor analysis with a sample of volunteer couples, they determined that this structure more adequately represented their data than either a unidimensional or a multidimensional model. Nevertheless, some items in their study on both the Consensus and Affectional Expression subscales had low loadings.

Summary from the Literature Review

A number of conclusions were summarized from the existing research on the DAS.

First, marital adjustment was originally defined (Spanier &Cole, 1976) as a multidimensional phenomenon which the DAS was reported to measure adequately (Spanier, 1976). Second, the DAS was presented as a scale that could be used both as a summary measure by utilizing the total score and as a measure of the separate components of marital adjustment by using the subscales separately (Spanier, 1976, 1979; Spanier &Thompson, 1982) . Third, factor analyses of the DAS demonstrated that there were problems with the validity of the Dyadic Satisfaction and Affectional Expression subscales (Crane et al., 1991; Kazak et al., 1988; Sharpley &Cross, 1982). Fourth, the samples used to evaluate the DAS often consisted of separated or divorced individuals rather than

distressed and nondistressed couples (Spanier, 1976; Spanier &Thompson, 1982). Fifth, one promising method of eliminating the unidimensional/multidimensional argument is to consider the DAS a hierarchical measure with first- and second-order constructs. Sixth, future studies on the DAS should revise and improve the DAS rather than just criticize it (Spanier, 1988; Thompson, 1988).

Construct Hierarchy and Multidimensional Scales

From the literature review, it is evident that the controversy surrounding the DAS centers on the question of whether the scale is a unidimensional global measure or a multidimensional instrument. This controversy has occurred because of the empirical work that failed to validate the subscales of the DAS. Since the original definition of dyadic adjustment was multidimensional in nature (Spanier &Cole, 1975; Spanier, 1976) and the subscales of the DAS are what make it unique among instruments in the field, it seems more appropriate to revise the instrument so that the subscales are valid than to suggest that it is now only a global instrument.

The problems with the subscales of the DAS can be understood and corrected by considering issues of construct hierarchy. The accepted method of developing scales in the social sciences follows a number of specific steps (American Psychological Association, 1985; Comrey, 1988):

1. The researcher decides what concept is to be measured.

2. The researcher must then make an assessment of the construct hierarchy of the proposed concept that is to be measured. Is the concept at the first level of the hierarchy, such as a four-item measure of family finances? Concepts at the First level are measured by creating questions that are homogeneous or alternate forms of one another. If the concept being measured is a second-order concept like dyadic consensus, it is a combination of a number of first-order concepts such as finances, leisure activities, and decision making.

3. Once the construct hierarchy is delineated, homogeneous questions can be written for each of the first-order concepts.

4. The internal consistency of each first-order subscale is then demonstrated by a series of factor analyses which show that all homogeneous items have substantial factor loadings on their respective subscales.

5. To evaluate the ability of the first-order constructs to measure the second-order constructs, a second-order factor analysis would be conducted.

An example of this process follows: Suppose that Dyadic Consensus was defined as stated in Spanier's article (1976): "consensus on matters of importance to marital functioning" (p. 128). Suppose that there were four areas of importance, such as finances, affection, decision making, and values. These four areas are first-order subscales and could each be measured by four homogeneous items, two positive and two negative. An example of a positive item could be the following: My partner and I have similar opinions about how we should handle finances in our marriage. A negative item could be: My partner and I do not agree about financial issues in our marriage. The questions would be answered on a 5-point scale with responses ranging from most of the time to never. Once the items were written and data were collected, a factor analysis could be conducted to see if the four items of each concept factored correctly on their respective subscales (Comrey, 1988).

To assess the ability of these four subscales (finances, affection, decision making, and values) to measure the second-order concept of consensus, a second-order factor analysis would be conducted (Comrey, 1988). The

important point that these conventions of scale construction emphasize is that first-order constructs are measured by homogeneous items; second-order concepts are measured by first-order subscales that are correlated by not alternative forms of each other.

The primary reason that the structure of the DAS was not confirmed in previous studies may have been that more recently developed methods of scale construction were not available to Spanier in 1976. Spanier(1976) was careful to create a conceptual definition of marital adjustment that could be used to operationalize test items; nevertheless, when questions were selected to measure the conceptual definition, the rules of construct hierarchy were apparently not followed. In other words, the subscales contained both first- and second-order items in that some of the items were homogeneous and others were related but were not alternate forms of each other; hence hybrid subscales were formed that spanned the first two hierarchical levels.

