THE CONSTRUCT VALIDITY OF
THE DIFFERENTIAL EMOTIONS SCALE FOR CHILDREN AND ADOLESCENTS

WILLIAM E. KOTSCH   DAVID W. GERBING

LYNNE E. SCHWARTZ

BAYLOR UNIVERSITY

MARCH 25, 1980

FINAL DRAFT OF A CHAPTER TO APPEAR IN
C.E. IZARD, ED.
MEASURING EMOTIONS IN INFANTS AND CHILDREN
CAMBRIDGE UNIVERSITY PRESS

I. INTRODUCTION

The Differential Emotions Scale for Children (DES-III) is a brief self-report inventory which assesses ten fundamental emotions in adolescents and in children. Izard developed the DES-III from both his theory of fundamental emotions (Izard, 1971, 1977) and through extensive empirical research on the adult version of the scale, the Differential Emotions Scale (DES) (Izard, 1972; Izard, Dougherty, Bloxom and Kotsch, Note 1). The aim of the chapter is to present the DES-III and evaluate the underlying measurement model through a series of exploratory and confirmatory factor analyses.

A. DIFFERENTIAL EMOTIONS THEORY

Differential emotions theory (Izard, 1971, 1977) emphasizes discrete emotions as distinct experiential motivational processes. Emotion is defined as a complex process with neurophysiological, neuromuscular, and phenomenological aspects. The theory postulates ten fundamental emotions: interest, joy, surprise, sadness, anger, disgust, contempt, fear, shame/shyness, and guilt. Each emotion has (a) a specific innately determined neural substrate; (b) a characteristic facial display or neuromuscular-expressive pattern; and (c) a distinct subjective or phenomenological quality. No one of these three facets constitutes emotion; each is a component of emotion. Thus each fundamental emotion is a system made up of its three components and their interactions. Internal or external events can elicit more than one emotion, and through both innate and learned relationships one emotion can elicit another. The fundamental emotions are interrelated in dynamic and relatively stable ways. The relatively stable combinations of emotions that occur simultaneously or sequentially are called patterns. In day-to-day human experience it appears that these patterns occur more frequently than do pure emotions.

B. THE DIFFERENTIAL EMOTIONS SCALE (DES)

The Differential Emotions Scale (DES) is a self-report instrument designed for use in the assessment of an individual's experience of fundamental emotions or patterns of emotions. The DES was originally conceived as a "state" measure of one's emotions, but variations in the instructions allows the same set of scales to be used in the assessment of emotions experienced over an extended period of time. The frequency with which an emotion is experienced over time may be viewed as an "emotion trait." This variation of the DES is labeled the Differential Emotions Scale II.

The DES and DES II both consist of thirty adjectives (items), three adjectives for each of the ten fundamental emotions listed above. The DES items are presented in Table 1. The usual DES instructions ask an individual to rate, on a single five-point intensity scale, the extent to which each word describes the way he or she feels at the present time. The DES-II, on the other hand, is composed of a five-point frequency scale. The DES-II instructions ask the individual to consider a specified time period in his or her day-to-day life and to rate the frequency with which he or she experiences each emotion during the time considered. The time period may be a specified period of time in the past (i.e., the past week, month, experiment, therapy hour) or may refer generally to the individual's entire life (i.e., "how often do you have these feelings in your day-to-day living?").

Insert Table 1 about here.

Development of the Scale. The vocabulary in the DES items was derived from an analysis of verbal labels of facial expression. Extensive research has demonstrated that subjects of different ethnic origins, languages, and cultures can differentiate among the facial expressions of the fundamental emotions and label them with a high degree of concenses (Ekman, 1973; Izard, 1971). As part of this research, American, English, French, and Greek subjects were asked to give verbal descriptions of the emotions displayed in a series of cross-culturally standardized facial photographs of the fundamental emotions. This work provided the background for the development of the DES by generating a set of words for each of the emotions that could be viewed as transcultural definitions. Exploratory factor analyses were employed as a guide for the final selection of the thirty DES items.

Validity. Several types of validity studies have been conducted with the DES (Izard, 1972, 1977). The DES has been subjected to repeated exploratory factor analytic studies which have shown the emotion factors to be highly stable. In all of these analyses, the factors obtained corresponded to the theoretically defined factors.

The DES also has been used in studies of anxiety and depression. Results of these investigations have shown that anxiety and depression can be conceptualized as patterns of the fundamental emotions. One such factor analytic study investigated the structure of anxiety. "Anxiety" was operationally defined by both the Speilberger-Gorsuch-Luchene (1970) State-Trait Anxiety Inventory (STAI) and the fundamental emotions measured by the DES. The two scales were combined and administered to college students who had been instructed to visualize a situation which made them highly anxious. The students used the DES and STAI to describe their emotions and feelings in that situation. An exploratory factor analysis produced eight factors which represented the fundamental emotions. All STAI items which had clear emotional connotation separated and had their primary loading on factors representing four of the fundamental emotions--fear, sadness, guilt and joy, the latter factor including negatively-weighted STAI items.

