

## Reliability and Validation of Value Scale for Primary School Students: Exploratory and Confirmatory Factor Analyses

Duygu Anıl

Department of Educational Science, Faculty of Education, Hacettepe University, Ankara, Turkiye

**Abstract:** Conducting better planned and purpose-driven activities are of paramount importance in order to apply value education effectively. There is an increasing need to develop a reliable, valid and genuine scale to determine primary school students' values so that value education can be shaped and directed accordingly. This research aims at developing a scale to identify the values of primary school students. Developed by the researchers, "Primary School Students' Value Scale" (PSVS) has been analyzed through content and construct validity. Exploratory and confirmatory factor analyses have been utilized during construct validity examination. As for the reliability assessment, PSVS has been studied through internal consistency (Cronbach Alpha Coefficient) and split-half methods. Findings have revealed that the scale is both reliable and valid. Following reliability and validity studies, "Primary School Students' Value Scale" (PSVS) has been developed; it is a single factor scale containing 34 items; and the lowest and the highest scores that can be obtained from this scale are 34 and 170, respectively.

**Key words:** Value · scale development · confirmatory factor analysis · exploratory factor analysis

### INTRODUCTION

Human behaviours lacking tolerance, affection, respect and mercy as well as actions of violence and illegal activities have been on the increase all over the world in recent years; thus the standards of behaviour have been falling and consequently, some of the values have been disappearing. The fact that global warming has been continuously rising due to man's damage to the environment-which in turn is threatening humanity and the whole world-and that humanity suffers from wars, terrorism and inequality directs our attention to value education [1, 2].

One of the most important factors keeping individuals in a society in integration is shared, common values. Both Turkey and the world as a whole have suffered from values crisis recently; and thus shared values have been diminishing while values opposing the societal interests have been on the rise [3].

It may be asserted that individuals not being able to acquire such values as affection, respect, tolerance, honesty and self control are more likely to perform negative behaviours. Therefore, students' acquisition of certain values is considered more important than their academic success at school. Piaget *et al.*

emphasise that children develop a quick understanding towards such values as studiousness and honesty and that those values can be instilled at primary school age.

Shaver ve Strong (1976) define value as the scale for deciding about whether or not objects, thoughts, actions and situations are good, valuable, desirable or bad, undesirable, valueless [4]. In other words, value means the principles with regard to attaching meaning to the world. Welton and Mallan [5] suggest that values are the thoughts serving to identify the measures of behaviour, goodness, beauty, efficiency and worth.

On investigating a number of definitions and explanations concerning values, it is found that certain properties related to values are available. The properties are listed as what follows. Values [4, 6]:

- Are an internal force affecting individual behaviours,
- Facilitate decision making when one is in dilemmas,
- Affect individual perception and interpretation of reality,
- Assure controlling behaviours,
- Are chosen by individuals,
- May depend on experiences and individuals concerned,

- Are representative of individuals' or societies' desires and aims,
- Are loaded with emotions, strengthen thoughts and understanding and lead behaviours,
- Include physical, intellectual, moral, aesthetic, socio-economic, political and religious dimensions of personality.

Moral education forms a considerable part of value education [6]. These concepts, which are previously known as value education or moral education, are also called character education in relevant literature. The three concepts appearing in field literature under different labels include educating students in schools so as to give them abilities to implement moral thoughts and behaviours in democratic ways and to raise them as morally mature individuals [5].

Although there is controversy over these concepts, the debating parties agree that schools have a vital role in young people's moral and value education. Therefore, it should be stressed that students must acquire values through primary school subjects, especially through such subjects as Social Studies [7].

Today it is considered compulsory to conduct value education at schools through programmes so as to assure that such education is not coincidental and to build a society composed of individuals who have acquired various values. It could be said that people are unlikely to have a firm character when they have very few or no values at all; which will certainly have a negative influence over all the societies and the world. Planned value education conducted at schools may assist students individually or in groups to develop positive behaviours [1].

Although value education dates back to the formation of history, it may be discovered-on examining the primary school programmes that have been implemented since 1924-that value education is not considered directly and that attempts have been made only through covert programmes to offer students values. However, in the syllabus of Life Sciences and Social Studies, which are the parts of the primary school programmes enforced in 2005-2006 academic year, values are considered within the programmes and explanations as well as examples are provided.

As individuals understand life, they consider their previous values and select and adopt their own values among the multiple value perspectives. Therefore, value classifications can be made on the basis of the social, political, economic and moral structure of a nation or a family [7].

The criteria concerning the properties of a decent and mature individual were developed by Programme Development and Inspection Association in the United States in 1988. The criteria are as what follows [8]:

- Respect for people,
- Considering others' peacefulness,
- Having social responsibilities,
- Being reliable,
- Being tolerant.

A close examination of the classification of values shows that the classifications include such values as mercy, courtesy, tolerance, honesty, studiousness, responsibility, courage, consistency, self control and honour-on which criterion development work was also based [7-10].