THE PURPOSE OF THIS STUDY

The purpose of this study was to revise the subscales of the DAS by adhering to the conventions of construct hierarchy. This could provide two possible benefits: (a) an improved multidimensional scale of dyadic adjustment and (b) an example for other researchers to employ in order to improve or develop additional assessment instruments.

The adequacy of the conceptual definition of marital adjustment that was proposed by Spanier (1976) was not at issue here. The purpose was to improve the DAS so that the existing subscale items measured the components of marital adjustment as defined by Spanier in a more reliable and valid way. Before more items or subscales can be added to the instrument (Spanier &Thompson, 1982; Thompson, 1988), the items that are currently contained in DAS should be analyzed and dropped if they are not appropriate. When the existing scale has been improved, other subscales can be added to measure the domains of marital adjustment that may be missing.

METHOD

Subjects

Data were collected from clinical programs in marriage and family therapy at Brigham Young University (couple n = 155) and Montana State University (couple n = 87). Of the 242 couples (484 individuals) involved in this study, 65 were seeking marital therapy at Brigham Young University and 33 were seeking therapy at Montana State University. The data were collected at the beginning of each couple's first therapy session. Spouses were asked to complete the measures independently and not to discuss their responses with one another. The remaining couples, 90 from Brigham Young and 54 from Montana State, were nonclinical volunteers.

The volunteer data were collected by graduate students at the two universities as part of a requirement for a class. They were asked to administer the inventories to nonclinical couples (not currently in treatment for marital problems).

Although the sample was one of convenience, there was no evidence to suggest that it was not typical of the average marital dyad that would respond to a study of this type. Still, a broader based probability sample would have been preferable.

In order to ensure that the clinical sample was distressed and the nonclinical sample was nondistressed, clinical individuals who scored above 107 on the total DAS and nonclinical individuals who scored below 107 were

eliminated from the study. The cutoff score of 107 was suggested by Crane et al. (1990) as the appropriate cutoff between distressed and nondistressed samples, as well as the score that was equivalent to 100 on the MAT, another commonly used marital adjustment test (Locke &Wallace, 1959). This procedure only eliminated 18 volunteers and 12 clinical individuals from the Montana State data. The resulting sample consisted of 271 nondistressed and 183 distressed individuals (N = 454).

The total sample was composed of predominantly young, Caucasian (95%) middle-income, and first-married couples. The average age of the respondents was 30.7 years; mean years of education was 15.2. The mean number of children was 1.68 and the mean income was $23,393. A one-way analysis of variance showed that the nondistressed sample was significantly better educated, had higher incomes, and had been married longer than the distressed sample. Nevertheless, although statistically significant, the mean differences were not large: the nondistressed sample had an average of 1.83 years more education, 2.2 years longer marriage, and $9,813 more income. These differences were probably attributable to the fact that the clinics from which the distressed sample was drawn serviced many married students.

An additional analysis of variance revealed no significant differences between the samples from the two geographical regions. This implies that the couples from each location did not score differently on the DAS.

Data Analyses

The following specifications were made in the LISREL program to perform all the confirmatory factor analyses. The correlation matrix was used for each factor analysis. Factor loadings for the items were left free to vary on their respective subscales and were fixed at zero for the other subscales. The variances of the factors were fixed at unity to scale the solution; the correlations between factors were left free to vary. The error variances for each item were free to vary; the covariances between the error terms were fixed at zero.

Three common statistics were used for evaluating the fit of the model to the data in LISREL, chi-square, goodness of fit index (GFI), and root mean square residual (RMR) (Joreskog &Sorbom, 1989). To demonstrate adequate fit to the data, the chi-square should be nonsignificant >.05), the GFI should be greater than .90, and the RMR should be small (less than .075). It is rare, with larger samples, to obtain results where all three statistics are within an acceptable range (Lavee, 1988). In this study if two of the three statistics indicated adequate fit to the data, the model was accepted.

RESULTS/DISCUSSION

The principle of parsimony is crucial in instrument construction and research in general. Parsimony refers to the necessity of testing simpler models before advancing to higher order or more complicated models (American Psychological Association [APA], 1985). The first step, then, in revising the DAS would be to test a simple single-factor DAS with all 32 items in the model using a confirmatory factor analysis.

The statistics evaluating the fit of the model to the data for the single-factor 32-item DAS were the following: chi-square 1350.65 (p = .000), GFI .83, and RMR .044. Two of the three statistics suggested that the model did not fit the data. This implies that an alternative model would most likely represent a better fit to the data.