A similar factor analytic study related the fundamental emotions to the concept of "depression." Nineteen depression items of Zuckerman's MAACL (Zuckerman, 1960) were combined with the DES items. An exploratory factor analysis produced eight factors representing the fundamental emotions. Most of the negative MAACL items had their primary factor loadings on the DES factor of sadness. Several of the positive MAACL items loaded on the DES factor of interest. These studies show the utility of DES in measuring anxiety and depression and demonstrate that the STAI and MAACL are both multidimensional scales consisting in part of several DES factors.

Mosher and Toedter (Note 2) used the DES to determine and compare the typical patterns of emotions involved in love and jealousy. College students were asked to describe a situation in which they had experienced love and another in which they had experienced jealousy. The subjects completed the DES while visualizing each situation. The love situation was characterized by interest, joy sexual arousal, and mild surprise. Disgust, anger and contempt were rarely experienced. In contrast, the experience of jealousy was characterized by sadness and anger, followed by surprise, contempt, and disgust. In addition, a pre-test measure of attitudes toward jealousy added a significant increment of prediction to the occurence of anger, contempt, and disgust.

Criterion-related validity data has come from a study of the test-anxiety situation (Barlett and Izard, 1972). Subjects took the STAI and the DES just prior to taking a major examination in their introductory psychology course. The students were divided into high and low anxiety groups on the basis of their STAI scores. High anxiety subjects had significantly higher DES means for fear, shame/shyness, guilt, sadness, and anger and a significantly lower DES mean for joy.

In a study of the relationship of subjective sexual arousal and emotion, Mosher and O'Grady (1979) showed a sex film to women rated either high or low on the Mosher Forced-Choice Guilt Inventory (Mosher, 1968). The women reported their level of sexual arousal and genital sensations both during and after the film. After the film the women completed the DES and a third measure of sexual arousal. Results showed moderate positive correlaions (.48 to .84) between the emotions of interest and joy and the three measures of sexual arousal, and positive correlations between fear and guilt (.65) and between guilt and shame (.72). Women who rated themselves high on the trait measure of sex-guilt reported feeling more guilt, disgust, and fear after viewing the sex film than did women who rated themselves low on sex guilt.

A second type of criterion-related validity data has come from the study of various psychodiagnostic groups (Izard, et.al., Note 1). The DES II was admininstered to psychiatric patients of various psychodiagnoses and to normal college students. The instructions for this administration of the DES II asked the patients to rate their emotions as experienced over the "past year or two." Results showed that subjects diagnosed as neurotic reported experiencing less joy and more sadness, fear, and guilt in their daily lives than did normal subjects. Subjects diagnosed with personality disorders reported experiencing less joy than normal subjects. Subjects with adjustment reactions reported more disgust than normals. Schizophrenics reported significantly more disgust and contempt than normal subjects.

The majority of the research on the DES used college students as subjects. Izard (1972) reported a factor analytic study of depression involving high school students from middle and lower-middle income groups. The factor structure of the DES was similar to that found in an identical study with college students. However, the magnitude of the factor loadings ranged .2 to .3 below those found in the college sample. The explanation which most likely accounts for this drop in factor loadings is that some of the high school students did not know the meaning of some of the DES adjectives. If some individuals do not understand the meaning of some items, their responses on those items do not contribute to the measurement of the emotion but rather add error. The additional measurement error would attenuate the item-factor correlations. Since this measurement error could be reduced by simplifying the items, a new form of the scale, the DES-III, was developed.

C. THE DIFFERENTIAL EMOTIONS SCALE FOR CHILDREN (DES-III)

Izard (Note 3) developed the DES-III for use with children and adolescents, and with adults having limited education. The DES-III was constructed by translating and expanding the single adjectives of the DES into phrases that describe an aspect of the experience of the fundamental emotions. The items of the DES-III are presented in Table 2.

The DES-III is scored by summing the scores for the three items that measure each emotion. This results in one score for each of the ten fundamental emotions. The resulting emotion profile then can be inspected to determine the relative strength (state instructions) or frequency (trait instructions) of individual emotions or of emotion patterns.

Insert Table 2 about here.

The investigations presented here were aimed at determining the construct validity of the DES-III an analysis of the internal structure of the scale. The investigation of the dimensionality of the DES-III consisted of three separate studies. The first study determined if college age individuals who are clearly capable of understanding the meaning of the items on the DES and the DES-III would interpret the corresponding items similarly. Given the favorable results from the first study, the focus of the second and third studies was the analysis of the responses to the DES-III of children and adolescents. While we felt that the youngest children in our study were capable of experiencing all ten emotions, the question remained as to whether these children had the cognitive and verbal ability to report the subjective experiences of all their emotions using a self-report inventory such as the DES.