Although values are a significant field to be developed and matured via education, the field has not received sufficient importance and attention. In a survey conducted on families, Johnson ve Immerwahr (1995) found that 71% of the families attached more importance to schools' teaching of values than of academic subjects [11].

In the current research, which is directed to identifying the effects of value education over the third and fourth graders of primary schools, it was found that through value education students behaved more honestly in dealing with conflicts and in assessing the situations, that they displayed fewer careless behaviours and that they valued others' opinions more [12].

In order to take a glance into the future of a society of individuals and to make attempts at improving the present situation, it is considered important to measure the extent to which students possess certain values. The development of an authentic, valid and reliable tool of measurement to identify primary school students' values is required for the purposes of performing an effective value education in primary schools and of orienting school-parents cooperation in terms of value education. Thus, values possessed by primary school students can be revealed and efforts to provide students with values can be made in a more planned, effective way and purposefully.

This research aims at developing a Likert-type scale measurement tool to identify the values that primary school students possess. In Likert type scales, estimating a person's level of attitude towards an issue of attitude based on the classical test theory is performed by summing up all the responses the person gives to the

items on the attitudes scale. Thus, it can be said that a student achieving low scores in Primary School Students' Value Scale (PSVS), Which was developed in the form of Likert type scales, has insufficient values while a student who achieves high scores in the same scale has sufficient values. The accuracy of judgements about individuals based on the scores gained from Likert type attitudes scale is closely related with the psychometric properties of the scale. Hence, assessments made with measurement tools having no validity or reliability at the desired level are not likely to be accurate. Consequently, estimating the psychometric properties of value scale designed for primary school students bears paramount importance for the accuracy of assessment of students.

## MATERIALS AND METHODS

**Research group:** Owing to the fact that this study was basic research and that it did not aim to make generalisations based on the information obtained, universe and sample groups were not identified. The research was conducted on a total of 242 8<sup>th</sup> graders in the spring semester of 2006-2007 academic year.

**Scale development:** In the first phase of Primary School Students' Value Scale (PSVS) development work, relevant literature on the concept of value, value classifications, criteria concerning the values in the classifications was reviewed. The values forming the basis of the scale in relevant literature are listed as:

- Mercy,
- Courtesy,
- Tolerance,
- Honesty,
- Studiousness,
- Responsibility,
- Courage
- Consistency
- Self-control,
- Honour

In an attempt at measuring the values that were based on the relevant field literature, firstly 62 items half of which representing negative and half of which representing positive attitudes were prepared by the researcher. Then, 50 of the items were selected following experts opinion and a trial form was formed out of these items. So as to assure fitness with students' level and clarity, the trial form was applied to 20 primary school

students; in consequence the 20 students and 3 teachers were interviewed. In line with the feedback received, the scale was conducted with a trial group of 242 respondents.

It took students approximately 35 minutes to complete the scale. Following the application, the students' measurement tools were checked and four students were found not to answer the scale appropriately, hence the scales completed by them were not considered for the analyses.

**Data analysis:** Following the application of "Primary School Students' Value Scale", which was based on classical test theory and was designed in the form of Likert type scale, the responses obtained from the students were scored in line with the following guidelines and were used for analyses on the computer (SPSS 11.5).

The responses given to the positive items of the Likert type scale and the scores were: A: Totally Agree: 5, B: Agree: 4, C: Indecisive: 3, D: Disagree: 2, E: Totally Disagree: 1; and the responses to the negative items and the scores were: A: Totally Agree: 1, B: Agree: 2, C: Indecisive: 3, D: Disagree: 2, E: Totally Disagree: 1. Excel 2003 and SPSS 11.5 programmes were employed to predict the scale features of Likert form used in PSVS based on classical test theory following the scoring step.

The psychometric properties of the items composing the PSVS based on the classical test theory were first identified in the analysis of the data. In primary school value scale, as in Likert type scales that are used for measuring attitudes, Psychometric properties were considered in two groups as the psychometric properties of the items (item discrimination and the average attitude levels measured by the items) and as the psychometric properties of the scale (reliability and validity). The methods and techniques mentioned below were employed in the classical test theory-based prediction of the psychometric properties of PSVS, which had been designed in the form of Likert type scale.

**Item discrimination:** The power of items comprising a tool of measurement to distinguish between the high level of features and low level of features is called item discrimination. Item discrimination is also regarded as the measure of validity of that item. Two separate methods of analysis suggested by Likert were used in this research in order to determine the distinguishing power of the items available in Likert type scales: 1) correlation-based analysis, 2) "internal consistency-based (t-test) analysis [13].

**Correlation-based analysis (r<sub>ij</sub>):** Item discrimination is represented with the correlation coefficient between item scores obtained for each value sentence and scale scores. This correlation coefficient is interpreted both as the item discrimination index and as the item validity coefficient. The stated coefficient receives values between -1 and +1. Of the obtained item-scale correlation coefficients, those values with a minus sign, those whose coefficient value is zero or around zero are the items which fail to measure the issue of value concerned. On the other hand, those with a positive sign or high value are the items which are capable of measuring the values. Therefore, it is desirable for the item-scale correlation coefficient of the items to have positive signs or higher values because this is important for the scale to be one dimensional and to achieve high validity and reliability.

