

Adapting Middle Grades Self-Efficacy to Teach Statistics Instrument into Turkish

Aslıhan Batur^a, Bülent Güven^b and Kemal Akoğlu^c

^a Artvin Çoruh University, Education Faculty, Artvin/Turkey (ORCID: 0000-0002-4461-0615)

^b Trabzon University, Fatih Faculty of Education, Trabzon/Turkey (ORCID: 0000-0001-8767-6051)

^c Boğaziçi University, Education Faculty, İstanbul/Turkey (ORCID: 0000-0002-5688-1316)

Article History: Received: 22 September 2020; Accepted: 12 December 2020; Published online: 31 December 2020

Abstract: This study is intended to conduct validation and reliability studies by adapting Middle Grades Self-Efficacy to Teach Statistics (SETS-MS) Instrument into Turkish within the framework of the Turkish Education System structure. Accordingly, 421 prospective middle grades mathematics teachers who are currently educated in five different state universities in Turkey have participated in the study. The data of the study is collected through Turkish form of the instrument. The findings toward the validation of construct of the instrument are performed through confirmatory factor analysis and findings toward its reliability are performed through analysis on the Cronbach Alpha coefficient, which evaluates the internal consistency of the instrument item, independent sample t-test based on the difference between the mean scores of the lower 27% and upper 27% groups and item-total correlations. The results of the analysis demonstrate that fit indices calculated for the two-dimensional model which is reading the data and reading between the data are fitting with the original instrument enough. The reliability of scores achieved from the instrument is determined to be enough. In line with this, Turkish adaption of the instrument is considered to be a valid and reliable instrument which can be used for examining prospective middle grades teachers' self-efficacy to teach statistics.

Keywords: Statistics Teaching, Self-efficacy, Instrument Adaption, Validation-Reliability, Prospective Teachers

Öz: Bu çalışmada Ortaokul Düzeyinde İstatistiği Öğretmeye yönelik Öz yeterlik (Middle Grades Self-Efficacy to Teach Statistics- SETS-MS) Ölçeği'nin Türk Eğitim Sisteminin yapısı çerçevesinde Türkçeye uyarlanarak geçerlik ve güvenilirlik çalışmalarının yürütülmesi amaçlanmıştır. Bu amaç doğrultusunda, Türkiye'deki beş farklı devlet üniversitede öğrenim görmekte olan toplamda 421 ilköğretim matematik öğretmen adayı araştırmaya dahil edilmiştir. Araştırmanın verileri ölçeğin Türkçe çevirisinin yapıldığı form ile toplanmıştır. Ölçeğin yapı geçerliğine ilişkin bulgular doğrulayıcı faktör analizi, güvenilirliğine ilişkin bulgular ise Cronbach Alfa iç tutarlık katsayısı, %27'lik alt ve üst grup ortalamaları farkına dayalı bağımsız örneklem için t testi ve madde toplam korelasyonu analizleri ile elde edilmiştir. Analiz sonuçları, veri okuma ve veriler arası okuma şeklindeki iki faktörlü model için hesaplanan uyum indeks değerlerinin orijinal ölçekle yeterli düzeyde uyum sağladığını ortaya koymuştur. Ayrıca ölçekten elde edilen puanların güvenilirliğinin ise, yeterli düzeyde olduğu tespit edilmiştir. Bu doğrultuda Türkçeye uyarlanan ölçeğin, öğretmen adaylarının ortaokul düzeyinde istatistiği öğretmeye yönelik öz yeterliklerinin incelenmesinde kullanılabilecek geçerli ve güvenilir bir ölçme aracı olduğu düşünülmektedir.

Anahtar Kelimeler: İstatistik Öğretimi, Öz yeterlik, Ölçek uyarlama, Geçerlik-güvenilirlik, Öğretmen adayları

[Türkçe sürüm için tıklayınız.](#)

1. Introduction

Especially with social media getting more widespread, statistical concepts have become more prominent in our lives and have been affecting it. People of the 21st century have been living their life with a shower of data. Most of the time, living our lives in a stream of data and the feeling of our thoughts being manipulated through these data are becoming a natural part of today people (Garfield, & Ben-Zvi, 2008). Developing skills to handle these data affecting our lives can be deemed to be becoming one of the prior goals of education. For enabling to develop such skills and cultivate statistically equipped individuals, qualifications that teachers have are considerably important. Specifically, these teachers who were not exposed to intense statistics education during their student life are known to have serious anxiety to teach statistics (Steinberger, 2020). Though the source of the anxiety seems to be the content knowledge and pedagogical content knowledge on the surface, in deep down teacher's self-efficacy to teach statistics is frequently able to get beyond the content knowledge and may affect teacher's teaching practices in classroom.

Self-efficacy is a significant concept founded on Bandura's social learning theory and emphasizing that it is necessary for individuals primarily to have self-confidence toward the field to use skills they have in the related field effectively (Pajares, 2002). Bandura (1997) defines self-efficacy as belief in one's capability to organize and execute a plan of action to succeed prospective situations and underline that this belief has more effect on individual's behaviour than one's skills in the field. Thus, individuals who have high self-efficacy are known to behave more willingly to carry out tasks and to be more durable and persistent by not giving up in comparison to those who have low self-efficacy (Arseven, 2016). Similarly, it is observed that teachers whose self-efficacy is

Corresponding Author: Aslıhan Batur  **email:** aslihanbatur729@artvin.edu.tr

Citation Information: Batur, A., Güven, B. & Akoğlu, K. (2020). Adapting middle grades self-efficacy to teach statistics instrument into Turkish. *Turkish Journal of Mathematics Education*, 1(1), 10-30.

high behave more interrogatively in classroom and apply more student-centered approach whereas those whose self-efficacy is low behave in a teacher-centered (Czerniak,1990). It is highly probable that such situations happen to be caused from teachers' self-efficacy to teach.

Self-efficacy to teach which can be seen as a dimension of self-efficacy is defined as teacher's belief regarding his/her teaching skills (Wheatley, 2005; Zee, & Koomen, 2016). Addressing this concept as peculiar to a field and teacher's focus on efficiency to teach in that field is rather important to enable a noticeable development (Bandura, 2006). In line with this, teachers' self-efficacy to teach mathematics is qualified as a pre-condition to reach targeted learning outcomes specifically in mathematics in which failure sense of students is high (Dede, 2008). Yet, teachers' self-efficacy of mathematics and self-efficacy to teach mathematics are used in each other's place, thus, it is emphasized that the difference between these two concepts is necessary to be made clear (McGee, & Wang, 2014). Namely, teachers' self-efficacy of mathematics explains a belief in skills to carry out tasks related to mathematics whereas self-efficacy to teach mathematics represents a belief in skills to teach other people mathematics (Kahle, 2008). Aside from content knowledge and pedagogical content knowledge that a teacher has for understanding the prospective situation that may occur, making an interpretation, and producing solution suggestions during the teaching mathematics, his/her self-efficacy to teach mathematics are expected to be developed. (Mersquita, & Drake, 1994). For instance; Göloğlu-Demir and Çetin (2010) think the time which teachers set aside for teaching mathematics, endeavours they put, their willingness to apply new methods and techniques and endeavours to prepare a plan for special students as reflection of their self-efficacy to teach mathematics. Thus, examining levels of teachers' self-efficacy to teach mathematics is becoming more important and various instrument are developed to achieve it (Enochs, Smith, & Huinker,2000; McGee, & Wang, 2014; Gerez-Cantimer, Şengül, & Akçin, 2020). These few instruments are important in the aspect of offering a framework for teachers' self-efficacy to teach mathematics even though they are less in number.

1.1. Self-Efficacy to Teach Statistics

We come across statistics as a part of mathematics education, however, further knowledge of mathematics is not enough on its own to understand fundamental concepts of this discipline (Begg, & Edwards, 1999). Many researchers agree on the opinion that statistics takes procedural foundations from mathematics, yet it is differentiated from mathematics with its own thinking system (delMas, 2004; Groth, 2007; Garfield, & Ben-Zvi, 2008). It is only natural for mathematics teachers who have a thinking formation that is centred around deductive structure of mathematics as a nature of the education they've been through to have hard time both understanding and teaching statistics which has data-based understanding and makes interpretation based on the context. Proofs that reveal mathematics teachers have a hard time understanding statistics are available in literature (Shaughnessy, 2007; Fitzmaurice, Leavy, & Hannigan, 2014). Reasons of their hardship are derived from the nature of statistics different from mathematics, statistics knowledge that mathematic teachers have and interpretation depending on self-efficacy in the field (Irakleous, & Panaoura, 2015). Thus, these teachers have less confidence in themselves to teach statistics (Watson, 2013). For instance; Fitzmaurice, Leavy, and Hannigan (2014) conclude that prospective mathematics teachers think of statistics as hard and it affects their self-efficacy to teach statistics as a result of interviews with them. Teachers not thinking of themselves as efficacy to teach statistics becomes an important ground for passing it over by putting it into few lessons at the end of the school term (Callingham, & Burgess, 2014). Gal (2004) points out that one of the biggest obstacles for integrating statistics into curricula is teachers' tendency (specifically their self-efficacy) of statistics. This tendency affects teachers' interpretation, motivation and endeavours regarding practicing statistics efficiently (Pierce, & Chick, 2011; März, & Kelchtermans, 2013). Consequently, examining teachers' self-efficacy to teach statistics is thought to be important for teaching practices and students' success.

