VOL 6(Special Issue)

e-ISSN 2602-3733

Research on Education and Psychology (REP)

Research on Education and Psychology (REP)
Volume: 6 • Number: Special Issue • 2022 • 1-91

Editorial Board

URL: http://dergipark.org.tr/rep
Email: editor.journalrep@gmail.com

Table of Contents

RESEARCH ON EDUCATION AND PSYCHOLOGY (REP)

Received: January 6, 2022
Accepted: January 26, 2022
http://dergipark.org.tr/rep
e-ISSN: 2602-3733

April 2022 ◆ 6(Special Issue) ◆ 1-26
Research Article
Doi: 10.54535/rep.1054235

Development of Mindset Theory Scale (Growth and Fixed Mindset): A Validity and Reliability Study

Ercan Yılmaz[1]

Necmettin Erbakan University

Abstract

This study aims to develop a valid and reliable measurement instrument to measure the quality of Mindset Theories of students aged 14- 22. A systematic approach was followed to develop the measurement tool. 1145 students participated in the study (48% were female and 52% were male). Exploratory and confirmatory factor analysis were applied to determine the scale's construct validity. As a result of exploratory factor analysis, the scale was determined to consist of 19 items and four sub-dimensions. In line with the literature, these dimensions are called Procrastination, Immutability of Belief, Belief in Improvement and Effort. The four-factor structure of the scale was confirmed by confirmatory factor analysis. In addition, it was found that the differences between the averages of the upper and lower groups that make up 27% of the scale items are significant. When the results of the reliability analyses were examined from the perspective of the factors of the mindset theory scale, the values of 0.724 for the Procrastination sub-dimension and 0.805 for the Immutability of Belief sub-dimension of the Fixed Mindset dimension were found. It was found 0.701 for the sub dimension of Effort and 0.771 for the sub dimension of Belief of Mindset Theory Scale's Growth Mindset dimension. The internal consistency coefficient was found 0.723 for the Fixed Mindset dimension and 0.714 for the Growth Mindset dimension of the Mindset Theory Scale. These results shows that the Mindset Theory Scale measures students' mindset theories in a valid and reliable way.

Keywords

• Fixed Mindset • Growth Mindset • Mindset • Mindset Theory

[1] **Correspondence to:** Ahmet Keleşoğlu Faculty of Education, Nrcmettin Erbakan University, Konya, Turkey. E-mail: ercanyilmaz70@gmail.com **ORCID:** 0000-0003-4702-1688

Citation: Yılmaz, E. (2021). Development of Mindset Theory Scale (Growth and Fixed Mindset): A validity and reliability study. *Research on Education and Psychology (REP), 6*(Special Issue), 1-26.

One of the most basic skills that education should provide to students is the awareness that learning is a task that needs to be continued all the time. The individual's ability to carry out learning actions continuously is only possible with the belief that he or she can change and improve himself or herself. In addition, the individuals' self-improvement is associated with the desire to learn and belief in their abilities. In line with these thoughts, non-cognitive skills have started to be considered as an important problem research area in educational studies. Non-cognitive skills generally refer to skills such as persistence, courage, endurance, resilience, self-sufficiency, mindset, effort, motivation, collaboration, and work habits (Farrington, et. al., 2012 as cited in Yilmaz, 2019). Many of these skills are also thought to be closely related to students ' academic achievement (Duckworth, 2009; Garcia, 2014). Hence, it is assumed that interest in studies to investigate the effects of individuals' non-cognitive skills on variables such as their cognitive learning will gradually increase (Fitzgerald &Laurian-Fitzgerald, 2016).

One of the non-cognitive skills is the Mindset Theory of individuals. Mindset Theory is the beliefs that individuals develop about their fundamental qualities (Dweck, 2016). Mindset Theory is associated with the mindset of individuals. Mindset includes thoughts, beliefs, emotions, motives and intentions (Ormrod, 2017). The mindset has been defined in different dimensions. First one of these is that mindset is cognitive activities performed to implement a particular task (Mather et al., 2013). According to the second dimension, mindset is the cognitive frameworks used to understand an event (French, 2016). Finally, the mindset can be defined as a belief that encompasses one's judgments about the flexibility of one's character or intelligence that one possesses (French, 2016). According to Dweck (2016) the mindset is the beliefs that people have their most basic qualities, such as their intelligence, ability, and personality. Students ' Mindset Theory is an explanation of their academic success, perseverance, persistence in the face of challenges, classroom participation, and feelings of academic involvement (Dweck et al., 2014). Based on all these explanations, the Mindset Theory can be defined/summarized as people's belief in the improvement of the abilities that individuals possess. Most of these abilities include intelligence. From this point of view, Mindset Theory is more associated with the belief that individuals' intelligence can be improved.

Students' mindsets are of critical importance because research has shown that students' mindsets affect the quality of learning processes, which in turn create different learning outcomes for students (Boaler, 2015). The Mindset Theory is not a cognitive skill, and there is an association between it and academic success (Laursen, 2015; Yilmaz, 2019). Considering all these, it can be argued that the Mindset Theory is a subject that needs to be explored with its different dimensions in the field of Education. Further research in this field can better describe students ' learning processes and achievements. These descriptions can help practitioners and policymakers. In order to do these, a valid and reliable measurement tool may be needed to describe students' Mindset Theory.

Six measurement instruments related to the mindset theory were found through a literature review. One of them is the 8-item single dimensional scale developed by Dweck (2006). The range of scores in this scale is 0-32. Scores between 0-16 indicate a fixed mindset, while scores between 16-32 indicate a growth mindset. The other scale developed by Dweck (2000) has two dimensions, fixed and growth mindset, and eight 7-point Likert type items. Midkiff et al. (2017), developed an 8-item, two-dimensional Growth Mindset scale. This scale is a psychometric scale developed in a sample of high school students in the UK. Ingebrigtsen's (2018) scale of growth mindset,

developed in his master's thesis with a sample of university students in Norway, has six items. The scale was found to be a valid and reliable psychometric scale. General Mindset Scale developed by Lottero-Purdue and Lachapelle (2019) for 10-11 age group is a 5-point Likert scale with six items and two dimensions. These dimensions are fixed and growth mindset theories. There is a 4-point Likert scale with 14 items developed by Abd-el-Fattah and Yates (2006) to identify individuals' beliefs about the nature of intelligence. This scale has two dimensions, Existence Theory and Incremental Theory.

Although it has been frequently cited in the literature in recent years (Busch, 2018; Rustin, 2016), there is a limited number of data collection tools for Mindset Theory that are exclusively for certain age groups and have been developed with particular focus on growth mindset. Thus, the lack of a comprehensive measurement tool that measures the Mindset Theory in a valid and reliable way is considered as a major gap in the field (Lüftenegger & Chen, 2017). In Turkey, there is no measurement tool that measures the mindset theory. This study aims to develop a data collection tool to measure the quality of adolescents' Mindset Theories. This scale aims to measure the mindset theories of students in adolescence. These years include formal education for most of the individuals. Adolescence comprises high school and university education periods. Adolescence is a period when psychological, mental and social development and maturation occurs and when individuals try to achieve the transition into adulthood (Ocakçı, 2015). Literature suggests adolescence in Turkey cover the age period between 11 and 20 (Çebi et al., 2016). Alternatively, Törüner and Büyükgönenç (2012) indicate that adolescence in Turkey starts at age 10-12 for girls and age 12-14 for boys and ends at age 21-24. Arnett (2004), states that the individual complete this period in late twenties and starts to feel as an adult. Considering these criteria about the adolescence, it was deemed suitable to design a measurement tool which would encompass individuals' high school and university education. At the same time, the planned measurement tool was expected to contribute to the research of various variables related to the individuals' educational process with regards to the mindset theory. In this context, the current research aims to develop a valid and reliable measurement tool for measuring the Mindset Theory qualities of students aged 14 and older.

Theoretical Framework

According to mindset theory, mindset has two dimensions. The first one is the Growth Mindset and the other is the Fixed Mindset (Dweck, 2008). Growth Mindset is the belief that one can improve one's intelligence, talents and skills (Fensterwald, 2015). Fixed Mindset is the belief that one's ability and intelligence have an invariant and unchanging structure. Individuals with this mindset believe they can accomplish a task to a certain extent with the characteristics they already have and therefore they think it is useless to strive (Dweck, 2016). The Mindset Theories of the individuals can be assumed to be an important predictor of their development and learning.

People whose Growth Mindset dimension of Mindset Theroy is developed assume that one's abilities and skills can develop through hard work (Laursen, 2015). These individuals believe that intelligence has improvable rather than a static structure (Claro et al., 2016). They believe that they can improve their intelligence, that learning is more important, that it is important to stand up to mistakes or failures. They tend to see mistakes as an opportunity to improve, make inferences from others' success and learn from their experiences. They think they should strive for

improvement by leaving their comfort zone (Beere, 2019). Growth Mindset is also associated with the individuals' belief in their capability to improve their intellectual abilities (Claro et al., 2016). Intellectual abilities refer to verbal and/or non-verbal mental skills, abstract reasoning, problem solving, mental speed or memory (Pfeiffer & Jarosewich, 2003). Individuals' beliefs that they can improve these intellectual competencies can enable them to develop a positive perspective towards themselves. According to Achor (2012), when individuals work with a positive mind set, their performance, productivity and creativity improve at almost all levels. Growth Mindset of Mindset Theory is also the belief that one can develop one's mental capacity like a muscle and make it stronger (Aronson et al., 2002). Growth Mindset calls for motivation and self-regulation to achieve a goal. For example, individuals with Growth Mindset of Mindset Theory are more likely to continue and repeat their efforts when faced with difficulties (Burnette et al., 2013).

Growth Mindset is a variable that positively affects students' academic achievement (Blackwell et al., 2007; Yeager et al., 2014). Growth Mindset is also associated with individuals ' traits of openness, responsibility, extraversion, compatibility and emotional balance (Lindgren et al., 2019). Students with Growth Mindset believe that they can learn the concepts, formulas, information etc. within a course with faith and determination. These students believe that it is possible to learn from mistakes and know that it is a way of learning. They are aware of the necessity of linking new information with past information in order to learn lesson topics and to make sense of the subject. Besides, they know the importance of taking time to think deeply and to really understand what they are learning instead of just rote memorization of course subject (Szpirglas & Saint-Onge, 2018). Growth Mindset may also require effort in the process of change and improvement. Effort is defined as struggle, zeal and desire to work (Turkish Language Society [TDK], 2020). Zealous individuals may be willing to do any work and they do not give up. They can struggle to achieve something and they believe they can change it. The diligence of individuals can make their growth mindset more dominant.

The other dimension of mindset theory is the Fixed Mindset. Individuals whose Fixed Mindset of Mindset Theory is developed believe that their mental abilities, characters and creativity skills are immutable traits given to them, and they do not make any effort to change or improve them. These people avoid taking the risk of engaging with new situations that require effort because they are afraid of making mistakes, perform under their real capacity (Güven &Yılmaz, 2017). Contrary to the people with Growth Mindset of Mindset Theory, people whose Fixed Mindset dimension is developed are more likely to escape from a difficult task and feel helpless in such a situation (Burnette et al., 2013).

Individuals with a developed fixed mindset dimension of Mindset Theory may be in procrastination. Procrastination can be defined as the tendency of the individual to delay his or her actions and plans for some reason, and to feel a stasis and laziness (Sekman, 2007). Individuals in procrastination may experience insufficient levels of motivation, lack of self-regulation, lack of empathy and socialization, and a sense of learned helplessness (Çankaya, 2010). People who are in procrastination may have tendencies such as stagnation, inaction, passivity, monotony, laziness, and they may be willing not to take action (Çankaya & Demirtaş, 2010). These people may not be open to

innovation, they may be reluctant to change the way they work and think. They may even see these changes unnecessary.

Method

Study Sample

The population of the study consists of 18778 high school students studying in a central district of Konya and 7524 university students from a faculty of education in the same district, which makes a total of 26302 students. Sample size needs to be at least five times or even almost ten times the number of items when developing a scale (Bryman & Cramer, 2001). Sample size of 100 is considered to be weak, whereas 200 is medium, 300 is good, 500 is very good and 1000 is perfect (Comrey & Lee, 1992). The draft version of the scale had 26 items. To perform exploratory and confirmatory factor analysis, a sample size between 260 and 500 was considered to be appropriate. The sample size was decided to be 700. Multi-stage sampling was used. The high schools in the universe were separated into stages according to their types (General, Science, Vocational), while university students were categorized according to their entry scores (Mathematics/Science, Social Science/Humanities, equally-weighted, General Ability). These students were divided into substages according to their grade/year. Grade level substages were accepted as clusters. 717 students comprised the sample of the scale development process via random cluster sampling. This sample was randomly divided into two to use one half for exploratory factor analysis and the other for confirmatory factor analysis. Of the participants, 18% were in 8^{th} grade, 17% were in 9^{th} grade, 16% were in 10^{th} grade, 18% were in 11^{th} grade, 11% were in 12^{th} grade and 20% were in university. The understandability of the items in the draft scale was established in a group of 102 students chosen via random cluster sampling from the high school first year substage of the universe. Test-retest reliability of the scale was tested with 124 students chosen from the high school third year substage of the universe via random cluster sampling. Finally, criterion validity of the scale was tested with 202 students chosen from high school second year substage of the universe via random cluster sampling. Clusters used in the stages of scale development were excluded from the universe to prevent them from being included in the other stages of development.

Data Collection Tools

The Short Grit Scale developed by Sarıçam et al. (2016) and The Psychological Hardiness Scale developed by Işık (2016) were used as criteria to determine the criterion validity of the Mindset Theory Scale (MTS).

The Short Grit Scale: Determination and effort constitutes a part of the Mindset Theory's scope of includes determination and effort. For this reason, the researcher decided to use The Short Grit Scale developed by Sarıçam et al. (2016). The Short Grit Scale is a 2-dimensional scale with 8 items. Overall Cronbach Alpha internal consistency reliability coefficients were calculated as .83 for the scale, .80 for the consistency of interest sub-dimension, and .71 for the persistence in effort sub-dimension.

The Psychological Hardiness Scale: The Mindset Theory also includes the belief that intelligence can improve and one can improve oneself diligently. Considering that these beliefs are related to dedication, control and challenge, the researcher decided to use the The Psychological Hardiness Scale developed by Işık (2016) as a

criterion fort the MTS developed in this study. The Psychological Hardiness Scale consists of 21 items and three sub-dimensions. These dimensions are Dedication, Control and Challenge. While the overall Cronbach alpha reliability coefficient is .76 for the scale, the Cronbach alpha reliability coefficient for each sub-dimension is between .62 and .74.

Development of the draft scale: Draft form of MTS and personal information form were used to collect the necessary data that would be used in the research. In the preparation of the draft version of MTS, the scale development stages such as creating a pool of items, obtaining expert opinions, pilot implementation, determining validity and reliability were followed (Şeker & Gençdoğan, 2014; Tavşancıl, 2005). The theoretical structure and research results in the relevant literature were utilized in the development of the draft items of the MTS (Anderson & Glover, 2017; Baruch-Feldman, 2017; Beere, 2019; Boaler, 2015; Breuning, 2015; Dweck, 2016; Ricci, 2013a; Ricci, 2013b; Ricci & Lee, 2016; Sternberg et al., 2011; Szpirglas & Saint-Onge, 2018; Yılmaz, 2019).

Mindset Theory qualities were considered with their dimensions and items were prepared in a way suitable for the study. Experts who have conducted studies in the theories of thought, intelligence, brain and mindset and who have academical studies in these fields (7 experts in their fields) were consulted before the finalization of MTS. The experts expressed their views on the content, structure, applicability and meaning of the items. Moreover, a questionnaire was developed to evaluate the experts' evaluation of the items in the draft scale. The experts were asked to score every item on scale of 1-4 (1: not suitable, 2: needs major correction, 3: needs minor correction, 4: very suitable). The conformity of the expert opinions was tested with Kendall's coefficient of concordance. Kendall's analysis performed showed that there was not a statistically significant difference between expert opinions (Kendall's W = .160, p = .308). In line with the written opinions of the experts, the necessary changes, corrections and item removals were made. The draft scale was evaluated by three linguists in terms of language, expression and statement. Taking into account the opinions of linguists, changes and corrections were made in some items in terms of spelling, form, language and expression. As a result of these revisions, the draft scale was reduced from 33 to 26 items.

Finally, the 26-item draft form of MTS was evaluated by assessment and evaluation experts and answering format and possible choices were decided. It was decided to use a 5-grade likert type and the choices were: "It is very suitable for my thoughts (5 points)", "It is suitable for my thoughts (4 points)", "It is partially suitable, partially not suitable for my thoughts (3 points)", "not suitable for my thoughts (2 points) and "Not suitable for my thoughts at all (1 point)". In accordance with measurement and evaluation experts' suggestions, the items in the Mindset Theory Scale were scored as 5-4-3-2-1. Then, the draft scale was finalized by adding instructions.

The draft scale was applied to 102 students studying in 9th grade During the application, students were asked to mark the items they did not understand. Immediately after the implementation of the draft scale, students' opinions on the intelligibility of each item of the draft scale were received. At the end of the application, problems in the implementation was considered and some revisions in terms of spelling and typos were made in the scale, in line with the opinions of the students. Finally, the draft scale was examined by linguistics experts in terms of language, expression and statement and the draft of the MTS was finalized.

The final draft scale was applied to 717 students. It was observed that the scale could be answered by participants in 10-11 minutes. During the application process, students were given information about the personal information form and how to implement MTS, and explanations were made about the issues that were not understood by the students.

Data Analysis

The existence of missing data in the data collected in the research process was examined. Substracted data can be tolerated if its ratio is below 5% and it has normal distribution (Acuna and Rodriguez, 2004). Therefore, two missing data was excluded from the data set. There is a possibility of outliers when the assumption of normality is examined in a data set. Unidirectional outliers in the data set can be checked by converting item scores to Z values (Tabachnick and Fidell, 2007). 4 data points with Z values outside of +3 / -3 interval were accepted as unidirectional outliers and excluded from the data set. Then, outliers and multivariate normality assumption in the dataset was examined through Mahalanobis distance values and 6 data points were not included in the analyses based on their possibility of being outliers ($p < 0.01$). Moreover, univariate normality assumption was evaluated by checking skewness and kurtosis coefficients. The skewness and kurtosis values of the data set were found to be between +1 and -1, which is the interval for normality assumption (Morgan et al., 2004). After these stages, the dataset was reduced to 705 participants.

During the analysis of the items of MTS, item analyzes and exploratory and confirmatory factor analyzes were carried out to determine the construct validity. According to Doğan et al. (2017), if the necessary minimum sample size for conducting both EFA and CFA with the data obtained from the same sample is reached, that sample can be divided into two to perform these. In this regard, data obtained from the same sample was randomly divided into two for EFA and CFA. 353 people randomly selected from the half of the same sample was used for EFA while 352 people were used for CFA. Before conducting EFA the suitability of the data set for factor analysis was checked by Kaiser-Meyer Olkin (KMO) and Barlett tests. Item-total score correlation values were analyzed to determine whether the draft scale items were related to the scale. The structural validity of MTS was tested through factor analysis. The correlation coefficients between the scores of MTS's subdimensions were found. 4 factor structure resulting from EFA was tested by single level CFA. Two latent variables were added to the single level factor analysis to test two level factor model. As a result of the test of two level factor model, Cronbach alpha internal consistency coefficient method was used to determine the reliability of the scales. To check whether the items have differentiated in lower and upper groups, t-Value, average of the scores received from the items and their standard deviation were calculated. Correlation coefficients between the scores of the sub-dimensions of MTS were found. With the latest version of the scale, Pearson Product-Moment Correlation coefficient was used in the process of the test-retest method to estimate stable measurements within the scope of the scale and its reliability. In order to estimate the criterion validity of MTS, the Short Grit Scale developed by Sarıçam et al. (2016) and the Psychological Hardiness Scale developed by Işık (2016) were used as criteria. In order to estimate the relationship between these scale scores, the Pearson Product-Moment Correlation coefficient correlation technique was used. To test whether the total score

from MTS can be used, three level multifactor analysis was conducted. A scoring guideline for MTS was created by gradient sum technique.

Ethical Aspect of the Research

In this study, all the rules stated in the "Higher Education Institutions Scientific Research and Publication Ethics Directive" were followed. None of the actions specified under the second section of the Directive, "Scientific Research and Publication Ethics Actions" have been carried out.

Ethics committee permits

Ethical evaluation committee: Necmettin Erbakan University Scientific Research Ethic Commission

Date of the ethical evaluation decision: 08.05.2020

Ethical evaluation document number: 2020/13

Finding

The validity and reliability estimation processes of MTS are explained below.

Item Analysis

First of all, in order to determine whether the draft scale items are related to MTS, item analysis was performed. When the item-total score correlations of MTS are examined, it can be seen that the relevant values vary between 0.117 and 0.451. Most of these values were found to be above 0.30. In general, it is stated that item-total score correlation is suitable for items above 0.30, but if the values between 0.20-0.30 are deemed appropriate, they can be used for testing (Büyüköztürk, 2015). For this reason, it is accepted that the items in the scale are suitable for the analysis.

Construct Validity

Exploratory factor analysis (EFA) of the mindset theory scale (MTS): While developing psychometric tests, the stages of analyzing the suitability for factor analysis, performing factor analysis, determining the factors and naming the factors should be followed in order to determine the construct validity (Kalaycı, 2014). The stages suggested by Kalaycı (2014) were followed in the Exploratory Factor Analysis of MTS.

In order to determine the suitability of the data collected for the MTS development process to factor analysis, the Kaiser-Meyer-Olkin (KMO) coefficient, Bartlett's test of sphericity result and the diagonal values of the data matrix were analyzed. KMO coefficient was calculated as 0.866 for the collected data. In addition, the Bartlett's test chi-square value was statistically significant (X^2 = 5982,354; p <0.01). All of the diagonal values in the anti-image matrix are greater than 0.50. In line with these results, it has been accepted that the data collected for the Growth Mindset scale provide the necessary conditions for factor analysis. Because, according to Büyüköztürk (2015), it is sufficient that the KMO coefficient is above 0.60 and the Bartlett test results are significant, whereas according to Pett et al. (2003), the main diagonal elements of the anti-image matrix is an indicator of the suitability of the items for factor analysis.

Literature shows that mindset theory is comprised of mental dimensions that are independent from each other and open to development (Dweck, 2016). Therefore, varimax rotation was used when the factor loads for EFA was calculated via Principal Component Analysis. Because varimax method reveals simple meaningful factors by rotating factor variances to maximum with few variables (Tavşancıl, 2005).

According to Büyüköztürk (2015), the factors with eigenvalue greater than 1 are important factors when interpreting the data obtained by Exploratory Factor Analysis. It is a good standard to have a factor load of at least 0.45 in order for a substance to be summed under a factor. This value can be lowered to 0.30 if necessary. Based on these explanations, it is considered that their eigenvalues are greater than 1 and factor loads are at least 0.40 for items to be included in a factor.

As a result of EFA, Scree Plot (Figure 1) was analyzed. According to Scree Plot (Figure 1), the number of X-axis components is four at the breaking point where the slope disappears. It was therefore decided that the number of important factors could be at least four.

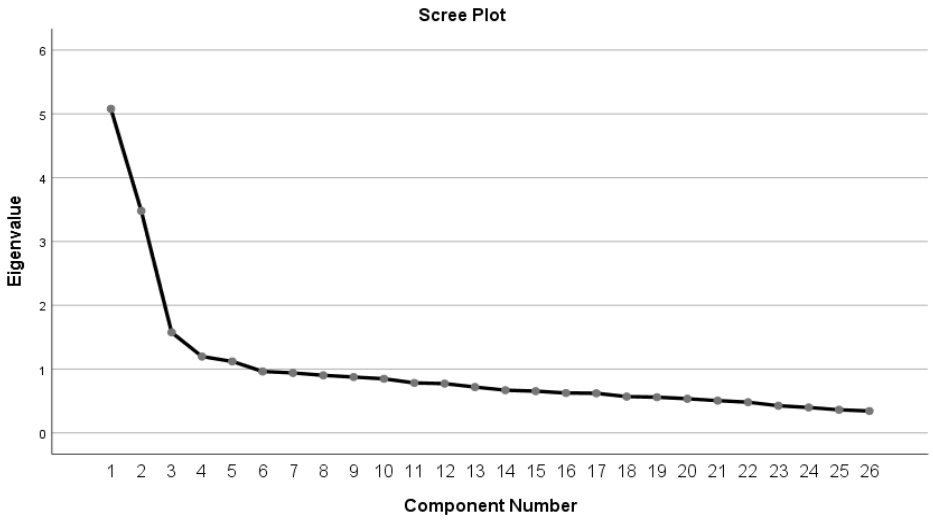


Figure 1. Scree plot of the Mindset Theory Scale

In this study, it is accepted that a four-factor structure may be possible considering the factors and theoretical foundations with eigenvalues greater than 1.

Of the 26 items belonging to MTS, 7 overlapping items (i9, i4, i2, i3, i3, i7 and i26) with a factor load below 0.40 or loaded into multiple factors were respectively excluded from the analysis. EFA was performed again in each item extraction. Experts were counselled at each step of item extraction. Based on experts' opinions it was decided that said items did not have cohesion with the other items composing a factor and that extraction of said items would not cause the scale to be inadequate in measuring its scope. Consequently, a structure with 4 sub-factors and a total explained variance ratio of 51,320% was obtained. The factor loads obtained from the EFA and the factor loads of the items are shown in Table 1.

Table 1

Rotated component matrix of the Mindset Theory Scale (varimax)

Components				
Items	1	2	3	4
i1				,708
i5				,683
i20				,525
i21			,683	
i8			,755	
i12			,713	
i15			,623	
i18			,502	
i22		,513		
i6		,763		
i10		,773		
i23		,777		
i16		,440		
i24	,602			
i7	,640			
i11	,411			
i14	,626			
i25	,736			
i19	,738			

Variance Total: 51,320, Factor 1: 22,750%, Factor 2: 15,224%, Factor 3: 7,757%, Factor 4: 5,589

When Table 1 is examined, it can be seen that factor loads obtained in Exploratory Factor Analysis are between .411 and .777. The factor loads are above 0.40 which is accepted as the lower limit. When the factor load values of the items are analyzed, it is clear that the items loaded on the factors measure the desired structure appropriately. The factors obtained are named by considering the items they contain. Accordingly, the first factor consisting of 6 items (i24, i7, i11, i14, i25 and i19) was named as “Procrastination”, the second factor consisting of 5 items (i22, i6, i10, i23 and i16) was named as “Belief in Improvement”, the third factor consisting of 5 items (i22, i6, i10, i23, ve i16) was named as "Effort", and the fourth factor consisting of 3 items (i1, i5 and i20), was named as "Immutability of Belief".

Table 2

Correlation coefficients between the factors of Mindset Theory Scale

Mindset Theory Scale's Dimensions		Procrastination	Immutability of Belief	Belief in Improvement
Immutability of Belief	r	,548**		
Belief in Improvement	r	-,174**	-,287**	
Effort	r	-,184**	-,140**	,393**

**: p<.01; *: *Bayram, 2013; Hu and Bentler, 1999; Schumacker and Lomax, 2004*

The correlations between the sub-dimensions of the scale ranged from -.287 to .548 and had a significant relationship at the level of .05 (Table 2).

The model resulted from EFA was tested in CFA to check construct validity (Kline, 2011).

Confirmatory factor analysis (CFA) of mindset theory scale: Confirmatory Factor Analysis (CFA) is used with EFA in scale development studies. Using CFA, the extent to which the existing theoretical structure overlaps with the available data can be tested (Schumacker & Lomax, 2004). The MTS model obtained from EFA, which includes 19 items with 4 factors, was subjected to confirmatory factor analysis. CFA analysis revealed that model fit values of 19 statements in the scale were not at an acceptable level. In the model created as a result of CFA error variance of 6 items were quiet high while their regression weights were very low, which made the authors think that these items were not fitting the structure of the scale. These six items were extracted from the model one by one and the model was tested again. Accepted values for fit indices were achieved at the last stage of repeated fit index calculations. Experts were counselled each time an item judged to be unfit was extracted from the model due to the concerns about the content validity of the scale. In the opinions received from the experts, it was noted that there were other items measuring the quality measured by the problematic items more comprehensively and that the scale would not lack content validity if the said items were removed. In light of these opinions, item numbers of MTS were revised as the number of items was down to 13 due to the removal of 6 items. Another model was created with these revised numbers and fit index calculations for the constructed model had adequate values. The diagram obtained from CFA for the validity study of MTS, after all the revisions, is given in Figure 2.