**Internal consistency criterion (t-test)-based analysis:** In predicting the distinguishing power of the items of the scale on the basis of internal consistency (t-test), each individual's scale scores are determined by summing up all the responses they gave to the items in the scale. Having individuals' scale scores, the scores are ordered from the highest to the lowest. 27% of the respondents at the top of the distribution of scale scores are chosen as the top group while the 27% at the bottom are selected as the bottom group. Those individuals in the top group are said to represent the individuals having the positive attitudes to be measured by the scale whereas those in the bottom group are said to represent the negative attitudes. After the top group and bottom group individuals' item score averages for each item are found, whether or not the differences between averages are statistically significant is tested through t statistics which is used in the comparison of independent groups' arithmetic averages. The items with significant t values are regarded as the items capable of distinguishing between those who have positive values and those who don't. That is to say, such items are the ones distinctive at the desired level. However, the items whose t values are not significant are the items that cannot distinguish appropriately between those who have positive values and who don't—that is, items which are not distinctive.

**The average attitude levels measured by the items:** While individuals are responding to the items of Likert type value scale, they also attach value scores to each item. The arithmetic average of the value scores attributed to each item by individuals is called "the average attitude level measured by the item". In consequence, if the

average value calculated for each item is high, it means that individuals have positive values with regard to the content of the item; and the reverse result means individuals have negative values.

Having determined the discrimination of the items, Kaiser-Meyer-Olkin (KMO) coefficient and Bartlett Sphericity test were utilized in checking the appropriateness of the data for factor analysis [14].

To determine the structure of the factors, exploratory and confirmatory factor analyses were done separately and thus construct validity for PSVS was checked.

Cronbach Alpha reliability coefficient as well as Spearman-Brown reliability coefficient value, a method of split half test and Guttman internal consistency coefficient were used for the identification of PSVS reliability. The level of significance was regarded as .05 in data analysis

**Findings and interpretations:** Determining the psychometric properties of PSVS in the classical test theory of Likert type scales was considered in two categories in this research. Findings concerning the psychometric properties of the items and of the scale were presented under two sub-headings in efforts for determining the psychometric properties of PSVS:

1. the psychometric properties of the items
  - 1.1. Item discrimination
  - 1.2. The Average Attitude Levels Measured by The Items
2. the psychometric properties of the scale
  - 2.1. Reliability
  - 2.2. Validity

#### **Findings and interpretations concerning the psychometric properties of PSVS items**

**1.1. Item discrimination:** The power of items comprising a tool of measurement to distinguish between the high level of features and low level of features is called item discrimination. Item discrimination is also regarded as the measure of validity of that item. Two separate methods of analysis suggested by Likert were used in this research in order to determine the distinguishing power of the items available in Likert type scales:

1. Correlation-based analysis,
2. "Internal consistency-based" (t-test) analysis [13].

**1. Correlation-based analysis (r<sub>ij</sub>):** Item discrimination is represented with the correlation coefficient between item scores obtained for each attitude sentence and