One explanation for this finding and for previous research findings that have failed to verify the four factors of the DAS is that the subscales are hybrids (Crane et al., 1991; Kazak et al., 1988; Sharpley &Cross, 1982). To investigate whether a subscale is hybrid the researcher can look at the items to see if some of them are alternate form items

and others are not. Additionally, factor loadings and modification indices from LISREL can be examined to explore the issue of homogeneity (Joreskog &Sorbom, 1989). Figure 1 presents the items of the DAS grouped by subscale.

Without looking at any statistical results, it was apparent from Figure 1 that some of the items within each subscale were homogeneous and others were not. Question 3 (religious matters) and Question 7 (conventionality, correct or proper behavior) seemed to be homogeneous; neither Question 5 (friends) nor Question 9 (ways of dealing with in-laws) appeared to have any other questions with which they were similar. In the Dyadic Satisfaction subscale, Questions 21 (how often quarrels occur) and 22 (how often partners get on each other's nerves) seemed to be alternate form items; Question 23 (Do you kiss your mate?) was unrelated to the other questions. The same problem was true for other items of the DAS. To make subscales that were not hybrids it was necessary to use questions like 3 and 7 as a first-order subscale and eliminate Questions 5 and 9.

Figure 2 represented a revised factor structure of the DAS. (Figure 2 omitted) The first-order subscales in Figure 2 were made up of two questions that were relatively homogeneous. The first-order subscales combined to form 3 second-level subscales of Consensus, Satisfaction, and Cohesion, which in turn combined to form the highest level concept of dyadic adjustment. This Revised DAS (RDAS) was consistent with the original definition of dyadic adjustment proposed by Spanier (1976) and conformed to the conventions of scale construction in that the lowest order subscales were made up of homogeneous items and the higher order concepts were made up of related but unique concepts.

There were a number of important differences between the revised structure and the original four- factor structure proposed by Spanier (1976): First, since Items 4 and 6 were questions assessing consensus on matters of affection, they were hypothesized to be a first-order subscale of the consensus construct rather than a part of an affectional expression construct. Second, items which were hybrids in that they were not homogeneous with at least one other item were removed (1, 5, 9, 11, 13, 17, 18, 19, 23, 26 ). Third, when more than two items appeared to be homogeneous (this only occurred in one instance with Items 3, 7, 8, 10), the principle of parsimony was followed in that only the two which appeared to be the most similar in content were selected. This procedure helped keep the DAS short and provided convenient two-item subscales that could be split in half to create two forms of the DAS for the purposes of pre- and posttests or for multiple testing of the same subject. Fourth, Items 29 and 30 were removed because they had a dichotomous yes/no response scale which made them incompatible with the other items that had a 4- or 5-point response scale (an alternative to removing Items 29 and 30 would be to expand the dichotomous scale to a 5-point scale). Fifth, Items 31 and 32 were removed since they were global subjective items that contaminated the scale (Sabatelli, 1988). This means that it is inappropriate to combine specific questions such as Item 12 (the amount of consensus on decision making) with global questions such as Item 31 (how happy the partner is in the relationship). Questions such as 31 and 32 are akin to asking respondents to rate their own IQ on an intelligence test. The respondents' answers to the specific items are likely to influence their responses to the global items, or vice versa.

Evaluating the RDAS

Prior to evaluating the more elaborate model of the RDAS, it was necessary to explore the most parsimonious model with all 16 items loading on a single factor. A confirmatory factor analysis was conducted to explore how well this single-item RDAS fit the data. The statistical results were as follows: chi square 938. 54 (p = .000), GFI .77, and RMR .098. In this instance none of the three statistics demonstrated an adequate fit to the data. Therefore, it was appropriate to explore the fit of the hierarchical model proposed in Figure 2.

It was hypothesized that the RDAS in Figure 2 would be an improvement over the original DAS if the following

criteria were met:

1. Confirmatory factor analyses results with several samples substantiated the hypothesized structure of the RDAS. This would be evidence of construct validity.

2. Additional construct validity would be demonstrated if the correlation coefficient between the RDAS and the MAT was similar to or higher than the coefficient between the DAS and MAT.

3. The RDAS was as successful as the DAS at discriminating between distressed and nondistressed samples. This would be evidence of one type of criterion validity often called discriminant validity.