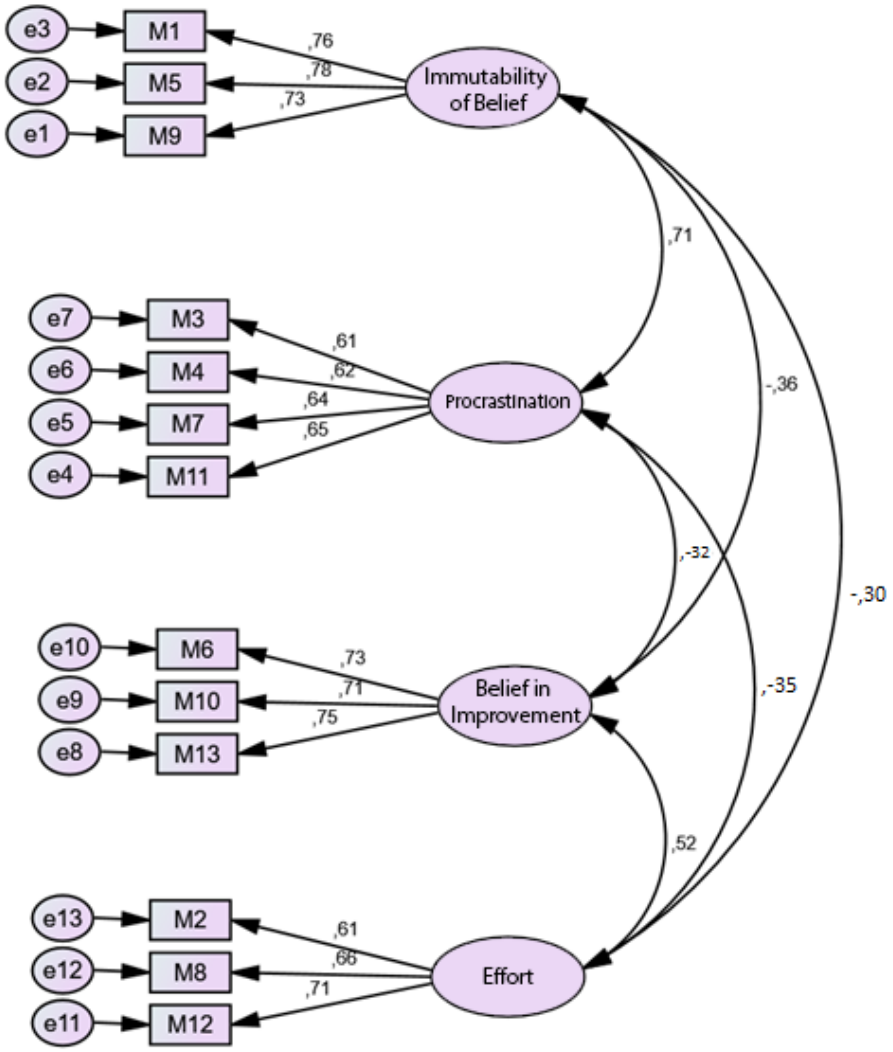


Figure 2. CFA results of Mindset theory scale: Standardized path diagrams (path analysis)

When the path diagrams of MTS related to CFA in Figure 2 are inspected, it is seen that the standardized path coefficients of the items vary between 0.61 and 0.78 (Figure 2). Kline (2005) states that items having standardized path coefficients of 0.50 or more represent the relevant variable. When the standardized path coefficient of the items in model are inspected, it can be thought that these items have adequate predictive value. Also, fit index values for this model are presented in Table 3.

Table 3

Compliance index values and comparison of the CFA results of the Mindset Theory Scale

Model	χ2/sd	GFI	CFI	IFI	AGFI	NNFI	RMSEA
	144,874./59=2,455	,966	,964	,964	,947	,941	0,047
Fit comment*	Perfect fit	Perfect fit	Perfect fit	Perfect fit	Perfect fit	Acceptable fit	Perfect fit

It is seen that the fit indices of the scale obtained with 4-factor by CFA of MTS generally have good values. The ratio of chi square value to degree of freedom is (χ2/sd=2,455. GFI (Goodness of Fit Index), CFI (Comparative Fit Index), IFI (Incremental Fit Index) and NNFI (Non-Normed Fit Index) compliance indices are close to 0.95 and RMSEA value is less than 0.05, which indicate that the model fits well with the data. The fit indices obtained for the scale in this study can be accepted as evidence that the proposed model matches the data at hand (Bayram, 2013; Hu and Bentler, 1999; Schumacker and Lomax, 2004).

Meydan and Sesen (2011) state that second-level multifactor models of multidimensional scales must also be tested when applying confirmatory factor analysis. Mindest Theory is primary interpreted as two dimensional. These dimensions are Fixed and Growth Mindset (Dweck, 2016). Based on literature, the subdimensions of Mindset Theory revealed by EFA and confirmed at single level, "Procrastination", "Immutability of Belief", "Effort" and "Belief in Improvement", were expected to be related to a higher level variable. The relationship of these 4 subdimensions with two latent variables was tested in a two-level multifactorial model. The second-level CFA results for this four-subdimensional, two-dimensional model is shown in Figure 3.

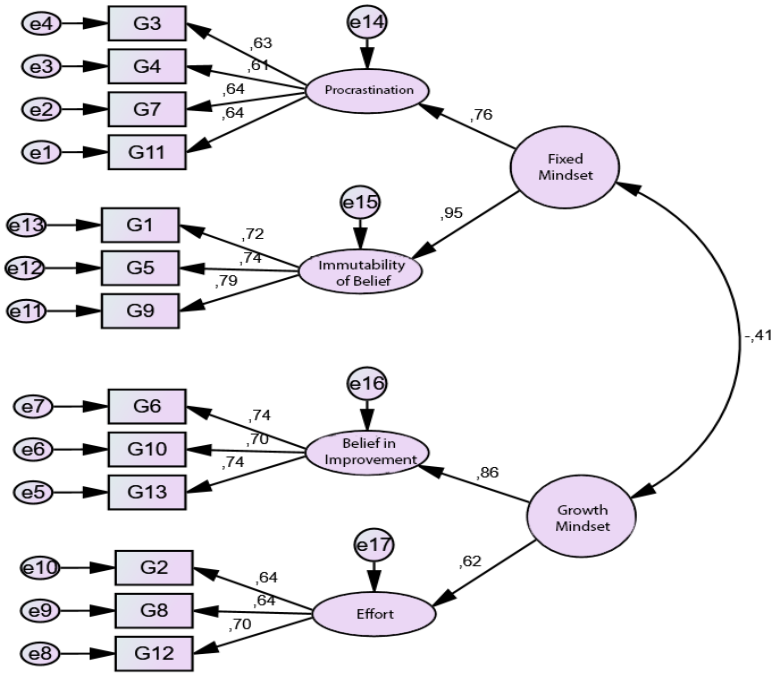


Figure 3. Second level CFA results of Mindset Theory Scale: Standardized path diagrams

A second-level confirmatory factor model was created to show that four factors obtained in the first level CFA represent theoretically proposed Mindset Theory in two higher level dimensions. More specifically, Procrastination and Immutability of Belief are loaded on one higher level dimension while Effort and Belief in Improvement on another (Figure 3). The second level factor model was tested by adding two latent variables named Fixed Mindset and Growth Mindset to the first level confirmatory structure tested with 4 latent and 13 indicator variables. The goodness of fit values is given in Table 4.

Table 4

Mindset Theory Scale second-level CFA results fit index values and their comparisons

Model	χ2/sd	GFI	CFI	IFI	AGFI	NNFI	RMSEA
	155,805/60=2,597	,964	,960	,960	,945	,937	0,050
Fit comment *	Acceptable fit	Perfect fit	Perfect fit	Perfect fit	Perfect fit	Acceptable fit	Perfect fit

* (Bayram, 2013; Hu and Bentler, 1999; Schumacker and Lomax, 2004)

Results of first level and second level CFA indicate that MTS is a valid measurement tool that can be used to determine the total Fixed Mindset levels, the Procrastination and Immutability of Belief levels, and the total Growth Mindset levels and Effort and Belief in Improvement levels.

Table 5 presents the item-total score correlation values of the items of the scale, factor loads, t-value for the lower and upper group difference, the mean and standard deviation of the scores received.

Table 5

The rotated components matrix of the Mindset Theory Scale (varimax)

	Item No	EFA factor load	T value for upper and lower group difference	Item average	Standard deviation
Procrastination	G3	-,602	24,90**	2,84	1,29
	G7	-,640	27,89**	2,71	1,35
	G4	-,736	25,27**	2,47	1,36
	G11	-,738	29,59**	2,36	1,41
Immutability of Belief	G1	,708	38,14**	2,90	1,37
	G5	,683	38,58**	2,76	1,35
	G9	,525	31,54**	2,59	1,32
Effort	G8	-,755	22,89**	4,10	1,00
	G12	-,713	28,11**	4,06	,99
	G2	-,623	23,55**	4,03	,98
Belief in Improvement	G6	,763	21,46**	4,46	,79
	G10	,773	26,02**	4,31	,86
	G13	,777	31,08**	4,17	,94

**: p<.01

Factor loads of MTS are over 0.525 and there is a significant difference between the upper and lower group item mean scores. Hence, it can be said that MTS differentiates individuals scoring high and individuals scoring low on the scale.

Criterion-related validity of mindset theory scale: In order to examine the criterion validity of MTS, Short Grit Scale and the Psychological Hardiness Scale were used. The correlation coefficients were calculated between the total scores obtained from the scales (Table 5).

Table 6

Criterion validity results of Mindset Theory Scale

		Grit	Psychological Hardiness
Growth Mindset	r	,399**	,556**
Fixed Mindset	r	-,312**	-,358**

**: p<.01

A positive and significant relationship was found between Growth Mindset and Psychological Hardiness scores (p <05) (Table 6). A significant positive relationship was found between Growth Mindset scores and Grit scores (p <05). A significant negative relationship was found between Fixed Mindset and Psychological Hardiness scores (p <05). A significant negative relationship was found between Fixed Mindset scores and Grit scores (p <05). These results show that the Mindset Theory Scale has criterion validity. Evans (1996) states that Pearson correlation coefficient r values of <0.40-059 is medium and -.20 - -.39 is weak. The relationship between MTS and criterion measures of Psychological Resilience and Perseverance can be said to be weak and medium.

Findings on the reliability of the mindset theory scale: To determine the reliability of MTS, item properties were determined using item analysis. Total scores of the items with their correlations were calculated. To further examine the reliability of the dimensions and the sub-dimensions of the MTS, the internal consistency of the items that make up the scale was estimated by the Cronbach-Alpha coefficient method. All these results are given in Table 7.

Table 7

Some reliability analysis values of the scale items of the Mindset Theory Scale

		Item no	Item-total score correlation	Cronbach alpha internal consistency coefficient	
Fixed Mindset	Procrastination	G3	,475	0,724	0,723
		G7	,506		
		G4	,522		
		G11	,547		
	Immutability of Belief	G1	,654	0,805	
		G5	,679		
		G9	,595		
Growth Mindset	Effort	G8	,519	0,701	0,714
		G12	,552		
	Belief in Improvement	G2	,474	0,771	
		G6	,609		
		G10	,600		
		G13	,615		
Whole of the scale				0,803	

Item-total score correlation of items of MTS varies between 0.474 and 0.667. The Cronbach Alpha internal consistency coefficient was examined to determine the scale's reliability. The internal consistency coefficient of MTS is 0.803 the reliability values of the four factors of MTS, Procrastination, Immutability of Belief, Effort and Belief in Improvement, were 0.724, 0.805, 0.701 and 0.771 respectively. The internal consistency coefficient of MTS is 0.803.

Test-retest results: In order to determine the reliability of the scale with test-retest method, MTS was applied to 102 students (8th, 9th, 10th, 11th, 12th and university students) with an interval of four weeks. The relationship between their scores was calculated using the Pearson Product-Moment Correlation coefficient. The results are presented in Table 8.

Table 8

Results on the test-retest results of the Mindset Theory Scale

			2nd Applıcatıon scores	
			Growth mindset	Fixed mindset
1st Application Scores	Growth Mindset	r	,380**	-
	Fixed Mindset	r	-	,470**

**: p<.01

A positive weak and medium significant relationship was found between the first and second application scores of the dimensions of MTS (Evans, 1996). These results indicate that MTS does not make strongly stable measurements. Overall, all of the analyzes conducted to determine the reliability of the MTS indicate that the reliability of the scale is sufficient.

In conclusion, the EFA results revealed that the scale has four dimensions is a four-dimensional structure. The model created by EFA was tested with CFA and the compliance values have been sufficient and acceptable. In the CFA analysis conducted at the second level, it was found that the sub-dimensions of Procrastination and Immutability of Belief were represented in the Fixed Mindset upper dimension, and the Effort and Belief in Improvement sub-dimensions in the Growth Mindset. The scale was observed to have criterion validity. It was found that the internal consistency coefficients were at a good level and made stable measurements in the results related to the reliability of MTS. It can be said that the scale is a reliable and valid scale by considering all values related to MTS.

Whether Fixed and Growth Mindset dimensions of MTS are components of Mindset Theory was tested in a three level factor analysis. To this end, the results of a three level CFA to test the possibility of 4 subdimensional and 2 dimensional structure to be a component of a higher dimension and the suitability of the scale for the use of its total score are shown in Figure 4.

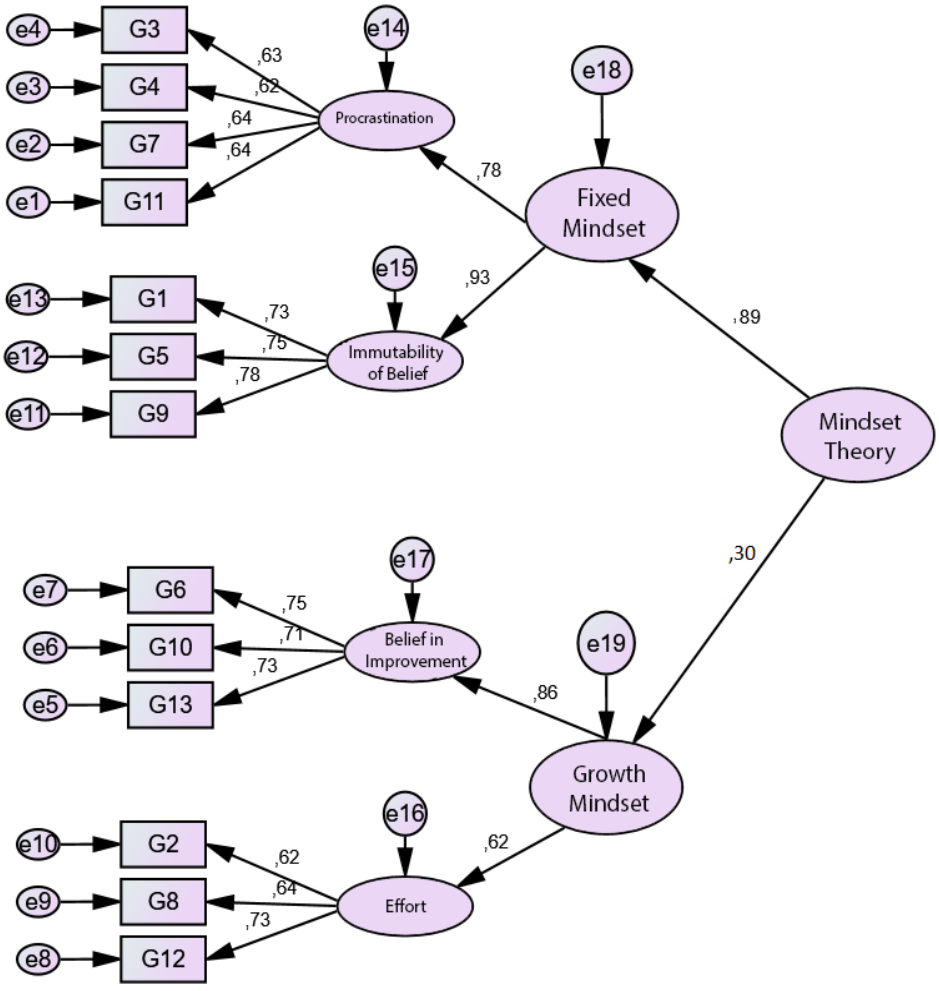


Figure 4. Three level multi factor model for Procrastination, Immutability of Belief, Effort and Belief in Improvement subdimensions of MTS's Fixed and Growth Mindset dimensions.

To test the possibility of using total scores for MTS a three level multi factor model was created. When the values related to the model in the Table 9 were reviewed, it was understood that fit indices were at an acceptable level. These values show that items represent a higher level with the factor they belong to. According to these results, it may be acceptable that the total score of the scale can be used.

Table 9

Fit indices of three level CFA results of MTS scale and their comparison

Model	χ2/sd	GFI	CFI	IFI	AGFI	NNFI	RMSEA
Fit comment*	155,805/60=2,597	,964	,960	,960	,945	,937	0,050
	Acceptable fit	Good fit	Good fit	Good fit	Good fit	Acceptable fit	Good fit

Scoring of the mindset theory scale: The score evaluations regarding the dimensions of the developed MTS are given in Table 10.

Table 10

Scoring table of Mindset Theory Scale

Dimensions and sub-dimensions of mindset theory scale	Items	The lowest possible score	The highest possible score
Procrastination	3,4,7 and 11	4	20
Immutability of Belief	1,5 and 9	3	15
Effort	2,8 and 12	3	15
Belief in Improvement	6,10 and 13	3	15
Fixed Mindset	3,4,7,11,1,5 and 9	7	35
Growth Mindset	2,8,12,6,10 and 13	6	30
Growth Mindset	1*, 2, 3*, 4*, 5*, 6, 7*, 8, 9*, 10, 11*, 12, 13	13	65

(*: Items requiring reverse scoring at the total score stage of the scale)

MTS was accepted to be scored in accordance with gradient sum technique. The score obtained from a scale suitable for gradient sums technique is generally is the sum of scores of responses to the items in the scale (Tezbaşaran, 1996). In the Likert type scale like MTS, the participants state the degree to which they agree with the attitude element contained in the each statement in the scale. Information regarding to whether the participant's attitude is positive or negative based on self-judgements of the participant's total score from the scale by the gradient sums technique (Tavşancıl, 2005; Tezbaşaran 1996). When MTS is applied, the participant responses to each item in the scale and states the degree of his/her attitude towards the item. The score corresponding to this degree is the participant's score on that item. The total score of MTS is obtained by summing the scores of subdimensions belonging to the dimensions.

MTS can be scored separately with its dimensions and sub-dimensions. Points can be obtained in the range of 7 to 35 from the Fixed Mindset dimension of the MTS. While the Procrastination sub-dimension of this dimension can be scored in the range of 4-20, the Immutability of Belief sub-dimension can be scored in the range of 3-15. Points from 6 to 30 can be obtained from the MTS's Growth Mindset dimension. A score of 3-15 can be obtained from the Effort sub-dimension of MTS's Growth Mindset dimension, while the other sub-dimension Belief in Improvement can be scored from 3 to 15. When the literature is examined, the Mindset Theory is more often called the Growth Mindset. For this reason, it is thought that there may be researchers who want to get a total score from the scale. In order to get total points from the scale, items number 1, 3, 4, 5, 7, 9 and 11 of the scale are reverse scored. These items are items that are collected under the Fixed Mindset dimension. These items are reverse scored only if one wants to obtain total scores from the entire scale.

The high score obtained from each dimension and sub-dimensions indicates that the person has a high level of competence in the relevant dimension and sub-dimension, and a low score indicates that the competence in the relevant dimension and sub-dimension is low. This evaluation is also valid for the score obtained from the whole scale.

Discussion, Conclusion and Recommendations

Fixed and Growth Mindset dimensions of the Mindset Theory have become important in terms of students' learning processes. Education stakeholders need to understand the importance of the Mindset Theories and be aware that this effects their performance (Baldwin, 2019). Therefore, the Mindset Theory has become an important field of study. For the studies to be carried out in this area, a measurement tool may be required to describe the Mindset Theories. In this context, in this study, it was aimed to develop a measurement tool to describe the mentality theories of students over the age of 14.

No Turkish language scale was found in the literature to describe the students' Mindset Theories. The scales developed in different countries had some limitations because the scales are generally focused on the growth mindset theories of individuals and do not cover any information about the qualities of the fixed mindset. In addition, these scales do not give information about individuals' efforts and motivations for the development of their talents.

There are two dimensions of MTS, each with two sub-dimensions. It is a scale that determines the students' own qualities according to their own perceptions. The subdimensions of the Growth Mindset dimension of MTS are Belief in Improvement and Effort. The MTS's Fixed Mindset dimension consists of the sub-dimensions of Procrastination and Immutability of Belief. The Growth Mindset dimension of MTS includes items 2, 6, 8, 10, 12 and 13. Some of the items in this dimension are as follows: "It is up to me to develop my intelligence", "I try to learn lessons from my mistakes". MTS's Fixed Mindset dimension includes items 1,3,4,5 7,9 and 11. Some of the items in this dimension are: "I believe my intelligence level will not change.", "I feel threatened while doing something/a job".

The Procrastination sub-dimension of the Fixed Mindset dimension of the MTS measures the students' tendency to delay their actions, their stasis, their laziness, their lack of effort. The Immutability of Belief subdimension of the

Fixed Mindset dimension of MTS is related to the idea that the students have a structure that does not change their intelligence and that they think of their intelligence as constant. The Belief in Improvement subdimension of MTS's Growth Mindset dimension measures students' beliefs that they think of their intelligence as a muscle and that they can improve it when they try to do so. The Effort sub-dimension of the Growth Mindset dimension includes the desire of the students to understand their mistakes while doing something, to learn from them and to make an effort.

When the validity and reliability values of MTS were reviewed, it was determined that the scale items could measure the quality they aimed to measure and differentiate the participants' level of quality aimed to be measured. Considering the expert opinions and its content validity MTS can be said to represent the universe aimed to be measured. The factor loads of the model can be accepted to be adequate based on the values obtained from exploratory factor analysis performed to test the structural validity of MTS. t values for the difference between top and bottom groups of the scale show that MTS can measure the structure it measures with enough differentiating power. The structure resulting from the EFA result MTS was tested with one, two and three level CFA, Fit indices of one level CFA showed that there is an adequate fit between the data and the model structure. Belief in Improvement, Effort, Procrastination and Immutability of Belief, revealed in EFA, was determined to be related to a higher level dimension in the literature. These dimensions are Growth and Fixed Mindset. It was determined by the second level CFA that it is a measurement tool that can be used to determine the levels of Fixed Mindset and Procrastination and Immutability of Belief and to determine the levels of total Growth Mindset and levels of Effort and Belief in Improvement. The good level of internal reliability coefficients related to the dimensions and subdimensions of MTS indicate that items of the dimensions and subdimensions are consistent with each other. MTS was found to have a reliable and valid structure as a result of this study.

The fact that MTS gives sufficient fit indices in the first, second and third level CFAs shows that it can be performed on the scores obtained from both dimensions and the subdimensions of the scale and it also reveals that the total score related to the Growth Mindset and Fixed Mindset dimension can be obtained from the scale. The increase in the scores obtained from the dimensions and subdimensions of the MTS means that there is high quality in that area.

The scales measuring the quality MTS aims to measure are generally labelled as Growth Mindset. The literature indicates that Growth Mindset consists of growing mentality and fixed mentality dimensions. This was judged to be an inconsistency between the subdimensions of the scales and their general name. However, when the literature is reviewed, it can be seen that the individuals' qualities of growing mentality and fixed mentality are related to their mental structures. Also in literature these are considered to be dimensions of mental theories. Because of these reasons, the scale developed in this study was called Mindset Theory Scale.

The age range of the study group is 14-22 during the development of the MTS. The validity and reliability studies of the scale for different developmental periods may be needed.

Ethic

I declare that the research was conducted in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

Funding

The authors received no financial support for the authorship, research, and publication of this article.

References

Abd-El-Fattah, S. M., & Yates, G. C. R. (2006). *Implicit theory of ıntelligence scale: testing for factorial ınvariance and mean structure.* Paper presented at the Australian Association for Research in Education Conference, Adelaide, South Australia.

Achor, S. (2012). *Positive intelligence.* Harvard Business Review.

Acuna, E., & Rodriguez, C. (2004). A meta analysis study of outlier detection methods in classification. *Technical paper, Department of Mathematics, University of Puerto Rico at Mayaguez*, 1-25.

Anderson, L. & Glover, D. R. (2017). *Building character, community, and a growth mindset in physical education: Activities that promote learning and emotional and social development.* Human Kinetics, Inc.

Arnett, J. J. (2004). *Emerging adulthood: The winding road from the late teens throughthe twenties.* Oxford University Press, New York.

Aronson, J. M., Fried, C. B. & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology, 38*(2), 113–125.

Baldwin, A. (2019). Impacting growth mindsets in student employment. *New Directions for Student Leadership, 162*, 49-60. https://doi.org/10.1002/yd.20333

Baruch-Feldman, C. (2017). The grit guide for teens: a workbook to help you build perseverance, self-control, and a growth mindset. New Harbinger Publications.

Bayram, N. (2013). Yapısal eşitlik modellemesine giriş AMOS uygulamaları [Introduction to structural equation modeling: Amos Applications]. Ezgi Publications.

Beere, J. (2019). *The growth mindset edge: your guide to developing grit*. Ulysses Press.

Blackwell, L. S., Trzesniewski, K. H. & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: a longitudinal study and an intervention. *Child Development, 78*(1), 246–263.

Boaler, J. (2015). *Mathematical mindsets: Unleashing students' potential through creative math, inspiring messages and innovative teaching.* John Wiley & Sons.

Breuning, L. G. (2015). *Habits of a happy brain: retrain your brain to boost your serotonin, dopamine, oxytocin, & endorphin levels.* Simon & Schuster.

Bryman, A. & Cramer, D. (2001). *Quantitative data analysis with SPSS release 10 for Windows.* London: Routledge.

Burnette, J. L., O'Boyle, E.H., VanEpps, E.M., Pollack, J.M. & Finkel, E.J. (2013). Mind-sets matter: a meta-analytic review of implicit theories and self-regulation. *Psychological Bulletin, 139*(3), 655–701.

Busch, B. (2018). *Research every teacher should know: growth mindset.* Retrieved from http://www.theguardian.com/teacher-network/2018/jan/04/research-every-teacher on the 06.04.2020.

Büyüköztürk, Ş. (2015). Sosyal bilimler için veri analizi el kitabı. [Data analysis for social sciences hand book] Ankara: Pegem Publications.

Claro, S., Paunesku, D. & Dweck, C. S. (2016). Growth mindset tempers the effects of poverty on academic achievement. *Proceedings of the National Academy of Sciences, 113*(31), 8664-8668.

Comrey, A. L & Lee, H. L. (1992). *A first course in factor analysis*. New Jersey: Erlbaum, Hillsdale.

Çankaya, İ. & Demirtaş, Z. (2010). Öğretmen adaylarının görüşlerine göre üniversite iklimi ve atalet arasındaki ilişki [The relationship between university climate and ınertia according to preservice teachers' view]. *Pamukkale University Journal of Education, 28*(28), 1-9.

Çankaya, İ. H. (2010). İlköğretim okul yöneticilerinin vicdan odaklı yaklaşım düzeyleri ile atalet algıları arasındaki ilişki [The relationship between elementary school managers' conscience-oriented approach levels and ınertia perceptions]. *Ahi Evran University Journal of Education, 11*(2), 65-74.