Table 1: Items Analysis Results Concerning Primary School Students' Value Scale

| Items | Mean | Standard deviation | r      | t      |
|-------|------|--------------------|--------|--------|
| 1     | 4.62 | 0.73               | 0.16*  | 2.09*  |
| 2     | 4.15 | 0.81               | 0.48*  | 7.30*  |
| 3     | 4.06 | 1.23               | 0.24*  | 4.48*  |
| 4     | 3.52 | 1.27               | 0.24*  | 4.46*  |
| 5     | 4.38 | 1.02               | 0.35*  | 3.80*  |
| 6     | 4.02 | 1.00               | 0.28*  | 5.08*  |
| 7     | 4.11 | 0.96               | 0.54*  | 8.55*  |
| 8     | 4.29 | 0.90               | 0.52*  | 6.58*  |
| 9     | 3.95 | 1.05               | 0.36*  | 4.83*  |
| 10    | 4.76 | 0.73               | 0.41*  | 3.35*  |
| 11    | 3.35 | 1.31               | 0.22*  | 3.30*  |
| 12    | 3.90 | 1.00               | 0.53*  | 8.61*  |
| 13    | 3.45 | 1.29               | 0.32*  | 5.11*  |
| 14    | 4.44 | 0.92               | 0.46*  | 5.33*  |
| 15    | 4.07 | 1.37               | 0.42*  | 7.30*  |
| 16    | 3.64 | 1.31               | 0.42*  | 7.75*  |
| 17    | 4.31 | 0.97               | 0.50*  | 7.56*  |
| 18    | 3.26 | 1.47               | 0.40*  | 7.54*  |
| 19    | 4.66 | 0.76               | 0.42*  | 4.99*  |
| 20    | 4.06 | 1.29               | 0.36*  | 6.16*  |
| 21    | 3.71 | 1.20               | 0.36*  | 6.08*  |
| 22    | 4.50 | 0.94               | 0.54*  | 5.97*  |
| 23    | 4.44 | 0.84               | 0.51*  | 7.33*  |
| 24    | 4.20 | 0.99               | 0.55*  | 8.33*  |
| 25    | 4.06 | 1.36               | 0.37*  | 5.75*  |
| 26    | 3.90 | 1.20               | 0.46*  | 7.71*  |
| 27    | 3.64 | 1.46               | 0.34*  | 6.80*  |
| 28    | 3.99 | 1.08               | 0.38*  | 6.05*  |
| 29    | 3.41 | 1.35               | 0.41*  | 6.56*  |
| 30    | 4.24 | 3.44               | 0.13*  | 1.92   |
| 31    | 4.33 | 0.97               | 0.53*  | 7.37*  |
| 32    | 3.69 | 1.24               | 0.38*  | 6.26*  |
| 33    | 3.43 | 1.29               | 0.41*  | 6.59*  |
| 34    | 4.26 | 1.16               | 0.56*  | 8.31*  |
| 35    | 3.79 | 1.30               | 0.33*  | 4.90*  |
| 36    | 4.43 | 0.93               | 0.60*  | 8.48*  |
| 37    | 3.56 | 1.25               | 0.43*  | 7.73*  |
| 38    | 4.29 | 1.09               | 0.25*  | 2.72*  |
| 39    | 4.48 | 0.99               | 0.61*  | 7.81*  |
| 40    | 4.30 | 1.16               | 0.31*  | 5.26*  |
| 41    | 4.40 | 1.01               | 0.43*  | 5.81*  |
| 42    | 2.23 | 1.17               | -0.26* | -3.15* |
| 43    | 3.95 | 1.21               | 0.48*  | 8.07*  |
| 44    | 4.07 | 1.17               | 0.35*  | 4.90*  |
| 45    | 4.08 | 1.10               | 0.26*  | 4.41*  |
| 46    | 4.46 | 0.90               | 0.52*  | 6.27*  |
| 47    | 3.78 | 1.38               | 0.35*  | 7.33*  |
| 48    | 4.32 | 1.05               | 0.42*  | 5.10*  |
| 49    | 4.03 | 1.11               | 0.42*  | 6.43*  |
| 50    | 3.87 | 1.34               | 0.34*  | 6.17*  |

\*p<0.05

scale scores (item-scale correlation) in correlation-based analysis. It is considered important for item-scale correlation coefficient of the items to have positive signs and high values to assure that the scale is one dimensional and also that validity and reliability are high. Items generating significantly high level of correlation (0.05) between scale scores and item scores in correlation-based analyses were included in the final form. Items generating a correlation value higher than 0.30 holding between scale scores and item scores were regarded as the ones that made the most contributions to the measurement and hence the items generating lower correlations than this value were not included in the final form. The correlation values between scale scores and item scores are shown in Table 1.

According to Table 1, it is clear that all of the 50 items coming from PSVS trial form generate significant levels of correlation values. The table shows that item 42 generates negative correlation values with scale scores while items 1, 3, 4, 6, 30, 38 and 45 have a lower correlation value ( $r < 0.30$ ) with scale scores.

**2. Internal consistency criterion (t-test)-based analysis:**

In determining the discrimination of items, a method of items analysis based on the difference between top and bottom group averages was employed in this analysis. Top and bottom groups of 27% were formed through this method and the significance of the difference between the two groups was checked with t-test. The items with significant t values are regarded as the items capable of distinguishing between those who have positive values and those who don't. On the other hand, the items whose t values are not significant are the items that cannot distinguish appropriately between those who have positive values and who don't-i.e., items which are not distinctive enough. The "t" values obtained from the scale items are shown in Table 1. The items with a significant difference ( $p < 0.5$ ) between the top and bottom group averages were ordered from the highest to the lowest t value, then items were selected for inclusion in the final form.

A close examination of Table 1 showed that 49 out of the top and bottom groups of 27% from 50 items were significant at the level of 0.05; however, item 30 was not found to be significant. Apart from that, a study of Table 1 demonstrated that items generating low correlation value ( $r < 0.30$ ) in correlation-based analyses also received low t values in items analysis results which were based on the difference between top and bottom group averages.