4. Internal consistency or reliability estimates of the RDAS would be .80 or above on the subscales and the total scale.

5. Split-half reliability coefficients would be equal to or superior to those obtained for the DAS.

Construct Validity

To evaluate the construct validity of the RDAS, factor analyses were conducted with the LISREL program (Joreskog &Sorbom, 1989). The first confirmatory analysis was to evaluate the first-order subscales simultaneously. It was hypothesized that the revised model proposed in Figure 2 would fit the data in this study.

The LISREL program produced factor scores and t- values for each item. The standardized factor loadings and t- values for each of the items of the Consensus, Satisfaction, and Cohesion subscales were large. Each item had t- values over 10 in addition to factor loadings above .50.

The factor scores and t- values produced in LISREL were similar to the t- tests and beta coefficients produced in most regression programs. The factor scores provided an assessment of the strength of the relationship between each question and its respective factor; the t- values assessed whether or not the factor scores were significantly different from zero.

Although the statistics for the individual items were large, statistics for the overall model were not acceptable. The chi-square was 279.48 (p = .000), the GFI was .93, and the RMR was .119. Two of the three statistics demonstrated that the model did not fit the data. Some hints as to why the model was not a good fit to the data were gleaned from the modification indices. The modification indices suggested that there was a considerable amount of overlap between the first-order Leisure and Activities subscales (see Figure 2). A modification index in LISREL was a statistic that gave an evaluation of which item loadings, if adjusted (or left free to vary), would improve the fit of the model to the data. In this instance the modification indices were used to assess whether there was overlap or redundancy between first-order subscales. If there were a high degree of redundancy between first-order subscale scores, the implication would be that one of the scales should be removed since the standards of construct hierarchy suggest that first-order subscale scores should be related to each other but not equivalent.

The Leisure and Activities subscales should have been related but not homogeneous, and the modification indices provided evidence that they were homogeneous; therefore, the Leisure subscale was eliminated, leaving a Consensus scale with three subareas (Decisions, Values, and Affection). The first-order Leisure subscale was eliminated rather than the first-order Activities subscale because the Consensus scale had 4 first-order subscales and the Cohesion scale had only 2 first-order subscales (see Figure 2). Therefore, the removal of the Leisure

subscale helped prevent the inadvertent weighting mentioned by Norton (1983) that results when one subscale has too many items relative to the other subscales.

Figure 3 contains the first-order confirmatory factor analysis results for the RDAS. (Figure 3 omitted)

The results in Figure 3 demonstrate one type of construct validity for the first-order factors of the RDAS. The t-values were robust for each item since all were larger than 10. Two of the three statistics evaluating the overall model demonstrated adequate fit to the data. The modification indices were relatively small (almost all were below five), and it did not make theoretical sense to adjust the few items that were larger than five.

Figure 4 presents the second-order factor analysis with the loadings of the first-order subscales on the higher order construct of consensus, satisfaction, and cohesion. The first-order loadings are not listed in the interest of space. (Figure 4 omitted)

The factor loadings in Figure 4 were all robust. The statistics evaluating the fit of the model to the data were the following: chi-square 149.44 (p = .000), GFI .95, RMR .050. Again, two of the three statistics demonstrated an adequate fit to the data. The t- values were all large, and the modifications indices were small.

Figure 3 and Figure 4 provide evidence that when the conventions of scale construction were followed, at least some of the items of the DAS formed robust subscales that offered measures of consensus, satisfaction, and cohesion, as well as 7 first-order constructs. Whereas the original DAS was not validated with other samples, the RDAS should be relatively consistent across samples.

Construct Validation of the RDAS with Other Samples

When Crane et al. (1991) performed factor analyses on the DAS with the distressed and nondistressed samples separately, there were significant differences between the two groups. Over one third of the items had low factor loadings for the nondistressed sample; the distressed sample produced much better factor loadings. This suggested that there was a different factor structure for the nondistressed sample than for the distressed sample. To assess whether this was the case with the RDAS, the LISREL program was used to run a stacked model with the nondistressed and distressed samples.

The stacked model estimated the models simultaneously for both subsamples. When the models were estimated simultaneously, it was possible to confirm that the factor structures were equivalent for the two groups. If the resulting chi-square was nonsignificant, the two groups were considered to have equivalent factor structures. The chi-square for the stacked model was 31.21 (22, p = .092) and the GFI was .97, providing evidence that the factor structure of the RDAS was the same for both the distressed and nondistressed samples.