Çebi, M., Yamak, B., & Öztürk, M. (2016). 11-14 yaş çağındaki çocukların spor yapma alışkanlığının duygu kontrolü üzerine etkilerinin incelenmesi. *Amasya University Journal of Education, 5*(2), 468-482.

Doğan, N., Soysal, S. & Karaman H. (2017). Aynı örnekleme açımlayıcı ve doğrulayıcı faktör analizi uygulanabilir mi? Küreselleşen Dünyada Eğitim [Can exploratory and confirmatory factor analysis be applied to the same sampling? Education in a Globalizing World*]*. (Demirel, Ö. ve Dinçer, S., Ed.). Ankara: Pegem Publications.

Duckworth, A. L. & Quinn, P.D. (2009). Development and validation of the short Grit Scale (Grist-S). *Journal of Personality Assessment, 91*(2), 166-174.

Dweck, C. (2008). Mindsets and math/science achievement. *Paper Prepared for the Carnegie-IAS Commission on Mathematics and Science Education.* Retrieved from http://www. opportunityequation.org/resources/commissioned-papers/dweck on the 06.04.2020

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality, and development.* Philadelphia: Psychology Press.

Dweck, C. S. (2006). *Mindset: The new psychology of success.* NY: Random House Publishing Group, New York.

Dweck, C. S., Walton, G. M. & Cohen, G. L. (2014). *Academic Tenacity: Mindsets and Skills that Promote Long-Term Learning*. Bill & Melinda Gates Foundation.

Dweck, C. S. (2016). *Mindset: the new psychology of success*. Ballantine Books.

Evans, J. D. (1996). *Straightforward statistics for the behavioral sciences*. Thomson Brooks/Cole Publishing Co, California Üniversitesi.

Farrington, C. A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T.S., Johnson, D.W. & Beechum, N.O. (2012). *Teaching adolescents to become learners. The role of noncognitive factors in shaping school performance: A critical literatüre review.* University of Chicago Consortium on Chicago School Research.

Fensterwald, J. (2015, November 23). *There's more to a growth mindset than assuming you have it.* Retrieved from http://edsource.org/2015/theres-more-to-a-growth-mindset-thanassuming-you-have-it/90780 on the 06.04.2020.

Fitzgerald, C. J., & Laurian-Fitzgerald, S. M. (2016). Helping students enhance their grit and growth mindsets. *Journal Plus Education / Educatia Plus, 14*(2), 52-56.

French, R. P. (2016). The fuzziness of mindsets. *International Journal of Organizational Analysis*, 24(4), 673-691. https://doi.org/10.1108/ijoa-09-2014-0797

Garcia, E. (2014*). The need to address noncognitive skills in the educational policy agenda. Economic Policy Institute*. Retrieved from https://www.epi.org/publication/the-need-to-address-noncognitive-skills-in-the-education-policy-agenda/ on the 06.04.2020.

Güven, Z., & Yılmaz, E. (2017). Üst düzey öğrenme (meta öğrenme) becerileri [High-level learning (meta learning) skills]. In E. Yılmaz & H. Yılmaz (Ed.), *Eğitimden Damlalar [Drops from Education]* içinde (s. 105-107). Palet Publications.

Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit ındexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55.

Ingebrigtsen, M. (2018). *How to measure a growth mindset. a validation study of the ımplicit theories of ıntelligence scale and a novel norwegian measure.* Master's thesis, UiT Arktiske Universitet, Norges. Retrieved from https://munin.uit.no/bitstream/handle/10037/12904/thesis.pdf?sequence=2&isAllow d=y

Işık, Ş. (2016). Psikolojik dayanıklılık ölçeği'nin geliştirilmesi: geçerlik ve güvenirlik çalışması. *The Journal of Happiness & Well-Being, 4*(2), 165-182.

Kalaycı, Ş. (2014). *Spss uygulamalı çok değişkenli istatistik teknikleri (6.baskı).* Asil.

Kline, R. B. (2011). *Principles and practice of structural equation modeling.* Guilford Publications.

Kline, R. B., (2005). *Principles and practice of structural equation modeling.* Guilford Press, New York.

Laursen, E. K. (2015). The power of grit, perseverance, and tenacity. *Reclaiming Children and Youth, 23*(4), 19.

Lindgren, K. P., Burnette, J.L., Hoyt, C.L., Peterson, K.P. & Neighbors, C. (2019). Growth mindsets of alcoholism buffer against deleterious effects of drinking identity on problem drinking over time. *Alcoholism: Clinical And Experimental Research.*

Lottero-Perdue, P. S., & Lachapelle, C. P. (2019). Instruments to measure elementary student mindsets about smartness and failure in general and with respect to engineering. *International Journal of Education in Mathematics, Science and Technology, 7*(2), 197-214.

Lüftenegger, M. & Chen, J. A. (2017). Conceptual issues and assessment of implicit theories. *Zeitschrift fur Psychologie, 225*(2), 99-106. https://doi.org/10.1027/21512604/a000286

Mather, M., Cacioppo, J. T. & Kanwisher, N. (2013). How fMRI can inform cognitive theories. *Perspectives on Psychological Science, 8*(1), 108-113. https://doi.org/10.1177%2F1745691612469037

Meydan, C. H. & Sesen, H. (2011). *Structural equation modeling AMOS applications.* Detay Yayıncılık.

Midkiff, B., Langer, M., Demetriou, C., & Panter, A. T. (2017). An IRT analysis of the growth mindset scale. *In The Annual Meeting of the Psychometric Society,* 163-174, Springer, Cham.

Morgan, G. A., Leech, N. L. Gloeckner, G.W., & Barrett, K. C. (2004). *SPSS for introductory statistics: Use and interpretation.* Psychology Press.

Ocakçı, A. F. (2015). Ergenlik dönemi özellikleri [Puberty characteristics]. In Ocakçı, A. F. & Üstüner Top, F. (Ed.), *Yetişkinliğe bir adım kala ergen sağlığı [Adolescent health one step away to adulthood]* (ss. 1-7).

Ormrod, J. E. (2017). *How we think and learn: Theoretical perspectives and practical implications.* Cambridge University Press.

Pett, M. A., Lackey, N. R. & Sullivan, J. J. (2003). *Making Sense of Factor Analysis. The Use of Factor Analysis for Instrument Development in Health Care Research.* Sage Publications

Pfeiffer. S., & Jarosewich, T. (2003). *Gifted rating scales.* Pearson.

Ricci, M. C., & Lee, M. (2016). *Mindsets for parents: Strategies to encourage growth mindsets in kids.* Sourcebooks, Inc.

Ricci, M. C. (2013a). *Mindsets in the classroom: Building a growth mindset learning community.* Sourcebooks, Inc.

Ricci, M. C. (2013b). *Mindsets in the classroom.* Prufrock Press.

Rustin, S. (2016). *New test for 'growth mindset', the theory that anyone who tries can succeed.* Retrieved from http://www.theguardian.com/education/2016/may/10/growth-mindset-research-ukschools-sats o the 06.04.2020.

Sarıçam, H., Çelik, İ., & Oğuz, A. (2016). Kısa azim (sebat) ölçeğinin Türkçeye uyarlanması- geçerlik ve güvenirlik çalışması [Turkish adaptation of the short grit scale (Grit-s): Validity and reliability study]. *International Journal of Turkish Literature Culture Education, 5*(2), 927-935. http://dx.doi.org/10.7884/teke.622

Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling*. Psychology Press.

Sekman, M. (2007). Kişisel ataleti yenmek [Overcoming personal inertia]. Alfa Publications.

Sternberg, R. J., Zhang, L. F. & Rayner, S. (Eds.). (2011). *Handbook of intellectual styles: Preferences in cognition, learning, and thinking.* Springer Publishing Company.

Szpirglas, J. & Saint-Onge, D. (2018). *Train your brain: How your brain learns best.* Crabtree Publishing Company.

Şeker, H., & Gençdoğan, B. (2014). *Psikolojide ve eğitimde ölçme aracı geliştirme. (2. Basım).* Nobel Yayınevi.

Tabachnick, B. G. & Fidell, L. S. (2007). *Using multivariate statistics*. Boston: Allyn and Bacon.

Tavşancıl, E. (2005). *Tutumların ölçülmesi ve spss ile veri analizi [Measuring attitudes and data analysis with SPSS]*. Nobel Publications.

Tezbaşaran, A. A. (1996). *Likert tipi ölçek geliştirme kılavuzu [Likert type scale developing guide]*. Ankara: Publications of the Association of Psychologists.

Törüner, E. & Büyükgönenç, L. (2011). *Çocuk sağlığı hemşireliği: çocuk sağlığı temel hemşirelik yaklaşımları [Child health nursing: basic nursing approaches to child health]*. 172-182, 1. Edition, Ankara: Gökçe Offset.

Turkish Language Society (TDK). (2020). Türkçe sözlük [Turkish dictionary]. Retrieved from https://sozluk.gov.tr/ on the 06.04.2020.

Yeager, D. S., Johnson, R., Spitzer, B. J., Trzesniewski, K. H., Powers, J. & Dweck, C. S. (2014). The far-reaching effects of believing people can change: Implicit theories of personality shape stress, health, and achievement during adolescence. *Journal Of Personality And Social Psychology*, *106*(6), 867.

Yılmaz. E. (2019). *Öğrenmede domino etkisi.* Palet Yayıncılık..

RESEARCH ON EDUCATION AND PSYCHOLOGY (REP)

Received: October 24, 2021
Accepted: February 18, 2022
http://dergipark.org.tr/rep
e-ISSN: 2602-3733

April 2022 ◆ 6(Special Issue) ◆ 27-37
Research Article
Doi: 10.54535/rep.1014296

Turkish Adaptation of a Scale to Measure Three Modes of Motivational Regulation Strategies: Self-, Co-, and Socially Shared Regulation of Motivation for Collaborative Activity

Nilüfer Atman Uslu[1]
Manisa Celal Bayar University

Hatice Yıldız Durak[2]
Bartın University

Abstract

In this study, the scale of "Assess Self-regulation, Co-regulation, and Socially Shared Regulation of Intrinsic Motivation for Collaborative Activity (SCSRM)" developed by Ito and Umemoto (2021) was adapted into Turkish, and it was aimed to perform the validity and reliability studies of the scale. The original scale consists of 7 Likert and 15 items. The purpose of the scale was to identify experiences related to group activities in school tasks and to examine how intrinsic motivation affects the three modes of regulation. This research was carried out on 215 university students, who had completed group activities and learning tasks, using a purposive and convenient sampling method. The scale adaptation stages were followed in the study. At the last stage, the validity and reliability of the scale were calculated. In the analysis of the data collected in the study, various analyzes were used for the validity and reliability studies of the scale. As a result of the research, a valid and reliable scale that can be used to determine experiences related to group activities and to examine how intrinsic motivation affects the three regulation modes has been brought to the literature.

Key Words

Collaborative learning • Regulation • Intrinsic motivation • Motivational regulation strategies • University students

[1]Manisa Celal Bayar University, Faculty of Education, Manisa, Turkey. E-mail: atmanuslu@gmail.com **ORCID:** 0000-0003-2322-4210
[2]**Correspondence to:** Bartın University, Faculty of Science, Bartın, Turkey. E-mail: hatyil05@gmail.com **ORCID:** 0000-0002-5689-1805

Citation: Atman Uslu, N. & Yıldız Durak, H. (2022). Turkish adaptation of a scale to measure three modes of motivational regulation strategies: Self-, Co-, and Socially Shared Regulation of Motivation for collaborative activity. *Research on Education and Psychology (REP), 6*(Special Issue), 27-37.

Collaborative learning involves the processes in which they work together to solve a problem, complete a task, or create a product (Laal & Ghodsi, 2012). Collaborative learning is seen as a source of cognitive development and as one of the foundations of learning (Stahl & Hakkarainen, 2021). Collaborative learning includes students' ability to share the responsibility of being the constructor of active and critical learning processes (Winne et al., 2010). It is important to understand regulation in collaborative learning to enable participation in social learning contexts (Volet & Summers, 2013). Self-regulated learning (SRL) is a process in which learners set their own learning goals, choose the appropriate strategy to achieve these goals, use these strategies, and make sense of their learning process (Schunk & Zimmerman 2008; Zimmerman, 2011). SLR includes cognitive, behavioral, and especially motivational processes in students' learning processes (Pintrich, 2000). In these processes, regulation of motivation is an important component (Boekaerts & Cascallar, 2006). Motivation is critical in biological, cognitive and social regulation (Ryan & Deci, 2000) and also plays an important role in collaborative learning contexts where social interaction is central. (Serrano-Cámara et al., 2014).

According to Järvelä and Hadwin (2013), regulation in collaborative learning takes place in three modes: (a) self-regulation (SR), (b) co-regulation (CoR), (c) socially shared regulation (SSR). Collaborative learning includes a process in which each group member organizes his/her own learning (SR), other members' learning (CoR), and also where all members collectively organize their learning (SSR) (Zheng, 2017). SR is defined as a process in which students set goals, monitor, and evaluate their cognitions, emotions, and behaviors (Pintrich, 2000). According to Usher and Schunk (2018), SR is the process of organizing an individual's thoughts, feelings, and actions to achieve their goals. Successful collaboration requires students' SR, both individually and, provides a rich context for learning (Lai, 2021). At the same time, SR is an important cornerstone for exploring more forms of social regulation such as CoR and SSR (Hadwin et al., 2018). CoR learning focuses on the mentoring relationship between the individual and a student, and the importance of giving and receiving support in peer interactions (Ito & Umometo, 2021; McCaslin 2009). CoR highlights the social interactions that occur between two or more group members (Zheng & Yu, 2016). SSR refers to the processes by which group members regulate their collective activities, and this type of regulation refers to regulatory processes, beliefs, and knowledge (e.g., strategies, monitoring) linked to a co-created or shared process or shared collectively. (Hadwin et al., 2011; Järvelä et al., 2015). SSR is important in terms of contributing to productive collaborative learning (Järvelä et al., 2019). Also, it is a strategic activity that involves more active participation than task-oriented interaction in general, occurs during high engagement and plays a role in engagement dynamics (Isohätälä et al., 2017). It is critical for students make consistent efforts to regulate their learning and participation in the emergence of SSR in collaborative work. (Järvelä & Järvenoja, 2011).

Emotion and motivation affect regulation processes in complex ways, and more research is needed on how this effect happens. (Järvenoja et al., 2020). In order to design effective practices that will support motivation and collaboration, scientific evidence is needed in this regard, and this will prevent problems in commitment to the task (Tateno et al., 2016). However, the number of studies focusing on self-regulation (SR), co-regulation (CoR), and socially shared regulation (SSR) of motivation are limited in the literature (Ito & Umemoto, 2021). This study aims to introduce a data collection tool to Turkish literature to understand motivational regulation and antecedent factors for university students.

Purpose of the Study

The aim of the current research is to adapt the "SCSRM" scale into Turkish and to conduct validity and reliability studies of the scale. This scale is based on three modes of motivational regulation strategies (SR, CoR, and SSR of intrinsic motivation) and can be used for collaborative group processes. Especially the Covid-19 epidemic process has highlighted the importance of new approaches to increase the online education process and the effectiveness of this education. Collaborative activity and motivational regulation strategies are critical concepts in this context. However, there are not enough studies in the literature on this subject. In addition, it is thought that studies on this subject will make important contributions in the context of learning in higher education. In this context, the following research question has been considered in the context of the research purpose.

How is the validity and reliability of the "SCSRM Scale" adapted into Turkish?

Method

Research Design

This research, which was carried out to determine three basic motivational regulation strategies of university students in collaborative group processes, is scale adaptation research.

Participants

The participants of this research consist of 215 university students who have experienced collaborative group processes in various state universities in Turkey and studying in different classes online. The age range of the participants is between 18-24. 62.8% of the students are female and 37.2% are male. 18.1% of the university students participating in the research are freshman, 47.4% are sophomore, 20% are junior and 14.5% are seniors.

Research Instruments and Processes

SCSRM Scale: This scale was originally developed by Ito and Umemoto (2021). The adaptation of this scale to Turkish was carried out in the context of this study. This scale, which aims to determine the scope of internal motivation regulation in collaborative activities with the self-reported method, was developed for university students and adult learners. The original scale consists of 3 sub-dimensions and 15 items. There are five items on the scale for each of the three modes of intrinsic motivational regulation strategies. These items are in a 7-point Likert structure ranging from 1 (not at all true for me) to 7 (very true for me). As a result of the validity and reliability studies conducted by Ito and Umemoto (2021), it was concluded that the original scale is a valid and reliable data collection tool that can be used in the field. Within the scope of this study, the results regarding the adaptation process of the scale are presented in the findings section. The Cronbach's alpha coefficient of the scale calculated within the scope of this study is 0.905.

In the scale adaptation process, firstly, permission was obtained from the authors who developed the scale via e-mail. Later, the items of the "SCSRM Scale", which was translated into Turkish, were conveyed to the language experts. It was presented to two different language experts and their opinions and suggestions were received about the translation of each article. Adjustments were made according to the suggestions.

Data Analysis

In this study, data were collected through online and face-to-face interviews, by giving information about the study, and through online and printed forms. The data collected in printed form was then transferred to digital media. 246 university students participated in the research. After the data were cleared from extreme values, 215 cases were included in the analysis.

For the adaptation study of the scale, construct validity and item analyzes were analyzed. The data collected at this stage of the study were only used for scale adaptation analyses. LISREL 8.72 was used to analyze the data. The construct validity of the SCSRM scale was examined within the scope of validity studies. Construct validity, exploratory and confirmatory factor analyzes were performed. $\chi 2/df$ ratio in evaluating the fit of the model tested in confirmatory factor analysis, Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean Square Residual (SRMR), Normed Fit Index (NFI), Non-Normed Fit Index (NNFI), Comparative Fit Index (CFI), Incremental Fit Index (IFI) values were examined. NFI, NNFI, CFI, and IFI values above 0.95 in the literature indicate a perfect fit (Hu & Bentler, 1999; Sümer, 2000). It is recommended that RMSEA and SRMR values should be less than 0.08 (Şimşek, 2007). A $\chi 2/df$ ratio below three indicates a perfect fit (Kline, 2011).

Cronbach's alpha internal consistency coefficient values were calculated for the reliability studies of the scale. In addition, item-total correlations were examined, and independent samples t-test analysis was performed to determine whether there was a significant difference between the scores of the upper 27% group and the lower 27% group according to the total score obtained from the scale.

Results

Descriptive Findings

The mean, standard deviation, skewness, and kurtosis values of the items in the scale adapted to Turkish are given in Table 1.

Table 1

Mean, Standard Deviation, Skewness, and Kurtosis Values of Scale Items

	Mean	**Standard Deviation**	**Skewness**	**Kurtosis**
Item1	5.3994	1.32209	-0.538	-0.177
Item 2	5.4076	1.23798	-0.463	0.145
Item 3	5.4787	1.27445	-0.703	0.378
Item 4	5.5723	1.24596	-0.685	0.301
Item 5	5.8007	1.14375	-0.840	0.282
Item 6	5.3233	1.31447	-0.594	0.047

Item 7	5.3450	1.18180	-0.172	-0.888
Item 8	5.4822	1.20304	-0.387	-0.677
Item 9	5.5469	1.07429	-0.373	-0.568
Item 10	5.3957	1.27838	-0.577	-0.132
Item 11	5.2344	1.42114	-0.708	0.092
Item 12	5.4127	1.21925	-0.512	0.088
Item 13	5.3784	1.22550	-0.452	-0.378
Item 14	5.5364	1.08783	-0.368	-0.399
Item 15	5.3782	1.30316	-0.452	-0.379

According to Table 1, the average scores of the items between 5.2344 and 5.8007, and their standard deviations between 1.07429 and 1.32209. The skewness and kurtosis values were found to be between +1 and -1 values. These findings regarding skewness and kurtosis show that the scores obtained from the items are in a normal distribution (Kline, 2011).

Confirmatory Factor Analysis

Confirmatory factor analysis was applied for the model consisting of three factors and 15 items for the factorial validity of the SCSRM scale. As a result of the analysis, the fit indices were found as [$\chi 2(87, N=215)=621.40$, RMSEA= 0.169, SRMR= 0.069, NFI= 0.91, NNFI=0.91, CFI=0.93, IFI=0.93]. Some of the items tested by confirmatory factor analysis were excluded because the fit indices were not in the recommended range and the factor loadings estimated above one for some items might be associated with multicollinearity (Item 1, Item 4, Item 7, Item 9, Item 12, Item 14). The values obtained after removing these items are as follows: [$\chi 2(24, N=215)=70.12$, RMSEA= 0.095, SRMR= 0.036, NFI= 0.96, NNFI=0.96, CFI=0.98, IFI=0.98]. These values indicate that the model has an acceptable and/or perfect fit. As a result of confirmatory factor analysis, standardized factor loads, and item structure parameters are presented in Figure 1.

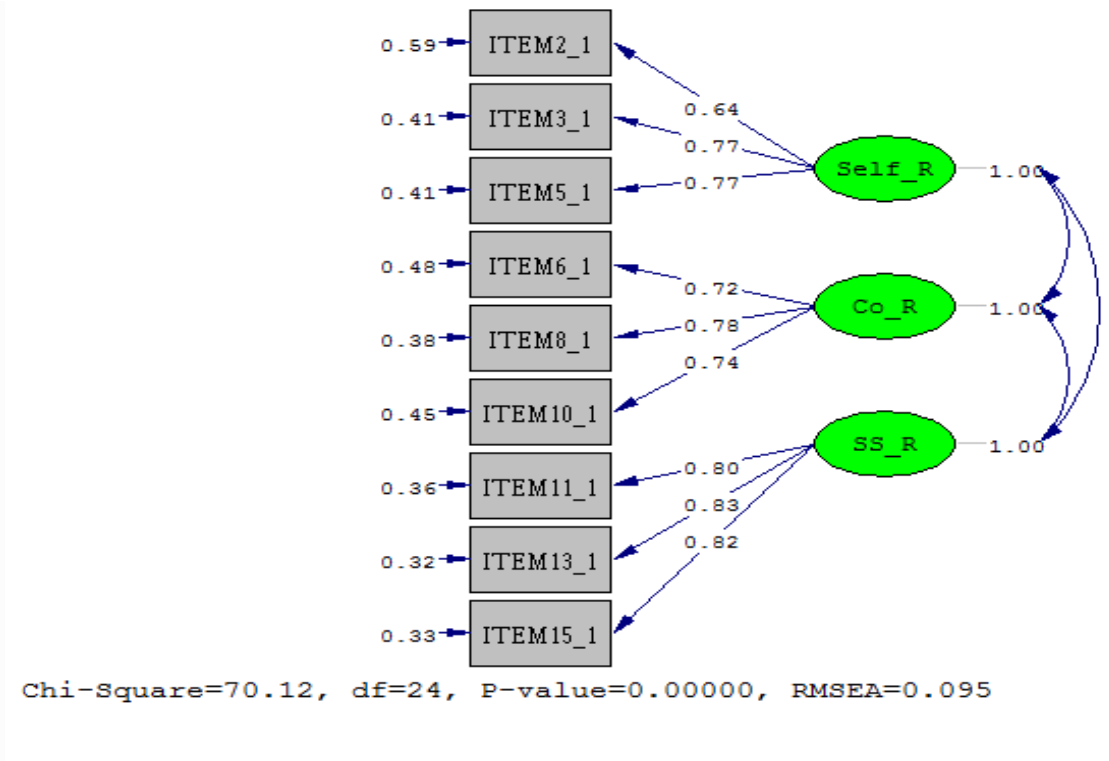


Figure 1. CFA Results

According to Figure 1, factor loadings are between 0.64 and 0.83 and it is statistically significant according to the t-test findings. These findings provide sufficient evidence for factorial validity.

Reliability

The reliability of the measurement tool in terms of internal consistency was tested with the Cronbach alpha coefficient. The Cronbach alpha internal consistency coefficient of nine items in the scale was calculated as 0.905. The Cronbach alpha internal consistency coefficient for the SR factor was 0.768, 0.792 for CoR, and 0.853 for SSR. The fact that these values are higher than 0.70 can be stated to provide evidence of reliability (Hair et al., 1998).

Item-total correlations, which are used to mean the relationship between the score for each item and the total score from the scale, were calculated. In addition, independent samples t-test analysis was performed to examine whether there was a statistically significant difference between the scores of the group in the upper 27% group (N=116) and the group in the lower 27% group, according to the total score obtained from the scale. Item analysis t-values and item-total score correlations between the lower and upper groups are presented in Table 2.

Table 2

Item-total correlations

Factor	Item	Lower 27%-Upper 27% t-value	(rjx)
Self-regulation	Item 2	10.953	0.514
	Item 3	14.666	0.633
	Item 5	12.635	0.630
Co-regulation	Item 6	12.945	0.682
	Item 8	16.572	0.745
	Item 10	14.317	0.701
Socially shared regulation	Item 11	15.210	0.722
	Item 13	14.655	0.749
	Item 15	16.197	0.745

When Table 3 is examined, it is seen that there is a significant difference between the independent samples between the upper 27% and the lower 27% groups according to the t values. Item-total correlation values were found to be between 0.514 and 0.749. In the literature, it is stated that 0.30 and higher item-total correlations for items have distinctiveness in terms of measured characteristics (Büyüköztürk, 2004). In this case, it is seen that the total score correlations of the items are sufficient.

Discussion, Conclusion and Suggestions

The aim of this research is to adapt the SCSRM Scale into Turkish and to carry out validity and reliability studies of the scale. The aim of the developed scale is to measure three basic motivational strategies of university students who are in collaborative group processes. In the context of this study, "SCSRM" developed by Ito and Umemoto (2021) has been adapted into Turkish and sufficient evidence has been obtained regarding the validity and reliability of the adapted scale.

Opinions of two language experts were sought to translate the scale into Turkish and bring it into an understandable form. Then, field experts were consulted for the items of the scale. Confirmatory factor analysis of the scale was performed to gain evidence for construct validity. For reliability, Cronbach's Alpha internal consistency coefficients, item-total correlation, and 27% lower-upper group discrimination were checked. As a result of the adaptation study of the scale, some items (Item 2, Item 4, Item 7, Item 9, Item 12, Item 14) were excluded from the scale because the fit indices were not in the recommended range and the factor loadings estimated above one for

some items might be related to the multicollinearity situation. As a result of the adaptation study, a 3-dimensional scale with 9 items and 5-point Likert Type was added to the literature. Scale sub-dimensions are as follows: SR, CoR, and SSR of Intrinsic Motivation.

The scale in this study can be used in different studies to measure the modes of three basic motivational strategies of university students in collaborative group processes. The scale adapted in this study can also be used to measure three modes of motivation in technology-enhanced colaborative learning experiences. Effective use of various digital tools is necessary to increase student performance and support participation in technology-rich environments (Saritepeci & Durak, 2016; Saritepeci & Yildiz, 2014). According to Jeong et al. (2019), supporting collaborative knowledge building and problem-solving by digital technologies improves collaborative learning. Therefore, studies in which motivation was examined as a dependent variable were conducted in the literature on computer supported collaborative learning. (e.g. Järvelä et al., 2008). Therefore, research with the CSCL, which examines three modes of motivation, could benefit from this scale, for which evidence of validity and reliability was presented in this study.

Ethic

In this study, all scientific ethical rules were followed.

Author Contributions

All stages of the study were organized and conducted by the authors.

Conflict of Interest

In addition, the authors declare that they have no conflict of interest.

Funding

This work has not received funding from any institution.

Acknowledge

This study is an expanded version of the oral presentation presented at the ERTE Congress.

References

Boekaerts, M., & Cascallar, E. (2006). How far have we moved toward the integration of theory and practice in self-regulation? *Educational Psychology Review, 18*, 199–210.

Büyüköztürk, Ş. (2004). *Veri analizi el kitabı* [Data analysis handbook]. Pegem A Yayıncılık, Ankara.