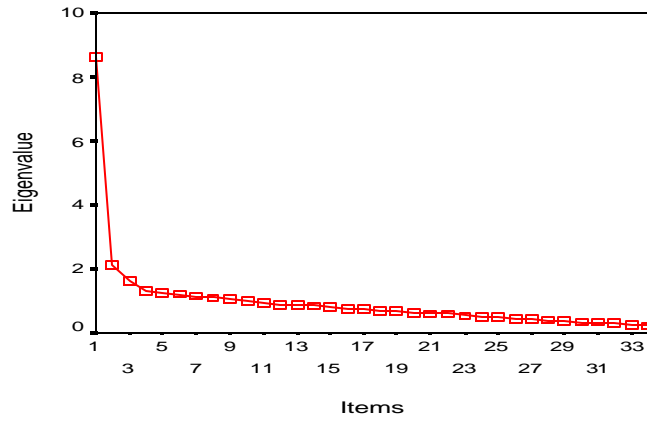
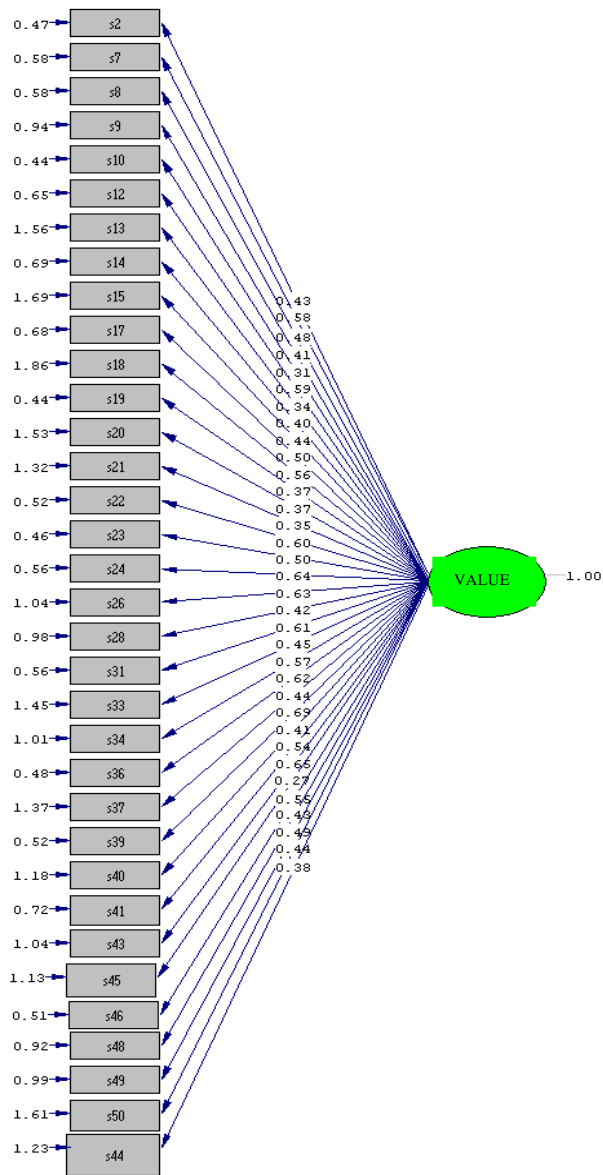


Fig. 1: Eigen value graph concerning the scale items



Chi-square=1033.10, df=527, P-value=0.00000, RMSEA=0.063

Fig. 2: Confirmatory factor analysis diagramme concerning the items in PSVS

### **1.2. The average attitude levels measured by the items:**

The arithmetic average of the value scores attributed to each item by individuals is called “the average attitude level measured by the item”. Accordingly, if the average value calculated for an item is high, it means that individuals have positive values with regard to the content of the item; if it is low, they have negative values in terms of the content of the item. Considering the average values and standard deviation values in Table 1, it was found that the arithmetic average values of the items were in the 2.23 (item 42) and 4.76 (item 10) interval ; whereas the standard deviation values of the items received values between 0.73 (item 1) and 3.44 (item 30). Besides, with a scale score average of 4.016, the primary school students were found to have a high level of values. And the average for scale scores standard deviation values was found as 1.16.

**Findings and interpretations concerning the psychometric properties of PSVS:** This part of the research provides explanations on validity and reliability of PSVS.

**Findings and interpretations concerning the validity of the scale:** Within the framework of validity studies of PSVS, efforts were made towards content and construct validity. For this purpose, two separate approaches of factor analysis were used and both exploratory and confirmatory factor analyses were checked in identifying the construct validity of PSVS.

### **1. Studies directed to determining content validity:**

Content validity is related with the extent to which a tool of measurement contains the behaviours it attempts at measuring. Although there are efforts to determine content validity quantitatively (Roozeboom 1966; Crocker ve Algina 1986), expert opinions are often consulted [15].

For the purposes of PSVS validity studies, items based on relevant field literature were prepared and experts working in the field of value education and scale development as well as teachers teaching Social Studies to the 8<sup>th</sup> graders were consulted for their views. Besides, as mentioned above, the trial scale was applied to 20 8<sup>th</sup> graders. Thus, receiving the target group’s opinions on the items prepared, it was checked whether there were any items probable to be misunderstood. Hence, firstly the content validity of the scale was assured in the study of PSVS validity.

**2. Studies directed to determining the construct validity of PSVS:** As for the construct validity of PSVS, both exploratory and confirmatory factor analyses were performed.

**2.1. Exploratory factor analysis:** Factor analysis was utilized in determining the construct validity of “primary School Students’ Value Scale”. Since factor analysis is not compatible with all data structures, the fitness of data was checked prior to the application of factor analysis through such procedures as Kaiser-Meyer-Olkin (KMO) coefficient and Barlett Sphericity test. In consequence, KMO coefficient value was identified as 0.84 and Barlett test was found to be significant. As stated by Büyüköztürk (2006), the case where KMO is higher than 0.60 and Barlett test is significant shows that data are appropriate for factor analysis: having seen this, the step was taken towards factor analysis application.