D. THE MODEL

The a priori measurement model specified by the DES and the DES-III defines 10 emotions that are conceptualized as latent variables or factors. Each cluster of observed variables is hypothesized as unidimensional; that is, each of the items within a cluster is specified as an alternate indicator of a single, common underlying emotion. The ten emotions are specified as distinct entities although they may be correlated as predicted by the theory.

The equations for this multiple indicator measurement model are the equations of classical reliability theory. Let

X_j be the j^{th} observed variable or item
F_k be the k^{th} latent variable or factor
e_j be the error term of the j^{th} item
β_j be the regression weight of X_j on F_k

That is,

$$X_j = \beta_j F_k + e_j$$

In the language of factor analysis β_j is, by definition, the factor pattern coefficient of item X_j on factor F_k. Since F_k is the only "predictor" in this equation, by implication, β_j is also the factor loading of X_j on F_k, i.e., $\beta_j = r_{X_j F_k}$.

The complete 30 x 10 factor pattern matrix for the model underlying the DES-III represents the limiting form of simple structure. Each of the 30 items is a function of only one of the ten factors, and each factor is represented by a small subset of the 30 items.

The goal of the analysis was to determine the extent to which the items of the DES-III conform to the equations of the measurement model. The first analysis primarily investigated the equivalence of the factors measured by the DES and the DES-III for college age subjects. Given this equivalence, the second and third studies addressed the dimensionality of the DES-III for children and adolescents.

II. ANALYTIC TECHNIQUE

A. COMPUTATIONAL ANALYSES

The analysis strategy consisted of a series of exploratory or blind factor analyses followed by a set of confirmatory factor analyses. The blind analyses, principal axis with communalities followed by an orthogonal (varimax) or oblique (direct oblimin) rotation , are used to suggest the factorial composition of the observed variables. However, these blind analyses can only provide crude tests of the hypothesized internal structure of the DES-III. As Joreskog (1978) writes,

> "That exploratory factor analysis may be quite useful in the early stages of experimentation or test development is widely recognized. . . The results of an exploratory analysis may have heuristic and suggestive value . . .and may generate hypotheses which are capable of more objective testing by other multivariate methods. . . It is highly desirable that a hypothesis which has been suggested by mainly exploratory proceedures should subsequently be confirmed, or disproved, by obtaining new data and subjecting these to more rigorous statistical techniques" (pp. 443, 444).

Thus the confirmatory factor analyses, which provide a direct test of the specified model, were considered primary in determining the construct validity of the DES-III. Both the exploratory and confirmatory analyses may be used, in conjunction with the analysis of the item content, to suggest revisions of the model.

Since the a priori model predicts a partitioning of the observed variables into mutually exlusive clusters, the initial confirmatory analysis is provided by a multiple groups solution in which each group factor is defined by a cluster or group of observed variables. Discussions of this technique are found in Spearman (1904, 1912), Burt (1908, 1917) and Holzinger (1944) and more recently by Nunnally (1978) and Hunter and Gerbing (Note 4). The analysis computes estimates for (a) the correlations of each item with each of the group factors and (b) the correlations of the group factors with each other. The computations of a multiple groups analysis can be accomplished with either a centroid factor analysis using a computer program such as PACKAGE (Hunter and Cohen, 1969; Hunter and Gerbing, Note 4) or a full information maximum likelihood analysis with LISREL (Joreskog and Sorbom, 1978).

LISREL is the more general program for a confirmatory analysis in the sense that it is not confined to a multiple groups solution. LISREL accommodates limited multiple factor models in which some of the observed variables may be expressed as a function of several of the latent variables. The statistical fit of a multiple groups model may sometimes be improved by permitting some of the items to have multiple antecedents.

Since each item is conceptualized as an imperfect indicator of the underlying emotion, the measurement model for each item contains an error term. According to this model, the emotion is distinct from the error introduced by the measurement process. The appropriate statistical adjustment for this error is to factor the correlations with communalities (e.g., Harmon, 1976). The use of communalities insures that the latent variable F and not the observed factor or cluster score is the entity of interest. That is, the use of communalities implies that the computed item-factor and factor-factor correlations are corrected for the attenuation due to the unwanted influence of the error term.

For example, in a centroid multiple groups analysis, each factor is defined as an unweighted sum of the corresponding items. If the analysis is conducted without communalities, the observed score variances and covariances are analyzed. Thus the computed item-factor and factor-factor correlations are also observed correlations. The problem is that the observed scores contain both true and error score components. Thus the computed correlations are attenuated due to the inclusion of this error in the computations. The error component is removed from the analysis by using communalities (e.g., Harmon, 1976; Hunter and Gerbing, Note 4). The resultant item-factor and factor-factor correlations are still defined as the sum of the corresponding items, but they are only estimated since the factors are now defined as latent variables. However, these are the correlations of interest since they contain only the true score component. In general, these correlations will be larger than the corresponding observed correlations.