In comparison to mathematics, teachers' self-efficacy to teach statistics is indicated to have more complicated structure (Harrel-Williams, Sorto, Pierce, Lesser, & Murphy, 2014). Because teachers are met with statistics curricula with the expectation to teach at the school grade even though they do not graduate from university as prepared and equipped to teach statistics (Wessel, 2014). Consequently, it is not reasonable for results concluded from measurements of teachers' self-efficacy to teach mathematics to be validated for statistics as well. In other words, teachers' self-efficacy teach to statistics can be different from teachers' self-efficacy to teach mathematics. Measuring self-efficacy of these two fields having many differences in the aspect of their own nature and teachings are required to have different instruments. On that note, as looking into instrument development studies for statistics, self-efficacy of statistics is generally observed to be the focus (Finney, & Schraw, 2003; Lane, Hall, & Lane, 2004; Mercimek, & Pektaş, 2013). Additionally, by being more into specific, instrument development studies examining self-efficacy of statistical literacy are available (Carmichael, & Hay, 2009; Lin, & Huang, 2013; Batur, Yiğit, & Baki, 2019). As a result of these instruments, significant ideas on students, teachers and prospective teachers' self-efficacy of statistics are achieved, however, only few instruments are developed for examining of self-efficacy to teach statistics. As one of these limited studies, Watson (2011) developed an instrument for revealing teachers' skills and self-efficacy to teach statistics and probability. As a result of studies conducted with the participation of 43 teachers who are at primary and middle

grades, it is revealed that teachers' self-efficacy to teach is the highest for graphics and the lowest for probability. In a sense Watson's instrument is believed to lay foundation for examining teachers' self-efficacy to teach statistics, yet the items included in this instrument which do not focus entirely on this (additionally including information such as teaching experience, teaching practices) cause an important deficit.

1.2. SETS Instrument

SETS Instrument whose studies started in 2008 is designated to examining prospective mathematics teachers' self-efficacy to teach statistics. Resulting from complex nature of statistics and hardships of measuring self-efficacy of the field, it is considerably differentiated from existing instruments focusing on self-efficacy to teach statistics. It is known that teacher's self-efficacy develops during the period prior to entering classroom (Smith, 1996) and significantly resistant to changes to occur after this period (Woolfolk Hoy, & Burke-Spero, 2005). Consequently, SETS Instrument intended for prospective teachers is considered to be an important advantage in the aspect of examining teachers' self-efficacy to teach statistics beforehand and studying to develop it. Thus, Tschannen-Moran and Woolfolk Hoy (2001) point out to the importance of support for improving teachers' self-efficacy to cultivate efficient, decisive and enthusiastic teachers. In this sense, it can be concluded that information achieved through SETS Instrument is important for efficiency of statistics teaching to be developed in years to come. Additionally, self-efficacy structure of Bandura (2006) who claims that teachers' self-efficacy should be measured as a specific to certain field lays foundation for SETS Instrument (Lovett, 2016). In this sense, the instrument becomes prominent in the aspect of setting certain boundaries for statistics differentiated from mathematics due to its nature and specifically focusing on self-efficacy to teach statistics.

SETS Instrument is parted into two versions as middle school (SETS-MS; Harrell-Williams vd., 2014) and high school (SETS-HS; Harrell-Williams, Lovett, Lee, Pierce, Lesser, & Sorto, 2017) grades. While SETS-MS has a two-dimensional construct, SETS-HS has a third dimension in addition to these two dimensions (Akoğlu, 2018). As determining these dimensions in SETS Instrument, it is based on three levels (A, B and C) suggested in Guidelines for Assessment and Instruction in Statistics Education [GAISE] (Franklin vd., 2007) report which contains important recommendation on how to teach statistics as efficiently. These levels included in the instrument are paired with Friel, Curcio and Bright's (2001) graphical literacy levels (reading the data, reading between the data and reading beyond the data), which have a key role in examining many skills such as individuals' knowledge on graphics, statistical thinking and reasoning. Putting together these structure that are considered as two source documents on their own in statistics education points out that researches based on SETS Instrument to be conducted will enable detailed information on self-efficacy to teach statistics. In this sense, the basic framework of the instrument including dimensions built upon a strong foundation can be schematized as follows.

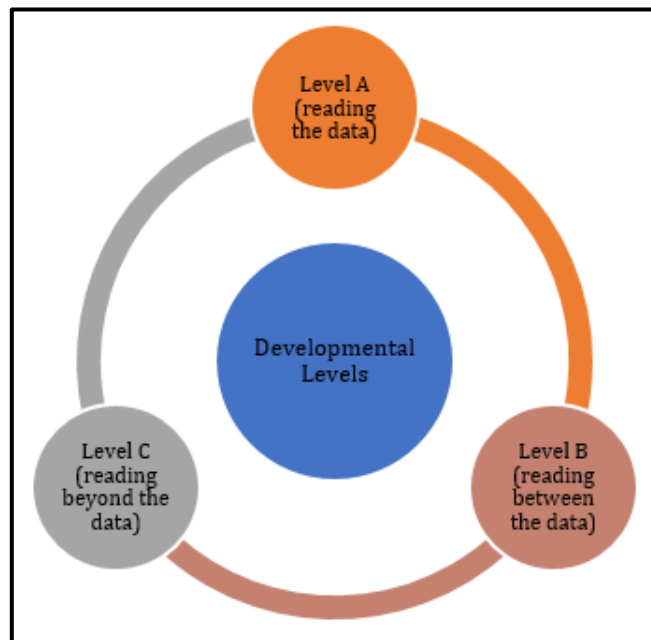


Figure 1. The basic framework of the SETS Instrument (URL-1, 2020)

Level A (reading the data) reflects self-efficacy to teach statistics concepts (for example, using table and graphics to summarize a data group and to recognize variable among data to be collected) at a basic level. At Level B (reading between the data), self-efficacy to teach prospective situations in which an association is made among concepts of statistics and put forwards comparisons (for example, to compare distribution of different

data groups and interpret the correlation coefficient among variables) is expected. Level C (reading beyond the data) focuses on self-efficacy to teach prospective situations in which require making statistical inferences at advanced level and generalization over statistical results (for example to make estimation regarding population mean through sample and to decide whether the significant difference between sample means). In this sense, starting from basic skills at Level A and extending to examining self-efficacy to teach advanced skills respectively at Level B and C, a developmental trajectory can be observed in SETS Instrument. Two dimensions in SETS-MS Instrument are comprised in a way to cover target behaviours that are paired with Level A and B at the basic framework by taking existence of such trajectory into consideration. Additionally, SETS HS Instrument have three dimensions examining skills regarding Level C as well as these levels.

It is especially need for students to develop more statistical skills in middle school in which adult individuals start to form their statistical understanding (Callingham, & Watson, 2017). In achieving such a development, examining levels of prospective teachers' self-efficacy to teach statistics at this grade can be guiding step (Akoğlu, 2018). In this sense, SETS-MS Instrument is considered to serve for such target. This instrument has a pioneering role for laying foundation of SETS Instrument and is paired up with SETS-HS Instrument that is developed followingly on psychometric properties (Harrell-Williams, Lovett, Pierce, Sorto, Lee, & Lesser, 2017). In other words, SETS Instrument can be used in studies to be conducted at high school grade. Consequently, levels of prospective middle school and high school grades teachers' self-efficacy to teach statistics can be achieved easily by using SETS Instrument. It is understood here that SETS-MS has a strong construct to meet information to be collected from the entire SETS Instrument. Thus, validation and reliability studies are intended to be made by adapting this instrument into Turkish within framework of structure of Turkish Education System in current research. Considering statistics concepts that are taught intensely at middle school grades, this research is considered to be important to examining prospective mathematics teachers' self-efficacy who is to come across these concepts excessively.

2. Method

2.1. Sample

Participants of the study were 421 prospective middle grades mathematics teachers from five different state university. Participants were chosen from 4th grade prospective teachers of their undergraduate programme with the criteria that is to have knowledge and skills regarding statistics. %69,1 (n= 291) of participants are female and %30,8 (n=130) of them are male. According to Kline (2011), the sample size should be as many as ten times the number of items and should not be worked with less than 200 cases. Accordingly, it can be expressed that participants of this study are considerably more than the defined criteria for twenty-six items of SETS-MS Instrument which is planned to be made as adaption.

2.2. Data Collection

Turkish adaption of SETS-MS Instrument is targeted to be made in this study. In line with this, primarily information regarding original form of SETS-MS Instrument is provided as follows. Subsequently information regarding translation process of Turkish form on which validation and reliability studies of this instrument is conducted is provided.

2.2.1. SETS-MS Instrument

Original form and psychometric properties of the instrument focusing on statistics teaching at middle-school grade are accessed through an article publishing items of the instrument (Harrell-Williams, Sorto, Pierce, Lesser, & Murphy, 2015). SETS-MS is consisted of two dimensions as reading the data (11 items) and reading between the data (15 items) and twenty-six items. Based on features reflected by items, items are addressed toward certain statistics concepts taught in middle school and they are rated on 6-point ranging from not at all confident (1) to completely confident (6). The instrument is developed through data collected from 309 prospective teachers who participated in the study and construct validation of instrument is examining via exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) based on Rasch model. The result of analysis shows that the instrument has a two-dimensional construct (reading the data and reading between the data). The correlation coefficient between sub-dimensions is estimated as 0.85. Additionally, interpretations toward the reliability of the instrument are made based on the Cronbach Alpha coefficient. This coefficient is estimated as 0.94 for the whole instrument, 0.87 for reading the data and 0.91 for reading between the data.

2.2.2. Translation of SETS-MS Instrument into Turkish

Prior to adaption process, Leigh M. Harrell-Williams who is the corresponding author is contacted via e-mail and necessary permissions are taken to adapt the instrument into Turkish. The instrument whose original language is English is translated into Turkish separately by three independent translators who are professionals in English. One of the translators works at Mathematics Education, one at English Language and Literature and the other one at Educational Sciences. Common evaluation of three different translations made by translators is

made by two mathematics educators who have content knowledge on statistics. In case of any disparity during the evaluation, it is consulted with opinions of authors of the instrument. Thus, Turkish draft form is created. Followingly, to enable suitability of language and expression of the draft form, two Turkish language educators are consulted. In line with opinions of professionals, Turkish draft form is rearranged. Rearranged draft form is presented to two statistics educators to back-translation. The back-translations made by two professionals are rated with 3-point Back-Translation Agreement Form arranged by researchers by a third professional on English. An example part of the form is presented as follows.

Original Item	Agreement Point			Back-Translation
	1	2	3	
Identify the association between two variables from scatterplots.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Determine the relationship between two variables in a scatter plot.
Generalize a statistical result from a small group to a larger group such as the whole class.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Generalize a statistical results from a small group to bigger one.