Because establishing whether the measured scale was one dimensional or not was important in determining the construct validity, factor analysis was used and it was checked if the scale items clustered on the same dimension. Basic components analysis factorizing technique, one of the factor analysis techniques, was used in application. In consequence of conducting factor analysis on the data obtained, factor loads were observed to cluster in three groups. The genuine values of the determined factors were 19.738 for the first factor, 5.984 for the second factor and 4.378 for the third factor. The fact that the genuine value for the first factor was bigger than that of the second factor and there were no big differences between the values for the other factors can be regarded as the indicator of one dimensionality of the measured property [16]. Of the 30.1% variance that the obtained three factors accounted for, 12.98% was explained by the first, 4.61% was explained by the second and 4.534% was explained by the third factor. In addition, on examining the factor analysis results, it was found that 34 items of PSVS out of 50 items were within the first factor load. Moreover, an examination of the graph about genuine values revealed that after the first factor load there was a rapid fall; which meant that the scale could have a general factor. The graph that was drawn following the results of factor analysis which was done to determine whether the measured property was one dimensional or not is given in Figure 1.

The 34 items (except for items 1, 3, 4, 5, 6, 11, 25, 27, 29, 30, 32, 33, 35, 38, 42 and 47) of PSVS that was designed in consequence of analyses performed were

Table 2: Values concerning the goodness-of-fit indices for PSVS

| Chi-square | DF  | P-Value | CFI  | NFI  | AGFI | IFI  | SRMR  | RMSEA | 90% C.I.RMSEA |
|------------|-----|---------|------|------|------|------|-------|-------|---------------|
| 1033.10    | 527 | P <.05  | 0.94 | 0.87 | 0.80 | 0.94 | 0.081 | 0.063 | .057; 0.069   |

found to be within the first factor load. Besides, those items were observed to generate high item scale scores in items analysis studies; and PSVS that was made up of those items was seen to be one dimensional and to achieve construct validity with its high validity items.

**2.2. Confirmatory factor analysis:** After exploratory factor analysis, confirmatory factor analysis was also carried over for PSVS and the results were presented in diagram in Figure 2. While doing the study of confirmatory factor analysis, the results of exploratory factor analysis were considered, and the 34 items of the first factor in PSVS were taken into consideration.

The most commonly used statistics for confirmatory factor analysis and for model-data fitness are chi-square ( $\chi^2$ ),  $\chi^2/sd$ , RMSEA, RMR, GFI and AGFI. Jöreskog and Sorbom (1993) & Marsh and Hocevar (1988) point out that there is a perfect model-data fitness when the calculated  $\chi^2/df$  proportion is smaller than 5, GFI and AGFI values are bigger than 0.90 and RMR and RMSEA values are smaller than 0.05. However Anderson and Gerbing (1984), Cole (1987) & M. Balla and McDonald (1988) believe that it is an acceptable minimum limit for model-data fitness when GFI is bigger than 0.85, AGFI is bigger than 0.80, RMR and RMSEA are smaller than 0.10 [17]. The statistics regarding the fitness of confirmatory factor analysis results are shown in Table 2.

The model that was constructed for the theoretical structure of PSVS can be seen in Figure 2. The results of fit indices that were obtained from the confirmatory factor analysis applied to test the fitness of the model showed that the model-data fitness was high. Chi-square, one of fitness statistics, was found to be significant. As McDonalds and M.Ho (2002), S.Engel, Moosbrugger and Müller (2003) & Thompson (2000) point out chi-square value is almost always significant especially with very big samples since the chi-square value is quite sensitive to sample size. Therefore, suggestion is that chi-square value should be divided by the degree of freedom and the resultant coefficient should be used. If the mentioned proportion is 2 or below, this is said to be a good model; if the proportion is 5 or a value below that, this model is said to have acceptable goodness of fit [17]. The  $\chi^2/df$  proportion concerning PSVS is 1.96, which proves this is a good model. AGFI (0.80), one of fitness statistics, is good for model-data fitness. Besides, values such as NFI (0.87), SRMR (0.81) and RMSEA (0.063) demonstrate that values are acceptable in terms of model-data fitness. On the other hand, CFI (0.94) and IFI (0.94) values that have

been found through analyses show that data fitness is high. Taking all the values for model-data fitness into consideration, it is clear that data fitness of the model established is good. Thus, the items prepared for PSVS are capable of measuring the value covert variable; in other words, the scale was shown to achieve construct validity.

**3. Validity based on a criterion:** Validity which is based on a criterion involves the comparison of a group of individuals' scores that they receive from a measurement tool with criterion scores, classifications, or other ability and skill measures [18]. No validity studies based on a criterion were conducted owing to the fact that no value scales targeting primary school students were found in our literature review against which PSVS scores could be checked or tested.