B. EVALUATING THE MODEL

The fit of the measurement model to the data is based on the evaluation of the item correlations, i.e., the item covariance structure. The model fits if the residuals defined by subtracting the observed correlations from the correlations predicted by the model are zero to within sampling error. Given standardized data, the covariance structure of a multiple groups model is described by the product rules for internal and external consistency (Hunter and Gerbing, Note 4). This covariance structure was first presented by Spearman in 1904 and 1912.

The product rule for internal consistency states that, for items X_1 and X_2 which are indicators of the same factor F,

$$r_{X_1X_2} = r_{X_1F}\, r_{X_2F} = \beta_1 \beta_2$$

The product rule for external consistency states that for factor G and item X, which is an indicator of factor F,

$$r_{XG} = r_{XF}\, r_{FG}$$

External consistency implies that the correlations of two items within the same cluster are proportional across factors. That is,

$$\frac{r_{X_1G}}{r_{X_2G}} = \frac{r_{X_1F}\, r_{FG}}{r_{X_2F}\, r_{FG}} = \frac{r_{X_1F}}{r_{X_2F}}$$

A heuristic test of the fit of the model is based on the size of the residuals, i.e., on the number of predicted item correlations not within a specified number of standard errors from the corresponding observed correlations. If this number exceeds the number expected by chance, for a given significance level, the model may be rejected.

LISREL also provides a χ^2 test of fit of the model under the null hypothesis that the observed covariances conform to the constraints imposed by the model (e.g. the constraints of internal and external consistency). The test, which is based on the likelihood ratio derived from the maximum likelihood analysis, is a function of the number of cases analyzed and the value of the likelihood function at maximization. The likelihood ratio is a function of the residuals such that the χ^2 value is zero when the residuals are zero (e.g. Joreskog, 1978).

Joreskog (1978) has warned, however, that " the values of χ^2 should be interpreted very cautiously because of the sensitivity of χ^2 to various model assumptions such as linearity, additivity, multinormality, etc., but also [because]...if a sufficiently large sample were obtained, the test statistic would, no doubt, indicate that any [proposed] model is untenable" (p.448). Thus the evaluation of the overall fit of the model is probably best made by a direct examination of the residuals.

LISREL also provides two methods for examining the significance of individual parameters. First, LISREL computes the asympotic standard errors of each parameter. Second, some of the constraints of a model might be relaxed. "If the drop in χ^2 (between the original and "relaxed" model) is large compared to the difference in degrees of freedom, this is an indication that the change made in the model represents a real improvement" (Joreskog, 1978, p. 448).

III. STUDY I

A. METHOD

Two-hundred and six college students, enrolled in an introductory psychology course at Baylor University, completed a combined 60 item DES + DES-III immediately following the final examination. The students were instructed to indicate how they felt at the present moment. The 60 items were presented in a random order. All items were measured on the following 5-point Likert scale: very slightly or not at all (0), slightly (1), moderately (2), considerably (3), and very strongly (4).

B. RESULTS AND DISCUSSION

Although the primary emphasis is on testing the proposed model with confirmatory factor analyses, most factor analytic studies have been exploratory. To provide continuity with these exploratory analyses, a traditional exploratory factor analysis is presented before the confirmatory analyses.

Exploratory Analysis. The 60 x 60 item correlation matrix was analyzed with a ten factor principle axis factor analysis followed by a varimax rotation. The resulting 10 factor x 60 item matrix of factor loadings was inspected and each item was grouped with the factor on which the item had its highest loading. This procedure identified nine of the ten emotions. For five of the a priori six-item clusters--interest, joy, surprise, anger, and guilt--all six items loaded highest on a single factor. All twelve items of the sadness and fear clusters grouped together on one factor.

The shame/shyness cluster was defined by five items. The sixth item of the shame/shyness cluster, DES-III item 25, loaded highest on the guilt factor and had its second highest loading on the shame/shyness factor.

The three items of the DES-III contempt cluster had reasonably high loadings on a unique factor. However, two items of the DES contempt cluster loaded highest on the anger factor, the third item had identical loadings on the contempt and the sadness-fear factor.

The disgust factor was defined by two items of the DES-III and one item of the DES. The other three items in the disgust clusters had loadings above .3 on this factor.