Figure 2. Sample from the Back-Translation Agreement Form

The professional who rated back-translations compared items of the original instrument with back-translated items of Turkish draft form arranged by professionals and examined their agreement. Data collected as result of all this process for all items is analysed through Cohen’s Kappa coefficient. This coefficient measures the reliability of agreement between two raters by removing chance factor and it is in range of -1 and +1 (Karagöz, 2017). In accordance with this, -1 represents negative perfect agreement and +1 does positive perfect agreement. it is concluded that estimated Kappa coefficient is agreement with back-translations of professionals in the research ($p < 0,001$). In addition, percentage of agreement is (%70,2) and it proves that back-translations of professionals are highly agreement.

A pilot study was conducted with 20 prospective mathematics teachers to examining its comprehensibility of Turkish draft form by participants. As a result of the pilot study, prospective teachers’ comprehend toward items are evaluated with Turkish language educators and comprehensibility of items can be increased by expressing items which have – m(a)k verb root ((to) verb) with –(e)bilirim (to be able to). In line with this, items in Turkish draft form are rearranged in the way stated before (for example; the item that is written as to describe numerically the variability between individuals within the same group is rearranged as I can describe numerically the variability between individuals within the same group). 6-point of instrument are rearranged as stated as well. Thus, finalized version of Turkish draft form called Ortaokul Düzeyinde İstatistiği Öğretmeye yönelik Öz yeterlik [İÖÖ-OD] Ölçeği can be used to conduct validity and reliability studies.

2.3. Ethics and Study Process

Data used in the research is collected via online or in person with ÖÖ-OD Instrument in accordance with necessary permissions of five different state university. The fact that the research is conducted within ethics rules is confirmed by Trabzon University Social and Human Sciences Ethics Committee. During the research, prospective teachers are provided with necessary information on the research and it is stated to them that their personal information shall not be shared with third person or party and only be used with scientific aim. Prospective teachers who participated the research based on volunteering respond the instrument within approximately 20 minutes.

2.4. Data analysis

Within the research, CFA is applied to test construct validity of İÖÖ-DD Instrument which is adapted into Turkish. Namely, CFA is an analysis that explains that to the extent to which association models determined beforehand between factors and indicators are confirmed through new data collected (Çokluk, Şekercioğlu, & Büyüköztürk 2016). Particularly in intercultural instrument adaption studies, factor structure in the target culture of the instrument can be revealed directly through CFA rather than EFA. Since factor structure of the instrument mentioned in related culture is essentially proved as experimental in many quantitative and qualitative studies. Consequently, CFA may be enough at the point of whether determined factor structure is preserved under target culture. As for testing the model with CFA, many fit indexes and parameters that they have are taken into account. Aside from having no accuracy to use which fit indexes in studies conducted, 4 fit indexes are recommended to interpret in structural equation model (SEM). These indexes are as follows: Chi-Square Goodness of fit (χ^2), Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean Square

Residual (SRMR) and Comparative Fit Index (CFI) (Kline, 2011). Additionally, Ratio of Chi-Square Goodness of fit to the degree of freedom (χ^2/sd), Goodness of Fit Index (GFI) and Tucker Lewis Index (TLI) are generally preferred in analysis. All fit indexes mentioned are used in this study. On the other hand, Cronbach Alpha coefficient, which evaluates the internal consistency of items is applied for the reliability of the instrument. In this sense, the Cronbach Alpha coefficient is estimated separately for the whole instrument and each dimension in the study. Item-total correlation is estimated with the intention of determining to what degree items in the instrument are functional to discriminate individuals and an independent-sample t-test is conducted for the difference between mean scores of the lower 27% and upper 27% groups (Büyüköztürk, 2011).

3. Findings

3.1. Validation

CFA is applied to examine to which degree original construct of SETS MS Instrument confirm in Turkish culture and ensure construct validation of the instrument. Prior to analysis data set is checked on the aspect of missing data, outliers and normality.

As a result of CFA applied, it is determined that the factor loadings of i25 (I can distinguish between association and cause and effect) and i26 (I can recognize sampling variability in summary statistics such as the sample mean and the sample proportion) items in the İÖÖ-OD Instrument are low and not statistically significant. In addition, fit indices are determined to be not good as well. Thus, these items are decided to be excluded from the analysis. Excluding item that are not significant from the analysis one by one and examining fit indices and factor loadings of items each time is important to achieve more accurate results (Karagöz, 2017). Thus, exclusion of items process is conducted respectively starting from the item having the lowest factor loadings and fit indices are examined each time. As a result, all factor loadings of items in the instrument is observed to be statistically significant with the exclusion of the mentioned two items ($p < 0,01$). Followingly, fit indices are estimated for the two-dimensional model as it is done in the original instrument. Accordingly, fit indices reveal that fit of the model formed is not good. Fit of the model can be corrected through covariance between items with the condition to be in accordance with theoretical construct (Karagöz, 2017). In line with this, it is determined that some items (i21-i23, i22-i24, i20-i22, i21-i22, i22-i23) are placed under the same latent variable and error covariations between these items are correlated considerably. In accordance with the opinions of the professionals, high-level error correlations between these items couples are included in the model and the model is decided to be tested again. Thus, fit indices made as a result of modification reveal it to be fitting enough with the original instrument and confirm the factor structure. Information related to values prior to exclusion of items with criteria for fit indices and cut-off points for acceptable (CFA I), fit indices achieved following the exclusion (CFA II) and fit indices achieved as a result of modifications made model (CFA III) are presented comparatively as follows in Table 1.

Table 1. Cut-off points for acceptable and criteria for fit indices of the SETS-MS Instrument

Fit Index	Criteria	Cut-off Point for Acceptance	CFA I	CFA II	CFA III
χ^2	$p > 0.05$	-	$\chi^2=1010,735$ (sd=298, $p=0.00$)	$\chi^2=707,126$ (sd=251, $p=0.00$)	$\chi^2=537,987$ (sd=246, $p=0.00$)
χ^2/sd	-	$\leq 2=$ perfect fit $\leq 5=$ moderate fit	3.39	2.81	2.18
GFI	0 (none) 1 (perfect fit)	$\geq 0.90=$ good fit $\geq 0.95=$ perfect fit	0.84	0.86	0.90
CFI	0 (none) 1 (perfect fit)	$\geq 0.90=$ good fit $\geq 0.95=$ perfect fit	0.80	0.86	0.91
TLI (NNFI)	0 (none) 1 (perfect fit)	$\geq 0.90 =$ good fit $\geq 0.95=$ perfect fit	0.78	0.84	0.90
SRMR	0 (perfect fit) 1 (none)	$\leq 0.05 =$ perfect fit $\leq 0.08 =$ good fit	0.06	0.05	0.04
RMSEA	0 (perfect fit) 1 (none)	$\leq 0.05 =$ perfect fit $\leq 0.08 =$ good fit	0.07	0.06	0.05

Note: The fit indices and cut-off points for acceptable are retrieved from Çokluk et al. (2016).

As examining the fit indices of the model in Table 1, Chi-Square values are deemed to be significant in all three analyses ($p < 0.001$). In addition, these values are observed to get lower with repeating the analysis. Chi-

Square value is considerably affected by sample size and significant results are usually produced in great sample. Thus, χ^2 /sd ratio that is less affected from sample size is considered to be used as criteria instead of it (Şimşek, 2007). In Table 1, it is deemed that this value starts to get close to perfect fit with exclusion of the items. In fact, results of modifications made in the model cannot completely achieve perfect fit with the original instrument, however, the achieved value is considered to be a proof for the current fit to be close to perfect fit. Additionally, GFI, CGI and TLI fit indices are achieved to be good fit only following of results of modifications made in the model. Namely, it can be expressed for the current form of the model to meet cut-off points determined for fit in minimum and offer enough fitting with the original instrument. Similarly, RMSEA fit indices is observed to be good fit in analysis conducted prior to modifications and achieve perfect fit with modifications. As for SRMR fit indices in the analysis, they are achieved perfect fit only with exclusion of items. In fact, a certain fit can be expressed to some degree prior to exclusion of items, however, this fit is observed to be at a good and not to be able to reach perfect. Accordingly, as a result of CFA III which fulfils all criteria determined for fit with the original instrument, finalized version of İÖÖ-OD Instrument is obtained. Path diagram regarding the model formed is presented as follows.