To further validate the RDAS with another sample, the data from the study published by Spanier and Thompson (1982) were used to perform a confirmatory factor analysis. Since the correlation matrix was provided in the 1982 article, it was possible to test the factor structure with their sample. The chi-square for the RDAS with 7 first-order scales was 78.73 (56, p = .024) and the GFI was .95. A close investigation of the modification indices showed that Item 20 (Do you ever regret that you married?) had large loadings on more than one subscale. It was understandable that this item would not be consistent when it was analyzed using the sample from Spanier and Thompson's (1982) study since in Spanier and Thompson's sample all of the respondents were separated from their partners. By the time individuals separate from their spouses, it is likely that the majority of them regret having married, whereas this would not be the case with a nonseparated sample. When Item 20 was set free to

vary on other subscales, the chi-square for the model was 69.48 (55, p = .091) and the GFI was .96, which represented an acceptable fit to the data.

## Construct Validity with the MAT

In previous studies with the DAS researchers have attempted to provide additional evidence of construct validity by calculating a correlation coefficient with another popular measure of marital adjustment, the Locke-Wallace Marital Adjustment Test (MAT; Crane et al., 1990; Spanier, 1976). It was hypothesized in this study that the RDAS would be an improvement over the DAS if the correlation coefficient between the RDAS and the MAT was similar to or higher than the coefficient between the DAS and the MAT. With this sample the correlation coefficient between the RDAS and the MAT was .68 (p <. 01); the correlation coefficient between the DAS and the MAT was .66 (p <.01). In addition, the correlation coefficient between the DAS and the RDAS was .97 (p <.01). These statistics provide evidence that the RDAS is an excellent representation of the domains of the DAS with less than half the items and that it is as highly correlated with the MAT as the 32-item DAS.

## Criterion Validity

The previous analyses have demonstrated that there is initial evidence of construct validity for the RDAS. An additional type of validity that would be important for the RDAS is criterion validity. In this study it was hypothesized that the RDAS would be an improvement over the DAS if it was as successful as the DAS at discriminating between distressed and nondistressed samples. This would be evidence of one type of criterion validity often called discriminant validity.

The discriminant analyses comparing the RDAS and the DAS demonstrated that the RDAS and the DAS were equal in their ability to classify cases as either distressed or nondistressed. Both scales correctly classified 81% of the cases. These results were surprising considering that the DAS had 18 more items than the RDAS.

The RDAS was somewhat better at correctly classifying nondiscussed respondents than distressed respondents. With the nondistressed respondents the RDAS had an 86% accuracy rate(14% false negatives),whereas with the distressed respondents only 74% were correctly classified (26% false positives). The results for the DAS were very similar, falling within 2 percentage points for both samples. In clinical and research work it is probably better to misclassify distressed individuals as nondistressed (false positive) than to misclassify nondistressed individuals as distressed (false negative) since the stigma of being labeled distressed is worse than being labeled nondistressed.

Discriminant analyses with the subscales of the RDAS were also conducted. The standardized discriminant coefficients for the Consensus, Satisfaction, and Cohesion subscales were .34, .55, and .32, respectively. These coefficients imply that the Satisfaction subscale had a larger influence on the discriminant ability of the RDAS than the other two subscales. When each subscale was analyzed alone, the Consensus subscale correctly classified 74% of the cases, the Satisfaction subscale correctly classified 75% of the cases, and the Cohesion subscale correctly classified 73% of the cases. Although the overall percentage of cases which were correctly classified by each subscale was similar, the Satisfaction subscale had the lowest level of false negatives, 16%, whereas the Consensus had a false-negative rate of 27% and the Cohesion subscale had a false negative rate of 23%.

The discriminant analyses results provided evidence that there was some criterion validity for the RDAS. The results also suggested that it was possible to discriminate between distressed and nondistressed individuals as

well with the 14-item RDAS as it was with the 32-item DAS.

Internal Consistency and Split-Half Reliability Estimates

Table 1 contains the Cronbach's alpha (internal consistency) reliability coefficients and the Guttman and Spearman-Brown split-half reliability coefficients for each of the subscales and the total RDAS. (Table 1 omitted) The reliability coefficients are within acceptable ranges and demonstrate that the RDAS has internal consistency and split-half reliability. Test-retest reliability is an important type of reliability that was not available as part of this study.