Confirmatory Analysis. The DES + DES-III was analyzed in a twenty factor centroid multiple groups factor analysis provided by PACKAGE. One DES factor and one DES-III factor were postulated to measure each of the ten fundamental emotions specified by differential emotions theory: interest, joy, surprise, sadness, anger, disgust, contempt, fear, shame/shyness, and guilt. The loadings of each of the DES-III items on the corresponding DES factor and the correlations between the DES and the DES-III group factor scores for each emotion are presented in Table 3. The factor loadings of the DES-III items on the corresponding DES factors and the high correlations between the DES and DES-III factors indicate that the new DES-III items provide a good measure of the fundamental emotions as measured by the DES.

Insert Table 3 about here.

All correlations between corresponding group factors on the DES and DES-III were above .92 except interest and contempt which were .67 and .66 respectively. Examination of the item-factor intercorrelation matrix revealed that the DES-III interest cluster is not as strong as the interest cluster of the DES. Specifically, DES-III item three shares common variance with the joy factor. However, the lower correlations between the DES and DES-III contempt clusters are due to weakness in the DES items. The contempt items emerge as a strong factor on the DES-III.

The results of this study demonstrate that the DES and DES-III are equivalent measures of the fundamental emotions for individuals who comprehend the meanings of the DES items.

IV. STUDY II

A. METHOD

Eight hundred and fifty-seven junior and senior high school students ranging in age from 11 years 4 months to 17 years 10 months completed the DES-III during a class period two days before Christmas vacation. The DES-III was administered by the teachers following a demonstration by one of the investigators. Students were asked to indicate how they had felt during the previous week. The children were asked to reflect on their experiences during the previous week before completing the scale. The children were given no information concerning the meaning of words or phrases on the DES-III. Instructions and sample items of the DES-III are presented in Table 4.

Insert Table 4 about here.

B. RESULTS AND DISCUSSIONS

Exploratory Analysis. To provide continuity with the traditional factor analytic study, the 30 x 30 item correlation matrix was analyzed with a ten factor principle axis factor analysis followed by a varimax rotation. The resulting 10 factor x 30 item matrix of factor loadings was inspected and each item was grouped with the factor on which the item had its highest loading. This procedure identified nine of the ten emotions. The three items in seven of the a priori clusters defined seven distinct factors. The interest and sadness factors were defined by two items each. Item three in the a priori interest cluster was placed with the items defining the joy factor and item twelve of the a priori sadness cluster was placed with the items defining the anger factor. The items of the a priori disgust cluster failed to define a single factor. Instead, the items of the disgust cluster loaded moderately on the anger, contempt, and joy factors.

The results of this exploratory factor analysis provided preliminary evidence supporting the measurement model. However, an orthogonal solution is unnecessarily restricting since there is no theoretical reason to expect the underlying emotions to be uncorrelated.

A moderately oblique rotation of the initial principle axis factor solution did indicate even stronger support for the a priori measurement model. (A direct oblimin rotation with δ = 0 (Harmon, 1976) was used.) For eight of the a priori clusters, all three items loaded highest on a single, unique factor. The six items in the anger and disgust clusters merged to define a single cluster. Thus, for each of the three item factors, in no instance did any of the items split apart on separate factors.

Insert Table 5 about here.

Table 5 presents the factor pattern matrix of the oblique principle axis factor solution. The pattern coefficients are the weights that would be assigned to each factor in expressing the item as a linear function of the factors. When these weights are very small, the factor makes a negligible contribution to the variable. Inspection of the table shows that the three items of an a priori cluster are primary in estimating the score on the corresponding factor. This finding provides justification for defining a latent variable as a function of the items in each cluster and performing a multiple groups confirmatory analysis on the 30 x 30 item correlation matrix.

Confirmatory Analysis. The a priori measurement model specifying 10 factors each consisting of three items was analyzed with the centroid factor multiple groups factor analysis provided by PACKAGE. The average within-cluster item correlation (excluding self correlations) equalled .40 (Range = .23 to .57). In contrast 26% of the between-cluster item correlations were below .1, 58% were below .2, and 84% were below .3. Only 2% of the between-cluster item correlations exceed the average within-cluster value of .40.

A residal matrix was constructed by subtracting the obtained item-item correlations from the value predicted by the measurement model. A heuristic evaluation of the overall fit of the model can be made by examining the magnitudes of the residuals. A standard error for each observed correlation was computed by using the observed correlation as an estimate of the corresponding population correlation. Of the 435 obtained correlations, 110 or 25% exceeded their predicted value by more than two standard errors.

Insert Table 6 about here.

An additional test of the model is the evaluation of the parallelism of the 30 items with the ten group factors. The matrix of factor loadings is presented in Table 6. An examination of the columns of Table 6 reveals that the item-factor correlations generally conform to the assumption of external consistency. Seventy percent of the boxed loadings exceed .6; the lowest is .37. The relatively large size of the correlations of the items with their own group factor in contrast to the items' correlations with the other group factors provides evidence for the validity of the a priori model.