**Findings and interpretations concerning scale reliability:** Internal consistency as well as split half methods were also included in PSVS reliability studies.

**1. Studies directed to determining internal consistency (Cronbach Alpha Coefficient):** Reliability is the power of a tool of measurement to yield sensitive, consistent and decisive measurement outcomes each time. Measurement results are reliable to the extent that they are free from random errors. The measurement results obtained through unreliable scales do not lead to consistent results both in themselves and from one measurement to the next. Thus, inconsistency of measurement results have a negative influence over the accuracy of decisions to be made about individuals on the basis of the measurement results. Therefore, high reliability of scale scores is desirable.

In order to predict the reliability of PSVS on the basis of classical test theory, Cronbach Alpha reliability coefficient was obtained by making use of the item variances of the items in the scale in addition to scale variance. Cronbach Alpha reliability coefficient regarding the scale was found as 0.90. That reliability coefficient for the scale was a value of 0.90 can be regarded as the indicator of consistency holding between scale items. In other words, the items composing the scale were designed for the purpose of measuring the same goal. Considering the size of coefficient obtained and the fact that Cronbach Alpha coefficient is a measure of internal consistency, it can be suggested that the scale performs measures consistent in themselves and that attains reliability in terms of internal consistency. That the items comprising



PSVS have internal consistency indicates that the scale is a reliable tool of measurement. Accordingly, it can be said that decisions to be made on the basis of measurement results obtained through PSVS are accurate and reliable.

**2. Studies directed to determining reliability through split half method:** On calculating the reliability for two halves, the reliability coefficient for the first half was found as 0.81, whereas it was found as 0.83 for the second half. Spearman Brown reliability coefficient regarding the overall scale, on the other hand, was found to be 0.86. Additionally, Guttman internal consistency coefficient-one of the split half methods-was estimated as 0.85. That reliability values calculated through split half methods, just like those calculated through Cronbach Alpha reliability coefficient, also received high values proved that PSVS was a reliable scale.

## CONCLUSIONS

Various activities of learning and teaching have been held for the purposes of instilling values on students through covert programmes and school subjects, particularly Life Sciences and Social Studies. To check whether students have acquired the offered values or not, the cases where they acquire values must be assessed. Although such methods and approaches as observations, interviews, value analyses, moral dilemmas and discussions provide critical information concerning the status of values absorbed by students, there is still need to develop a tool of measurement to measure primary school students' values.

This paper aims at developing a scale directed towards identifying primary school students' values. In determining the psychometric properties of PSVS, which had been designed in the form of Likert type scale, findings concerning the psychometric properties of both the items and the scale were utilized.

Within the framework of the study, item discrimination indices as well as item criterion average value levels were calculated in determining the psychometric properties of the items in PSVS, which had been developed by the researcher and applied to 242 primary school 8<sup>th</sup> graders. So as to determine item discrimination indices, both correlation-based analysis and internal consistency criterion (t-test)-based analyses were employed. The correlation-based analyses results concerning the 50-item PSVS demonstrated that item 42 produced negative correlations with scale scores whereas items 1, 3, 4, 6, 11, 30, 38 and 45 gave low correlation value ( $r < 0.30$ ) with scale scores. 49 t-values out of 50 that were calculated for the top and bottom groups of 27% were

found to be significant at the level of 0.05; however, item 30 was not found significant. Items producing low correlation values ( $r < 0.30$ ) in correlation-based item analyses were observed to receive low t-values also in consequence of analyses based on top-bottom groups difference averages.

Scale reliability and validity studies were conducted in determining the psychometric properties of PSVS. Scale validity studies included such issues as content validity, construct validity and criterion validity. Having received expert opinion concerning the content validity of the scale, it was applied to 20 primary school students attending the 8<sup>th</sup> grade as the trial scale and thus interviews were made as to understandability of the scale items. Factor analysis was utilized in determining the construct validity of "Primary School Students' Value Scale". Moreover, both exploratory and confirmatory factor analysis studies were performed in determining the construct validity of PSVS. In consequence of exploratory factor analysis, 34 items of PSVS, excluding items (1, 3, 4, 5, 6, 11, 25, 27, 29, 30, 32, 33, 35, 38, 42, 47.), were seen to be in the first factor load and those items were also found to produce high item scale scores in item analysis studies. Following exploratory factor analysis, confirmatory factor analysis was also performed; consequently, results were demonstrated. And confirmatory factor analysis led to such results as: comparative fitness index CFI=0.94, normed fitness index NFI=0.87, corrected goodness of fit index AGFI= 0.77; which meant that fitness indices concerning scale items were within acceptable limits even though they did not attain perfect fitness values. Besides, the fact that fitness index yielding the probability of SRMR value independently of the sample received a value as high as 0.94 can be interpreted as having good model-data fitness. Another indicator of the strength of model-data fitness was that SRMR value producing model fitness regarding standardised errors of the model was 0.08. On the other hand, although RMSEA value was bigger than 0.06, RMSEA value contained 0.06 value of 90% probability security period; which meant high degree of model-data fitness [19]. Taking the overall values regarding model-data fitness into consideration, it was concluded that the established model-data fitness was good; and hence it was able to measure the value covert variable of the items-that is, the scale achieved construct validity.