Hadwin, A. F., Järvelä, S., & Miller, M. (2011). Self-regulated, co-regulated, and socially shared regulation of learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 65–84). New York: Routledge.

Hadwin, A., Järvelä, S., & Miller, M. (2018). Self-regulation, co-regulation, and shared regulation in collaborative learning environments. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (pp. 83–106). Routledge/Taylor & Francis Group.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5th ed.). Upper Saddle River, NJ: Prentice-Hall (Upper Saddle River, NJ).

Hu L. T., & Bentler P. M. (1999). Cut off criteria for fit ındexes in covariance structure analysis: conventional criteria versus new alternatives. Structural Equation Modeling: *A Multidisciplinary Journal, 6*(1), 1-55

Ito, T., & Umemoto, T. (2021). Self- regulation, co- regulation, and socially shared regulation of motivation for collaborative activity: comparison between university students and working adults 1. *Japanese Psychological Research*.

Isohätälä, J., Järvenoja, H., & Järvelä, S. (2017). Socially shared regulation of learning and participation in social interaction in collaborative learning. *International Journal of Educational Research, 81*, 11-24.

Järvelä, S., Järvenoja, H., & Veermans, M. (2008). Understanding the dynamics of motivation in socially shared learning. *International Journal of Educational Research, 47*(2), 122-135.

Järvelä, S., & Järvenoja, H. (2011). Socially constructed self-regulated learning and motivation regulation in collaborative learning groups. *Teachers College Record, 113*(2), 350-374.

Järvelä, S., & Hadwin, A. F. (2013). New frontiers: Regulating learning in CSCL. *Educational Psychologist, 48*(1), 25-39.

Järvelä, S., Kirschner, P. A., Panadero, E., Malmberg, J., Phielix, C., Jaspers, J., ... & Järvenoja, H. (2015). Enhancing socially shared regulation in collaborative learning groups: Designing for CSCL regulation tools. *Educational Technology Research and Development, 63*(1), 125-142.

Järvelä, S., Järvenoja, H., & Malmberg, J. (2019). Capturing the dynamic and cyclical nature of regulation: Methodological Progress in understanding socially shared regulation in learning. *International Journal of Computer-Supported Collaborative Learning, 14*(4), 425-441.

Järvenoja, H., Järvelä, S., & Malmberg, J. (2020). Supporting groups' emotion and motivation regulation during collaborative learning. *Learning and Instruction, 70*, 101090.

Jeong, H., Hmelo-Silver, C. E., & Jo, K. (2019). Ten years of computer-supported collaborative learning: A meta-analysis of CSCL in STEM education during 2005–2014. *Educational research review*, *28*, 100284.

Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling* (3rd ed.). New York: The Guilford Press.

Laal, M., & Ghodsi, S. M. (2012). Benefits of collaborative learning. *Procedia-social and behavioral sciences*, *31*, 486-490.

Lai, C. L. (2021). Effects of the group- regulation promotion approach on students' individual and collaborative learning performance, perceptions of regulation and regulation behaviours in project- based tasks. *British Journal of Educational Technology*, *52*(6), 2278-2298.

McCaslin, M. (2009). Co-regulation of student motivation and emergent identity. *Educational Psychologist, 44*(2), 137–146.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.). *Handbook of self-regulation* (pp. 452–502). Cambridge, MA: Academic Press.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68.

Sarıtepeci, M., & Durak, H. (2016). Bilgi teknolojilerinin temelleri ünitesinin işlenmesinde dijital hikâye anlatımı kullanımının öğrenen motivasyonuna etkisi [Digital storytelling effect on students' motivation in subject of information technologies fundamentals]. *Journal of Research in Education and Teaching*, *5*, 258-265.

Sarıtepeci, M., & Yıldız, H. (2014). Harmanlanmış öğrenme ortamlarının öğrencilerin derse katılım ve derse karşı motivasyonları üzerine etkisinin incelenmesi [The effect of blended learning environments on students' engagement to course and motivation toward the course]. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi* [Ahi Evran University Journal of Kırşehir Education Faculty], *15*(1), 207-223.

Schunk, D. H., & Zimmerman, B. J. (Eds.). (2008). *Motivation and self-regulated learning: Theory, research, and applications*. New York: Lawrence Erlbaum Associates.

Serrano-Cámara, L. M., Paredes-Velasco, M., Alcover, C. M., & Velazquez-Iturbide, J. Á. (2014). An evaluation of students' motivation in computer-supported collaborative learning of programming concepts. *Computers in Human Behavior, 31*, 499-508.

Stahl, G., & Hakkarainen, K. (2021). Theories of CSCL. In *International handbook of computer-supported collaborative learning* (pp. 23-43). Springer, Cham.

Sümer, N. (2000). Yapısal eşitlik modelleri: Temel kavramlar ve örnek uygulamalar [Structural equation models: Basic concepts and sample applications]. *Türk Psikoloji Yazıları* [Turkish Psychology Articles], *3*(6), 49-73.

Şimşek, Ö. F. (2007). *Yapısal eşitlik modellemesine giriş; temel ilkeler ve LİSREL uygulamaları* [Introduction to structural equation modeling; basic principles and LISREL applications*]*. Ankara: Ekinoks Yayıncılık.

Tateno, Y., Nakahara, J., Kimura, M., Yasuda, E., Yoshimura, H., Tanaka, S., … Mizokami, S. (2016). Effect of undergraduate students' learning behavior and fulfillment on proactive behavior at work. *Japan Journal of Educational Technology*, 40, 1–11.

Usher, E. L., & Schunk, D. H. (2018). Social cognitive theoretical perspective of self-regulation. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (pp. 19–35). Routledge/Taylor & Francis Group.

Volet, S., & Summers, M. (2013). Interpersonal regulation in collaborative learning activities: Reflections on emerging research methodologies. In S. Volet, & M. Vauras (Eds.). *Interpersonal regulation of learning and motivation: Methodological advances* (pp. 204–220). Abingdon: Routledge

Winne, P. H., Hadwin, A. F., & Gress, C. (2010). The learning kit project: Software tools for supporting and researching regulation of collaborative learning. *Computers in Human Behavior*, *26*(5), 787-793.

Zheng, L. (2017). Analysis of co-regulation behavioral patterns by cluster and sequential analysis in CSCL. In *Knowledge building and regulation in computer-supported collaborative learning* (pp. 49-63). Springer, Singapore.

Zheng, L., & Yu, J. (2016). Exploring the behavioral patterns of co-regulation in mobile computer-supported collaborative learning. *Smart Learning Environments*, *3*(1), 1–20.

Zimmerman, B. (2011). *Handbook of self-regulation of learning and performance*. Routledge.

RESEARCH ON EDUCATION AND PSYCHOLOGY (REP)

Received: November 29, 2021
Accepted: March 1, 2022
http://dergipark.org.tr/rep

e-ISSN:2602-3733

April 2022 ◆ 6(Special Issue) ◆ 38-46

Research Article

Doi: 10.54535/rep.1029896

The Validity and Reliability of the Group Regulation Scale Turkish Form: A Study with the Rasch Model

Hatice Yıldız Durak [1]
Bartin University

Beyza Aksu Dünya[2]
Bartın University

Abstract

In this study, the group regulation scale was adapted to Turkish and validity was checked using Rasch Model. The original scale was created by adapting the form developed by Papamitsiou and Economides (2019) by Lai (2021). The scale used in this study was adapted into Turkish by the researchers with permission from Lai (2021). The original scale consists of 12 items scored on a five-point Likert type scale and 4 sub-dimensions (effort regulation, goal expectancy, help seeking and time management). This research was carried out on 170 university students based on voluntary participation using purposive sampling method. Rasch analysis was used for analyzing data in order to examine validity and reliability of the scores. According to the Rasch analysis results, it was concluded that the group regulation scale is a unidimensional measure of group regulation among university students.

Key Words

Group regulation • Rasch model • Regulation • University students

[1] **Correspondence to:** Bartın University, Faculty of Science, Department of Computer Technology and Information Systems, Bartın, Turkey, hatyil05@gmail.com **ORCID:** 0000-0002-5689-1805
[2] Bartın University, Faculty of Education, Educational Measurement and Evaluation, Bartın, Turkey, baksu@bartin.edu.tr **ORCID:** 0000-0003-4994-1429

Citation: Yıldız Durak, H., & Aksu Dünya, B. (2022). The validity and reliability of the group Regulation Scale Turkish Form: A Study with the Rasch Model. *Research on Education and Psychology (REP), 6*(Special Issue), 38-46.

Self-regulation is the use of the power of influence on the motivation, cognitive processes, emotional and behavioral states, and patterns of the individual (Bandura, 1994). The concept of self-regulation is the individual's awareness of his potential and inner abilities about the actions she/he will take (Bandura, 1991). Zimmerman and Schunk (2008) define the concept of self-regulation as the control of motivation-based behavior related to the next goal or ideal that a person sets of the individual.

The concept of self-regulation is grounded in the context of the Social Cognitive Theory (Bandura, 2001). According to the principle of mutual determination in this theory, external and internal factors are in constant interaction. Therefore, the individual regulates the processes of noticing, monitoring, and directing his behavior individually and environmentally. Environmental self-regulation involves regulating behavior in an environmental context. Internal self-regulation is the individual's awareness, monitoring, evaluation, and regulation of self-performance.

From a learning perspective, self-regulation is a motivational process in which students manage their emotions, thoughts, and behaviors in a planned manner in the context of learning goals and make adaptations when needed (Pintrich, 2000; Zimmerman, 2000). According to Zimmerman (1989), self-regulation is a process in which learners manage their own learning. This process consists of three cyclical phases. These phases are pre-thought, performance, and self-reflection. Pre-thought phase, pre-performance preparation; performance phase, operational processes; and self-reflection phase includes post-performance processes.

In learning environments, more important outcomes can be achieved, especially in the interaction of students with their peers. Especially successful cooperative group work requires self-regulation of students, both individually and, this provides a rich context for learning (Lai, 2021). Therefore, for group learning to be successful, students need to regulate their own learning as well as the learning of other members of the group as a group. As a matter of fact, according to Chan (2012), the work of group members is also reflected in their individual performances.

On the other hand, according to Kwon, Liu, and Johnson (2014), a group must coordinate its efforts and resources effectively to achieve common goals. According to Saab (2012), the concept of self and group regulation has common features in terms of task definition, process tracking, strategy development and evaluation. However, according to self-regulation, students who work collaborative with group members need extra group regulation behaviors. In addition, if group regulation behaviors are not coordinated in a planned way, positive results may not be obtained from the group-based learning process.

In order to design effective practices that will support students' collaborative behavior in a learning environment, it is necessary to define group regulation behaviors. As a matter of fact, the recently used learning environments (e.g. flipped learning, blended learning, social networking) allow group work as they offer various interaction opportunities (Durak, 2019; Durak, 2020; Sarıtepeci & Yıldız, 2014). To understand group regulation behaviors of university students, this study focuses on adapting a data collection tool to Turkish, which aims to measure these behaviors.

Aim of the Study

The purpose of this study is to examine if Group Regulation Scale reflects a single dimension of group regulation that can yield a single summary score according to Rasch Model. The dimensional structure of the scale was analyzed using Classical Test Theory (CTT) framework before (Lai, 2021). This study adds on the existing literature by utilizing advantages of Rasch framework which is theoretically more robust and feasible than CTT (Humphrey et al., 2011).

Method

Participants

The research data were obtained from 170 university students studying in different departments. Participants formed in accordance with the convenience sampling (Fraenkel, Wallen & Hyun, 2012) technique. 50 (29.4%) of the participants were female and 120 (70.6%) were male students. The mean age of the participants was 21.31 (SD=4.72). Most of the students were in their first years in university.

Research Instruments and Processes

For the adaptation of the scale, firstly, permission was requested from Chiu-Lin Lai via e-mail from the correspondence address in the scale article. After obtaining permission, two field experts who knew both Turkish and English were determined. One of these experts translated the scale items into Turkish. Other experts translated the scale items back into English. In this process, the meaning and comprehensibility of the items of the scale were examined and the consistency of the three item lists was examined. Necessary arrangements have been made. Then, the scale form was examined by two field experts in terms of meaning and comprehensibility and the final form was created.

Group Regulation Scale: This scale was originally developed by Papamitsiou and Economides (2019). Lai (2021), on the other hand, created an adapted form using this form. In the scope of the research, this scale was adapted to Turkish language. The original scale consists of 12 items and 4 sub-dimensions. The sub-dimensions are as follows: effort regulation, goal expectancy, help seeking and time management. The scale follows a 5-point Likert-type rating scale structure. In the original scale, the Cronbach's alpha values of this dimension were 0.78, 0.87, 0.80 and 0.85, respectively. In the study conducted by Lai (2021), it was concluded that the scale is a data collection tool with a high degree of validity and reliability. Within the scope of this study, the validity and reliability were re-investigated using a different methodology; Rasch Model.

Data Analysis

The analyses were conducted by employing Rasch Rating Scale Model (RSM, Andrich, 1978). The original instruments' dimensional structure was investigated by using Classical Test Theory (CTT) approach and exploratory factor analysis results yielded four factor structure (Lai, 2021). Rasch model is a member of Item Response Theory (IRT) Models that is frequently used to evaluate validity of the instruments by transforming ordinal data to an interval scale. The advantages of IRT models over CTT have repeatedly been cited in instrument development

literature. The most important benefits of IRT approach over CTT can be summarized as follows: (1) CTT uses sum scores which are ordinal in nature and threats them as interval level measures to make inferences. Yet, IRT models transform raw scores to logit scores using a logistic equation to estimate item difficulty and person ability measures. (2) In IRT models, each item and person can be analyzed individually via fit statistics to detect any aberrant person-level response behaviors or misfitting item (Hambleton, Swaminathan and Rogers, 1991). The logistic equation to estimate person and item parameters using Rasch RSM is presented and explained below:

$$ln\left(\frac{P_{nix}}{P_{ni\,x-1}}\right) \equiv B_n - D_i - T_k.$$

In this formula, the probability of selecting category x over selecting category x-1 depends on the threshold parameter *T*k which represents the transition point between two rating scale categories, *Bn*, which represents person's group regulation level and *Di* represents the difficulty of endorsing an item (Andrich, 1978). Parameter estimation for the RSM was conducted on Winsteps software (Linacre, 2021).

Messick's (1995) validity framework was utilized to assess construct validity of the scale. Under the framework content, structural, substantive and generalizability aspects of the validity were examined. The statistical indices that were used to evaluate each aspect of the construct validity using Rasch Model estimates are presented on Table 1.

Table 1

Statistical indices for validity inquiry

Statistical index	Type of validity	Cut score-Decision rule	Interpretation
Item fit	Content	Standardized unweighted mean-squared item fit (MNSQ) indices >2.00	If the indices are greater than 2.00,it may indicate lack of predictability in the responses.
Item fit	Content	Point-measure correlation values for items < .40	If the point measure correlation values are smaller than .40, this may indicate inconsistent scores on particular item with the scores of rest of items
Unidimensionality	Structural	Eigenvalues of the residuals >2.00	If the eigenvalues of the contrasts exceed 2.00, the contrast may imply a dimension in the data

Rating scale category thresholds (*Tk*)	Substantive	Category thresholds should increase at least 1.4 logits	Monotonic increase of category thresholds from one category to another implies consistency between observed responses and theoretical construct being measured.
Person separation reliability	Generalizability	A reliability value >.80	The value indicates how precise and replicable are the measures across different settings and applications within the same population

Content validity refers to the degree of items' representativeness of the construct being measured. For this, two types of item fit indicators were checked: point-measure correlation values and standardized unweighted mean-squared fit values. The point-measure correlation values quantify the relation between responses to a particular item and total score of the respondents on the instrument (Wolfe & Smith, 2007). Standardized unweighted mean-squared fit values display degree of fit between theoretical measurement model and response set to each item. For each item on the instrument, a fit value is calculated and expected to be between 0.6 and 1.4 logits (Linacre, 2002).

Another important validity aspect that was investigated in the study is structural validity which refers to the dimensional structure of the data. Winsteps program provides analysis of residuals via principal component analysis (PCA). We expected to obtain a unidimensional structure to proceed with Rating scale model (Andrich, 1978). The decision criteria was to obtain eigenvalue smaller than 2.00 for assuming unidimensionality (Linacre, 2002). After checking unidimensionality assumption for applying Rating Scale Model, we proceed to the analyses of rating scale categories as an evidence for substantive validity. The functionality of the rating scale for capturing the group regulation behavior supports if the observed responses and theoretical rationale that the instrument was grounded on is consistent (Messick, 1995). For this, we examined if the category thresholds advanced monotonically (at least by 1.4 logits, Linacre, 2002) as the values of rating scale categories increase. Lastly, we examined generalizability aspect of validity by checking the person separation reliability values which indicates if the responses are replicable and consistent across settings.

Results

Summary statistics for 12 items were outlined in Table 2. As seen in the Table, mean item difficulty measure was .00 while mean person measure was 1.81. It is promising to see that the range of item fit statistics ranged within the expected value for this statistics (min=. 82 logits and max=1.39 logits). The mean value of the standardized unweighted mean-squared fit statistics was 1.02. The fit values obtained from the data supported content validity aspect of the instrument. In addition, point measure correlation values were found substantially high, ranging

between .77 and .85, indicating instrument items functioned in accordance with each other and with the instrument as a whole. The item fit and correlation values indicated that content validity of the adapted group regulation scale has been established.

Table 2

Summary Statistics for 12 Items

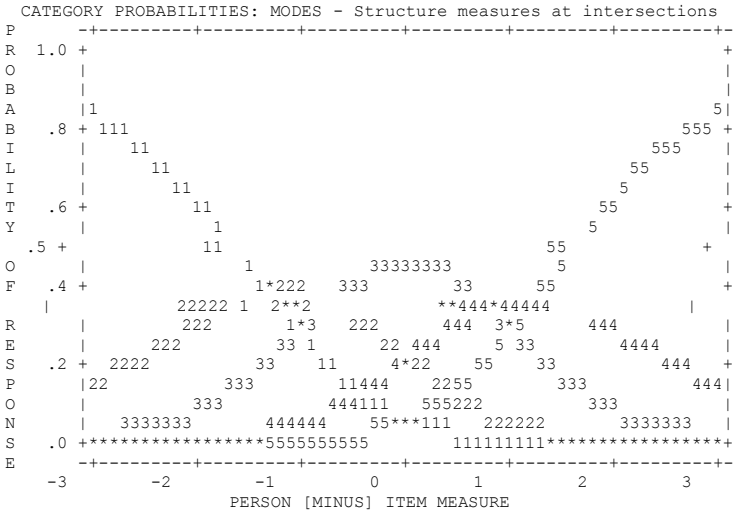
	Logit value	**Point-measure correlation**	**Standardized unweighted mean-squared fit index**
Mean	.00	.72	1.02
Standard Deviation	.44	.12	.16
Maximum value	.92	.85	1.39
Minimum value	-.63	.77	.82

Structural validity evidence was obtained based on PCA of the residuals results. The first contrast had an eigenvalue of 2.00 which means that it does not indicate a secondary dimension. Rasch measures explained 57% of the variance in the data. Overall findings supported that the instrument reflects a single measure of group regulation behavior. The instrument was unidimensional based on Rasch Rating Scale Analysis results.

The degree of effective functioning of rating scale provided evidence for substantive validity. Following Linacre's (2002) guidelines for rating scale effectiveness, category threshold measures were examined. The results suggested that category thresholds which advance monotonically and larger than 1.4 from one category to another. The visual inspection of rating scale category curves on Figure 1 also suggest that respondents were distinguished by the rating scale categories effectively

Figure 1

Category Probability Curves for the Full Scale



As seen on the above figure, each category point has a unique peak although category 4 was utilized as effective as other medium categories.

Lastly, the person separation reliability value was found as .87. This value was above the cut value for the reliability coefficient which indicates the person measures are replicable over different instrument administrations.

Discussion, Conclusion & Suggestions

In this study, it is aimed to adapt to Turkish language the "Group Regulation Scale" developed by Lai (2021). It is thought that the results obtained will contribute significantly to the studies on explaining and improving the group regulation performance of students. In addition, it is expected that studies on this topic will make significant contributions to the improvement of group-based learning activities and will support effective outcomes in higher education.

The psychometric properties of the Group Regulation Scale, which aims to measure group regulation for students for group work, were tested on a Turkish sample. Differently, Rasch model was used to investigate psychometric properties of the adapted scale. Summary of findings from Rasch analysis displayed satisfactory level of reliability and evidence of construct validity.

As a result of this study, it has been determined that the unidimensional and 12-item version of the Group Regulation Scale can be used as a valid and reliable measurement tool in Turkish culture. However, there are some limitations of the study. In this study, it was not examined whether the students participating in the research experienced participating in group work or whether they took an active role in group work. Due to this situation, the effects of experience and active performance on students' self-regulation activities in group work cannot be determined. On the other hand, in studies examining regulation in the literature (e.g. Lin & Tsai, 2016), it is seen that self-regulation behaviors are generally measured. Again, in the literature, beyond self-regulation, regulation behaviors within the group have been discussed in the context of co- and socially shared regulation (e.g. Quackenbush & Bol, 2020; Ito & Umometo, 2021; Hadwin, Järvelä, & Miller, 2011; Uslu, & Yildiz Durak, 2022). This study, unlike previous studies, focused on the term group regulation. In future studies, taking into account the risk factors in the selection of the sample, the students who are active and experienced in group work and those who are not can be compared. In addition, this scale was adapted for university students in the context of group work.

Ethic

In this study, all scientific ethical rules were followed.

Author Contributions

All stages of the study were organized and conducted by the authors.

Conflict of Interest

In addition, the authors declare that they have no conflict of interest.

Funding

This work has not received funding from any institution.

References

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, *43*(4), 561-573.

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50*(2), 248-287.

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behavior* (pp. 71-81). New York: Academic Press.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, *52*, 1-26.

Chan, C. K. (2012). Co-regulation of learning in computer-supported collaborative learning environments: A discussion. *Metacognition and Learning*, *7*(1), 63-73.

Durak, H. Y. (2019). Examining the acceptance and use of online social networks by preservice teachers within the context of unified theory of acceptance and use of technology model. *Journal of Computing in Higher Education*, *31*(1), 173-209.

Durak, H. Y. (2020). Modeling different variables in learning basic concepts of programming in flipped classrooms. *Journal of Educational Computing Research*, *58*(1), 160-199.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). How to design and evaluate research in education (8th ed.). New York: Mc Graw Hill.

Hadwin, A., Järvelä, S., & Miller., M. (2011). Self-regulated, co-regulated, and socially shared regulation of learning." In *Handbook of Self-Regulation of Learning and Performance*, edited by B. Zimmerman, and D. Schunk, 65–84. New York: Routledge.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Sage: California

Humphrey, N., Kalambouka, A., Wigelsworth, M., Lendrum, A., Deighton, J., & Wolpert, M. (2011). Measures of social and emotional skills for children and young people: A systematic review. *Educational and psychological measurement, 71*(4), 617-637.

Ito, T., & Umemoto, T. (2021). Self-regulation, co-regulation, and socially shared regulation of motivation for collaborative activity: Comparison between university students and working adults. *Japanese Psychological Research*. https://doi.org/10.1111/jpr.12337

Kwon, K., Liu, Y. H., & Johnson, L. P. (2014). Group regulation and social-emotional interactions observed in computer supported collaborative learning: Comparison between good vs. poor collaborators. *Computers & Education*, *78*, 185-200.

Lai, C. L. (2021). Effects of the group- regulation promotion approach on students' individual and collaborative learning performance, perceptions of regulation and regulation behaviours in project- based tasks. *British Journal of Educational Technology*. https://doi.org/10.1111/bjet.13138

Lin, J. W., & Tsai, C. W. (2016). The impact of an online project-based learning environment with group awareness support on students with different self-regulation levels: An extended-period experiment. *Computers & Education*, *99*, 28–38.

Linacre, J. M. (2002). *A user's guide to Winsteps: Rasch model computer programs*. Chicago, IL: MESA Press.

Linacre, J. M. (2021). *Winsteps® _Rasch measurement computer program (version 4.7.3)*. Oregon: Beaverton. Winsteps.com.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 5*, 741-749.

Papamitsiou, Z., & Economides, A. A. (2019). Exploring autonomous learning capacity from a self-regulated learning perspective using learning analytics. *British Journal of Educational Technology*, *50*(6), 3138–3155.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). Academic Press. https://doi.org/10.1016/b978-012109890-2/50043-3

Quackenbush, M., & Bol, L. (2020). Teacher support of co-and socially-shared regulation of learning in middle school mathematics classrooms. *Frontiers in Education*, 5, https://doi.org/10.3389/feduc.2020.580543

Saab, N. (2012). Team regulation, regulation of social activities or co-regulation: Different labels for effective regulation of learning in CSCL. *Metacognition and Learning*, *7*(1), 1-6.

Sarıtepeci, M., & Yıldız, H. (2014). Harmanlanmış öğrenme ortamlarının öğrencilerin derse katılım ve derse karşı motivasyonları üzerine etkisinin incelenmesi. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi*, *15*(1), 207-223.

Uslu, N. A., & Yildiz Durak, H. (2022). Predicting learner autonomy status in collaborative learning: The role of group metacognition, self-, co-, and socially shared of regulation of motivation. *Learning and Motivation*, *78*, 101804.

Wolfe, E. W., and Smith, E. V. (2007). Instrument development tools and activities for measure validation using Rasch models: Part I – Instrument development tools. *Journal of Applied Measurement*, *8*, 97-123.

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of educational psychology*, *81*(3), 329-339.

Zimmerman, B. J. (2000). Attaining self-regulation: A social-cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation*, (pp. 13–39). San Diego: Academic.

Zimmerman, B., & Schunk, D. (2001). *Self-regulated learning and academic achievement* (2nd ed.). Mahwah, NJ: Erlbaum.

Zimmerman, B.J., & Schunk, D.H. (2008). Motivation: An essential dimension of selfregulated learning. In D.H. Schunk & B.J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 1-30). Boca Raton, FL: Erlbaum/Tay

RESEARCH ON EDUCATION AND PSYCHOLOGY (REP)

Received: February 28, 2022
Accepted: April 6, 2022
http:// dergipark.org.tr/rep
e-ISSN: 2602-3733

April 2022 ◆ 6(Special Issue) ◆ 47-56
Research Article
Doi: 10.54535/rep.1080132

Adaptation of T-STEM CT Scale to Turkish: Teacher Self-Efficacy and Outcome Expectancy for Teaching Computational Thinking

Mustafa Sarıtepeci[1]
Necmettin Erbakan University

Aykut Durak[2]
Ministry of National Education

Abstract

Computational thinking (CT) skills are accepted as fundamental literacy. Although the idea that K-12 teachers should teach students CT skills in an interdisciplinary context is heavily expressed, there is a need for a measurement tool in Turkish that measures teachers' self-efficacy in this regard. This study aims to adapt the T-STEM CT scale, developed by Boulden et al. (2021), into Turkish and to carry out validity and reliability studies of this scale. The original scale consists of a 5-point Likert scale and 13 items. The participants of this study consisted of 168 teachers from different branches working in K-12 schools. It was carried out by selecting for application purposes and a convenient sampling method. Various validity and reliability methods were used to validate the scale. According to the results, the two-factor (Factor1: T-STEM CT self-efficacy, Factor2: T-STEM CT outcome expectancy) and thirteen-item structure had an acceptable fit with the data. Consequently, the validity and reliability of a Turkish tool measuring teaching efficacy beliefs for computational thinking skills were confirmed.

Key Words

In-service teachers • Scale adaption • Computational thinking • Self-efficacy

[1] **Correspondance to:** Faculty of Ereğli Education, Necmettin Erbakan University, Konya, Turkey. E-mail: mustafasaritepeci@gmail.com **ORCID:** 0000-0002-6984-0652
[2] Ministry of National Education, Bartın, Turkey. E-mail: aykdur78@gmail.com **ORCID:** 0000-0001-7070-9048
Citation: Sarıtepeci, M., & Durak, A. (2022). Adaptation of T-STEM CT Scale to Turkish: Teacher self-efficacy and outcome expectancy for teaching computational thinking. *Research on Education and Psychology (REP)*, *6*(Special Issue), 47-56.