However, notable exceptions to external consistency are present for items 3, 21, 24 and 25. In several instances loadings are not parallel with the loadings of the remaining two items in their cluster. For example, in the first box in Table 6 , the correlation of item 3 with interest is approximately 70% the size of the correlation of items 1 and 2 with interest. The correlation of item 3 with joy should be about 70%, the size of the correlations of items 1 and 2 with joy, or about .28. The observed value of .48 is much larger than the predicted value.

The a priori multiple groups model was also analyzed with the full information maximum likelihood method of LISREL. The PACKAGE and LISREL solutions agreed fairly well since the largest single difference in the two estimated factor patterns was .17 and only four of the 30 estimated pattern coefficients differed by more than .10. The estimated factor-factor correlations in the two solutions were almost identical; the largest discrepancy was only .08.

The overall test of fit of the multiple groups model was $\chi^2 = 965.90$ for $df = 360$. Since $p > .01$, a literal interpretation of this statistic indicates that the model does not fit the data. A heuristic, but more reasonable evaluation of fit is provided by the residuals. The magnitude of the observed correlations varied from .00 to .57. The respective sampling errors of population correlations for these values for 857 subjects are .034 and .028. Using .03 as an average value, a predicted correlation is within two standard errors of the corresponding observed correlation if the magnitude of the residual is no larger than .06. Of the 435 residuals, 70 or 16.1% were larger than .06 and the largest residual was .160.

Thus both the centroid and full information maximum likelihood analyses of the multiple groups model suggests that the fit of the model could be improved statistically. To improve the statistical fit of the model it is necessary to generalize the multiple groups model to a multiple factor model, in which some of the items are linked to more than a single factor. A variety of respecified models were analyzed with each respecification based on the information provided by the residuals. The factor pattern for the final analysis is presented in Table 7. Although the χ^2 test of fit was negative, (χ^2 = 571.49, df =329, p > .01), the analysis of the residuals indicated that the fit of the model was quite good. Only 15 of 435 or 3.4% of the residuals were larger than the approximate two standard errors of .06. Moreover, the magnitude of the largest residual was only .086.

Insert Table 7 about here.

The statistical importance of the additional factor pattern coefficients in the revised model, which is a multiple factor generalization of the original multiple groups model, is evaluated in two different ways. First, the difference in the χ^2 values for the original and revised models was 394.42. This large drop in the χ^2 value relative to the loss of 31 degrees of freedom indicates the significance of the additional pattern coefficients as a set. Second, the magnitudes of all but four of the estimated pattern coefficients were larger than twice their standard errors. The smallest ratio of a parameter estimate to its standard error was 1.61.

Despite the improved fit obtained with the revised model, the original multiple groups model remains valid. The size of most of the pattern coefficients which were not part of the original model is small. Item 3 is the only item for which a pattern coefficient off the diagonal block is larger than the corresponding diagonal block coefficient. And only for items 3, 21, 24, and 25 does the size of an off diagonal block coefficient approach the size of the coefficient on a diagonal block.

A second indication of the utility of the original multiple groups model is derived from an analysis of factor scores. Although a set of differential weights for computing factor scores could be derived from a multiple groups analysis, the simplest method is to define each factor score as the unweighted sum of the items which define the group. In the multiple factor model, each factor score is computed as a weighted linear composite of all thirty items. The weighting reflects the size of the pattern coefficients as well as the correlations among factors.

The correlations between the factor scores computed by unit weighting and from the factor score regression weights derived from the multiple factor confirmatory analysis are presented in Table 8. The magnitude of these correlations provides strong evidence for the validity of the original multiple groups model. The ten latent variables defined by the more "sophisticated" multiple factor model are almost identical to the ten latent variables defined by the a priori model.

Insert Table 8 about here.

A separate confirmatory factor analysis using data from the youngest group of children (11.4 to 13.11) was performed to determine if the measurement model would show signs of failure. Both the factor loadings from the centroid multiple groups analysis and the factor pattern from the confirmatory multiple factor analysis were very similiar to the solutions presented in Tables 6 and 7. The analysis of the residuals from the multiple factor model indicated that 3.8% of the residuals exceeded their predicted values by two standard deviations, which is less than the 5% predicted by chance. Thus the measurement model provides a uniform fit across the age range represented in this study.

The correlations among factors derived from the centroid multiple groups solution are presented in Table 9. Factor means for each of the emotions were computed as an unweighted composite of the corresponding three DES-III items. These means for subgroups of the sample are presented in Table 10. Inspection of the table shows that the emotion profiles are similar across both age and sex.

Insert Tables 9 and 10 about here.

V. STUDY III

The purpose of this study was twofold: (a) to determine if the measurement model would fit the responses of children between the ages of eight and twelve, and (b) to determine if the DES-III can be improved through revision of the weakest items.