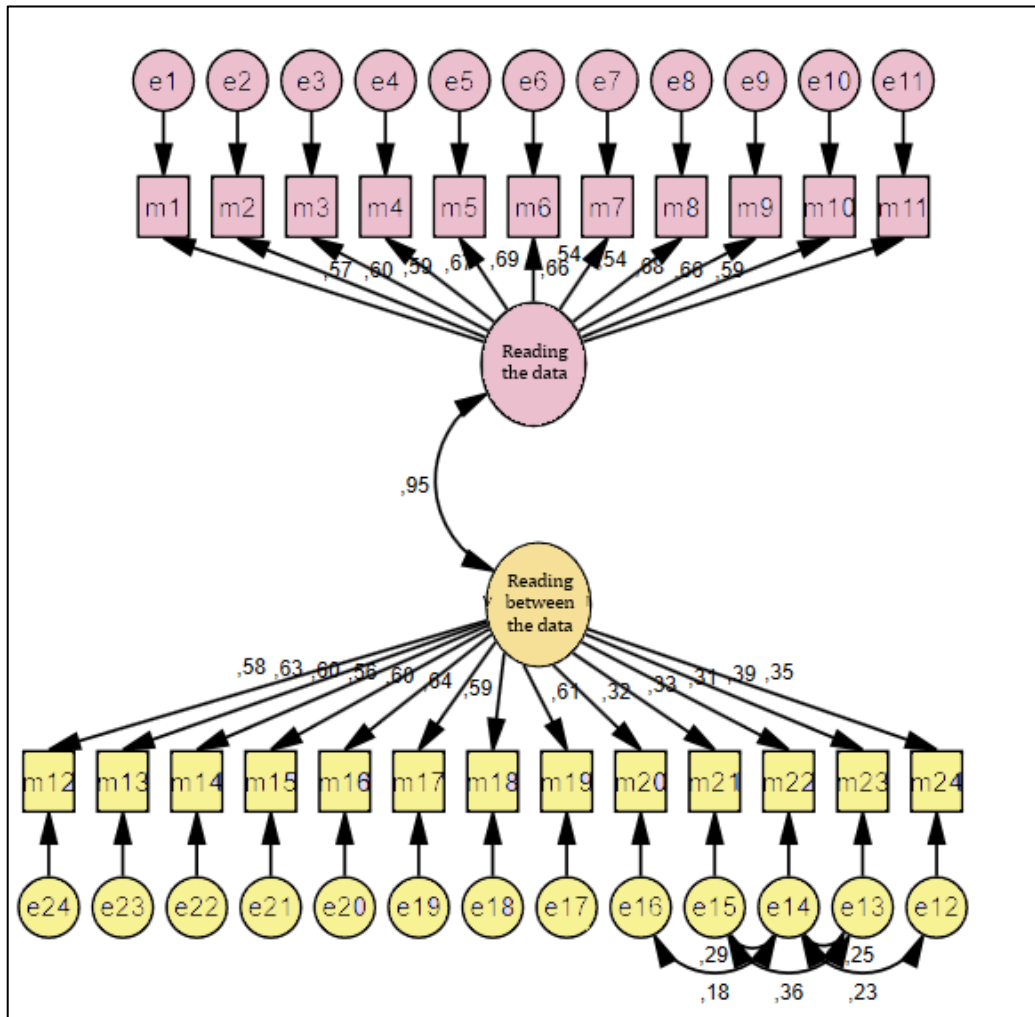


Figure 3. Path diagram for the two-dimensional model $n=421$, $\chi^2=537,987$ $sd=246$, $p<0.001$

As examining the model, it is observed that the factor loadings of items within reading the data dimension vary from 0.54 to 0.69, the factor loadings of items within reading between the data dimension vary from 0.31 to 0.64 and the factor loadings of all items is statistically significant ($p<0.01$). The factor loadings of the items are usually preferred to be 0.45 or above, yet it is stated that items whose factor loadings are 0.30 can be included in the instrument as well (Kline, 2011). It can be observed that the factor loadings of the items included in the model are not below 0.30. This situation reveals that items have fit factor loadings with the criteria and reflects the factors that they are based on enough. As for correlation coefficient between factors, the value is determined to be 0.95. It is preferred a condition that the correlation between factors which is not in range of extreme values and is statistically significant in instrument (Gündüz, & Coşkun, 2012). Consequently, it can be expressed that correlation between factors of the instrument ($r=0.95$, $p<0.01$) is not placed in the range of extreme values and statistically significant.

3.2. Reliability

Cronbach Alpha coefficient calculated for each factor and composite score of the instrument to determine the reliability of İÖÖ-OD Instrument. Accordingly, it is estimated as reliability coefficient of the first factor (reading the data) which includes eleven items of the instrument is 0.87; the reliability coefficient of the second factor (reading between the data) which includes thirteen items is 0.82. The reliability coefficient calculated for the whole the instrument is determined to be 0.91. Preferred in context of internal consistency in scientific studies, Cronbach Alpha coefficient is accepted to be 0.70 or above (Şencan, 2005). In line with this, İÖÖ-OD Instrument is deemed to have a reliable construct. On the other hand, (i) independent t test based on the difference between item mean scores of lower 27% and upper 27% groups and (ii) item-total correlation values are calculated to examine to what degree items in the instrument can discriminate individuals. Achieved results are presented in Table 2.

Table 2. Independent t test for lower 27% and upper 27% groups and Results and Item-total Correlations

Factor	Item	Item-total Correlation	t (lower 27%- upper 27%)	Factor	Item	Item-total Correlation	t (lower 27%- upper 27%)
Reading the Data	m1	0.57	12,289	Reading Between the Data	m12	0.58	13,383
	m2	0.60	13,201		m13	0.62	15,486
	m3	0.58	12,069		m14	0.63	12,585
	m4	0.65	15,217		m15	0.58	12,651
	m5	0.68	15,956		m16	0.60	14,264
	m6	0.66	13,796		m17	0.63	13,605
	m7	0.55	12,303		m18	0.59	12,752
	m8	0.55	11,848		m19	0.60	13,631
	m9	0.68	17,284		m20	0.39	7,850
	m10	0.65	16,444		m21	0.42	9,324
	m11	0.61	14,834		m22	0.42	8,149
					m23	0.48	11,775
					m24	0.44	9,416

$p < 0.01$

As examining Table 2, item-total correlations of items covered in the instrument are observed to vary from 0.39 to 0.68. Item-total correlations being 0.30 and above indicates discrimination to be good (Büyüköztürk, 2011). Consequently, SETS-MS Instrument is expressed to be convenient to discriminate individuals. Additionally, significant difference is determined for all items according to results of t test conducted for difference between mean scores of lower 27% and upper 27% groups ($p < 0.01$). This situation indicates that items within each factor covers behaviors reflecting related factor. Accordingly, twenty-four itemed İÖÖ-OD Instrument is presented in Appendix 1.

4. Discussion, Conclusion and Suggestions

SETS Instrument is developed with the intention of measuring prospective teachers' self-efficacy to teach statistics. This study targets that the adaption of the instrument focusing on statistics teaching at middle school grade of SETS Instrument is made within the framework of structure of Turkish Education System by conducting reliability and validation studies for it. Accordingly, findings related to construct validation of the instrument are achieved through CFA and findings related to reliability are achieved through Cronbach Alpha coefficient, which evaluates the internal consistency of items, independent sample t-test based on the difference between the mean scores of the lower 27% and upper 27% groups and item- total correlations.

For the first part of the study, fit indices calculated according to results of CFA I are determined not to be good. The factor loadings of the i25 and i26 coded items covered in the instrument are deemed not to be significant. Namely, significant of factor loadings is a significant indicator that shows items are loaded into factors rightly. Thus, not significant items are required to be excluded from the analysis (Karagöz,2017). Consequently, these items are decided to be excluded from the analysis by concluding that the items do not reflect enough the dimension (reading between the data) covered in the original instrument. As a result of CFA II conducted over 24 items, it is determined that factor loadings of all items covered in the instrument are significant, yet fit indices of the Turkish version of the instrument are not fitting with the required level in the original instrument. Thus, modifications that are recommended based on analysis result are decided to be made. Namely, arrangements are important for a better fit of the model. However, these arrangements are required to

be made in accordance with theoretical construct. Correlation of error terms of items related to the same factor is made (Karagöz, 2017). Thus, analysis is repeated by forming covariances between suggested doubles with the condition of being in accordance with theoretical construct. Accordingly, fit indices achieved through CFA results ($\chi^2= 537,987$, $p<0.01$, $\chi^2/sd= 2,18$, $GFI=0,90$, $CFI=0,91$, $TLI=0,90$, $SRMR= 0,04$, $RMSEA=0,05$) demonstrate that values related to the model are within acceptable limit. Consequently, two-dimensional construct of İÖÖ-OD that covers 24 items is concluded to be confirmed. The factor loadings of the items in the instrument vary from 0.31 to 0.69 and covers behaviours reflecting the related factor. Correlation coefficient between factors is calculated to be 0.95. Including the value calculated for the original form of the instrument, this value is expressed to be high for both forms.

Result of the reliability analysis made in the study demonstrates its Cronbach Alpha coefficient to be 0.87 for the reading the data factor of the instrument and 0.82 for the reading between the data. The coefficient calculated for the whole instrument is determined to be 0.91. These are considered to be close to values estimated for the original form of the instrument. Results of independent sample t test which is calculated based on the difference between mean scores of the lower 27% and upper 27% groups reveal the difference between mean of scores to be significant. The item-total correlations of items in the instrument are observed to vary from 0.39 to 0.68. In line with this, these values are observed to be below the values estimated for the original form of the instrument, however, items covered in the Turkish form can be indicated to able to discriminate for individuals.

As a result, İÖÖ-OD Instrument which is Turkish adaption is deemed to consist of two dimensions as the original instrument does, fit with data collected through participators of the study and have enough levels for the instrument's internal consistency coefficient. Thus, this instrument can be used in studies measuring prospective teachers' self-efficacy to teach statistics at the middle grades. This study is considered to be important for supporting results related to the reliability and validation of the instrument. As it is in the original instrument, validation and reliability studies can be conducted in Turkish form to be used for different education grades as well. On the other hand, the study is considered to inspire researchers working in the field to conduct studies to develop a genuine instrument in accordance with statistics teaching provided at middle school grade in Turkey.

Ortaokul Düzeyinde İstatistiği Öğretmeye Yönelik Öz Yeterlik Ölçeği'nin Türkçe'ye Uyarlanması

1. Giriş

Özellikle sosyal medyanın yaygınlaşmasıyla birlikte istatistiğe ilişkin kavramlar, yaşantımız üzerinde daha fazla görünür ve etki eder hale gelmiştir. 21. yüzyıl insanı veri sağanağı altında hayatını sürdürmek durumundadır. Çoğu zaman bir veri akışı içinde yaşadığımız hatta düşüncelerimizin verilerle manipüle edildiği hissi, çağımız insanının günlük yaşamının doğal bir parçası haline gelmektedir (Garfield, & Ben-Zvi, 2008). Günümüz insanının yaşamında oldukça etkili olan bu verilerle baş edebilme becerilerinin geliştirilmesi, eğitimin öncelikli hedeflerinden biri olduğu rahatlıkla söylenebilir. Böyle becerilerin gelişiminin sağlanması ve istatistiksel açıdan donanımlı bireylerin yetiştirilmesinde öğretmenlerin sahip olduğu yeterlikler oldukça önemlidir. Özellikle kendi öğrencilik yaşantılarında yoğun bir istatistik eğitime maruz bırakılmamış olan bu öğretmenlerin istatistiği öğretmeye yönelik ciddi kaygıları olduğu bilinen bir gerçektir (Steinberger, 2020). Kaygının kaynağı görünürde alan ve alanı öğretme bilgisi gibi gözükse de, aslında daha derinlerde öğretmenin istatistiği öğretme ile ilgili kendine yönelik öz yeterliği çoğu zaman alan bilgisinin önüne geçebilmekte, öğretmenin sınıf ortamındaki öğretim uygulamalarına yön verebilmektedir.