Computational thinking (CT) is one of the fundamental skills that all individuals need to learn and develop today (Wing, 2006, 2008). There are many definitions, learning contexts, and assessments studies on this concept (Aho, 2012; Durak & Saritepeci, 2018; Saritepeci, 2020; Wing, 2014; Yadav et al., 2014; Yildiz Durak et al., 2021). Although there is no consensus definition in the literature, CT is concerned with the effective use of information technology concepts and procedures in solving complex problems (Hsu et al., 2018; Shute et al., 2017). CT skills help individuals understand issues in various fields and use their solutions to cope with the challenges posed by the complicated digital world (Zhao et al., 2022). In summary, CT skill is the problem-solving process and way of thinking, in which designs produced with communication technologies supported or unplugged activities to the solution of problem situations (ISTE, 2016b; Wing, 2014). CT refers to a context that includes the use of various high-level skills (algorithmic thinking, problem-solving, abstract thinking, creative thinking, critical thinking, etc.) (Basogain et al., 2012; Sarıtepeci & Durak, 2017). Bundy (2007) claims that CT is used in many learning areas through problem-solving processes and is indispensable for every discipline. Barr and Stephenson (2011) emphasized that CT is a skill associated with self-confidence and perseverance in problem-solving skills.

The interest of policymakers and educators regarding CT and the view that CT should be included in the curriculum (Boulden et al., 2021; Grover & Pea, 2013; Lai et al., 2021; Lindberg et al., 2019; Mohaghegh & McCauley, 2016) and with the wide acceptance of this view, it has become important to develop standards for teachers and students for the use of technology in learning and teaching processes (e.g. ISTE, 2016a, 2016b) and to include CT among the basic skills that students should acquire.

Sanford and Naidu (2016) accentuate the expectations of today's society and the need to train individuals who are competent in terms of CT skills to solve complex 21st-century problems. While Boulden et al. (2021) state that CT is an integral part of the 21st-century life skill required for digital citizenship, Zhao et al. (2022) highlight CT as an essential skill for the daily life of every citizen in the age of information technology. The concurrence that CT is a core skill to acquire requires (Atmatzidou & Demetriadis, 2016; Barr et al., 2011; Papadakis, 2022; Saritepeci, 2020; Wing, 2006, 2008) that K-12 level in-service teachers have the competence to teach and integrate CT into their classrooms (Saritepeci, 2021). In this context, teachers' self-efficacy in integrating technology into the learning and teaching processes is of great importance (Özgün & Saritepeci, 2021; Yildiz Durak, 2021). Therefore, assessing teachers' self-efficacy in this subject is critical for effective teaching of CT. Indeed, self-efficacy is the level of belief that one has the competence to perform a task. We consider that this tool will provide practical benefits for teacher education policymakers and pre-service teacher training.

Purpose of the Study

The current study aims to adapt the T-STEM CT scale developed by Boulden et al. (2021) into Turkish and carry out validity and reliability studies of the scale. In this context, we sought a response to the following research question.

- How is the validity and reliability of the "T-STEM CT Scale" adapted into Turkish?

Method

Participants

The participants consist of 168 in-service teachers who are actively working in K12 schools in various regions of Turkey. All of the participants work in public schools. 48.80% of the participants are female, and 51.20% are male. The average age is 34.80, the age range varies between 23 and 61, and the average seniority is 11.1 years.

The T-STEM CT scale consists of 13 items. It is substantial to determine the item-responder ratios in determining the sample size. For this reason, we reviewed the suggestions in the literature. There are different suggestions in the literature regarding the number of respondents for each item in the scale: For example, three-six respondents according to Cattell (2012), at least five respondents according to Gorsuch (1983), five-ten respondents according to Bryman and Cramer (2002) are enough. In this context, we found it sufficient for each item in the scale to be answered by 12.92 respondents. Considering the relevant literature, the determined number of respondents means a sufficient and generalizable sample size for the current study.

Research Instruments and Data Analysis

The T-STEM CT scale, developed by Boulden et al. (2021), was adapted into Turkish in this study. The original scale consists of 2 factors (teachers' self-efficacy and outcome expectancy beliefs for teaching CT) and 13 items. There are seven items on the scale for the CT self-efficacy factor and six items on the CT outcome expectancy factor. These items are in a 5-point Likert structure: Strongly disagree (1), disagree (2), neither agree nor disagree (3), agree (4), and strongly agree (5).

To adapt the scale, we first requested permission to use the scale from Danielle Cadieux Boulden's correspondence address via e-mail. Following this, two field experts who know both Turkish and English languages translated the scale into Turkish, and two different translators translated the Turkish version back into English. In this process, we discussed the contextual meaning and intelligibility of each item of the scale with experts. We compared the three lists of items that emerged with the translation-re-translation processes, made the necessary adjustments, and created a draft form. After this process, two experts and two teachers who had experience with the subject area evaluated the scale form in terms of meaning and intelligibility. As a result of these reviews and evaluations, we created the final scale form.

We used confirmatory factor analysis (CFA) in the analysis of the data in the study. CFA is an analysis that tests a model related to an existing theory or a predefined structure (Çokluk et al., 2014; Hair, 2009). In this study, since T-STEM CT is a scale with a previously defined factor structure, we decided to test the factorial validity of the scale with CFA in adapting it to Turkish. In the study, descriptive statistics and CFA analyzes were performed in Jamovi 2.2.5 (R Core Team, 2020; Rosseel, 2018; The Jamovi Project, 2021).

Results

We tested the factorial structure of the CT T-STEM scale, which consists of two factors and 13 items, with CFA. According to the CFA results, (x^2 / df=2.469, CFI=.948, TLI=.938, SRMR=.0398, RMSA=.0938) the RMSA values were outside the acceptable range (Browne & Cudeck, 1993). Thereupon, residual covariance -

modification indices review, we combined the error variances of item pairs SE03-SE06 and SE05-SE07. Following this, the goodness of fit values (x^2 / df=1.984, CFI=.966, TLI=.957, SRMR=.0397, RMSA=.0767) shows that the measurement model, which includes the relationship between scale factors and items, has an acceptable fit or/and a perfect fit (Bentler & Bonett, 1980; Browne & Cudeck, 1993; Kline, 2016; Tabachnick & Fidell, 2007). In addition, item factor loads are between .66-.91 (see Figure 1).

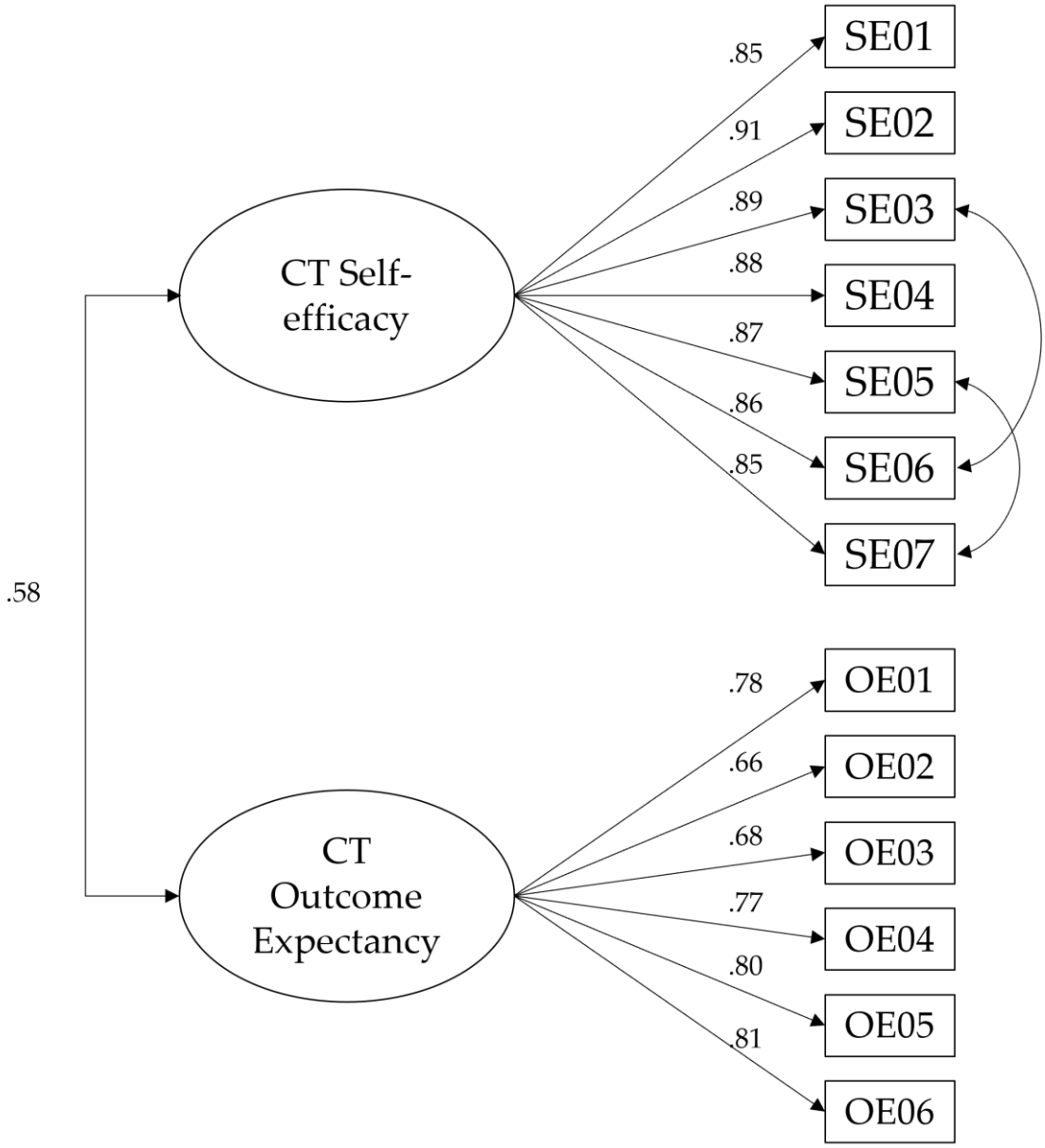


Figure 1. CFA Model for CT T-STEM

To determine the reliability and convergent validity level of the scale, we reviewed the average variance extracted (AVE), composite reliability, and Cronbach alpha values (see Table 1). For both factors, AVE is higher than .50, composite reliability greater than .70, and Cronbach Alpha greater than .70. The results indicate that the scale has good reliability, and convergent validity is achieved (Bagozzi & Yi, 1988; Gefen et al., 2000; Hair, 2009).

According to descriptive statistics, the mean score of the CT T-STEM scale is 46.10. CT Self-efficacy factor mean score is 24.80, and CT outcome expectancy factor is 21.30 (see Table 1). Accordingly, the participants' perceptions of CT self-efficacy and CT outcome expectancy are relatively high.

Table 1

Factor loading, AVE and reliability

Sub-scale	M	Sd	Factor loading	AVE	Composite Reliability	Cronbach Alpha
CT Self-efficacy	24.80	6.68		0.761	.957	.957
SE01			0.847			
SE02			0.905			
SE03			0.894			
SE04			0.884			
SE08			0.869			
SE06			0.859			
SE07			0.846			
CT Outcome Expectancy	21.30	3.55		0.567	0.886	.883
OE01			0.782			
OE02			0.659			
OE03			0.678			
OE04			0.768			
OE05			0.802			
OE06			0.814			

Notes: The CT T-STEM scale is a 5-point Likert scale ("strongly disagree" .... "strongly agree") structure.

According to Table 2, the diagonal values (square roots of AVEs) are higher than the value in the rows and columns. These results show that discriminant validity is provided.

Table 2

Discriminant Validity

		[1]	**[2]**
CT Self-efficacy	**[1]**	**.872**	
CT Outcome Expectancy	**[2]**	.550	**.753**

Discussion

Integration of learning-teaching activities into different courses for the teaching and development of CT skills is included in the literature as a considerable requirement (Grover & Pea, 2013; Lee et al., 2014; Qualls & Sherrell, 2010; Weintrop et al., 2016). As a matter of fact, in recent years, studies on the integration of CT-related concepts and skills in different disciplines into the curriculum have found more space in the literature (Bell & Bell, 2018; Gadanidis, 2017; Rubinstein & Chor, 2014; Wolz et al., 2011). One of the substantial elements of such integration activities in learning-teaching processes is teachers. In this context, teachers' self-efficacy and outcome expectancy beliefs are two critical factors in integrating CT skills into course processes. The level of self-efficacy in any subject is one of the most fundamental indicators of whether the individual will fulfill the task related to this subject. The weak self-efficacy belief in integrating CT into the course processes

will cause the teacher to be distant from such integration efforts. Another determinant of the successful performance of a task is the expectations regarding the results (Guo et al., 2015). The expectancy-value theory considers the expectation about the outcome and the values attributed to this task as the main ingredients in explaining the motivation of individuals to perform a task (Wigfield & Eccles, 2000). Accordingly, teachers' self-efficacy and outcome expectancy levels are essential determinants in integrating CT into their course processes. In this context, in this study, we aimed to adapt the T-STEM CT scale, developed by Boulden et al. (2021), into Turkish and to carry out validity and reliability studies of the scale.

We used CFA for the Turkish adaptation of the T-STEM CT instrument since it has a predefined structure. According to the results, the two-factor and thirteen-item structure had an acceptable fit with the data. This structure provided convergent and discriminant validity and had high internal consistency.

CT is a much newer concept for educators, especially at the K12 level (Li et al., 2020). Therefore, in the usage of this instrument, we recommend providing a CT description in the data collection tool for participants to refer to it. For this purpose, we prefer to consider the CT indicators descriptions in the ISTE (2016) student standards.

Ethic

In this study, all scientific ethical rules were followed.

Author Contributions

All stages of the study were organized and conducted by the authors.

Conflict of Interest

In addition, the authors declare that they have no conflict of interest.

Funding

This work has not received funding from any institution.

References

Aho, A. V. (2012). Computation and computational thinking. *The Computer Journal*, *55*(7), 832-835. https://doi.org/10.1093/comjnl/bxs074

Atmatzidou, S., & Demetriadis, S. (2016). Advancing students' computational thinking skills through educational robotics: A study on age and gender relevant differences. *Robotics and Autonomous Systems*, *75*, 661-670. https://doi.org/10.1016/j.robot.2015.10.008

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, *16*(1), 74-94.

Barr, D., Harrison, J., & Conery, L. (2011). Computational thinking: A digital age skill for everyone. *Learning & Leading with Technology*, *38*(6), 20-23.

Barr, V., & Stephenson, C. (2011). Bringing computational thinking to K-12: What is involved and what is the role of the computer science education community? *Acm Inroads*, *2*(1), 48-54. https://doi.org/10.1145/1929887.1929905

Basogain, X., Olabe, M., Olabe, J., Maiz, I., & Castaño, C. (2012). Mathematics Education through Programming Languages. 21st Annual World Congress on Learning Disabilities, In *21st annual world congress on learning disabilities* (pp. 553-559).

Bell, J., & Bell, T. (2018). Integrating computational thinking with a music education context. *Informatics in Education*, *17*(2), 151-166. https://www.doi.org/10.15388/infedu.2018.09

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological bulletin*, *88*(3), 588. https://doi.org/10.1037/0033-2909.88.3.588

Boulden, D. C., Rachmatullah, A., Oliver, K. M., & Wiebe, E. (2021). Measuring in-service teacher self-efficacy for teaching computational thinking: development and validation of the T-STEM CT. *Education and Information technologies*, *26*(4), 4663-4689. https://doi.org/10.1007/s10639-021-10487-2

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (Vol. 154, pp. 136). SAGE Publications. https://doi.org/10.1177/0049124192021002005

Bryman, A., & Cramer, D. (2002). *Quantitative data analysis with SPSS release 10 for Windows: A guide for social scientists*. Routledge. https://doi.org/10.4324/9780203471548

Bundy, A. (2007). Computational thinking is pervasive. *Journal of Scientific and Practical Computing*, *1*(2), 67-69.

Cattell, R. (2012). *The scientific use of factor analysis in behavioral and life sciences*. Springer Science & Business Media. https://doi.org/10.1007/978-1-4684-2262-7

Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2014). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları [Multivariate statistics for social sciences: SPSS and LISREL applications]*. Pegem Academy.

Durak, H. Y., & Saritepeci, M. (2018). Analysis of the relation between computational thinking skills and various variables with the structural equation model. *Computers & Education*, *116*, 191-202. https://doi.org/10.1016/j.compedu.2017.09.004

Gadanidis, G. (2017). Artificial intelligence, computational thinking, and mathematics education. *The International Journal of Information and Learning Technology*. https://doi.org/10.1108/IJILT-09-2016-0048

Gefen, D., Straub, D., & Boudreau, M.-C. (2000). Structural equation modeling and regression: Guidelines for research practice. *Communications of the association for information systems*, *4*(1), 7. https://doi.org/10.17705/1CAIS.00407

Gorsuch, R. L. (1983). *Factor analysis*. Lawrence Erlbaum.

Grover, S., & Pea, R. (2013). Computational thinking in K–12: A review of the state of the field. *Educational researcher*, *42*(1), 38-43. https://doi.org/10.3102/0013189X12463051

Guo, J., Marsh, H. W., Parker, P. D., Morin, A. J., & Yeung, A. S. (2015). Expectancy-value in mathematics, gender and socioeconomic background as predictors of achievement and aspirations: A multi-cohort study. *Learning and Individual Differences*, *37*, 161-168. https://doi.org/10.1016/j.lindif.2015.01.008b

Hair, J. F. (2009). *Multivariate data analysis: A global perspective*. Prentice Hall

Hsu, T.-C., Chang, S.-C., & Hung, Y.-T. (2018). How to learn and how to teach computational thinking: Suggestions based on a review of the literature. *Computers & Education*, *126*, 296-310. https://doi.org/10.1016/j.compedu.2018.07.004

ISTE. (2016a). *ISTE Standards for Educators*. Retrieved from https://www.iste.org/standards/for-educators

ISTE. (2016b). *ISTE Standards for Students*. Retrieved from https://www.iste.org/standards/iste-standards-for-students

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4 ed.). Guilford publications.

Lai, Y.-H., Chen, S.-Y., Lai, C.-F., Chang, Y.-C., & Su, Y.-S. (2021). Study on enhancing AIoT computational thinking skills by plot image-based VR. *Interactive Learning Environments*, *29*(3), 482-495. https://doi.org/10.1080/10494820.2019.1580750

Lee, I., Martin, F., & Apone, K. (2014). Integrating computational thinking across the K--8 curriculum. *Acm Inroads*, *5*(4), 64-71. https://doi.org/10.1145/2684721.2684736

Li, Y., Schoenfeld, A. H., diSessa, A. A., Graesser, A. C., Benson, L. C., English, L. D., & Duschl, R. A. (2020). On computational thinking and STEM education. *Journal for STEM Education Research*, *3*(2), 147-166. https://doi.org/10.1007/s41979-020-00044-w

Lindberg, R. S., Laine, T. H., & Haaranen, L. (2019). Gamifying programming education in K- 12: A review of programming curricula in seven countries and programming games. *British Journal of Educational Technology*, *50*(4), 1979-1995. https://doi.org/10.1111/bjet.12685

Mohaghegh, D. M., & McCauley, M. (2016). Computational thinking: The skill set of the 21st century. *International Journal of Computer Science and Information Technologies*, *7*(3), 1524-1530.

Özgün, Z., & Saritepeci, M. (2021). Determination of the factors affecting teachers' perceptions of classroom management competence in technology assisted courses. *Technology, Pedagogy and Education*, *30*(5), 673-691. https://doi.org/10.1080/1475939X.2021.1956579

Papadakis, S. (2022). Apps to promote computational thinking concepts and coding skills in children of preschool and pre-primary school age. In *Research Anthology on Computational Thinking, Programming, and Robotics in the Classroom* (pp. 610-630). IGI Global. https://www.doi.org/10.4018/978-1-6684-2411-7.ch028

Qualls, J. A., & Sherrell, L. B. (2010). Why computational thinking should be integrated into the curriculum. *Journal of Computing Sciences in Colleges*, *25*(5), 66-71.

R Core Team (2020). *R: A Language and environment for statistical computing*. (Version 4.0) [Computer software]. Retrieved from https://cran.r-project.org. (R packages retrieved from MRAN snapshot 2020-08-24).

Rosseel, Y., et al. (2018). *lavaan: Latent Variable Analysis*. [R package]. Retrieved from https://cran.r-project.org/package=lavaan

Rubinstein, A., & Chor, B. (2014). Computational thinking in life science education. *PLoS computational biology*, *10*(11), e1003897. https://doi.org/10.1371/journal.pcbi.1003897

Sanford, J. F., & Naidu, J. T. (2016). Computational thinking concepts for grade school. *Contemporary Issues in Education Research (CIER)*, *9*(1), 23-32. https://doi.org/10.19030/cier.v9i1.9547

Saritepeci, M. (2020). Developing computational thinking skills of high school students: design-based learning activities and programming tasks. *Asia-Pacific Education Researcher*, *29*(1), 35-54. https://doi.org/10.1007/s40299-019-00480-2

Saritepeci, M. (2021). Modelling the effect of TPACK and computational thinking on classroom management in technology enriched courses. *Technology, Knowledge and Learning*, 1-15. https://doi.org/10.1007/s10758-021-09529-y

Sarıtepeci, M., & Durak, H. (2017). Analyzing the eff ect of block and robotic coding activities on computational thinking in programming education. In G. D. Irina Koleva (Ed.), *Educational research and practice*. St. Kliment Ohridski University Press.

Shute, V. J., Sun, C., & Asbell-Clarke, J. (2017). Demystifying computational thinking. *Educational Research Review*, *22*, 142-158. https://doi.org/10.1016/j.edurev.2017.09.003

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (6 ed.). Pearson.

The jamovi project (2021). *Jamovi*. (Version 1.6) [Computer Software]. Retrieved from https://www.jamovi.org.

Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., Trouille, L., & Wilensky, U. (2016). Defining computational thinking for mathematics and science classrooms. *Journal of Science Education and Technology*, *25*(1), 127-147. https://www.doi.org/10.1007/s10956-015-9581-5

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68-81. https://doi.org/10.1006/ceps.1999.1015

Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, *49*(3), 33-35.

Wing, J. M. (2008). Computational thinking and thinking about computing. *Philosophical transactions of the royal society of London A: mathematical, physical and engineering sciences*, *366*(1881), 3717-3725. https://doi.org/doi/10.1098/rsta.2008.0118

Wing, J. M. (2014). Computational thinking benefits society. *40th Anniversary Blog of Social Issues in Computing*, *2014*, 26.

Wolz, U., Stone, M., Pearson, K., Pulimood, S. M., & Switzer, M. (2011). Computational thinking and expository writing in the middle school. *ACM Transactions on Computing Education (TOCE)*, *11*(2), 1-22. https://doi.org/10.1145/1993069.1993073

Yadav, A., Mayfield, C., Zhou, N., Hambrusch, S., & Korb, J. T. (2014). Computational thinking in elementary and secondary teacher education. *ACM Transactions on Computing Education*, *14*(1), 5. https://doi.org/10.1145/2576872

Yıldız-Durak, H., Sarıtepeci, M., & Aksu-Dünya, B. (2021). Examining the Relationship between Computational Thinking, Lifelong Learning Competencies and Personality Traits Using Path Analysis. *Bartın University Journal of Faculty of Education*, *10*(2), 284-294. https://doi.org/10.14686/buefad.888374

Yildiz Durak, H. (2020). The effects of using different tools in programming teaching of secondary school students on engagement, computational thinking and reflective thinking skills for problem solving. *Technology, Knowledge and Learning*, *25*(1), 179-195. https://doi.org/10.1007/s10758-018-9391-y

Yildiz Durak, H. (2021). Modeling of relations between K-12 teachers' TPACK levels and their technology integration self-efficacy, technology literacy levels, attitudes toward technology and usage objectives of social networks. *Interactive Learning Environments*, *29*(7), 1136-1162. https://doi.org/10.1080/10494820.2019.1619591

Yildiz Durak, H., Saritepeci, M., & Durak, A. (2021). Modeling of Relationship of Personal and Affective Variables with Computational Thinking and Programming. *Technology, Knowledge and Learning*, 1-20. https://doi.org/10.1007/s10758-021-09565-8

Zhao, L., Liu, X., Wang, C., & Su, Y. S. (2022). Effect of different mind mapping approaches on primary school students' computational thinking skills during visual programming learning. *Computers & Education*, 104445. https://doi.org/10.1016/j.compedu.2022.104445

RESEARCH ON EDUCATION AND PSYCHOLOGY (REP)

Received: March 17, 2022
Accepted: April 15, 2022
http://dergipark.org.tr/rep
e-ISSN: 2602-3733

April 2022 ◆ 6(Special Issue) ◆ 57-68
Research Article
Doi: 10.54535/rep.1089295

Turkish Adaptation Study of the Trust in Science and Scientists Scale: Validity and Reliability Study

Seher Esen[1]
Selcuk University

Menşure Alkış Küçükaydın[2]
Necmettin Erbakan University

Abstract

In this study, the scale of "Instrument to Measure Trust in Science and Scientists" developed by Nadelson et al. (2014) was adapted into Turkish, and it was aimed to perform the validity and reliability studies of the scale. The original scale consists of 5 point Likert-type, single factor, and 21 items. The study was carried out with 236 pre-service teachers selected according to the convenient sampling technique. Some steps were followed in the adaptation of the scale. Content validity, construct validity, convergent and divergent validity were tested and reliability analyses were made. After the adaptation, a 2-factor structure consisting of 10 items emerged. It was accepted that the scale, in which expert opinions were consulted, is a valid and reliable measurement tool. Thus, a measurement tool that can be used in studies of trust in science and scientists has been adapted to the Turkish language and brought to the literature.

Key Words

Scale adaptation • Science and scientist • Trust

[1] **Correspondence to:** Selcuk University, Faculty of Education, Konya, Turkey. E-mail: sesen@selcuk.edu.tr
ORCID: 0000-0002-3569-1185
[2] Necmettin Erbakan University, Eregli Faculty of Education, Konya, Turkey. E-mail: mensurealkis@hotmail.com
ORCID: 0000-0003-4410-1279
Citation: Esen, S., & Alkış Küçükaydın M. (2022). Turkish adaptation study of the Trust in Science and Scientists Scale: Validity and reliability study. *Research on Education and Psychology (REP), 6*(Special Issue), 57-68.

Emotions are of great importance in people's attitudes towards science and in their learning of science (Robinson et al., 2020). Likewise, emotions are involved in the relationship between the trustor and the trusted, and long-term interaction contributes to the formation of interpersonal interests and attachments (Rousseau et al., 1998). This is also true in the public's trust relationship with science and scientists, and different situations can be mentioned that affect this trust relationship. Because individuals' levels of trust towards science and scientists may vary emotionally, rational thoughts are at the forefront in the relationship of trust between scientists and the public (Larson et al., 2011). Trust in science and scientists can also vary according to personal, religious, and ideological views (Hamilton et al., 2015; McCright et al., 2013). In addition, ordinary people can establish a relationship of trust with science and scientists by considering the expertise, benevolence, and honesty of the scientist on a science-related issue (Hendriks et al., 2016; Mayer et al., 1995). The failure of scientists to carry out their studies in reliable ways may lead to the loss of public trust in science (Kennedy, 2008). Examples such as Diederik Stapel, working at Tilburg University, who imitated the data in many scientific articles he produced for 15-20 years and did not comply with ethical rules (Crocker & Cooper, 2011), can negatively affect the public's trust in science and scientists. Subjective decisions made on scientific issues such as genetically modified foods, hydraulic fracturing, vaccines, climate change, and applications of science away from objectivity and data-based decision-making may negatively affect trust in science and scientists (Nadelson et al., 2014). Hendriks et al. (2016) stated that the public's trust in science decreases when discussing specific issues such as nuclear energy and genetically modified foods. Because although people try to make personal and social meanings by associating their scientific ideas with their lived experiences, they can stay out of science without alienating themselves from their social contexts (Feinstein, 2011). However, when personal issues are discussed, individuals' perspectives towards science and scientists may change. For example, an individual who is sick may have increased interest and trust in science and scientists on issues related to health or an individual struggling with drought on issues related to climate change (Rousseau et al., 1998).