A. METHOD

The subjects were 457 elementary school children ranging in age from 8.1 to 12.2. Except for the scale modifications listed below, this study was identical to Study II. The instructions and the DES-III scale were modified accordingly: (1) The sample item, "Feel cheerful," which appeared in the instructions of the previous study (Table 4) was printed on the DES-III form. This allowed the children to complete the sample item. This item, which was not scored, appeared as example A on the form. (2) Four items of the scale were revised as:

Item 3A: Feel alert, kind of curious about something
Item 21A: Feel like you are better than somebody
Item 24A: Feel afraid
Item 25A: Feel ashamed, like you want to disappear

B. RESULTS AND DISCUSSION

An exploratory nine factor principle axis factor analysis followed by an oblique rotation revealed seven factors which accounted for nine of the a priori clusters. The interest and joy clusters merged to define one factor as did the anger and sadness clusters. The fear, surprise, guilt, shame/shyness, and contempt clusters each defined a unique factor. The disgust items failed to define a factor.

The a priori measurement model was analyzed with the confirmatory centroid multiple groups analysis. The average within-cluster item correlation (excluding self correlations) equaled .35 (Range = .16 to .64). Only 1.2% of the between-cluster item correlations exceeded the average within cluster value.

The matrix of factor loadings is presented in Table 11. Both the relatively large size of the correlations with their own group factor and the parallelism of clustered items across factors provides evidence for the validity of the a priori model. Table 11 shows that the revised items 3A, 21A, and 24A are good indicators of their respective latent factors. However, the size of the factor loading of item 21A is less than optimal. Item 25A is a poor indicator of shame/shyness for this sample.

Insert Table 11 about here.

An examination of the residual matrix showed that 16.5% of the observed correlations exceeded their predicted value by more than two standard errors. Many of these high residuals involve item 25. The multiple groups solution was computed omitting item 25 which resulted in a two item shame/shyness cluster. In this analysis, the number of observed correlations exceeding their predicted value by two standard deviations fell to 8.5%, indicating a good fit of the model.

The confirmatory multiple groups factor analysis of the 30 items using the maximum likelihood solution yielded a $\chi^2 = 671.9$ ($df = 360$, $p > .01$). Six percent of the observed correlations exceeded their predicted value by more than two standard errors. An analysis using a multiple factor model similiar to the one presented in Study II yielded a $\chi^2 = 395.17$ ($df = 321$, $p > .01$). Only two of the 435 observed correlations exceeded their predicted value by more than two standard errors.

The means and standard deviations of the ten factors are presented in Table 12. The correlations among the factors are presented in Table 13.

Insert Tables 12 and 13 about here.

This study provided additional support for the measurement model. First, the analysis of the model in study II was cross-validated in study III. Second, it was shown that the DES-III can be used to measure the ten fundamental emotions experienced by children as young as eight years. In addition, the revised Items 3A, 21A, and 24A are better indicators of thieir respective latent factors than the original items. These revised items should replace the original items in the DES-III.

VI. SUMMARY AND CONCLUSIONS

The DES-III is a 30 item self-report scale which measures the fundamental emotions defined by differential emotions theory. The scale can be administered to children above the age of eight years. Groups of three items are summed to obtain scores on ten emotions: interest, joy, surprise, sadness, anger, disgust, contempt, fear, shame/shyness, and guilt.

The construct validity of the model was assessed with exploratory factor analyses and with two multiple groups confirmatory factor analysis, using both centroid and the full information maximum likelihood methods. The exploratory analyses, expecially the factor pattern from the oblique rotation of Study II, suggested that the a priori model was appropriate for these data sets. The a priori model was directly tested with the confirmatory multiple groups analyses. Both multiple groups analyses of the DES-III provided reasonable support for the measurement model derived from differential emotions theory which asserts that the experience of emotion can be divided into ten discrete categories of emotion.

A multiple factor analysis, which permitted more causal links from the factors to the items, indicated that the modification of the a priori model improved the statistical fit of the model. In a statistical sense, the adjustments to the factor pattern of the common factor model provided by the full information maximum liklihood analysis resulted in a slightly better fit than the fit of the a priori theoretical solution provided by the multiple groups factor analysis. However, we conclude that the a priori model is superior in a substantive sense since (a) the a priori model is theoretically based and thus is more likely to cross validate, (b) most of the additional factor pattern coefficients in the multiple factor model were small in size, and (c) the factor scores obtained from unit weighting correlate very highly with the factor scores from the "improved" multiple factor solution.

There is, however, an important characteristic shared by both the multiple groups and multiple factor models. Both models provide evidence for the construct validity of the DES-III since both models retain ten distinct emotions as the underlying factors. The difference is that since the items of the multiple groups model are each an indicator of only a single factor, the analysis with the multiple groups model is easier to score and interpret.