PSVS reliability studies also included such procedures as internal consistency and split half methods. Here Cronbach Alpha reliability coefficient was calculated as 0.90. As pointed out by Tezba<sup>o</sup>aran (1997), the scales composed of items that are in relation to each other have high Alpha coefficients [13]. The highness of the scale

Alpha coefficient can be interpreted as consistency of scale items with each other and as testing the same properties. Similarly, this scale can be said to have high reliability due to the fact that Cronbach Alpha coefficient regarding the scale was 0.90. Apart from that, Spearman Brown reliability value was found to be 0.87.

In conclusion, this research was conducted so as to develop a scale aiming at identifying primary school students' values. The findings obtained revealed that the scale was both valid and reliable. Consequently, this study puts forward a one-dimensional, 34-item, 5-graded tool of measurement called "Primary School Students' Value Scale" (PSVS) which achieves a high level of validity and reliability and from which scores between 34 and 170 can be received.

As Arıcaç and Ilgaz (2007) point out, in order for a scale to be of high quality and beneficial, it should be used in various research attempts and for different samples on different occasions and validity and reliability investigations should be conducted; which will contribute not only to the scale but also to the field considerably [20]. Therefore, validity and reliability studies concerning the "Primary School Students' Value Scale" can be conducted repeatedly with separate samples. It is believed that the scale is useful in terms of assessing the values that 8<sup>th</sup> graders hold and that there is still need to test students of other levels at primary school to determine their values.

## REFERENCES

1. Ediger, M., 1998. Trends and issues in teaching elementary school social studies. *College Student Journal*, 32(3): 347-363.
2. Scapp, R., 2003. Teaching values critical perspectives on education, politics and culture. New York: Taylor&Francis Group.
3. Erden, M. (nd) Sosyal bilgiler öğretimi. Ankara: Alkım Kitapçılık Yayıncılık.
4. Kaltsounis, T., 1987. Teaching social studies in the elementary school. New Jersey: Prentice Hall.
5. Welton, D.A. and J.T. Mallan, 1999. Children and their world: Strategies for teaching social studies (6th ed.).-Boston: Houghton Mifflin Co.,
6. Quisumbing, L.R. and J. de Leo, 2005. Learning to do: Values for learning and working together in a globalized world. Report prepared by The Asia Pasific Network for International Education and Values Educaation in Partnership with UNESCO-UNVEOK International Centre for Technical and Vocational Education and Training Sourcebook, Bonn, Germany.
7. Allen, M.G. and R.L. Stevens, 1998. Middle grades social studies: Teaching and learning for active and responsible citizenship (6th ed.).-Boston: Allyn and Bacon.
8. Martorella, P.H., 2001. Teaching social studies in middle and secondary schools. New Jersey: Merrill Prentice Hall.
9. MEB. 2005. İlköğretim sosyal bilgiler dersi öğretim programı ve klavuzu (4. ve 5. sınıflar). Ankara: Devlet Kitapları Müdürlüğü.
10. Ohio State Department of Education, 1990. Character education in Ohio: Sample strategies, (Guides-Non Classroom Use 055-Test Evaluation Instruments). (ERIC Document Reproduction Service No. ED350 196).
11. Doğanay, A., 2006. Değerler eğitimi. (edit. C.Öztürk) Hayat bilgisi ve sosyal bilgiler öğretimi (ss.255-286). Ankara: PegemA Yayıncılık
12. Jones, C. and D. ve Gower 1994. Effects of values instruction on third and fourth grade students. (Paper presented at the Manual of the Mid-South Meeting Research Association). Nashville, TN, November 9-11, 1994. (ERIC Document Reproduction Service No. ED389 997).
13. Tezbağaran, 1997. Likert tipi ölçek geliştirme kılavuzu (İkinci baskı). Ankara: Psikologlar Derneği Yayınları.
14. Büyüköztürk, Ş., 2006. Sosyal bilimler için veri analizi el kitabı. Ankara: Pegem Yayınları.
15. Baykul, Y., 2000. Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması. Ankara: Ösym Yayınları.
16. Lord, F.M., 1980. Applications of item response theory to practical testing problems. Hillsdale, N.J:Lawrence erbaum.
17. Şimşek, Ö.F., 2007. Yapısal eşitlik modellemesine giriş: Temel ilkeler ve lisreal uygulamaları. Ankara: Ekinoks Yayınları
18. Aiken, L.R., 1985. Psychological testing and assesment. America.
19. Hu, L. and P.M. Bentler, 1999. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6: 1-55.
20. Arıcaç, T ve Ilgaz, G., 2007. Açıklayıcı ve doğrulayıcı faktör analizi ile biyoloji dersi tutum ölçeğinin yapı geçerliliğinin incelenmesi. *Eğitim Araştırmaları*, 7(28): 1-8.