Although it is difficult to build, restore and maintain trust between scientists and the public, establishing a sense of trust is an important social task (Irzik & Kurtulmus, 2021). Trust in science and scientists can be achieved through science education. There is a widespread view that science education is necessary for all people, including those who do not consider scientific and technical careers (Feinstein, 2011). The reason for this is the thought that behaviors and attitudes developed through education at an early age can affect individuals' future lives and perceptions (Blalock et al., 2008; Boyd et al., 2006; Nieswandt, 2007). Perceptions towards science and scientists can affect individuals and societies. For example, people who do not trust science and believe that the vaccine harms people may endanger the health of the public (Dyer & Hall, 2019; Tvrdy, 2021). As a matter of fact, it is stated that educated individuals are more likely to trust scientists about solutions to climate change or vaccines (Hamilton et al., 2015; Sleeth-Keppler et al., 2017). There are other factors that affect the issue of trust in science and scientists. For example, Keelan et al. (2010) argue that the public has different views on the benefits or harms of vaccines. It is stated that trust in vaccines varies and individuals who can approach situations objectively are more likely to trust vaccines and scientists (Larson et al., 2011). A high level of public confidence in science and scientists is very important in terms of being able to read and understand scientific information correctly (Hendriks et al., 2016).

It is necessary for the public to trust in science and scientists, to protect and maintain this trust, and to solve the problems of trust. In order to develop trust in science and scientists and to solve the problems of trust, there is a need to document these conditions. Nadelson et al. (2014) developed a trust scale for science and scientists in line with this need. When the relevant literature is examined, there is no national scale developed or adapted to Turkish that measures the public's trust levels towards science and scientists. In this respect, it is seen that there is a need for a scale developed or adapted to Turkish by ensuring its validity and reliability.

Rationale and Purpose of the Study

In this study, Nadelson et al. (2014) in the study titled "I Just Don't Trust Them: The Development and Validation of an Assessment Instrument to Measure Trust in Science and Scientists", it is aimed to determine the psychometric properties of the scale by adapting it to Turkish.

Instrument to Measure Trust in Science and Scientists

Nadelson et al. (2014) were developed a scale to measure the level of trust in science and scientists titled "Instrument to Measure Trust in Science and Scientists". The scale was developed by a team of scientists from six different disciplines (a geoscientist, a chemist, a biologist, a biochemist, a sociologist, and a mathematician). The original scale consists of 21 items and the 5-point Likert-type (1=strongly disagree…5=strongly agree). In the first stage, the data were analyzed and after reviewing, necessary changes were made. Then, the data collected for the second stage were analyzed and the Cronbach's alpha value of the scale was reported as .84.

For another verification of the validity of the scale, correlation analysis was performed by using personal characteristics thought to be related to trust in science and scientists, and the level of trust. As a result of the analysis, it was reported that trust in science is related to religiosity, and as religiosity increases, trust decreases. In addition, it was stated that trust in science is related to political philosophy. It was concluded that liberals trust science and scientists more than conservatives. Finally, it was stated that there is a positive relationship between the trust in science and scientists and the number of science courses and grade level at the university level. These associations found confirmed the validity of the scale (Nadelson et al., 2014).

Method

Research Design

In this study, scale adaptation was made. Therefore, it was tried to reach a large sample representing the universe. In this context, the survey model was adopted in the study. In the survey model, the current situation is tried to be revealed and a general picture is drawn about the existing situation (Sezgin Selçuk, 2019). Since a measurement tool was adapted to be used in studies to be carried out with preservice teachers, the survey model was found to be appropriate in the study.

Participants

In this study, in which a scale adaptation was made, a sample was created for exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). In the study, in which the convenience sampling technique was used

(Büyüköztürk et al., 2008), the data were collected from a total of 236 preservice teachers studying at the education faculty of a large university located in the Central Anatolia Region. In this context, demographic information about the sample of the study is presented in Table 1.

Table 1

Demographic Information of the Sample Group

Participants		
Gender	*f*	%
Female	171	72,5
Male	65	27,5
Total	236	100

Research Instruments

In this study, in which the scale was adapted, the data were collected with the help of a personal information form and a questionnaire containing the scale items. In the questionnaire, the genders of the pre-service teachers were asked, and then Turkish equivalents of the scale items were included. The questionnaire form was applied to the preservice teachers in the online environment. Preservice teachers used 10-15 minutes for the application.

Ethical Permissions of Research

For the adaptation of the scale, Louis Nadelson was first contacted via e-post, and permission to use the scale was requested. The ethics committee application was made in line with the permission obtained from the researcher. In this context, the form regarding the ethics committee permission, dated 15.03.2022 and decision number 25, obtained from the Selcuk University Faculty of Education Ethics Committee, was delivered to the preservice teachers who voluntarily agreed to participate in the research.

Adaptation Process of the Scale

The responsible author was contacted on behalf of the researchers who developed the scale and permission to adapt the scale was obtained. The corresponding author has informed us about the issues to be considered regarding adaptation in the e-post. After this stage, the translation process of the scale into Turkish was started. First of all, scale items were translated separately by 2 researchers in this study. After the translation, the researchers came together and tried to reach a common view on the translated scale items. In line with the opinion that emerged, the Turkish translation of the scale was prepared. Meanwhile, the scale was sent to 2 independent English teachers and they were asked to translate. Comparing the incoming translations with the translations made by the researchers, they were sent to an expert with a doctorate in science education and a good command of English, and they were asked to express their opinions in order to reach the appropriate expressions among the translations. It was decided that the form from the expert was the final form for translation into Turkish. Then, this form was sent to a specialist in English Language and Literature and a science educator who completed his doctorate abroad for back translation, and they were asked to translate it into English. It has been observed that there are small differences between the

incoming translations in terms of meaning not to cause differences. At this point, the researchers of the study made the decision together and the sentences in the final form were clarified. Before applying the items in the scale, a specialist who has a doctorate in Turkish teaching was asked to read it and review it for language, grammar, and expression disorders. In line with the feedback and suggestions of the expert, the final version of the scale items was created and transferred to the online environment, and delivered to the preservice teachers.

Data Analysis

In this study, in which the scale adaptation was made, some analyzes were carried out. In this context, descriptive analyzes and EFA and CFA were conducted. In descriptive analyzes, missing data and whether the data showed normal distribution were discussed. Then the assumptions for the EFA were checked. The relevant literature mentions a specific sample size for EFA. Accordingly, some literature states that the number of participants/items should be over 15 (Stevens, 2009) and some literature states that the number of participants should be five or 10 times the number of items (Field, 2000; Tabachnick & Fidell, 2001). With the stated sample size, the hidden structure of the scale is tried to be revealed (Tavşancıl, 2006). In this context, a total of 236 preservice teachers participated in the study. Since the latent structure was revealed in the scale adaptation process after EFA and CFA, convergent and divergent validity was tried to ensure construct validity.

SPSS 27.0 for EFA and AMOS 27.0 for DFA were used for data analysis. Chi-Square fit test (χ^2/Sd), goodness fit index (GFI), adjusted fit index (AGFI), comparative fit index (CFI), normized fit index (NFI), recommended by Marcoulides and Schumacher (2001) at the point of determining the model fit of the scale) and unnormed fit index (NNFI) and root mean square errors (RMSEA- Root Mean Square Error of Approximation) values were taken into account. χ^2/Sd < 2, p > .05, .85<AGFI<.99, .90<GFI <.99, .90< CFI<.99, .90<TLI<.99, .005<RMSEA < .009 values were considered as model fit index values (Dimitrov, 2012; Özdamar, 2017). At the point of deciding the reliability of the scale, the Cronbach alpha value was examined.

Results

Descriptive Analyzes and Exploratory Factor Analysis

The data collected for the adaptation study of the trust scale for science and scientists were first evaluated in terms of their suitability for EFA. At this stage, Kaiser Meyer Olkin (KMO) coefficient and Barlett sphericity test results were examined. In the examination, it was seen that the KMO value was .838 and the Barlett test was significant at the p<0.01 level (χ^2=722,750, df= 45, p=.00). Based on the relevant review, it has been seen that the amount of data collected is very good (Field, 2000), so EFA can be performed.

After moving to the EFA stage, the direct oblimin rotation option under the principal components analysis was preferred as the rotation technique (Kieffer, 1998), and the factor structure with an eigenvalue greater than 1 was taken into account (Zwick & Velicer, 1986). Then, it was checked whether the anti-image values of the items were above .50 (Hair et al., 2010), and the criterion of the contribution of each item to the common variance was .30 and above (Büyüköztürk, 2004). Before examining the total variance values for EFA, it was examined whether the difference between the correlation coefficients of the model and the actual correlation coefficients of the items was

above .05. In the meantime, it was observed that 38% of the values obtained exceeded the specified value and therefore remained below the 50% limit recommended in the literature (Field, 2000), and analysis table analysis was started.

Table 2

Factor Loading Values for the Scale and Item Analysis

Items	Factor Load Values		Item Total Correlation	Independent Samples t-Test Comparison of Lower and Upper 27%
	Factor 1	**Factor 2**		
M12	.671		.464	7.007
M14	.482		.313	4.912
M15	.679		.566	8.439
M21	.792		.651	8.434
M5		.499	.314	3.638
M6		.801	.643	8.432
M7		.645	.458	8.020
M9		.737	.549	8.005
M10		.850	.723	7.570
M11		.791	.627	9.124
Eigenvalue	3.919	1.387		
Explained Rate of Variance (%)	39.192	13.872		

In Table 2, independent samples t-test results of the answers given in the lower and upper groups according to the factor loading values, item-total correlation values , and the responses to the items are presented. It was observed that the factor loads of the scale collected under two factors ranged between .482 and .792 for the first factor, between .499 and .850 for the second factor, and the total variance ratio explained was approximately 53%. Considering the item-total correlation values that make up the scale, it was seen that these values ranged between .449 and .790. It was observed that the factor load values reached were above .45, the item-total correlations were above .30, and the items with less overlap between factor loads were excluded from the study (Büyüköztürk, 2004).

All these values obtained show that the items that make up the scale have integrity and competence within themselves. Finally, the scores obtained for each item that makes up the scale were ranked from largest to smallest,

and the lower 27% and upper 27% slices were determined. The response rate of the answers given to these slices and the items by the lower and upper groups was discussed. In other words, the power of the items to distinguish between those in the lower group and those in the upper group was tried to be determined. As a result of the analysis, it was seen that there was a significant difference (p<.01) between the lower and upper groups that made up the scale.

Confirmatory Factor Analysis

In the first stage of the study, CFA was performed for the scale. The resulting structure is presented in Fig. 1.

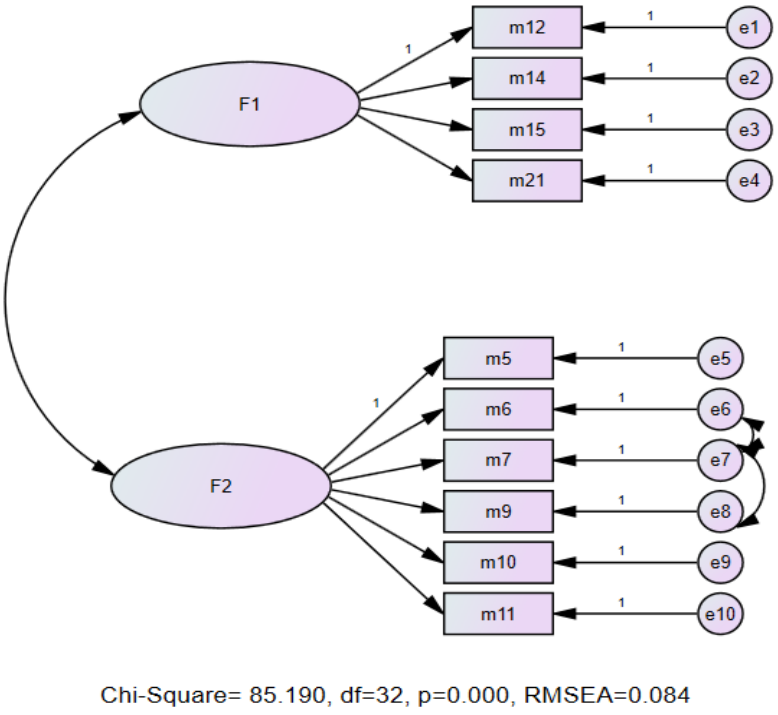


Figure 1. DFA Results

When Fig. 1 is examined, it is seen that the scale has a 2-factor structure, and covariance was formed between the 6th-7th and 8th items. When the values of the resulting structure are examined; χ^2/Sd= 2.662, AGFI = .886, RMR= .062, GFI =.934, CFI =.923, TLI =.892, RMSEA =.084, and these values were found to be in good agreement (Dimitrov, 2012; Kline , 2005). In the next step, the construct validity and reliability analysis of the items forming the scale were made.

Construct Validity and Reliability Analysis

Convergent and divergent validity methods were used to test the construct validity of the model revealed by CFA. At this stage, factor average variance extracted (AVE), composite reliability (CR), and Cronbach's alpha values were examined (Fornell & Larcker, 1981). The results obtained are summarized in Table 3.

Table 3

AVE, CR, and Cronbach's Alpha Values

Trust in Science and Scientist Scale	AVE	CR	Cronbach's Alpha
Trust in science	.914	.915	.658
Trust in scientist	.913	.916	.723

Table 3 presents the AVE, CR, and Cronbach's alpha values for the scale's construct validity and reliability analysis. Accordingly, it is seen that the AVE value is above .50 and the CR value is close to 1, thus ensuring the construct validity of the scale (Fornell & Larcker, 1981). In addition, it can be said that Cronbach's alpha values are .658 and .723 and therefore the measurement tool is reliable (Büyüköztürk, 2004). These results show that convergent validity is provided for the scale. The discriminant validity of the scale was also calculated in the study. Accordingly, the square root of the AVE values should be higher than the correlation between the factors. The discriminant validity values are presented in Table 4.

Table 4

Discriminant Validity

Factor	Trust in science	Trust in scientist
Trust in science	**.956**	
Trust in scientist	.595	**.955**

When the values in Table 4 are examined, it is seen that the square root of the AVE value is (.956) higher than the correlation value between factors (.595) for trust in science factor. Similarly, the value obtained from the factor of trust in scientists (.955) is found to be higher than the correlation value between factors. Thus, Fornell and Larcker's (1981) criterion was met.

Discussion, Conclusion & Suggestions

Negative attitudes towards science may cause a decrease in the number of scientists or weaken the public's trust in scientific knowledge and scientists (Gauchat, 2008). There are many reasons for the increase in negative attitudes towards science and scientists. Unethical behaviors in scientific studies, the role of scientists in issues such as climate change, genetically modified foods, vaccines, and nuclear weapons can negatively affect the public's trust in science and scientists (Crocker & Cooper, 2011; Kennedy, 2008). The level of trust towards science and scientists can be shaped according to the social groups or educational status of the individual, and changing personally

(Gauchat, 2012). It is an important issue to ensure the public's trust in science and scientists and to develop this trust. It was necessary to document the trust levels of the public in order to reveal whether this sense of trust varies according to individuals or social groups, or to reveal the role of education in this issue. In line with this need, Nadelson et al. (2014) developed the scale of trust in science and scientists. The scale developed in this study was adapted to Turkish.

As a result of the adaptation, a 2-factor structure emerged. Accordingly, these factors were named as "trust in science" and "trust in scientists", unlike the original. The eigenvalues of the resulting factors were found to be respectively 3.919 and 1.387. The explained total variance rate of the 2-factor scale is approximately 53%. The CR, AVE, and Cronbach's alpha values of the scale met the expectations. However, unlike the original structure, there has been a serious decrease in the number of items. This may be due to cultural and linguistic differences. however, a total of 11 items were removed from the scale in the analyses. This situation may also reduce the content validity of the scale. Against this risk, a total of 3 experts working in the field of science education and science or scientist were contacted and asked to give their opinions on the new scale structure that emerged. It was asked whether the items in the scale were actually sufficient to measure trust in science and scientists. Each of the 3 experts reported that the number of items was sufficient and that it was quite understandable in terms of Turkish. Therefore, it was accepted that the scale was successful in measuring the trust in science and scientists with this number of items. Based on this, trust in science and scientists can be measured in Turkey, especially during the pandemic period. With modeling studies, the factors affecting trust in science and scientists can be revealed or their predictors can be determined. In this context, the relations between the trust of different age groups in science and scientists and their ideological structures can be revealed.

Ethic

All procedures in this study involving human participants were carried out in accordance with the ethical standards of Selcuk University Faculty of Education Ethics Committee with date 15.03.2022 and number 25.

Author Contributions

All stages of the study were organized and conducted by the authors.

Conflict of Interest

The authors declare that they have no conflict of interest.

Funding

This work has not received funding from any institution.

References

Blalock, C. L., Lichenstein, M. J., Owen, S., Pruski, L., Marshall, C., & Toepperwein, M. (2008). In pursuit of validity: a comprehensive review of science attitude instruments 1935–2005. *International Journal of Science Education*, *30*(7), 961-977. https://doi.org/10.1080/09500690701344578

Boyd, D. J., Grossman, P., Lankford, H., Loeb, S., Michelli, N. M., & Wyckoff, J. (2006). Complex by design: Investigating pathways into teaching in new york city schools. *Journal of Teacher Education*, *57*(2), 155-166. https://doi.org/10.1177/0022487105285943

Büyüköztürk, Ş. (2004). *Veri analizi el kitabı* [Data analysis handbook]. Ankara: Pegem Academy.

Büyüköztürk, Ş., Kılıç-Çakmak, E., Akgün, Ö., Karadeniz, Ş., & Demirel, F. (2008). *Bilimsel araştırma yöntemleri* [Scientific research methods]. Ankara: Pegem Academy

Crocker, J., & Cooper, M. L. (2011). Addressing scientific fraud. *Science*, *334*(6060), 1182-1182.

Dimitrov, D. M. (2012). *Statistical methods for validation of assessment scale data in counseling and related fields.* Alexandria: American Counseling Association.

Dyer, K. D., & Hall, R. E. (2019). Effect of critical thinking education on epistemically unwarranted beliefs in college students. *Research in Higher Education*, *60*(3), 293-314.

Feinstein, N. (2011). Salvaging science literacy. *Science Education*, *95*(1), 168-185.

Field, A. (2000). *Discovering statistics using* SPSS. New York, NY: Sage Publication.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39-50.

Gauchat, G. (2012). Politicization of science in the public sphere: a study of public trust in the United States, 1974 to 2010. *American Sociological Review*, *77*(2), 167-187.

Gauchat, G. W. (2008). A test of three theories of anti-science attitudes. *Sociological Focus*, *41*(4), 337-357. https://doi.org/10.1080/00380237.2008.10571338

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Upper Saddle River, New Jersey: Prentice Hall

Hamilton, L. C., Hartter, J., & Saito, K. (2015). Trust in scientists on climate change and vaccines. *SAGE Open*, *5*(3), 2158244015602752. https://doi.org/10.1177/2158244015602752

Hendriks, F., Kienhues, D., & Bromme, R. (2016). Trust in science and the science of trust. In B. Blöbaum (Ed.), *Trust and communication in a digitized world: Models and concepts of trust research* (pp. 143-159). Springer, Cham. https://doi.org/10.1007/978-3-319-28059-2_8

Irzik, G., & Kurtulmus, F. (2021). Well-ordered science and public trust in science. *Synthese*, *198*(19), 4731-4748. https://doi.org/10.1007/s11229-018-02022-7

Keelan, J., Pavri, V., Balakrishnan, R., & Wilson, K. (2010). An analysis of the human papilloma virus vaccine debate on MySpace blogs. *Vaccine*, *28*(6), 1535-1540. https://doi.org/https://doi.org/10.1016/j.vaccine.2009.11.060

Kennedy, D. (2008). Science and its discontents: An evolutionary tale. Center for Studies in Higher Education. Berkeley, CA: University of California

Kieffer, K. M. (1998). *Orthogonal versus oblique factor rotation: a review of the literature regarding the pros and cons.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Kline, T. (2005). *Psychological testing: A practical approach to design and evaluation*. Thousand Oakk, CA: Sage.

Larson, H. J., Cooper, L. Z., Eskola, J., Katz, S. L., & Ratzan, S. (2011). Addressing the vaccine confidence gap. *The Lancet*, *378*(9790), 526-535.

Marcoulides G., & Schumacher R. (2001). *New developments and techniques in structural equation modeling*. London: Lawrence Erlbaum

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, *20*(3), 709-734.

McCright, A. M., Dentzman, K., Charters, M., & Dietz, T. (2013). The influence of political ideology on trust in science. *Environmental Research Letters*, *8*(4), 044029.

Nadelson, L., Jorcyk, C., Yang, D., Jarratt Smith, M., Matson, S., Cornell, K., & Husting, V. (2014). I just don't trust them: the development and validation of an assessment instrument to measure trust in science and scientists. *School Science and Mathematics*, *114*(2), 76-86.

Nieswandt, M. (2007). Student affect and conceptual understanding in learning chemistry. *Journal of Research in Science Teaching*, *44*(7), 908-937. https://doi.org/https://doi.org/10.1002/tea.20169

Özdamar, K. (2017). Eğitim, sağlık ve davranış bilimlerinde ölçek ve test geliştirme- yapısal eşitlik modellemesi [Scale and test development - structural equation modeling in education, health and behavioral sciences]. Ankara: Nisan Publishing.

Robinson, K. A., Beymer, P. N., Ranellucci, J., & Schmidt, J. A. (2020). Momentary emotion profiles in high school science and their relations to control, value, achievement, and science career intentions. *Motivation Science*, *6*(4), 401-412. https://doi.org/10.1037/mot0000174

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review*, *23*(3), 393-404.

Sezgin Selçuk, G. (2019). Tarama yöntemi [Survey mthod]. H.Özmen and O.Karamustafaoğlu (Ed.) İçinde *Eğitimde araştırma yöntemleri* [Research method in education] (pp.140-161). Ankara: Pegem Academy.

Sleeth-Keppler, D., Perkowitz, R., & Speiser, M. (2017). It's a matter of trust: American judgments of the credibility of informal communicators on solutions to climate change. *Environmental Communication*, *11*(1), 17-40.

Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences* (5th ed.). New York: Routledge.

Tabachnick, B.G., & Fidell, L.S. (2001). *Using multivariate statistics*. Needham Heights, MA: Allyn & Bacon.

Tavşancıl, E. (2006). *Tutumların ölçülmesi ve SPSS ile veri analizi* [Measuring attitudes and data analysis with SPSS]. Ankara: Nobel Publishing.

Tvrdy, F. (2021). Vice epistemology of believers in pseudoscience. *Filozofia*, *76*(10), 735-751.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, *99*(3), 432-442.

RESEARCH ON EDUCATION AND PSYCHOLOGY (REP)

Received: April 15, 2022
Accepted: April 30, 2022
http:// dergipark.org.tr/rep
e-ISSN:2602-3733

April 2022 ◆ 6(Special Issue) ◆ 69-91
Research Article
Doi:10.54535/rep.1104114

Development of Data Driven Decision Making Scale: A Validity and Reliability Study

Ercan Yılmaz[1]
Necmettin Erbakan University

Gulnar Jafarova[2]
Necmettin Erbakan University

Abstract

In this study, it was aimed at developing a valid and reliable evaluation tool with the purpose of evaluating the Data Driven Decision Making Skills of teachers who work in primary school, middle-school and high-school levels. 534 teachers were included in the study (256 for EFA and 278 for CFA) (63 % female and 37 % male). For the scale development process, 730 teachers constituted the whole study group. In order to determine the structural validity of the scale, exploratory factor analysis and confirmatory factor analysis were used. As a result of the exploratory factor analysis, it was determined that the scale consisted of 10 items and 2 sub-dimensions. In the light of the literature, these dimensions were titled "Data literacy" and "Decision making". The 2 sub-dimensional structure of the scale was subjected to the confirmatory factor analysis and as a result of the CFA, 1 item was excluded from the scale. The 2 sub-dimensional model created as a result of the EFA of DDDMS was tested with CFA and the adaptive values are at an acceptable level. In addition, the t values related to the high and low group difference of the scale showed that DDDMS is able to assess the structure in a distinctive manner. In order to determine the reliability of the scale, the Cronbach alpha internal consistency coefficients were calculated. When the reliability analyses results were viewed in the light of Data Driven Decision Making Scale's factors, 0,782 value was obtained for the "Data Literacy" sub-dimension and 0,672 value was obtained for the "Decision Making" sub-dimension. The inner consistency coefficient of DDDMS is 0,790. As a result of the findings, it was determined that Data Driven Decision Making Scale is a valid and reliable assessment tool to evaluate the DDDM skills of teachers.

Key Words

Data • Data driven decision making • Data literacy

[1] **Correspondance to:** Faculty of Ahmet Keleşoğlu Education, Necmettin Erbakan University, Konya, Turkey. E-mail: eyilmaz@konya.edu.tr **ORCID:** 0000-0003-4702-1688
[2] Faculty of Ahmet Keleşoğlu Education, Necmettin Erbakan University, Konya, Turkey. E-mail: aykdur78@gmail.com **ORCID:** 0000-0001-7070-9048
Citation: Yılmaz, E., & Jafarova, G. (2022). Development of Data Driven Decision Making Scale: A validity and reliability study. *Research on Education and Psychology (REP), 6*(Special Issue), 69-91.

Although numerous factors have an important effect on the realization of educational objectives, teachers are in a key position (Gujjar & Naoreen, 2009, Ozkan & Arslantas, 2013), because teachers are the individuals who plan, implement (Bahar, 2019) and evaluate (Sahin, 2011) the education process which makes students acquire the planned knowledge and skills (Calık & Arslan, 2019). As teachers manage these processes, they encounter certain problems and are forced to choose one of the different methods in terms of the selection of materials, methods, resources and this procedure requires the decision making skill (Eskiocak, 2005). In this respect, it can be stated that the decision making skill of teachers is important in terms of realizing educational objectives (Mazlumoglu, 2019). Decision making is choosing one of two or more alternatives in a given situation in a conscious manner (Gunduz et al., 2020). The developments experienced in the recent times in information and communication technologies have put data to the fore as an important factor in terms of the effectiveness of the decisions taken (Karabacak, 2019, Murrell, 2012). Data Driven Decision Making (DDDM) is the use data, instead of relying on intuition or incomplete information to make conscious decisions for the development of students (Jr, 2016).

Taking the results of scientific studies which show that students' characteristics influence their success (Arici, 2007; Considine & Zappala, 2002; Cubuk, 2019; Olufemioladebinu et al., 2018; Perger & Takacs, 2016, Yildiz, 2016) as a starting point, it can be stated that teachers' taking data into consideration when making decisions may give better educational results. In this regard, it has a great importance that teachers have Data Driven decision making skills. In order to determine teachers' level in these skills, there is a need for assessment tools which can be used in scientific studies.