The Guttman split-half reliability coefficient for the RDAS was .94, which was particularly strong. In contrast, the Guttman statistic for the DAS was .88. The strong split half reliability of the RDAS suggests that it may be appropriate to divide the instrument into two forms, taking one item from each of the 7 first-order subscales (see Figure 3). This process would create two 7-item forms of the RDAS which could be used for pre- and posttests in clinical work and outcome studies. Two forms of an instrument are especially useful since they control for the biased results that can occur when a respondent is required to take the same test more than once (APA, 1985; Campbell &Stanley, 1963).

To evaluate further the appropriateness of dividing the RDAS into two equivalent forms, Items 1, 2, 3, 7, 8, 11, and 12 from Appendix A (see Appendix B also) were combined to form one version, and Items 4, 5, 6, 9, 10, 13, and 14 were combined to form the second version. (Appendix A omitted) A correlation coefficient was calculated between the two versions of the RDAS and the result was a coefficient of .90. This large coefficient provides additional evidence that it may be appropriate to use the RDAS as two separate forms. Future studies should evaluate the appropriateness of using the two forms as if they were equivalent.

In the original DAS, questions were selected from an existing pool of items instead of being written to measure specific constructs, hence hybrid subscales were created that spanned the first and second hierarchical levels. This problem was corrected by selecting out items that were homogeneous; 7 first-order scales were created which were combined to create the 3 second-order concepts of consensus, satisfaction, and cohesion. The confirmatory factor analyses provided evidence for the construct validity of the RDAS structure with the distressed, nondistressed, and total samples of this study as well as with the sample from Spanier and Thompson's (1982) study.

The evidence in this study suggests that the RDAS is an improvement over the DAS for the following reasons: (a) The RDAS, unlike the DAS, has acceptable levels of construct validity as demonstrated by several confirmatory factor analyses with more than one sample; (b) the RDAS is as highly correlated with the MAT as the DAS; (c) although the RDAS has less than half the items of the DAS, it is as successful at discriminating between distressed and nondistressed individuals; (d) the RDAS and its subscales have adequate internal consistency estimates and excellent split-half reliability coefficients, estimates which were larger than those of the DAS; (e) at this time there is some evidence that the RDAS can be divided into two forms and used in situations where repeated testing is necessary.

The summary statistics for the RDAS and its subscales are presented in Table 2. (Table 2 omitted) These statistics can be used to develop standardized scores and for comparisons with other samples. The RDAS is presented in Appendix A; the individual items are presented by subscale in Appendix B.

Implications for the Field of Marriage and Family Therapy

During the last two decades, a plethora of instruments have been created in the family fields. Many of these instruments are multidimensional in nature and as such should follow the construct hierarchy guidelines outlined by Comrey (1988) and illustrated in this study. As a whole, the majority of the multidimensional instruments in the field have not adhered to the principles of construct hierarchy. Common measures such as the Marital Satisfaction Inventory (Synder, 1979), the Family Environment Scale (Moos, 1981), and the Personal Assessment of intimacy in Relationships (PAIR; Schaefer &Olson, 1981)are examples of instruments that contain hybrid subscales that span the first and second hierarchical levels. These scale developers did not delineate a clear construct hierarchy for their instruments and used exploratory factor analyses on hybrid subscales to determine "construct validity." This is the same way that the subscales of the original DAS were evaluated and is one of the reasons that the subscales have not been validated with other samples. Nevertheless, many of these instruments have confirmed criterion validity which, depending on how the instrument is used, can be more important than construct validity.

Assessment instruments can be valuable tools for clinicians, helping them determine levels of distress, gather important background information in an expedient way, understand clients' subjective views of the world, and evaluate outcome. Assessment instruments are also vital in research as independent and dependent variables. Nevertheless, if the instruments are going to continue to be used to answer such important questions as therapeutic effectiveness, they must be valid and reliable.

One of the principles that is related to both reliability and validity is construct hierarchy. Construct hierarchy is related to reliability because the consistency of a scale is closely tied to the homogeneity of the items within the scale. Subscale items which are homogeneous will have high correlations which will produce high reliability coefficients; hybrid subscales that contain more heterogeneous items will often have lower reliability coefficients.

Construct hierarchy is directly related to one of the forms of validity known as construct validity. Although it is not within the realm of this paper to discuss construct validity in detail (for comprehensive information on this issue readers are encouraged to read the manual on Standards for Educational and Psychological Testing, APA, 1985), it is necessary to mention a few major points. Construct validity is the type of validity that is concerned with assessing whether the instrument measures the concept of interest. The concept of interest for a particular instrument must be embedded in a clear conceptual framework that delineates what the construct means and how it is related to other variables. This conceptual framework is what the researcher uses to organize subscales and then develop items to fit within each subscale. The subscales can then be evaluated with statistical methods as they were in this study. As can be seen, the conceptual framework, instrument development, and evaluation of the validity of an instrument are all intricately tied to issues of construct hierarchy.