Öz yeterlik, temeli Bandura'nın sosyal öğrenme kuramına dayanan ve bireylerin ilgili alanda sahip olduğu becerileri etkin bir şekilde kullanmaları için öncelikle o alana ilişkin öz güven sahibi olmaları gerektiğine vurgu yapan önemli bir kavramdır (Pajares, 2002). Bandura (1997) öz yeterliği, bireyin bir işi yapmak için gerekli olan eylem serisini organize etme ve yürütme yeteneğine olan inancı olarak tanımlamış ve bu inancın bireyin davranışı üzerinde ilgili alandaki yeteneğinden daha fazla etkiye sahip olduğunun da altını çizmiştir. Nitekim güçlü bir öz yeterliğe sahip olan bireylerin zayıf olanlara kıyasla karşılaştıkları görevleri üstlenmede daha istekli davrandığı, zorluklar karşısında pes etmeyerek daha dayanıklı ve ısrarcı oldukları bilinmektedir (Arseven, 2016). Bu anlamda öz yeterliği yüksek olan öğretmenlerin de sınıf ortamındaki durumlara karşı daha sorgulayıcı davrandığı, daha çok öğrenci merkezli yöntemlere başvurduğu, düşük olanların ise genellikle öğretmen merkezli bir anlayış içerisinde olduğu gözlenmektedir (Czerniak, 1990). Pek muhtemeldir ki, bu tür durumlar aslında öğretmenlerin öğretmeye yönelik öz yeterliklerinden kaynaklı olarak ortaya çıkmaktadır.

Öz yeterliğin bir boyutu olarak görülebilecek öğretmeye yönelik öz yeterlik, öğretmenin öğretme yeteneğine ilişkin inancı olarak tanımlanmaktadır (Wheatley, 2005; Zee, & Koomen, 2016). Bu kavramın bir alana özgü olarak ele alınması ve öğretmenin o alanı öğretme yeterliğine odaklanması, gözle görülür bir gelişimin sağlanması adına oldukça önemlidir (Bandura, 2006). Bu doğrultuda özellikle öğrencilerde başarısızlık algısının ciddi ölçüde yüksek olduğu matematikte, öğretmenlerin matematiği öğretmeye yönelik sahip oldukları öz yeterlik, hedeflenen öğrenme çıktılarına ulaşılması için ön koşul niteliğindedir (Dede, 2008). Ancak, öğretmenlerin matematik öz yeterlikleri ile matematiği öğretmeye yönelik öz yeterliklerinin birbiri yerine kullanıldığı, dolayısıyla iki kavram arasındaki ayrımın netleştirilmesi gerektiği vurgulanmaktadır (McGee, & Wang, 2014). Şöyle ki; öğretmenlerin matematik öz yeterlikleri, matematikle ilişkili görevleri yerine getirme yeteneklerine olan inancı açıklarken; matematiği öğretmeye yönelik öz yeterlikleri ise, öğretmenlerin başkalarına matematik öğretme yeteneği ile ilgili inançları temsil etmektedir (Kahle, 2008). Bir öğretmenin matematik öğretim sürecinde oluşacak durumları anlama, yorum getirme ve çözüm önerisi üretmesinde sahip olduğu alan ve alanı öğretme bilgisinin yanında matematiği öğretmeye yönelik öz yeterliğinin de gelişmiş olması beklenmektedir (Mesquita, & Drake, 1994). Söz gelimi Göloğlu-Demir ve Çetin (2010) öğretmenlerin matematik öğretimine ayıracağı zaman, göstereceği çaba, yeni yöntem ve teknikleri uygulama istekliliği ve özel öğrencilere farklı bir program hazırlama gayretlerini matematiği öğretmeye yönelik öz yeterliklerinin bir yansıması olarak görmektedir. Bu sebeple, öğretmenlerin matematik öğretimine ilişkin öz yeterlik düzeylerinin belirlenmesi büyük önem kazanmakta ve bunun için çeşitli ölçme araçları geliştirilmektedir (Enochs, Smith, & Huinker, 2000; McGee, & Wang, 2014; Gerez-Cantimer, Şengül, & Akçin, 2020). Sayıca az olan bu ölçme araçları öğretmenlerin matematiği öğretmeye yönelik öz yeterliklerine bir çerçeve sunması bakımından önemlidir.

1.1. İstatistiği Öğretmeye Yönelik Öz Yeterlik

İstatistik her ne kadar matematik eğitiminin bir parçası olarak karşımıza çıksa da, bu disiplinin temel kavramlarını anlamak için iyi bir matematik bilgisi tek başına yeterli değildir (Begg, & Edwards, 1999). Çoğu araştırmacı istatistiğin işlemsel temellerini matematikten aldığı; buna karşın kendi düşünme sistemi ile matematikten farklılaştığı konusunda hem fikirdir (delMas, 2004; Groth, 2007; Garfield, & Ben-Zvi, 2008). Aldıkları eğitim gereği, matematiğin tümdengelimli yapısını merkeze alan bir düşünme formasyonuna sahip olan matematik öğretmenlerinin, veriye dayalı bir anlayışa sahip, yorumların bağlam üzerinden şekillendiği istatistik alanını hem anlama da hem de öğretilmede güçlük yaşamaları son derece doğaldır. Alan yazında matematik öğretmenlerinin istatistiği anlamada zorluk yaşadıklarını gösteren kanıtlar mevcuttur (Shaughnessy, 2007; Fitzmaurice, Leavy, & Hannigan, 2014). Bu zorlukların nedeni bir yandan istatistiğin matematikten ayrışan

doğasından kaynaklanırken bir yandan da öğretmenlerin sahip olduğu istatistik bilgisinin yanında bu alandaki öz yeterliklerine bağlı olarak yorumlanmaktadır (Irakleous, & Panaoura, 2015). Nitekim öğretmenler, istatistik alanını öğretme konusunda kendilerine daha az güven duyma eğilimindedir (Watson, 2013). Örneğin; Fitzmaurice, Leavy ve Hannigan (2014) gerçekleştirdiği mülakatlar neticesinde, matematik öğretmen adaylarının istatistiği zor olarak gördüğü ve bunun istatistiği öğretmeye yönelik öz yeterliklerine yansıdığı sonucuna ulaşmıştır. Öğretmenlerin istatistiği öğretmeye ilişkin kendilerini yeterli görmeyişi, bu konunun okul yıllarının sonunda ve kalan birkaç derse sığdırılarak (Callingham, & Burgess, 2014) geçiştirilmesine de önemli bir dayanak oluşturmuştur. Gal (2004) istatistiğin öğretim programlarına işlenmesinin önündeki en büyük engellerden birinin, öğretmenlerin istatistik alanına yönelik eğilimi (daha özel olarak öz yeterliği) olduğuna işaret etmektedir. Bu eğilim, öğretmenlerin istatistiği gerçekten etkili bir şekilde uygulamaya ilişkin yorumlarını, motivasyon ve çabalarını etkilemektedir (Pierce, & Chick, 2011; März, & Kelchtermans, 2013). Dolayısıyla öğretmenlerin istatistiği öğretmeye yönelik öz yeterliklerinin irdelenmesinin öğretim uygulamaları ve öğrencilerin başarısı adına önemli olduğu düşünülmektedir.

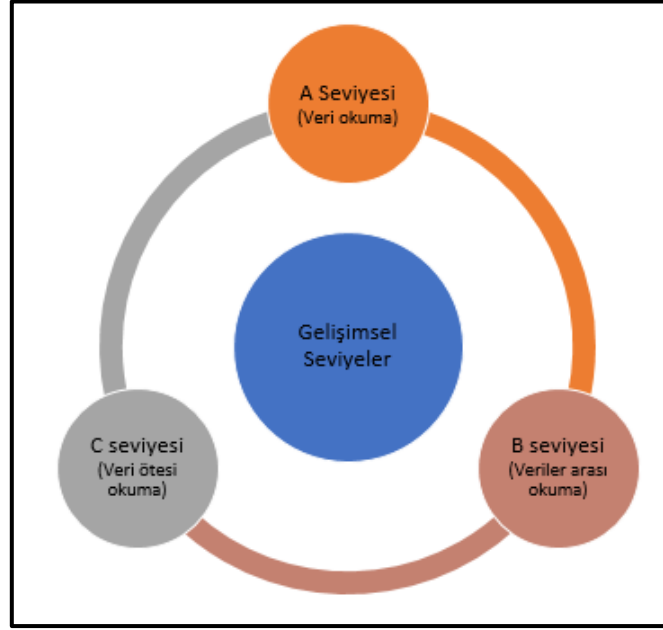
Matematiğe kıyasla öğretmenlerin istatistiği öğretmeye yönelik öz yeterliklerinin daha karmaşık bir yapıya sahip olduğu belirtilmektedir (Harrell-Williams, Sorto, Pierce, Lesser, & Murphy, 2014). Çünkü öğretmenler üniversiteden istatistik öğretimine hazır ve donanımlı olarak ayrılmamalarına (Wessel, 2014) rağmen okul düzeyinde öğretilmeleri beklenen bir istatistik programıyla karşı karşıya bırakılmaktadırlar. Bu nedenle öğretmenlerin matematiği öğretmeye yönelik öz yeterliklerinin ölçülmesi sonucunda elde edilecek olan sonucun, istatistik için de geçerli olacağı sonucuna varmak anlamlı değildir. Bir başka tabirle, öğretmenlerin istatistiği öğretmeye yönelik öz yeterlikleri hakkında fikir sahibi olunması, onların matematiği öğretmeye yönelik öz yeterliklerinden oldukça farklı olabilmektedir. Hem kendi doğaları hem de öğretimleri açısından çok sayıda farklılık içeren bu iki alana ilişkin öz yeterliğin ölçülmesi, farklı ölçme araçlarını gerektirmektedir. Bu anlamda istatistik alanında yapılan ölçek geliştirme çalışmaları incelendiğinde; genellikle istatistiğe yönelik öz yeterliğe odaklanıldığı görülmektedir (Finney, & Schraw, 2003; Lane, Hall, & Lane, 2004; Mercimek ve Pektaş, 2013). Bununla birlikte, daha özele inerek istatistik okuryazarlığına yönelik öz yeterliği irdeleyen ölçek geliştirme çalışmaları da mevcuttur (Carmichael, & Hay, 2009; Lin, & Huang, 2013; Batur, Yiğit, & Baki, 2019). Bu ölçekler neticesinde, öğrencilerin, öğretmenlerin ve öğretmen adaylarının istatistiğe yönelik öz yeterlikleri hakkında önemli fikirler elde edilse de, istatistik öğretimine ilişkin öz yeterliğin saptanması noktasında oldukça az sayıda ölçme aracı geliştirilmiştir. Bu sınırlı çalışmalardan biri olarak Watson (2001), öğretmenlerin istatistik ve olasılığı öğretmeye yönelik becerileri ve güvenlerini ortaya çıkaracak bir ölçme aracı geliştirmiştir. İlk ve ortaokul düzeyinden 43 öğretmenin katılımıyla gerçekleştirilen araştırmanın sonucunda, öğretmenlerin en fazla grafikler, en az ise olasılık konularını öğretmeye yönelik güvene sahip olduğu görülmüştür. Watson'un ölçme aracının bir bakıma öğretmenlerin istatistiği öğretmeye yönelik öz yeterliklerinin irdelenmesine temel oluşturduğu düşünülse de bu aracın içerdiği maddelerin tam olarak buna odaklanmayışı (ayrıca öğretmenlik deneyimi, öğretim uygulamaları vs. bilgiler içermesi) önemli bir eksiklik oluşturmaktadır.