When the literature is reviewed, it can be seen that there are numerous studies which analyze Data Driven decision making in the area of education from different angles (Anderson, 2015; Bouchard, 2018; Corey, 2016; Halverson et al., 2006; Harris, 2011; Luo, 2005; Mandinach, 2012; Markarian, 2009; Moriarty, 2013; Simpson, 2011; Starks, 2014; Teigen, 2009; Wagaman, 2015; White 2008; Yao, 2009). However, it was seen that the "Statewide Data-Driven Readiness Study: Teacher Survey" developed by McLeod and Seashore (2006) has been used in various studies in different ways with the purpose of collecting quantitative data. The items in the original assessment tool were separated into 4 parts as "State Assessments", "Acting upon data", "Support Systems" and "School Culture." The scale is a 6-point Likert type assessment tool (Anderson, 2015). When the Turkish literature was reviewed, it was observed that there is a limited number of studies on this subject. In addition, it was determined that a majority of these studies (Altun & Karasu, 2021; Dilekci et al., 2020; Demir, 2019; Tabak et al., 2020) were carried out through qualitative methods. Besides these studies, a study by Dogan (2021) which mixed method was used, was also found in the literature. Within the scope of this study titled "Evaluation of Data Driven decision making process in school administration in terms of the views of administrators", the "Data Driven Decision in Schools Scale" was developed with four dimensions, "chronological infrastructure and equipment", "data usage culture", "data usage purpose" and "data literacy". However, it was seen that this study involved school administrators and that there were no data collection tools which could be used in quantitative studies related to teachers' making data driven decisions. It can be stated that there is a need to develop an assessment tool to assess data driven decision making skills with the purpose of contributing to the literature. In addition, it is considered that the scale to be developed will facilitate the implementation of various studies on the subject. In terms of developing the scale, the use of vignettes which are typically used in other assessment tools and make it possible to obtain more reliable and valid participant

answers compared to "simple" abstract questions (Alexander & Becker, 1978) was found suitable for this purpose. Vignettes are designed as simulated texts (short stories), pictures, etc. and are presented to the participants to obtain answers related to the subject (Hughes & Huby, 2002). Vignettes (short stories) differ from question types in other assessment tools as they are able to concretize the context of the studied subject (Alexander & Becker, 1978). In this respect, it can be stated that vignettes can be more useful assessment tools to obtain quality data in the analysis of individuals' attitudes, perceptions and beliefs (Hughes & Huby, 2002). Since "data driven decision making" is a newly accepted concept in schools (White, 2008), it is considered that concretization of the questions to be asked to the teachers in the assessment tool through vignettes will allow obtaining participant answers with better quality. In the study carried out with this thought and the fact that there are no assessment tools based on vignettes in domestic literature, it was aimed at developing a valid and reliable assessment tool based on vignettes to assess teachers' data driven decision making skills.

Theoretical Framework

Data can be defined as information collected to be analyzed, considered, and used to help in decision making, in particular information such as facts, numbers, or electronic information which can be stored and used in an electronic environment (Cambridge Dictionary, 2022). Davenport and Prusak (1998) have defined data as basic raw material to create knowledge, "a series of separate, objective facts about events" (Luo, 2005). Recent technological developments have made data an indispensable part of the decision making process (Jr, 2016). Data in the field of education provides benefits in areas such as determining curriculum objectives (Ediger, 2010), determining education strategies, establishing effective communication, making assessments (White, 2008) and identifying the strong and weak characteristics of students (Starks, 2014). In this respect, it can be stated that making decisions based on data emerges as a need.

When the literature is reviewed, it can be seen that various definitions have been made on the data driven decision making process. Data driven decision making is the process of taking information and data into consideration when making decisions (Cemaloglu, 2019). Data driven decision making means the use of data in making conscious decisions on education (Wagaman, 2015). Data driven decision making in education means the process of collecting and analyzing data which will provide information to increase the success of students and schools (Marsh et al., 2006). In a wider sense, Data driven Decision Making (to be referred to as DDDM from this point on) is a concept related to the systematic collection, analysis, study and interpretation of data to guide policies and implementations in educational environments (Mandinach, 2012). According to another definition, data driven decision making is a process in which educators analyze assessment data to identify the strong and weak points of students, and use the obtained findings in educational applications (Mertler, 2014). In the current literature, there are various approaches which are similar to each other that related to the stages of the DDDM process. It can be stated that the DDDM process takes place in a cycle of data collection, analysis, decision making, implementation and assessment (Anderson, 2015). According to Ikemoto and Marsh's (2007) framework on the data driven decision making process, the initial stage of this process is the collection and organization of raw data. In this stage, different types of educational data are collected and included in the process. In the second stage, raw data are merged through related methods in the analysis and summary process to be transformed into information. In the third stage, data users turn information into knowledge which can be transformed into action and the implementation decision is taken. In the next stage, new data may be required to

be collected to assess the effectiveness of actions and this in turn creates a continuous cycle of collecting, organizing and synthesis of data. The data types to be used to realize this cyclical process in line with the aims have an important effect (Ikemoto & Marsh, 2007).

The realization of data driven decision making in the area of education at the state, school/local and classroom levels (Kaufman et al., 2014), diversifies the data types to be used in this process. Although the required data type is determined in line with the quality of the decision to be taken, in general standardized data, in other words data such as end of the year assessments, test data, and success scores are considered when making decisions on the basis of districts and schools (Ediger, 2010; Starks, 2014). At the classroom level, data obtained as a result of identification, observation (formation) and result assessments are underlined (Kaufman et al., 2014). However, many researchers (Corrigan et al., 2011; Starks, 2014; Turan, 2019) maintain that data driven decision making related to the success of students in education should not be limited with these data. Besides quantitative data such as report cards, tests, and comparison scores, qualitative data such as student portfolios, demographic information, surveys, observation, homework, interviews, etc. should be taken into consideration, as well. When the literature is reviewed, it can be seen that data needed to be collected to determine the current state of school success are grouped under 4 headings. These data types and groups can be expressed as below:

"student learning data"- notes, lessons taken, standard test scores, etc.

"student demographic information"- ethnic roots, gender, etc.

"perception data"- school personnel, feedback from parents or community, etc.

"school process data" – programs offered by schools, educational applications, strategies used, etc. (Bernhardt, 2001; Schwartz, 2002; Wagaman, 2015; White, 2008).

Besides all these, scientific studies show that skills such as emotional intelligence, motivation level (Erdogdu & Kenarli, 2008; Seyis, 2011), time management (Durmaz et al., 2016), critical thinking and problem solving (Sahin Kolemen & Erisen, 2017) also have positive effects on the success of students. In this respect, it can be stated that taking such data into consideration in the decision making process on students can give better educational results.

Data cannot make an organization successful on its own. For an organization to be based on data, the individuals in the organization need to be employees who can accurately interpret, use and infer data (Anderson, 2015). Ikemeto and Marsh (2007) have expressed that Data Driven Decision Making in education means teachers, school principals and other administrators who systematically collect and analyze data to make beneficial decisions with the aim of increasing the success of students and schools. When in particular it is taken into consideration that practical use of data in the classroom level is significant in terms of the success of improving education to gain continuity (Wagaman, 2015), it is important that teachers who work in schools, which are a type of organization, analyze and interpret student data accurately (White, 2008).

Method

Participants

The population of the study consists of 6000 teachers who work in central districts of the city of Konya. Bryman and Cramer (2001) have stated that sample size in the sample development process is related to the number of items and that the number of individuals in the sample should be at least five times, or even ten times the number of items in the scale. Whereas Comrey and Lee (1992) have stated that there should be 200 individuals in the medium level in the sample development process and that 300 individuals is a good number. There are 11 items in the draft scale. It was considered that the number of individuals for each of the exploratory factor analysis and confirmatory factor analysis could be achieved as a sample size of at least 250-300 individuals in the sample development process. Separate samples were determined for exploratory factor analysis and confirmatory factor analysis. It was decided that there should be 250 individuals in the sample for exploratory factor analysis. While determining samples for EFA and CFA, the multi-stage sampling method was used. The teachers in the population were separated into classes based on the types of primary, middle and high-schools they work in. Schools in each sub-class were accepted as clusters. Schools were determined with the random sampling method and the draft scales were applied to the teachers who work in those schools.

There are 256 teachers in the sample determined for EFA. 61,8 % of the participants is female and 38,2 % is male. 7,5 % of the teachers in the sample have work experience of 1-5 years, 19 % have work experience of 6-10 years, 30,4 % have work experience of 11-15 years, 22 % have work experience of 16-20 years, 15,9 % have work experience of 21-25 years and 5,2 % have work experience of 26 years and more. There are 278 teachers in the sample determined for CFA. 63,7 % of the teachers in this sample is female and 36,6 % is male.

In order to achieve the comprehensibility of the items in the draft scale form (25 teachers) and to determine the response time for the scale (11 teachers), 36 teachers who work in two of the selected school were chosen randomly from the sub-class of the population. For scale stability, 85 teachers chosen through the random sampling method from the sub-class of the population formed the study group in the test-retest process. The criteria validity of the final scale was tested in the group consisting of 75 teachers who were chosen through the random sampling method from the high-school sub-class of the population. The clusters included in the sample development process stages were excluded from the population to prevent them from being included in another study group in a different stage. Together with the participants in all of these stages, the study group consists of 730 teachers.

Data Collection Tools

In order to determine the criteria validity of the developed Data Driven Decision Making Scale (will be abbreviated as “DDDMS” from this point on), Digital Data Security Awareness Scale developed by Yilmaz et al. (2015) was used as criteria.

Digital Data Security Awareness Scale

Digital Data Security Awareness Scale was developed by Yilmaz et al. (2015) with the purpose of determining the digital data security awareness of teachers. As a result of literature review and focus group interviews with critical shareholders, an item pool consisting of 93 items was created. After receiving the views

of 12 field experts, the preliminary testing of the draft scale form was carried out with 79 teachers. For structural validity of the scale, exploratory factor analysis (EFA) was done using the data collected from 529 teachers, and a structure with a single factor, consisting of 32 items was created. The 5-point Likert scale which has acceptable internal consistency (α: 0.945) and explanatory variable (36.1 %) values was applied to 335 different participants and the confirmatory factor analysis (CFA) was done. The awareness expressions were rated through the 5-point Likert scaling. The Likert type ratings are: "Strongly agree (1)", "Agree (4)", "Neither agree nor disagree (3)", "Disagree (2)" and "Strongly disagree (1)". As the total score obtained from the scale increases, digital data security awareness also increases. All items in the scale consist of positive expressions. The structure with a single factor which reaches the ideal values through the help of modification indexes shows that Digital Data Security Awareness Scale (DDSAS) is valid and reliable.

Development of the Draft Scale

In the preparation of the draft form of DDDMS, the scale development process suggested by Seker and Gencdogan (2014) was followed. In this process, the literature, organizational structure and the results of the studies carried out in this field were analyzed (Anderson, 2015; Corey, 2016; Ikemoto & Marsh, 2007; Luo, 2005; Mandinach, 2012; Markarian, 2009; Moriarty, 2013; Simpson, 2011; Starks, 2014; Tabak et al., 2020; Wagaman, 2015; White, 2008; Yao, 2009). Based on these analyses, the first criteria on teachers' data driven decision making were created. Then, suitable vignettes were designed for these criteria. The evaluations of experts (3 experts in their fields) who have academic studies in the area of decision making, teacher competencies and data driven decision making were asked for the designed characteristics and vignettes. A questionnaire was developed for the experts to communicate their evaluations. The experts in question evaluated the characteristics and vignettes which constitute the structure of the developed scale in terms of content, structure, applicability and meaning, and have communicated their replies. The experts were asked to make their evaluations between scores from 1 to 4 for each item formed by a vignette in the draft scale (1=not suitable, 2=too many corrections are required, 3=small corrections are required, 4=very suitable). The existence of concordance in the experts' evaluations was predicted with the Kendall's Coefficient of Concordance. As a result of Kendall's analysis, a statistically significant difference was not found between the views of the experts. (Kendall's W =,333, p =.450). According to the evaluations of the experts, the necessary changes and corrections were made. The vignettes in the draft scale were analyzed by four linguistics experts in terms of language, narration, exemplification of the characteristics teachers need to have and expression style and their views were asked. According to the evaluations of the linguistics experts, the vignettes were rearranged. As a result, the 14 items of the draft scale were reduced to 11 items based on the expert evaluations and anticipated arrangements.

In terms of how the 11 vignette based items of the DDDMS should be answered, the views of assessment and evaluation experts were asked. It was decided to answer the vignette based items of the draft scale in the 5-point Likert style and to create and score the answers based on these expressions: "Highly reflects me (5 points)", "Reflects me (4 points)", "Neither reflects me nor does not reflect me" (3 points)", "Does not reflect me (2 points)" and "It does not reflect me at all" (1 point)". Then, the instruction part was prepared and the draft scale was created.

The draft scale was applied to the group consisting of 25 teachers. During the application, the teachers were asked to mark the items they did not understand and had difficulty reading and state their views. The indicated

problems were taken into consideration and the vignettes were corrected in terms of grammar and spelling mistakes within the scope of the teachers' views. Prior to finalizing the last version of the draft scale, it was analyzed once again by linguistics experts, all the necessary corrections were made, and the draft form of DDDMS was finalized. The final draft scale was applied to 11 teachers. As a result of this application, it was observed that the teachers were able to answer the draft scale in about 11-13 minutes.

Data Analysis

The SPSS and AMOS packaged software were used in the analysis of the data. Prior to the analyses, the lost data related to the data collected for both EFA and CFA, and 9 lost data were excluded from the data sets. According to Acuna and Rodriguez (2004), if the excluded data are less than 5 % of the data set, then this is a tolerable level. Therefore, the excluded data is at a tolerable level. Then, the outliers in the data set were excluded based on the Z scores, because the one-way outliers in the data set can be controlled by transforming the scores related to the items into Z scores (Tabachnick & Fidell, 2007). The Z value of the scores in the data sets; 6 data outside of the +3 and -3 range were accepted as outliers with a single variable and were excluded. Lastly, the outliers in the data set and whether the multiple variable normality assumption was met were tested with the Mahalanobis distance values and 2 data were not included in the analyses ($p < 0.01$). In addition, the value with the single variable was assessed based on the normality assumption, kurtosis and skewness coefficients. When the kurtosis and skewness values of the data sets are within the +1 and -1 range, it can be assumed that the data sets meet the conditions for normal distribution (Morgan, Leech, Gloeckner and Barrett, 2004). The kurtosis and skewness values of the data sets collected for EFA and CFA were found within the +1 and -1 range and it was assumed that these data had normal distribution. The analyzed data sets were 256 for EFA and 278 for CFA.

In the analysis of the items of the DDDMS, exploratory factor analysis and confirmatory factor analysis were done. The conformity of the collected data for the exploratory factor analysis was tested with the Kaiser-Meyer Olkin (KMO) and Bartlett tests. Then, the full conformity of the scale items which were created based on vignettes was analyzed with item total score correlation values. The structural validity of DDDMS was attempted to be determined through factor analysis. The correlation coefficients between the determined sub-dimensions' scores were calculated. The structure with 2 factors which emerged as a result of the exploratory factor analysis was tested with the first order confirmatory factor analysis. The reliability of the scale and its sub-dimensions which was tested with the first order CFA was calculated with the Cronbach alpha internal consistency coefficient method. In order to be able to determine whether DDDMS's items differed in the sub-group and higher group, the t value, average of the scores received from the items, and standard deviation were calculated. Whether accurate calculations were made or not within the reliability of the scale was determined with the test-retest method and the Pearson product moment correlation coefficient method was used in this process. In order to be able to predict the criteria validity of DDDMS, Digital Data Security Awareness Scale was used as criteria. In order to predict the relationship between the scores of this scale and DDDMS scores, Pearson product moment correlation coefficient method was used. To test whether a total score can be received from DDDMS, second order confirmatory factor analysis was done. The scoring instructions for DDDMS were created with the summated ratings technique.

Results

The stages of DDDMS's validity and reliability prediction process are presented below.

Item Analysis

During the development process of DDDMS, after the data set was structured and it was understood that it complied with the conditions of normality, items' item-total score correlations and anti-image correlation matrixes were calculated with the purpose of testing the relevance of the items created based on the vignettes to the scale.

Table 1

The Anti-Image Matrix and Item Total Correlation Values of Data Driven Decision Making Scale

	Items	**M1**	**M12**	**M3**	**M4**	**M5**	**M6**	**M7**	**M8**	**M9**	**M10**	**M11**	**İtems**	Item Total Correlation
Anti-İmage Correlation Values	**M1**	**,890a**	-,054	-,083	-,111	-,058	-,031	,070	-,091	-,046	-,066	-,226	M1	,449
	M2	-,054	**,846a**	-,257	-,097	,031	-,045	,051	,069	-,224	-,017	-,074	M2	,449
	M3	-,083	-,257	**,860a**	-,199	-,102	-,114	-,071	-,058	,105	-,110	-,088	M3	,545
	M4	-,111	-,097	-,199	**,842a**	-,274	-,031	,020	,023	-,167	-,231	,102	M4	,573
	M5	-,058	,031	-,102	-,274	**,840a**	-,095	-,063	,095	-,213	-,024	,061	M5	,458
	M6	-,031	-,045	-,114	-,031	-,095	**,918a**	-,123	-,121	-,119	-,023	-,024	M6	,466
	M7	,070	,051	-,071	,020	-,063	-,123	**,835a**	-,182	-,098	,002	-,212	M7	,357
	M8	-,091	,069	-,058	,023	,095	-,121	-,182	**,786a**	-,192	,002	-,345	M8	,409
	M9	-,046	-,224	,105	-,167	-,213	-,119	-,098	-,192	**,830a**	-,246	,077	M9	,609
	M10	-,066	-,017	-,110	-,231	-,024	-,023	,002	,002	-,246	**,882a**	-,140	M10	,559
	M11	-,226	-,074	-,088	,102	,061	-,024	-,212	-,345	,077	-,140	**,751a**	M11	,390

As it can be seen in Table 1, the scale items' item-total score correlations are between 0,357 and 0,609. In order to process the items of the developed scale, the item-total score correlation should be over 0.30 (Buyukozturk, 2015). The scale items' all diagonal values in the anti-image matrix are higher than 0.50. When the diagonal value of the anti-image correlation matrix of the developed scale is below ,5, then those items need to be excluded from the analysis (Can, 2018). The items of the draft scale can be included in the analysis based on these values and explanations.

Sample Size

In order to be able to decide on the sample size of the data collected for DDDMS, the Kaiser-Meyer-Olkin (KMO) and Bartlett test values were checked, as Alpar (2013) has suggested that sample size sufficiency in the scale development process should be checked with the Kaiser-Meyer-Olkin (KMO) value. In addition, Buyukozturk (2015) has stated that KMO coefficient being higher than 0.60 and the results of the Bartlett test being significant means that sample size is sufficient. The KMO coefficient calculated for DDDMS was

determined as 0.842. Additionally, the Bartlett test Chi-square value was statistically significant (X2= 687,354; p<0.01). According to these results, it was accepted that the data collected for Data Driven Decision Making Scale met the requirement for factor analysis.

Structural Validity

Exploratory factor analysis (EFA) of the Data Driven Decision Making Scale (DDDMS) (EFA):

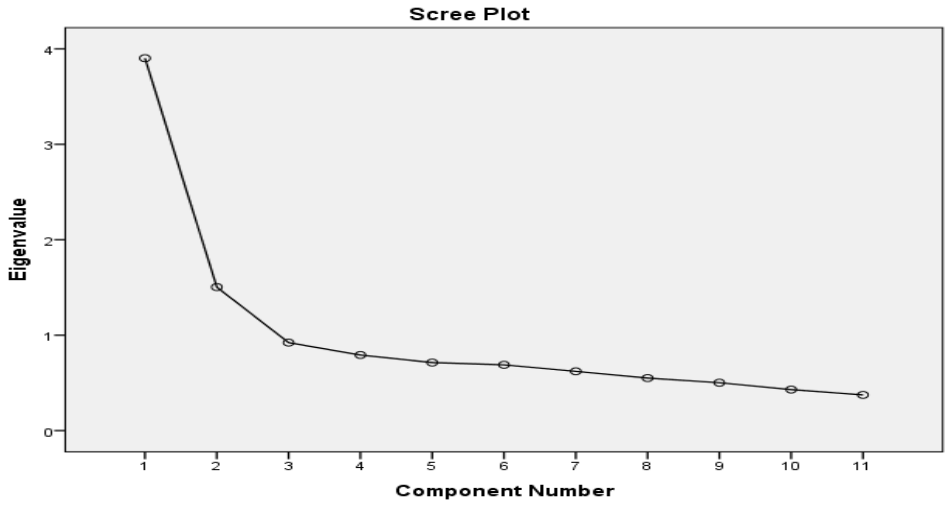
The primary statistical methods used to test the structural validity of a scale which is being developed are exploratory and confirmatory factor analyses. In order to be able to demonstrate DDDMS's structure and produce structural validity proof, the Exploratory Factor Analysis (EFA) was used. Factor analysis is a multivariate statistics which brings numerous variables which are related to each other and aims at presenting less, significant and new variables (Buyukozturk, 2015). Therefore, varimax, which is the principal components analysis in the calculation of factor loadings and the rotation technique, was Exploratory Factor Analysis (EFA), because this method reveals factors with simple meanings by rotating factor variances to be maximum with a small number of variables (Tavsancil, 2005).

As a result of the factor analysis, factor number was decided according to its eigenvalues. If a factor's eigenvalues are higher than 1, it means that the factor contains the required information about the structure which is to be assessed (Thompson, 2008). According to Guris and Astar (2015), the eigenvalues should be higher than 1.0 when deciding on the factor number. In addition, in order for an item to fall under a factor, at least 0.45 is a good value for the factor load. When necessary, this value can be lowered down to 0.30 (Buyukozturk, 2015). The factor loads were assumed to be at least 0.40 in order for DDDMS's factor eigenvalues to be higher than 1 and for its items to fall under a factor.

As a result of EFA, Scree-Plot (Figure 1) was analyzed. According to Scree-Plot (Figure 1), it can be seen that the X axis component number is two at the break point where the slope disappears. Therefore, it was decided that DDDMS's factor number was at least two.

Figure 1

Scree-Plot of Data Driven Decision Making Scale



As a result of the factor analysis, it was accepted that it can be a structure with two factors when the factors with eigenvalues higher than 1 and the theoretical bases were taken into consideration. Among DDDMS's 11 items, co-occurring item M1 which its factor load was higher than 0.40 but loaded on more than one factor was excluded from the analysis. The experts were consulted when item M1 of the still developing DDDMS was excluded from the analysis. It was decided that there was another item which assessed the same characteristic with the item related to the expert opinion assessed, the related item was ambiguous, and the exclusion of the related item would not make the scale insufficient in terms of evaluating the scope of the scale. After the exclusion of item M1, EFA was repeated. According to EFA results, a structure with 2 sub-factors with an explained total variance rate of 51,369 % was achieved. The factors achieved as a result of EFA and the items' factor loads are presented in Table 2.

Table 2

Varimax of Data Driven Decision Making Scale

	Components	
Items	1	2
M2	,600	
M3	,599	
M4	,807	
M5	,714	
M6	,436	
M9	,698	
M10	,648	
M7		,682
M8		,795
M11		,794

Total Variance Explained: 51,369, Factor 1: 36,402 % and Factor 2: 14,968 %

When Table 2 is analyzed, it can be seen that the factor loads achieved from the Exploratory Factor Analysis are between .436 and .807. It can be seen that the factor loads are higher than 0,40 which is indicated as a threshold. When the factor load values of the items were analyzed, it was assumed that the items loaded on the factors assessed the desired structure. As a result of the analyses, the factors were named in line with the items they contained. As a result of the analyses, the first factor consisting of 7 items (M2, M3, M4, M5, M6, M9 and M10) was named "Data Literacy" and the second factor consisting of 3 items (M7, M8 and M11) was named "Decision Making."

Table 3

Correlation coefficients between the factors of Data Driven Decision Making Scale

		Decision Making
Data Literacy	r	,369**

(**: p<.01)

If the correlation coefficient value of the relationship between the sub-dimensions of the scale is higher than .60, then it can be stated that all dimensions are dependent and all dimensions assess a conceptual structure (Sencan, 2005: 778). When Table 3 is analyzed, it can be seen that the correlation between the sub-dimensions of the scale have a .369 significant relationship. In this regard, it can be stated that the sub-dimensions of the scale do not assess the same conceptual structures and can be used independently.

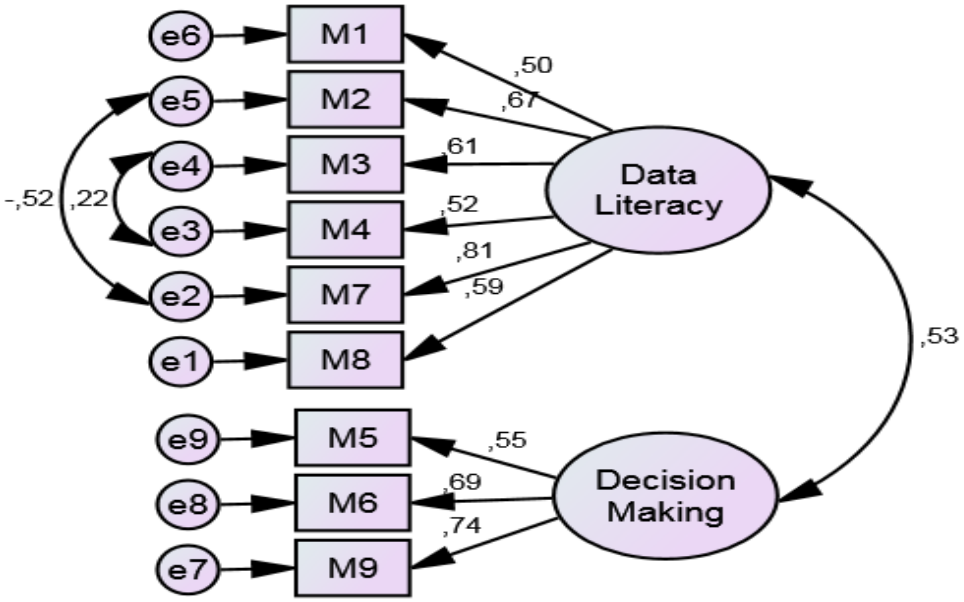
The confirmatory factor analysis (CFA) of the model, which emerges after the exploratory factor analysis (EFA), should be done to assess structural validity (Kline, 2011). Therefore, the model which emerged after EFA was tested with CFA.

Confirmatory Factor Analysis (CFA) of Data Driven Decision Making Scale (DDDMS)

Confirmatory factor analysis was used with the purpose of testing the accuracy of the two dimensional structures which was determined in accordance with the results of exploratory factor analysis to test the validity of DDDMS. As a result of the confirmatory factor analysis, it was seen that the model adaptive values of the 10 items in the scale were not at an acceptable level. Since the error variance of an item in the model created as a result of the CFA was quite high and its regression weight was very low, it was considered that it was incompatible with the scale structure of the items. This item was excluded from the model and the model was retested. Through the analysis with the exclusion of an item, the fit index calculations were renewed and the values accepted for the fit indexes were achieved. The experts were consulted when excluding this item. The experts stated that the related item could be excluded, because there were more inclusive items which assessed the characteristic that this item assessed, and the exclusion of the item would not cause the scale to be insufficient in assessing the scope of the scale. The item numbers of DDDMS, which now consisted of 9 items with the exclusion of an item in the light of these views, were revised. The model was created with the revised item numbers of DDDMS and the accepted fit index values were achieved in the renewed fit index calculations of the created model. As a result of all the revisions, the diagram achieved from CFA done for the validity of DDDMS is given in Figure 2.

Figure 2

CFA Results of Data Driven Decision Making Scale; Standardized path diagram



When DDDMS's path diagrams related to CFA are analyzed in Figure 2, it can be seen that the standardized path coefficients of the items range between 0.50 and 0.81. Kline (2005) stated that items' standardized path coefficients being 50 and over predictive quality means that they represent the variable. When the items' related path coefficients in the model are analyzed, it can be stated that the items have sufficient predictive quality. The fit index values related to this model are given in Table 3.