Criterion validity is an additional type of validity that is very important. Criterion validity evaluates how successful an instrument is at predicting some important outcome, such as membership in a distressed or nondistressed marital group. Criterion validity is closely related to construct validity and reliability since it is impossible to have a valid measure that is unreliable.

This study has important implications for the field of marriage and family therapy. The outcome of the study is an improved version of the DAS that can be used to evaluate dyadic adjustment in distressed and nondistressed relationships. The process used in the study can be implemented by professionals in the field to develop and revise installments so that they are more reliable and valid.

A reliable, valid, and short 14-item instrument with 7 first-order concepts and 3 second-order concepts is a useful addition to the existing measures in the family field. In addition, a marital instrument that can be divided into two

forms is a necessity for much of the repeated testing that occurs in the field. Nevertheless, there are certainly important areas of dyadic functioning that are not measured adequately by the RDAS, such as marital finances and communication. An improved theory of dyadic adjustment should be developed that includes the major concepts important to marital adjustment delineated into first- and second-order levels. Once this theory is developed, more questions can be added to the RDAS to measure the missing concepts.

The process of deleting items, adjusting subscales, and evaluating constructs contained in assessment instruments is a laborious process that takes years of effort. The RDAS is an improvement over the DAS, which was an improvement over the MAT. Future studies should continue to tamper with the instrument by verifying it with other samples, clarifying the conceptual definition underlying the concept of dyadic adjustment, adding missing concepts, and deleting poor questions.

Figure 1 The Items of the DAS (Spanier, 1976) Grouped by Subscale

Dyadic Consensus subscale*

1. Handling family finances

2. Matters of recreation

3. Religious matters

5. Friends

7. Conventionality

8. Philosophy of life

9. Ways of dealing with parents or in-laws

10. Aims, goals, and things believed important

11. Amount of time spent together

12. Making major decisions

13. Household tasks

14. Leisure time interests and activities

15. Career decisions

Dyadic Satisfaction subscale

16. How often do you discuss or have you considered divorce, separation, or terminating your relationship?

17. How often do you or your mate leave the house after a fight?

18. In general, how often do you think that things between you and your partner are going well?

19. Do you confide in your mate?

20. Do you ever regret that you married?

21. How often do you and your partner quarrel?

22. How often do you and your mate "get on each other's nerves"?

23. Do you kiss your mate?

31. Rate how happy you are in your relationship, ranging from extremely unhappy to perfect.

32. Rate your feelings about the future of the relationship, ranging from wanting the relationship to succeed at any cost to feeling that the relationship can never succeed.

Dyadic Cohesion subscale

24. Do you and your mate engage in outside interest together?

25. Have a stimulating exchange of ideas

26. Laugh together

27. Calmly discuss something

28. Work together on a project

Affectional Expression subscale

4. Demonstrations of affection

6. Sex relations

29. Being too tired for sex

30. Not showing love

* The Answer Scales for the items of the DAS are available in Spanier (1976).

Appendix B

The Questions of the RDAS Grouped by Subscale

Consensus

Decision Making

Item 3. Making major decisions

Item 6. Career decisions

Values

Item 1. Religious matters

Item 5. Conventionality (correct or proper behavior)

Affection

Item 2. Demonstrations of affection

Item 4. Sex relations

Satisfaction

Stability

Item 7. How often do you discuss terminating your relationship?

Item 9. Do you ever regret that you married?

Conflict

Item 8. How often do you and your partner quarrel?

Item 10. How often do you and your mate "get on each other's nerves"?

Cohesion

Activities

Item 11. Do you and your mate engage in outside interests together?

Item 13. How often do you work together on a project?

Discussion

Item 12. How often do you have a stimulating exchange of ideas?

Item 14. How often do you calmly discuss something?

REFERENCES

American Psychological Association. (1985). Standards for educational and psychological testing. Washington, DC: Author.

Campbell, D. T., &Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. Boston, MA: Houghton Mifflin.

Comrey, A. L. (1988). Factor-analytic methods of scale development in personality and clinical psychology. Journal of Consulting and Clinical Psychology, 56, 754-761.