1.2. SETS Ölçeği

2008 yılında çalışmalarına başlanan SETS Ölçeği, matematik öğretmen adaylarının istatistiği öğretmeye yönelik öz yeterliklerini değerlendirmek amacıyla tasarlanmıştır. Gerek istatistiğin doğasının karmaşıklığı gerekse de bu alana yönelik öz yeterliği ölçmenin güçlüğü nedeniyle hâlihazırda istatistiği öğretmeye yönelik öz yeterliğe odaklanan ölçme araçlarından önemli ölçüde ayrılmaktadır. Öğretmenin öz yeterliğinin sınıfa girmeden önceki dönemde geliştiği (Smith, 1996) ve bundan sonraki süreçte değişime oldukça dirençli olduğu bilinmektedir (Woolfolk Hoy, & Burke-Spero, 2005). Bu nedenle, SETS Ölçeği'nin öğretmen adaylarına yönelik olmasının, öğretmenlerin istatistik öğretimine yönelik öz yeterliğinin önceden belirlenmesi ve gelişimine çalışılması açısından önemli bir avantaj sağlayacağı düşünülmektedir. Nitekim Tschannen-Moran ve Woolfolk Hoy (2001) etkili, kararlı ve hevesli öğretmenler yetiştirmek için öğretmenlerin öz yeterlik algılarının gelişiminin desteklemesinin önemine işaret etmiştir. Bu bakımdan SETS Ölçeği'nden edinilecek bilgilerin, ileriki yıllarda gerçekleştirilecek olan istatistik öğretiminin etkililiği adına önem taşıdığı söylenebilmektedir. Bununla birlikte, öğretmen öz yeterliğinin belli bir alana özgü olarak ölçülmesi gerektiğini savunan Bandura (2006) öz yeterlik yapısı, SETS Ölçeği'nin temelini oluşturmaktadır (Lovett, 2016). Bu bakımdan ölçek, doğası gereği matematikten ayrılan istatistiğin sınırlarını net olarak çizmesi ve özel olarak istatistik öğretimine ilişkin öz yeterliğe odaklanması açısından da ön plana çıkmaktadır.

SETS Ölçeği, ortaokul (SETS-MS; Harrell-Williams vd., 2014) ve lise (SETS-HS; Harrell-Williams, Lovett, Lee, Pierce, Lesser, & Sorto, 2017) kademelerinde olmak üzere iki farklı ölçeğe ayrılarak ele alınmıştır (Akoğlu, 2018). SETS-MS Ölçeği iki faktörlü yapıya sahipken, SETS-HS Ölçeği bu faktörlere ilave olarak üçüncü bir faktöre daha sahiptir. SETS Ölçeği'nde yer alan bu faktörler belirlenirken, etkili bir istatistik öğretiminin nasıl yapılabileceğine ilişkin önemli tavsiyeler içeren Guidelines for Assessment and Instruction in Statistics Education [GAISE] (Franklin vd., 2007) raporunun önerdiği üç seviyenin (A, B ve C) göstergeleri esas alınmıştır. Ölçekte yer alan bu seviyeler, içerdikleri beceriler bakımından ayrıca bireylerin grafiklere dair bilgisi, istatistiksel düşünme ve muhakeme gibi birçok becerisinin sorgulanmasında anahtar rol üstlenen Friel, Curcio ve

Bright'ın (2001) grafik okuryazarlığı seviyeleri (veri okuma, veriler arası okuma ve veri ötesi okuma) ile eşleştirilmiştir. İstatistik eğitiminde başlı başına iki kaynak doküman olarak gösterilen bu yapıların bir araya getirilmesi, SETS Ölçeği'ne dayalı yapılacak incelemelerin istatistik öğretimine ilişkin öz yeterlik hakkında derinlemesine bilgi sağlayacağına işaret etmektedir. Bu anlamda içerdiği faktörlerin güçlü bir alt yapı üzerine inşa edildiği bu ölçeğin temel çatısı aşağıdaki gibi şematize edilebilir.



Şekil 1. SETS Ölçeği'nin Temel Çatısı (URL-1, 2020)

A seviyesi (veri okuma), temel düzeydeki istatistik kavramlarının (örneğin, bir veri grubunu özetlemek için tablo ve grafikleri kullanma, toplanacak veriler arasındaki değişkenliğin farkına varma gibi) öğretimine ilişkin öz yeterliği yansıtmaktadır. B seviyesinde (veriler arası okuma), istatistiğin kavramları arasında bir ilişkilendirmenin yapıldığı ve karşılaştırmaların ileri sürüldüğü durumların (örneğin, farklı veri gruplarının dağılımlarını karşılaştırmak, değişkenler arasındaki korelasyon katsayısını yorumlamak gibi) öğretimine ilişkin öz yeterlik beklenmektedir. C seviyesinde (veri ötesi okuma) ise, ileri düzeyde istatistiksel çıkarımların yapılması ve istatistiksel sonuçlar üzerinden bir genellemeye gidilmesini (örneğin; örneklemden hareketle popülasyon ortalamasına dair kestirim yapma, örneklem ortalamaları arasındaki farkın anlamlı olup olmadığına karar verme gibi) gerektiren durumları öğretmeye yönelik öz yeterliğe odaklanılmaktadır. Bu anlamda SETS Ölçeği'nde; A seviyesinde temel becerilerden yola çıkarak, sırasıyla B ve C seviyelerinde daha üst düzey becerilerin öğretimine yönelik öz yeterlik algısının sorgulanmasına uzanan gelişimsel bir çizginin olduğu görülmektedir. Böyle bir çizginin varlığı gözetilerek SETS-MS Ölçeği'ndeki iki faktör temel çatıdaki A ve B seviyelerine karşılık gelen hedef davranışları içerecek şekilde oluşturulmuştur. Bununla birlikte, SETS-HS Ölçeği ise, bu seviyelere ilave olarak C seviyesine yönelik becerileri de sorgulayan üç faktöre sahiptir.

Özellikle yetişkin bireylerin istatistiksel anlayışlarının oluşmaya başladığı ortaokul yıllarında öğrencilerin istatistiksel becerilerinin daha fazla geliştirilmesine ihtiyaç duyulmaktadır (Callingham, & Watson, 2017). Böyle bir gelişimin sağlanmasında ise, bu kademedeki öğretmen adaylarının istatistik öğretimine ilişkin öz yeterlik düzeylerinin belirlenmesi yol gösterici olabilir (Akoğlu, 2018). Bu anlamda SETS-MS Ölçeği'nin böyle bir amaca hizmet edeceği düşünülmektedir. Nitekim bu ölçek SETS Ölçeği'nin temellerinin atılmasında öncü bir role sahip olup daha sonra geliştirilen SETS-HS Ölçeği ile de psikometrik anlamda eşleşme sağlamaktadır (Harrell-Williams, Lovett, Pierce, Sorto, Lee, & Lesser, 2017). Bir diğer ifade ile, SETS-MS Ölçeği lise düzeyinde yapılacak araştırmalarda da kullanılabilir. Dolayısıyla bu ölçeğin kullanılması ile gerek ortaokul gerekse de lise kademesindeki öğretmen adaylarının istatistik öğretimine ilişkin öz yeterlik düzeylerine rahatlıkla ulaşılabilir. Buradan SETS-MS Ölçeği'nin SETS Ölçeği'nin tümünden elde edilecek bilgileri karşılayan bir güce sahip olduğu anlaşılmaktadır. Dolayısıyla mevcut araştırmada bu ölçeğin Türk Eğitim Sisteminin yapısı çerçevesinde Türkçeye uyarlanarak geçerlik ve güvenilirlik çalışmalarının yapılması amaçlanmaktadır. Nitekim ülkemizde istatistik kavramlarının yoğun olarak ortaokul düzeyinde ele alındığı göz önüne alındığında, bu kavramlarla fazlaca karşılaşacak olan matematik öğretmen adaylarının öz yeterliklerin irdelenmesi adına bu araştırmanın önemli olacağı düşünülmektedir.

2. Yöntem

2.1. Çalışma grubu

Bu çalışmanın katılımcılarını, beş farklı devlet üniversitesinde öğrenim görmekte olan toplamda 421 ilköğretim matematik öğretmeni adayı oluşturmaktadır. Katılımcılar, istatistik alanına ilişkin bilgi ve becerileri kazanmış olma kriteri gözetilerek, bağlı oldukları lisans programının 4. sınıf öğrencilerinden seçilmiştir. Buna göre, katılımcıların %69,1'ini (n= 291) kadınlar ve %30,8'ini (n=130) erkekler oluşturmaktadır. Kline'a (2011) göre, örneklemin madde sayısının 10 katı kadar olması ve 200'den az örnekleme çalışılmaması gerekmektedir. Bu doğrultuda, adaptasyonu yapılması düşünülen SETS-MS Ölçeği'nin yirmi altı maddesi için bu çalışmanın katılımcı sayısının belirlenen ölçüt değerini oldukça üzerinde olduğu ifade edilebilir.