Analysis of the factor loadings from the centroid multiple groups analysis or the confirmatory multiple factor solution of Study II did indicate that it may be desirable to modify some of the items which serve as indicators of the underlying emotions. Our analysis showed that items 3, 21, 24, and 25 of the DES-III could be rewritten to improve the multiple groups model. These improved alternatives are listed and analyzed in Study III.

Initially we suspected that the measurement model would begin to fail as the age of the respondent decreased due to lack of comprehension of the task or of specific items. However, this did not occur for the children as young as eight in our sample. The similarity between the item-factor correlations from Study II and Study III suggests that the psychological meaning of the items was essentially the same for our eight year olds as it was for our seventeen year olds. Thus we did not reach the lower age limit of the DES-III in the population under study. While the factorial composition of the scale is relatively invariant across the age range under study, the interrelationship among the factors could vary not only as a function of age, but also sex, and other variables such as the particular set of circumstances associated with a given administration of the test. The existence of stable patterns or profiles can be explored in detail through Q factor analysis. This is a logical "next step" in our work on the DES-III.

Given the cumulative evidence from the orthogonal and oblique exploratory analyses, and from the confirmatory factor analyses with both centroid multiple groups and full information maximum likelihood, the validity of the ten distinct emotions which are measured by the ten corresponding groups of DES-III items has been supported for three different data sets represented by college age adults, adolescents and children above the age of eight. The examination of the construct validity of these ten emotions must ultimately extend beyond the realm of paper and pencil measures and the present study encourages such future validation studies. The DES-III now can be used in a wide variety of research and applied applications which ultimately will determine its scientific and clinical utility.

REFERENCE NOTES

1. Izard, C. E., Dougherty, F. E., Bloxom, B. M., Kotsch, W. E. The differential emotions scale: A method of measuring the subjective experience of discrete emotions. Unpublished manuscript, Department of Psychology, Vanderbilt University, 1974.

2. Mosher, D. L., and Toedter, L. Jealousy in differential emotions theory. Unpublished manuscript, University of Connecticut, 1979.

3. Izard, C. E., The differential emotions scale for children (DES-III). Unpublished manuscript, Department of Psychology, University of Deleware, 1979.

4. Hunter, T. E., and Gerbing, D. W. Unidimensional measurement and confirmatory factor analysis. Occasionally paper No. 20, The Institute for Research on Teaching, College of Education, Michigan State University, E. Lansing, MI, 1979.

REFERENCES

Bartlett, E. S., Izard, C. E. A dimensional and discrete emotions investigation of the subjective experience of emotion. In C. E. Izard (Ed.), _Patterns of emotions: A new analysis of anxiety and depression_. New York: Academic Press, 1972.

Burt, C. Experimental test of general intelligence. _British Journal of Psychology_, 1909, 3, 94 - 177.

Burt, C. _The distribution and relations of educational abilities_. London: P. S. King and Son, 1917.

Ekman, P. _Darwin and facial expression._ New York: Academic Press, 1973.

Izard, C. E. _The face of emotion_. New York: Appleton-Century-Crofts, 1971.

Izard, C. E. _Patterns of emotions: A new analysis of anxiety and depression._ New York: Academic Press, 1972.

Izard, C. E. _Human Emotions_. New York: Plenum Press, 1977.

Harmon, H. H. _Modern factor analysis_ (Third edition). Chicago: The university of Chicago Press, 1976.

Holzinger, K. J. A simple method of factor analysis. _Psychometrika_, 1944, 9, 257 - 262.

Joreskog, K. G. Some contributions of maximum likelihood factor analysis. _Psychometrika_, 1967, 32, 443 - 482.

Joreskog, K. G. Structural analysis of covariance and correlation matrices. _Psychometrika_, 1978, 43, 443 - 477.

Joreskog, K. G. and Sorbom, D. _LISREL IV--A general computer program for estimation of a linear structural equation_ system by maximum likelihood methods. Chicago: National Educational Resources, 1978.

Mosher, D. L. Measurement of guilt in females by self-report inventories. Journal of Consulting and Clinical Psychology, 1968, 32, 690 - 695.

Mosher, D. L., and O'Grady, K. E. Sex guilt, trait anxiety, and females' subjective sexual arousal to erotica. Motivation and Emotion, 1979, 3, 235 - 249.

Nunnally, J. Psychometric theory. New York: McGraw-Hill Book Co., 1978.

Spearman, C. General intelligence, objectively determined and measured. American Journal of Psychology, 1904, 15, 201 - 293.

Spearman, C. The theory of two factors. Psychological Review, 1914, 21, 101 - 115.

Spielberger, C. P., Gorsuch, R. R., and Lushene, R. E. State-trait anxiety inventory test manuel for form X. Palo Alto: Consulting Psychologists Press, 1970.

Zuckerman, M. The development of an affect adjective check list for the measurement of anxiety. Journal of Consulting Psychology, 1960, 24, 457 - 462.