Crane, D. R., Allgood, S. M., Larson, J. H., &Griffin, W. (1990). Assessing marital quality with distressed and nondistressed couples: A comparison and equivalency table for three frequently used measures. Journal of Marriage and the Family, 52, 87-93.

Crane, D. R., Busby, D. M., &Larson, J. H. (1991) A factor analysis of the dyadic adjustment scale with distressed and nondistressed couples. American Journal of Family Therapy, 19, 60-66.

Joreskog, K. G., &Sorbom, D. (1989). LISREL 7: A guide to the program and applications. Chicago, IL: SPSS.

Kazak, A. E., Jarmas, A., &Snitzer, L. (1988). The assessment of marital satisfaction: An evaluation of the dyadic adjustment scale. Journal of Family Psychology, 2, 82-91.

Lavee, Y. (1988). LISREL in family research. Journal of Marriage and the Family, 50, 929-936.

Locke, H. J., &Wallace, K. M. (1959). Short marital adjustment and prediction tests: Their reliability and validity. Marriage and Family Living, 2, 251-255.

Moos, R. H. (1981). The family environment scale. Palo Alto, CA: Consulting Psychologists Press.

Norton, R. (1983). Measuring marital quality: A critical look at the dependent variable. Journal of Marriage and the Family, 49, 141-151.

Sabatelli, R. M. ( 1988). Measurement issues in marital research: A review and critique of contemporary survey instruments. Journal of Marriage and the Family, 50, 891-915.

Sabourin, S., Lussier, Y., Laplante, B., &Wright, J. (1990). Unidimensional and multidimensional models of dyadic adjustment: A hierarchical reconciliation. Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2, 333-337.

Schaefer, M. T., &Olson, D. H. (1981). Diagnosing intimacy: The PAIR inventory. Journal of Marital and Family Therapy, 7, 47-60.

Schumm, W. R., Paff-Bergen, L.A., Hatch, R. C., Obiorah, EC., Copeland. J. M., Meens, L. D., &Bugaighis, M. A. (1986). Concurrent and discriminant validity of the Kansas marital satisfaction scale. Journal of Marriage and the Family, 48, 381-87.

Sharpley, C. E, &Cross, D. G. (1982). A psychometric evaluation of the Spanier dyadic adjustment scale. Journal of Marriage and the Family, 44, 739-741.

Snyder, D. K. (1979). Multidimensional assessment of marital satisfaction. Journal of Marriage and the Family, 41, 813-823.

Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. Journal of Marriage and the Family, 38, 15-28.

Spanier, G. B. (1979). The measurement of marital quality. Journal of Sex and Marital Therapy, 5, 288-300.

Spanier, G. B. (1985). Improve, refine, recast, expand, clarify--don't abandon. Journal of Marriage and the Family, 47, 1073-1074.

Spanier, G. B. (1988). Assessing the strengths of the dyadic adjustment scale. Journal of Family Psychology, 2, 92-94.

Spanier, G. B., &Cole, C. L. (1975). Toward clarification and investigation of marital adjustment. International Journal of Sociology of the Family, 6, 121-146.

Spanier, G. B., &Thompson, L. (1982). A confirmatory analysis of the dyadic adjustment scale. Journal of Marriage and the Family, 44, 731-738.

Thompson, L. (1988). Women, men, and marital quality. Journal of Family Psychology. 2, 95-100.

Thompson, L., &Spanier, G. B. (1983). The end of marriage and acceptance of marital termination. Journal of Marriage and the Family, 45, 103-113.

Dean M. Busby, PhD, is an Assistant Professor in the Marriage and Family Therapy Program, Department of Child and Family Studies, Syracuse University, 008 Slocum Hall, Syracuse, NY 13244.

D. Russell Crane, PhD, is the Director of the Marriage and Family Therapy Programs,. Department of Family Sciences, Brigham Young University, Provo, UT 84602.

Jeffry H. Larson, PhD, is an Associate Professor in the Marriage and Family Therapy Program, Department of Family Sciences, Brigham Young University, Provo, UT 84602.

Clark Christensen, MS, is a Doctoral Student in the Marriage and Family Therapy Program, Department of Child and Family Studies, Syracuse University, 008 Slocum Hall, Syracuse, NY 13244.

## DETAILS

| | |
|---|---|
| Copyright: | Copyright American Association for Marriage and Family Therapy Jul 1995 |
| Last updated: | 2018-10-05 |
| Database: | Social Science Premium Collection,Research Library |

## LINKS

Linking Service