2.2. Veri toplama aracı

Bu çalışmada SETS-MS Ölçeği'nin Türkçeye uyarlanması amaçlanmıştır. Bu doğrultuda aşağıda ilk olarak SETS-MS Ölçeği'nin orijinal formuna ilişkin bilgilere yer verilmektedir. Daha sonra bu ölçeğin geçerlik ve güvenilirlik çalışmalarının yürütüldüğü Türkçe formunun çeviri sürecine ilişkin bilgiler yer almaktadır.

2.2.1. SETS-MS Ölçeği

Ortaokul düzeyindeki istatistik öğretimine odaklanan bu ölçeğin orijinal formu ve psikometrik özelliklerine ölçeğin maddelerinin yayınlandığı makaleden (Harrell-Williams, Sorto, Pierce, Lesser, & Murphy, 2015) erişilmiştir. SETS-MS Ölçeği, veri okuma (11 madde) ve veriler arası okuma (15 madde) olmak üzere iki faktör ve yirmi altı maddeden oluşmaktadır. Maddeler yansıtıkları özellik bakımından ortaokulda öğretilen belli istatistik kavramlarına yönelik olarak ele alınmakta ve hiç güvenmeme (1)'den tamamen güvenme(6)'ye uzanan altı 6 derece altında sıralanmaktadır. Ölçek, araştırmaya katılan 309 öğretmen adayından elde edilen verilerle geliştirilmiş olup, ölçeğin yapı geçerliliği Rasch modeli esas alınarak gerçekleştirilen açımlayıcı faktör analizi (AFA) ve doğrulayıcı faktör analizi (DFA) ile incelenmiştir. Analiz sonucu ölçeğin iki faktörlü (veri okuma ve veriler arası okuma) yapıya sahip olduğunu göstermiştir. Faktörler arasındaki korelasyon katsayısı ise, 0.85 olarak hesaplanmıştır. Ayrıca ölçeğin güvenilirliğine ilişkin yorumlar Cronbach Alpa katsayısına bakılarak yapılmıştır. Bu doğrultuda ölçeğin tümü için bu katsayı 0,94, veri okuma faktörü için 0.87 ve veriler arası okuma faktörü için 0.91 olarak hesaplanmıştır.

2.2.2. SETS-MS Ölçeği'nin Türkçeye Çevrilmesi

Uyarlama sürecine geçmeden önce, sorumlu yazarlardan Leigh M. Harrell-Williams ile elektronik posta kanalıyla iletişim kurulmuş ve ölçeğin Türkçe'ye uyarlanabilmesi konusunda gerekli izinler alınmıştır. Daha sonra orijinal dili İngilizce olan ölçek, İngilizce alanında uzman üç bağımsız çevirmen tarafından ayrı ayrı Türkçe'ye çevrilmiştir. Çevirmenlerden birisi matematik eğitimi, diğeri İngiliz dili ve edebiyatı ve bir diğeri de eğitim bilimlerinde görev yapmaktadır. Çevirmenler tarafından elde edilen üç farklı çevirinin ortak değerlendirmesi istatistik alanında bilgi sahibi iki matematik eğitimcisi tarafından yapılmıştır. Değerlendirme esnasında anlaşmazlık bulunması durumunda ölçeğin yazarlarının görüşlerine başvurulmuştur. Böylece Türkçe taslak bir form oluşturulmuştur. Daha sonra taslak formun dil ve anlatım açısından uygunluğunun sağlanması amacıyla iki Türkçe eğitimcisinin görüşlerine başvurulmuştur. Uzmanlardan alınan görüşler doğrultusunda Türkçe form üzerinde yeniden düzenleme yoluna gidilmiştir. Düzenlemesi yapılan form geri çeviri için iki istatistik eğitimcisi ile paylaşılmıştır. İki uzmanın geri çevirileri araştırmacılar tarafından hazırlanan üç dereceli Geri Çeviri Uygunluk Formu ile İngilizce alanında deneyimli üçüncü bir uzman tarafından değerlendirilmiştir. Aşağıda bu formdan örnek bir kesite yer verilmiştir.

Orijinal madde	Uygunluk derecesi			Geri Çeviri
	1	2	3	
Identify the association between two variables from scatterplots.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Determine the relationship between two variables in a scatter plot.
Generalize a statistical result from a small group to a larger group such as the whole class.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Generalize a statistical results from a small group to bigger one.

Şekil 2. Geri Çeviri Uygunluk Formu'ndan Örnek bir Kesit

Geri çevirileri değerlendiren uzman, ölçeğin orijinal maddeleri ile uzmanların Türkçe taslak formda yer alan maddelere dair geri çevirilerini karşılaştırmış ve uygunluk derecesini belirlemiştir. Tüm maddeler için sürdürülen

bu işlem sonunda elde edilen veriler Cohen'in Kappa katsayısına başvurulmuş ve analiz edilmiştir. Bu katsayı iki gözlemci arasındaki uyumun güvenilirliğini şans faktörünü ortadan kaldırarak ölçmekte ve -1 ile +1 arasında değerler almaktadır (Karagöz, 2017). Buna göre -1 negatif tam uyumu ve +1 pozitif tam uyumu simgelemektedir. Bu çalışmada hesaplanan Kappa katsayısı ile uzmanların geri çevirileri arasında bir uyum olduğu sonucuna varılmıştır ($p < 0,01$). Ayrıca hesaplanan uyum yüzdesi (%70,2) olup, bu değer uzmanların geri çevirilerinin yüksek dereceli bir uyuma sahip olduğunu göstermektedir.

Türkçe taslak formun katılımcılar tarafından anlaşılabilirliğini tespit etmek amacıyla 20 matematik öğretmen adayı ile pilot çalışma yürütülmüştür. Pilot çalışma sonucunda, öğretmen adaylarının maddelere dair anlamaları Türkçe eğitimcileri ile değerlendirilmiş ve -m(a)k fiil köküne sahip maddelerin -(e)bilirim şeklinde izah edilmesi ile ölçek maddelerinin anlaşılabilirliğinin artırılacağı sonucuna varılmıştır. Bu doğrultuda Türkçe taslak formda yer alan maddelerin belirtilen şekilde yeniden düzenlenmesi yoluna gidilmiştir (örneğin; Aynı grupta yer alan bireyler arasındaki değişkenliği sayısal olarak ifade etmek maddesi, Aynı grupta yer alan bireyler arasındaki değişkenliği sayısal olarak ifade edebilirim şeklinde değiştirilmiştir). Ayrıca ölçeğin 6 derecesi de belirtilen şekilde düzenlenmiştir. Böylece nihai şeklini alan Türkçe forma Ortaokul Düzeyinde İstatistiği Öğretmeye yönelik Öz yeterlik [İÖÖ-OD] Ölçeği adı verilerek geçerlik ve güvenilirlik çalışmalarının yapılması için hazır hale getirilmiştir.

2.3. Etik ve İşlemler

Araştırmadaki veriler, beş farklı devlet üniversitesinden alınan izinler doğrultusunda İÖÖ-OD Ölçeği vasıtası ile online ya da birebir olarak toplanmıştır. Ayrıca araştırmanın etik kurallara dikkate alınarak yapıldığı Trabzon Üniversitesi Sosyal ve Beşeri Bilimler Etik Kurulu tarafından onaylanmıştır. Araştırmada öğretmen adaylarına araştırmaya dair bilgiler verilmiş ve yanıtlarının sadece bilimsel amaçlı kullanıldığı, kişisel bilgilerin herhangi bir kişi ya da kurumla paylaşımının söz konusu olmadığı belirtilmiştir. Gönüllülük esasına dayalı olarak araştırmaya katılan öğretmen adayları yaklaşık 20 dakikalık bir süre içerisinde ölçeği yanıtlamıştır.

2.4. Verilerin analizi

Araştırma kapsamında Türkçe çevirisi yapılan İÖÖ-OD Ölçeği'nin yapı geçerliğinin sınanmasında DFA 'ya başvurulmuştur. Nitekim DFA faktörler ile göstergeleri arasında önceden belirlenmiş olan ilişki modellerinin elde edilen yeni veriler ile ne derece desteklendiğini açıklayan bir analizdir (Çokluk, Şekercioğlu, & Büyüköztürk 2016). Özellikle kültürlerarası ölçek uyarlama çalışmalarında, ölçme aracının hedef kültürdeki faktör yapısı CFA'dan ziyade doğrudan DFA ile ortaya konulabilmektedir. Çünkü bahsi geçen ölçeğin ilgili kültürdeki faktör yapısı esasen yürütülen birçok nicel ve nitel çalışmalarla deneysel olarak kanıtlanmaktadır. Dolayısıyla belirlenen bu faktör yapısının hedef kültür altında korunup korunmadığı noktasında DFA yeterli olmaktadır. DFA ile modelin test edilmesinde ise, çok sayıda uyum indeksi ve bu indekslerin sahip oldukları parametreler göz önünde bulundurulmaktadır. Öyle ki, yapılan analizlerde hangi uyum indeksine yer verilmesi gerektiğine dair kesinlik olmamakla birlikte, Yapısal eşitlik modellemesi (YEM) analizlerinde yorumlanması tavsiye edilen dört uyum indeksinden bahsedilmektedir. Bunlar: Ki-kare Uyum İyiliği İndeksi (Chi-Square Goodness of fit- χ^2), Yaklaşık Hataların Ortalama Karekökü (Root Mean Square Error of Approximation- RMSEA), Standardize Edilmiş Kök Ortalama Kare Artık (Standardized Root Mean Square Residual- SRMR) ve Karşılaştırmalı Uyum İndeksi (Comparative Fit Index- CFI) (Kline, 2011). Bunlara ilaveten Ki-kare Uyum Testinin Serbestlik Derecesine Oranı (χ^2/sd), Uyum İyiliği İndeksi (Goodness of Fit Index- GFI), Tucker Lewis İndeksi (TLI) uyum iyiliği indeksleri de analizlerde genellikle tercih edilmektedir. Bu çalışmada söz konusu uyum indekslerinin tümüne yer verilmiştir. Öte yandan ölçeğin güvenilirliği için Cronbach Alfa iç tutarlık katsayısına başvurulmuştur. Bu bağlamda çalışmada ölçeğin geneli ve her bir faktörü için Cronbach Alfa katsayısı ayrı ayrı hesaplanmıştır. Ayrıca ölçekteki maddelerin kişileri ayırt etmede ne derece işlevsel olduğunu belirlemek amacıyla madde toplam korelasyonları hesaplanmış ve %27 'lik alt ve üst grup ortalamaları farkına dayalı bağımsız gruplara yönelik t testi yapılmıştır. (Büyüköztürk, 2011).