Table 3

Fit Index Values of the CFA Results of Data Driven Decision Making Scale and Their Comparison

Model	χ2/sd	GFI	CFI	IFI	AGFI	NNFI	RMSEA
	59,877/24=2,495	,966	,958	,959	,936	,933	0,061
Fit comment*	Perfect fit	Perfect fit	Perfect fit	Perfect fit	Perfect fit	Acceptable fit	Acceptable fit

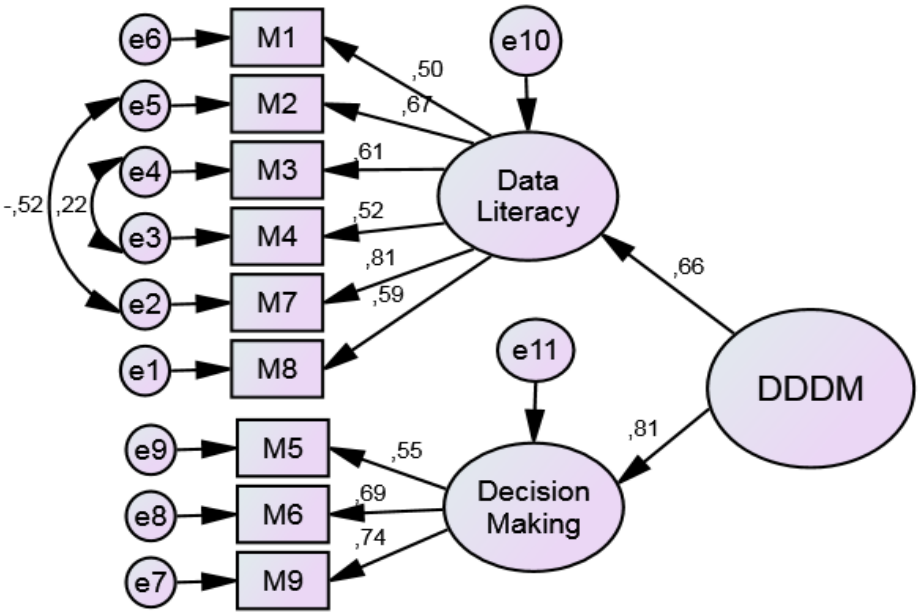
(*:SimSek, 2007; Yılmaz and Celik, 2009*)*

It can be seen that in general the fit indexes of the scale, which was obtained with 2 factors as a result of the CFA of DDDMS, have good values. The rate of Chi-square value to the degree of freedom was found as (χ2/sd=2,495. GFI (Goodness of Fit Index), CFI (Comparative Fit Index), IFI (Incremental Fit Index) and NNFI (Non-Normed Fit Index) fit indexes being close to 0.95 value and the RMSEA value being lower than 0.07 can be accepted as an indication that the model has good fit to the data. The fit indexes achieved for the scale in this study can be accepted as proof that the suggested model and the data in hand have good fit (Simsek, 2007; Yilmaz and Celik, 2009). When some calculated modification values were analyzed, correlation was found between (M2-M7; M3-M4) error covariances.

Meydan and Sesen (2011) state that the second order multi-factored models of multi-dimensional scales should also be tested when doing the confirmatory factor analysis. The components of a latent variable of the two factors of DDDMS were tested with the second order multi-factored model. The second order CFA results related to this two sub-dimensional and one dimensional model are shown in Figure 3.

Figure 3

Second order CFA results of Data Driven Decision Making Scale: Standardized path diagrams



With the purpose of showing that the Data Literacy and Decision Making dimensions of DDDMS, achieved through the first order confirmatory factor analysis, represent Data Driven Decision Making suggested theoretically in the next dimension, the second order confirmatory factor model was created (Figure 3). DDDMS was tested with the second order factor model, by adding a latent variable named Data Driven Decision Making to the first order confirmatory structure which was tested with two latent and 9 indicator variables. As a result of the testing of the second order factor model, the goodness of fit values are shown in Table 4.

Table 4

The fit index values of second order CFA results of Data Driven Decision Making Scale and their comparison

Model	χ2/sd	GFI	CFI	IFI	AGFI	NNFI	RMSEA
	59,877/24=2,495	,966	,958	,959	,936	,933	0,062
Fit Comment*	Perfect Fit	Perfect Fit	Perfect Fit	Perfect Fit	Perfect Fit	Acceptable Fit	Acceptable fit

(:Simsek, 2007; Yılmaz and Celik, 2009)*

When the first order and second order confirmatory factor analyses results are analyzed, it can be stated that, based on the structural validity result of the scale, DDDMS is an assessment tool which can be used to determine the total Data Driven Decision Making levels.

Table 5 shows the factor loads, t value for high and low group difference, average and standard deviation of the received scores.

Table 5

T value for high and low group difference, item averages and standard deviations of ***Data Driven Decision Making Scale***

	Madde no	T value for high and low group difference	İtem averages	Standard deviations
Data literacy	M1	12,358**	3,92	,843
	M2	15,154**	3,69	,918
	M3	16,135**	3,41	1,072
	M4	14,997**	3,15	1,035
	M7	19,476**	3,56	1,050
	M8	14,369**	3,45	,943
Decision making	M5	11,175**	3,80	,879
	M6	10,010**	4,36	,797
	M9	11,919**	4,33	,866

*(***: p<.01)

It was seen that there is a significant difference between the item score averages of DDDMS. According to this, it can be stated that DDDMS can distinguish individuals who receive high scores and who receive low scores from the scale.

Data Driven Decision Making Scale's scale related validity: In the scale validity study of DDDMS which was done using the Digital Data Security Awareness scale, the correlation coefficients between the total scores achieved from the scales were calculated and the results are given in Table 6.

Table 6

Scale validity results of Data Driven Decision Making scale

		Digital data security awareness
Data Driven Decision Making	r	,305**

(**: p<.01)

As it can be seen in Table 6, a positive significant relationship was found between Data Driven Decision Making Scale scores and Digital Data Security Awareness scores ($p<05$). These results show that Data Driven Decision Making Scale has scale validity.

Findings related to the reliability of Data Driven Decision Making Scale: Within the scope of DDDMS's reliability study, firstly the items' total scores and correlations were calculated. Within the scope of DDDMS' dimension and sub-dimension reliability study, the internal consistency of the items which constitute the scale were predicted with the Cronbach-Alpha Coefficient method. All the results are given in Table 7.

Table 7

Some reliability analysis values related to the scale items of Data Driven Decision Making Scale

	İtem no	Item-total score correlation	Cronbach alpha internal consistency coefficients	
Data literacy	M1	,453		
	M2	,527		
	M3	,567	,782	
	M4	,449		
	M7	,636		,790
	M8	,519		
Decision making	M5	,393		
	M6	,365	,672	
	M9	,416		

When Table 7 is analyzed, the calculated Cronbach alpha internal consistency coefficients to determine DDDMS's reliability can be seen. According to the reliability analyses, DDDMS's "Data Literacy" sub-dimension was calculated as 0,782 and "Decision Making" sub-dimension was calculated as 0,790.

Test-retest results: For DDDMS's stable assessment quality, the test-retest method was used. For this purpose, DDDMS was applied to 85 teachers twice with a four week interval. The relationship between the two application scores was calculated with the Pearson Product Moment Correlation Coefficient method. The results are presented in Table 8.

Table 8

Results related to the test-retest results of Data Driven Decision Making Scale

			2nd Application Scores		
			Data Literacy	**Decision Making**	**Data Driven Decision Making**
1st Application Scores	**Data Literacy**	r	,777**		
	Decision Making	r		,570**	
	Data Driven Decision Making				,738**

(**: p<.01)

In the analyses done with the purpose of determining the reliability of the scale through the test-retest method, a positive significant relationship was found between the first and second application scores of DDDMS's dimensions. According to these results, it can be stated that DDDMS can make strong, stable assessments. All of these analyses, done within the reliability of DDDMS, show that the reliability of the scale is sufficient.

According to DDDMS's structure with two sub-dimensions' (Data Literacy and Decision Making) being a component of an upper dimension (Data Driven Decision Making) and the second order CFA results related to the receivability of a total score from the scale (see, Table 4), it was seen that the fit indexes are at an acceptable level. These values show that the items with the factors are a part of represent a higher dimension. In the light of all these results, it can be stated that a total score can be received from the scale. The score evaluations on the scoring of Data Driven Decision Making Scale are shown in Table 9.

Table 9

Score table of Data Driven Decision Making Scale

Sub-dimensions of Data Driven Decision Making Scale	Items	Score which can be received from the whole of the sub-dimensions	
		Lowest score possible	Highest score possible
Data Literacy	1, 2, 3, 4, 7 and 8	6	30
Decision Making	5, 6 and 9	3	15
Data Driven Decision Making	1, 2, 3, 4, 7, 8, 5, 6 and 9	9	45

DDDMS is scored with the summated rating technique. According to the summated rating technique, a score received from a scale being scored is the sum of the scores given to the reactions to the items in the scale

(Tezbasaran, 1996). In the application process of DDDMS, the reaction given by the scorer to each item in the scale indicates the scorer's level of attitude towards the items. The score which indicates this level is the participant's score for that item. DDDMS's total and sub-dimension scores is calculated by adding the scores of the related items.

DDDMS can be scored separately with its sub-dimensions. While a score between 6 and 30 can be received from DDDMS's Data Literacy sub-dimension, a score between 3-15 can be received from its Decision Making sub-dimension. In order to receive a total score from the scale, the scores received from items 1, 2, 3, 4, 5, 6, 7, 8 and 9 of the scale are added to each other. There are no items in the scale which require reverse scoring. A high score received from each dimension and sub-dimensions shows that the individual has a high level of sufficiency in the related dimension and sub-dimension, whereas a low score indicates that the individual has a low level of sufficiency in the related dimension and sub-dimension. This assessment is also valid for the score received from the total scale.

As a result, the results of the EFA showed that the scale has a structure with two sub-dimensions. The model created by EFA was tested with CFA and it was determined that the fit values are at an acceptable level. It was determined with second order CFA that the Data Literacy and Decision Making sub-dimensions are represented in the Data Driven Decision Making higher dimension. It was observed that the scale achieved scale validity. In the results related to the reliability of DDDMS, it was determined that the internal consistency coefficients were at a good level and that the scale made stable assessments. When all values related to DDDMS are analyzed, it can be stated that it is a reliable and valid scale.

Discussion

The technological developments we experience around us have made the use of data an important factor in the decision making process. The results of scientific studies show that the characteristics of students have significant effects on their academic success. In this regard, it can be suggested that teachers' taking student data into consideration when making decisions related to students will improve educational results. There might be a need for assessment tools to be used in studies which deal with identifying teachers' state of taking data into consideration in the decision making process.

When the related literature was reviewed, it was determined that there was no Turkish scale which can describe teachers' Data Driven decision making skills. It was observed that a scale which was developed in another country involved an evaluation on a state scale (McLeod and Seashore, 2006). Taking the view that each country might have a different educational policy as a starting point, it was considered that developing another scale other than the scale in question might contribute to the literature.

With the purpose of concretizing subject content and achieving a higher quality of participant replies in Data Driven Decision Making Scale (DDDMS), which involves teachers' identifying their own characteristics according to their own perceptions, vignettes (short studies) were used in the study. The scale, which consists of a total of 9 vignettes, has two sub-dimensions. The first sub-dimension named "Data Literacy" consists of items based on vignettes 1, 2, 3, 4, 7 and 8. The items based on the vignettes in the first sub-dimension involve teachers' knowing the importance of data related to the DDDM process, data collection, description, knowing data types, separating, storing, analyzing, interpreting data, etc. skills. The second sub-dimension named

"Decision Making" involves items based on vignettes 5, 6 and 9. The vignettes under the "Decision Making" sub-dimension assess teachers' characteristics such as making Data Driven assessments rather on the education process, establishing communication and identifying goals.

When the validity and reliability values of DDDMS were analyzed, it was determined that the scale items were able to assess the characteristic as they aimed at. According to the views of the experts and scope validity, it can be stated that DDDMS represents the population which is desired to be assessed. According to the values of the exploratory factor analysis done to test the structural validity of DDDMS, it was assumed that the items loaded on the factors assessed the desired structure. The t values related to the scale's high and sub group difference proved that DDDMS is able to assess the structure in a distinguishing manner.

The model with two sub-dimensions created as a result of DDDMS's EFA results was tested with CFA and it was seen that the fit values are at an acceptable level. With the second order CFA, it was determined that Data Literacy and Decision Making sub-dimensions are represented in the Data Driven Decision Making higher dimension. It was observed that the scale was able to provide scale validity. According to the results related to DDDMS's reliability, it was determined that the internal consistency coefficients are at a good level and the scale is able to make stable assessments. In the light of all the values related to DDDMS, it can be concluded that it is a reliable and valid scale.

It is considered that DDDMS can be used by field experts who wish to study the DDDM skills of teachers who work in primary, middle and high-school levels.

Ethic

In this study, all rules indicated within the scope of "Higher Education Institutions Scientific Research and Publication Ethics Regulation" were followed. None of the actions indicated in the second part of the regulation under the heading, "Actions Contrary to Scientific Research and Publication Ethics" were carried out.

Author Contributions

All stages of the study were organized and conducted by the authors.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

Acuna, E., & Rodriguez, C. (2004). A meta analysis study of outlier detection methods in classification. *Technical paper, Department of Mathematics, University of Puerto Rico at Mayaguez*, 1-25.

Alexander, C.S. & Becker, H. J. (1978). The use of vignettes in survey research. *Public Opinion Quarterly, 42*(1), 93–. https://doi:10.1086/268432

Alpar, R. (2013). Çok Değişkenli İstatistiksel Yöntemler (Multivariate Statistical Methods). Ankara: Detay Yayıncılık.

Altun, N. & Karasu, N. (2021). Risk grubu öğrenciler için gönderme öncesi süreçte veriye dayalı karar verme. [Data-Driven Decision Making in the Pre-Referral Process for Risk Group Students] *TEBD, 19*(1), 593-612. https://doi.org/10.37217/tebd.906636

Anderson, C. (2015). *Creating a Data-Driven Organization*. USA: O'Reilly Media.

Anderson, D.L. (2015). *Teacher perceptions of data driven decision making for school improvement* . (Unpublished doctoral dissertation). Marian University. USA: Fond du Lac, Wisconsin.

Arici, İ. (2007). *İlköğretim din kültürü ve ahlak bilgisi dersinde öğrenci başarısını etkileyen faktörler (Ankara örneği).* (Doktora tezi). [*The Effective factors on the success of the students in the religious culture and ethics course* (Ankara example)]. (Unpublished doctoral dissertation). Ankara University Institute of Social Sciences, Ankara.

Bahar, H.H. (2019). Temel kavramlar. Kağan, M.& Yalçın, S.(Ed.), *Eğitime Giriş* içinde, (s.1-36). [Basic consepts]. Kagan, M & Yalcin, S. (Ed.), in İntroduction to Education, (p.1-36). Ankara: Pegem Academy.

Bernhardt, V.L. (2001). Intersections. *Journal of Staff Development, 21*(1), ss. 33-36.

Bouchard, H. (2018). U*sing data-driven decision-making to drive ınstructional decisions with high school mathematics teachers* (Unpublished doctoral dissertation). Concordia University. USA:Portland.

Bryman, A. & Cramer, D. (2001). *Quantitative data analysis with SPSS release 10 for Windows.* London: Routledge.

Buyukozturk, S. (2015). Deneysel desenler: Öntest-sontest, kontrol grubu, desen ve veri analizi. [Experimental designs: Pretest-posttest, control group, design and data analysis]. Ankara: Pegem A.

Calık, T. & Arslan, M.M. (2019). *Eğitime Giriş*. [*İntroduction to Education*] Ankara: Pegem Academy.

Cambridge Dictionary (2022). https://dictionary.cambridge.org/dictionary/english/data

Can, A. (2018). *SPSS ile bilimsel araştırma sürecinde nicel veri analizi* [Quantitative data analysis in scientific research process with SPSS.]. Ankara: Pegem Yayıncılık.

Cemaloglu, N. (2019). *Veriye dayalı yönetim* [Data driven management]. Ankara: Pegem Academy.

Comrey, A. L & Lee, H. L. (1992). *A first course in factor analysis.* New Jersey: Erlbaum, Hillsdale.

Considine, G. & Zappala, G. (2002). The influence of social and economic disadvantage in the academic performance of school students in Australia. *Journal of Sociology, 38*(29), 129-148.

Corey, M.M. (2016). *School leadership and data-driven ınstruction: a small city school case study* (Unpublished doctoral dissertation). University at Buffalo. USA: New York.

Corrigan, M. W., Grove, D., & Vincent, P. F. (2011). *Multi-dimensional education: A common sense approach to data-driven thinking*. Thousand Oaks, CA: Corwin Publishing.

Cubuk, A. (2019). *Ortaokul öğrencilerinin internet bağımlılığı, fiziksel aktivite düzeyleri ve akademik başarı arasında ilişkinin incelenmesi*. (Yüksek lisans tezi). [Internet addiction, physical activity level and academic achievement status of middle school students (master thesis)], Marmara University Institute of Educational Sciences, Istanbul.

Demir, K. (2009). İlköğretim okullarında verilere dayalı karar verme. [Data-Driven Decision-Making in Primary School.] Educational Administration: Theory and Practice 15 (59), 367-398. Retrieved from https://dergipark.org.tr/en/pub/kuey/issue/10338/126675

Dilekci, U., Nartgun, S. & Nartgun, Z. (2020). Okullarda Veriye Dayalı Yönetim [Data-Driven Management in Schools]. *International Pegem Conference on Education (IPCEDU-2020)* . pp. 232-243.

Dogan, E. *(2021). Okul yönetiminde veriye dayalı karar verme sürecinin yönetici görüşlerine göre değerlendirilmesi.* [Evaluating the data-driven decision making process in school management according to the views of the administrators] *(*Doctoral dissertation) Gazi Unıversity Graduate School Of Educatıonal Sciences. Ankara.

Durmaz, M., Huseyinli, T. & Güclü, C. (2016). Zaman yönetimi becerileri ile akademik başarı arasındaki ilişki. [The relation between time management and academic success]. *Journal of Humanities and Social Sciences Research,, 5*(7). pp. 2291-2303.

Ediger, M. (2010). Data Based Instruction in Reading. *Reading Improvement, 47*(4), 175-178.

Erdoğdu M. Y. & Kenarlı Ö (2008). Duygusal zeka ile akademik başarı arasındaki ilişki. [Relations between emotional quotient with academic success.] *Milli Eğitim, 37*(178), 297 - 310.

Eskiocak, S.(2005). *Sınıf öğretmenlerinin öğretimi planlama aşamasında karar verme sürecine etki eden etmenlerin analizi* (Yüksek Lisans Tezi). [The analyzes of the factors which affected the primary school teachers' making decision process in their plannıing for education. (Master Thesis)] Cukurova University Department of Educational Sciences, Adana.

Gujjar, A.A. & Naoreen, B (2009). Role of teacher as classroom manager. İ-*manager's Journal on Educational Psychology, 2*(4), 65-73. Erişim adresi: https://files.eric.ed.gov/fulltext/EJ1097655.pdf

Gunduz, S., Atas, H. & Elkovan, C.G. (2020). *Çok kriterli karar verme* [Multi-criteria decision making]. Ankara: Gazi Kitabevi.

Guris, S., & Astar, M. (2015). *Bilimsel araştırmalarda SPSS ile istatistik* [Statistics with SPSS in scientific research]. *Istanbul: Der Publications.*

Halverson, R., Grigg, J., Prichett, R. & Thomas, C. (2006). The new instruction leadership: Creating data-driven instructional systems in schools. *Journal of School Leadership 25(*3). ss.3-58. https://doi:10.1177/105268461502500305.

Harris, N. (2011). *The Impact of professional development in data based decision making on the teaching practices of educators* (Unpublished doctoral dissertation). Walden University. USA: Minnesota.

Hughes, R. & Huby, M. (2002). The application of vignettes in social and nursing research. *Methodological İssues in Nursing Research*, *37*(4), 382–386. https://doi:10.1046/j.1365-2648.2002.02100.x

Ikemoto, G. S., & Marsh, J. A. (2007). Cutting through the "data-driven" mantra: Different conceptions of data-driven decision making. In the *Yearbook of the National Society for the Study of Education* (Vol. 106, pp. 105-131). https://doi: 10.1111/j.1744-7984.2007.00099.x

Jr.S.C. (2016). *An investigation of primary and secondary teachers' beliefs in the usefulness of data-drıven decision-making (*Unpublished doctoral dissertation). Pace University. USA: New York.

Karabacak, M.E. (2019). *Data driven decision making in public administration: an assessment based on business intelligence maturity model.* (Master Thesis). Ankara Yildirim Beyazit University The Institute Of Social Sciences. Ankara.

Kline, R. B., (2005). *Principles and practice of structural equation modeling.* Guilford Press, New York.

Kaufman, T., Graham, C. R., Picciano, A. G., Wiley, D., & Popham, J. A. (2014). Data-driven decision making in the k12 classroom. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), Handbook of research on educational communications and technology (4th ed., pp. 337–346). New York, NY: Springer.

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. Guilford Publications.

Luo, M.N. (2005). *AN empirical study of high school principals' data-driven decision-making practices and their relationships to contextual variables* (Unpublished doctoral dissertation). University of Nebraska. USA: Omaha.

Mandinach, E.B. (2012). A Perfect time for data use: using data-driven decision making to inform practice. *Educational Psychologist, 47*(2), ss.71-85. http://dx.doi.org/10.1080/00461520.2012.667064

Markarian, B.S.(2009). *Holding on and holding out: why some teachers resist the move toward data-driven decision making* (Unpublished doctoral dissertation). University Of Southern California. USA: California

Marsh, J.A.,, Pane, J.F. & Hamilton, L.S. (2006). *Making sense of data-driven decision making in education. https://www.rand.org/content/dam/rand/pubs/occasional_papers/2006/RAND_OP170.pdf*

Mazlumoglu, M. (2019). *Sınıf öğretmenlerinin karar verme becerileri ile öz düzenleme becerileri arasındaki ilişkinin incelenmesi* (Yüksek lisans tezi). [Investigation of the relationship between decision making skills and self-regulation skills of classroom teachers (Master Thesis)] Ataturk University İnstitute Of Educational Sciences , Erzurum.

McLeod, S., & Seashore, K. (2006, Spring). *Data driven decision-making readiness survey: Teachers.* Minneapolis, MN: University of Minnesota.

Mertler, C.A.(2014). *The Data-Driven Classroom: How do I use student data to improve my instruction? (ASCD Arias).* USA: ASCD.

Morgan, G. A., Leech, N. L. Gloeckner, G.W. ve Barrett, K. C. (2004). *SPSS for introductory statistics: Use and interpretation.* Psychology Press.

Meydan, C. H. & Sesen, H. (2011). *Structural equation modeling AMOS applications.* Detay Yayıncılık.

Moriarty, T.W. (2013). *Data-driven decision making: teachers' use of data in the classroom* (Unpublished doctoral dissertation). University of San Diego. USA: San Diego.

Murrell, M.A. (2012). *Principals can improve student achievement with Data driven decision making (*Unpublished doctoral dissertation). University of Houston. USA: Houston.

Olufemioladebinu, T., Adediran, A.A., Oyediran, W.O. (2018). Factors ınfluencing the academic achievement of students' in colleges of education in Southwest, Nigeria. *Journal of Education and Human Development, 7*(3). 109-115. https://doi: 10.15640/jehd.v7n3a12

Ozkan, M.& Arslantas.(2013). Etkili öğretmen özellikleri üzerine sıralama yöntemiyle bir ölçekleme çalışması. [A study of scalıng wıth rankıng judgment method on characterıstıc of effectıve teacher]. *Trakya University Journal of Social Sciences , 15*(1), 311-330.

Perger, M. ve Takacs, İ. (2016). Factors contributing to students' academic success based on the students' opinion at bme faculty of economic and social sciences. *Periodica Polytechnica Social and Management Sciences, 24*(2). 119-135.

Sahin Kolemen, C. & Erisen, Y. (2017). Mesleki ve Teknik Ortaöğretim öğrencilerinin problem çözme ve eleştirel düşünme becerileri ile akademik başarıları arasındaki ilişkinin incelenmesi. [An İnvestigation on the relationship between problem solving and critical thinking skill, and academic achievement of vocational and technical high school students]. *Turkey Education Journal, 2*(1), 42-60.

Sahin, A. (2011). Öğretmen algılarına göre etkili öğretmen davranışları. [*Effective teacher's attitudes according to teacher's perceptions*]. *KEFAD, 12*(1), 239-259.

Schwartz, W. (2002). Data-driven equity in urban schools. *Eric Clearinghouse on Urban Education,* 177, 1-7.

Seker, H. & Gencdogan, B. (2014). *Psikolojide ve eğitimde ölçme aracı geliştirme. (2. Basım).* [Developing a measurement tool in psychology and education. (2nd Edition)]. Nobel Yayınevi

Sencan, H. (2005). *Sosyal ve Davranışsal Ölçümlerde Güvenirlik ve Geçerlik.* [Reliability and Validity in Social and Behavioral Measurements. Ankara: Seçkin Publishing.] Ankara: Seçkin Yayıncılık.

Seyis, S. (2011). *Ortaöğretim öğrencilerinin motivasyonları ve duygusal zekaları ile akademik başarıları arasındaki ilişki* (Yüksek Lisans Tezi). [Relationship between high school students' motivation, emotional intelligence and academic achievement) (Master Thesis).] Karadeniz Technical University, Institute of Educational Sciences, Trabzon.

Simpson, G.H. (2011). *School leaders' use of data-driven decision-making for school improvement: a study of promising practices in two California charter schools* (Unpublished doctoral dissertation). University Of Southern California. USA: California

Starks, C.V. (2014). *Data driven decision-making in midwest charter schools: teachers' experiences in the classroom.* (Unpublished doctoral dissertation). University of Missouri Kansas City. USA: Missouri.

Tabachnick, B. G. & Fidell, L. S. (2007). *Using multivariate statistics*. Boston: Allyn and Bacon.

Tabak, H. , Sahin, F. & Yavuz Tabak, B. (2020). Okul yöneticilerinin karar alma yaklaşımları: veriye dayalı karar almaya geçiş [School administrators decision making approaches: transitıon to data-based decision making]. *The Journal of Turkish Social Research, 24* (3), pp. 713-725. Retrieved from https://dergipark.org.tr/en/pub/tsadergisi/issue/58254/727254

Tavsancıl, E. (2005). *Tutumların ölçülmesi ve spss ile veri analizi* [Measuring attitudes and data analysis with spss. Nobel Publishing]. Nobel Yayıncılık.

Teigen, B.N. (2009). *A Systematic Examination of Data-Driven Decision-making within a School Division: The Relationships among Principal Beliefs, School Characteristics, and Accreditation Status (*Unpublished doctoral dissertation). Virginia Commonwealth University. USA: Richmond.

Tezbasaran, A. A. (1996). *Likert tipi ölçek geliştirme kılavuzu*. (*Likert type scale development guide).* Ankara: Publications of the Association of Psychologists.

Thompson, B. (2008). *Foundations of behavioral statistics: An insight-based approach.* New York, NY: Guilford Press

Turan, H. (2019). Data-Driven decision making. N.Cemaloglu in. Data-Driven Management *(p.2-17).* Ankara: Pegem Academy.

Wagaman, R.M. (2015). *Understandıng the prıvate school: how do prıvate school teachers understand and descrıbe data-drıven decısıon makıng?* (Unpublished doctoral dissertation). Capella University. USA: Minnesota.

White, V.C. (2008). *Relationships among principals' beliefs about data-driven decision making, principal and school characteristics, and student achievement in elementary schools* (Unpublished doctoral dissertation). University Of Florida. USA: Florida.

Yao, J.M. (2009). *Examining the applications of data-driven decision making on classroom instruction (*Unpublished doctoral dissertation). University Of Southern California. USA: California

Yildiz, B. (2016). *İlkokul öğrencilerinin akademik başarılarının arttırılmasında öğretmen, okul yönetimi ve öğrenci veli görüşlerinin incelenmesi* (Yüksek lisans tezi). [The review of teacher, school management and students' parents to improve the success of elementary students' academic successes (Master thesis).] Mersin University Institute of Educational Sciences. Mersin.

Yilmaz, E., Sahin, Y.L & Akbulut, Y. (2015). Dijital veri güvenliği farkındalığı ölçeğinin geliştirilmesi. [Development of the digital data security awareness scale] *Online Academic Journal of Information Technology, 6*(21). 23-40. https://doi: 10.5824/1309-1581.2015.4.002.x