3. Bulgular

3.1. Geçerlik

SETS-MS Ölçeği'nin orijinal yapısının Türk kültüründe ne derece geçerli olduğunu incelemek ve ölçeğin yapı geçerliğini sağlamak amacıyla DFA uygulanmıştır. Analiz öncesinde veri seti kayıp veriler, uç değerler ve normallik açısından kontrol edilmiştir.

Uygulanan DFA sonucunda, İÖÖ-OD Ölçeği'nde yer alan m25 (İlişkisel ve neden-sonuç arasındaki ayrımı yapabilirim) ile m26 (Örneklem ortalaması ve örneklem oranı gibi özet istatistiklerdeki örnekleme değişkenliğini fark edebilirim) maddelerinin faktör yük değerlerinin oldukça düşük çıktığı ve istatistiksel olarak anlamlı olmadığı tespit edilmiştir. Ayrıca elde edilen uyum indeks değerlerinin de iyi olmadığı saptanmıştır. Dolayısıyla bu maddelerin analizden çıkartılmasına karar verilmiştir. Anlamlı olmayan maddelerin analizden birer birer çıkarılması ve her defasında uyum indeks değerlerinin ve faktör yüklerinin incelenmesi daha doğru sonuçlar

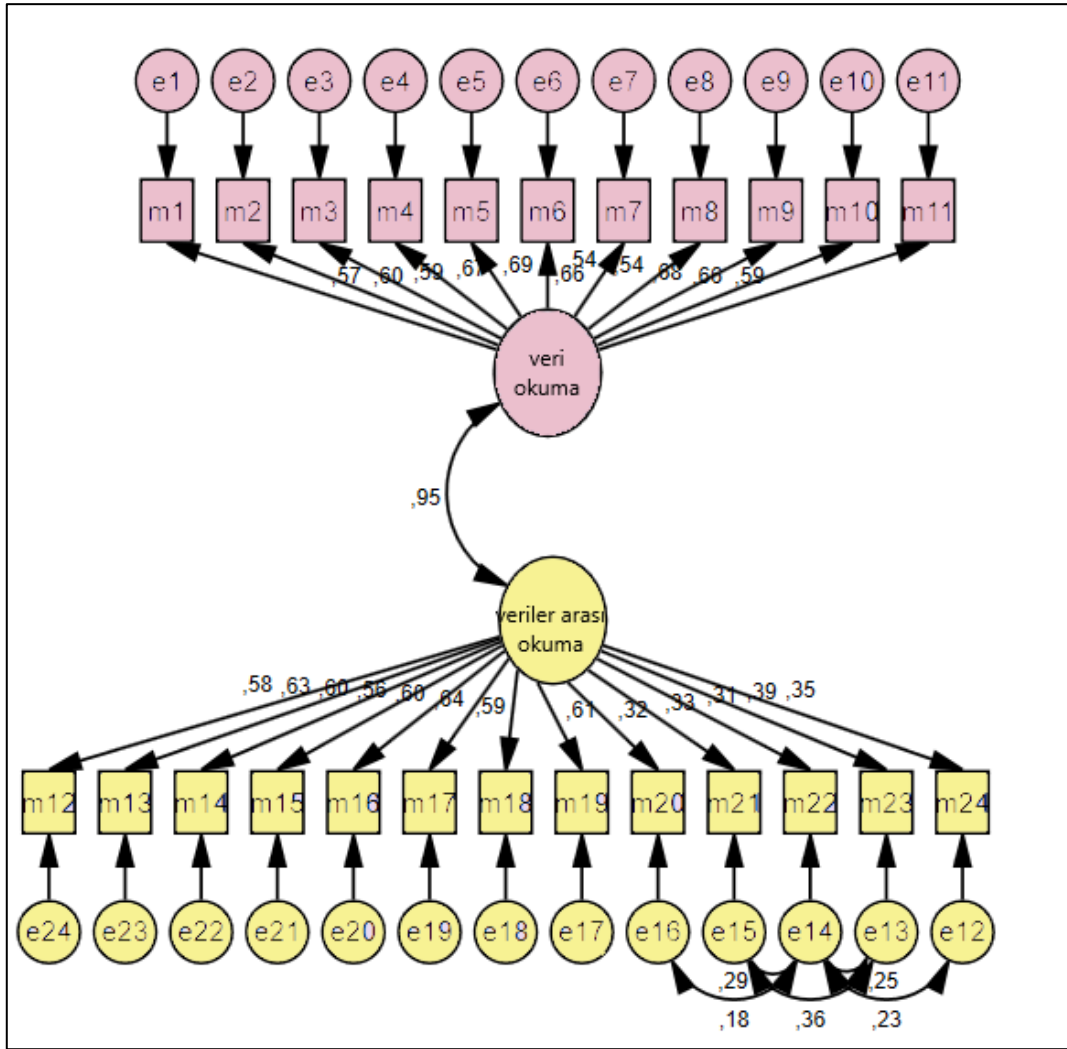
alınması adına önemli olmaktadır (Karagöz, 2017). Bu nedenle maddelerin çıkarılması işlemi en düşük yük değerine sahip olan maddeden başlayarak sırasıyla yapılmış ve her defasında uyum indeks değerleri incelenmiştir. Sonuçta söz konusu iki maddenin çıkarılması ile ölçekteki tüm maddelerin faktör yüklerinin istatistiksel olarak anlamlı olduğu görülmüştür ($p < 0,01$). Daha sonra orijinal ölçekte olduğu gibi iki faktörlü model için uyum indeks değerleri hesaplanmıştır. Buna göre elde edilen uyum indeks değerleri oluşturulan modelin uyumunun iyi olmadığını göstermiştir. Kuramsal yapıya uygun olması şartı ile maddeler arasında kovaryanslar oluşturularak model uyumu düzeltilebilmektedir (Karagöz, 2017). Bu doğrultuda analizdeki modifikasyon değerlerine bakılarak bazı maddelerin (m21-m23, m22-m24, m20-m22, m21-m22, m22-m23) ölçekte aynı gizil değişken altında yer aldığı ve hata kovaryansları arasında kayda değer ölçüde ilişki olduğu saptanmıştır. Uzman görüşleri doğrultusunda bu madde çiftleri arasında görünen yüksek düzeyli hata korelasyonları modele dahil edilmiş ve modelin tekrar test edilmesi yoluna gidilmiştir. Buna göre yapılan modifikasyon sonucu elde edilen uyum indeks değerleri modelin orijinal ölçekle yeterli uyum içerisinde olduğunu ve faktör yapısının doğrulandığını ortaya koymuştur. Uyum değerleri kriterleri ve kabul için kesme noktaları ile birlikte, maddeler çıkarılmadan önce elde edilen değerler (DFA I), çıkarıldıktan sonra elde edilen değerler (DFA II) ile modelde yapılan modifikasyon sonucu elde edilen değerlere (DFA III) ilişkin bilgiler Tablo 1'de karşılaştırmalı olarak yer almaktadır.

Tablo 1. SETS-MS Ölçeği'nin Uyum İndeks Değerleri ve Uyum İndekslerinin Kabulü için kesme noktaları

Uyum indeksi	Kriterler	Kabul için kesme noktaları	DFA I	DFA II	DFA III
χ^2	$p > 0.05$	-	$\chi^2=1010,735$ (sd=298, p=0.00)	$\chi^2=707,126$ (sd=251, p=0.00)	$\chi^2=537,987$ (sd=246, p=0.00)
χ^2/sd	-	≤ 2 =mükemmel uyum ≤ 5 =orta düzeyde uyum	3.39	2.81	2.18
GFI	0 (uyum yok) 1 (mükemmel uyum)	≥ 0.90 =iyi uyum ≥ 0.95 =mükemmel uyum	0.84	0.86	0.90
CFI	0 (uyum yok) 1 (mükemmel uyum)	≥ 0.90 =iyi uyum ≥ 0.95 =mükemmel uyum	0.80	0.86	0.91
TLI (NNFI)	0 (uyum yok) 1 (mükemmel uyum)	≥ 0.90 =iyi uyum ≥ 0.95 =mükemmel uyum	0.78	0.84	0.90
SRMR	0 (mükemmel uyum) 1 (uyum yok)	≤ 0.05 = mükemmel uyum ≤ 0.08 = iyi uyum	0.06	0.05	0.04
RMSEA	0 (mükemmel uyum) 1 (uyum yok)	≤ 0.05 = mükemmel uyum ≤ 0.08 = iyi uyum	0.07	0.06	0.05

Not: Uyum indeks değerleri ve bu değerlerin kabulü için belirlenen kesme noktaları için, Çokluk ve arkadaşları (2016) göz önünde bulundurulmuştur.

Tablo 1'den modelin uyum indeks değerleri incelendiğinde, her üç analizde de yer alan ki-kare değerinin anlamlı olduğu görülmektedir ($p < 0.01$). Ayrıca analizlerin tekrarlanmasıyla birlikte, bu değerde bir düşmenin olduğu da göze çarpmaktadır. Ki-kare değeri örneklem büyüklüğünden fazlaca etkilenmekte ve büyük örneklemelerde genellikle anlamlı sonuçlar üretmektedir. Bu nedenle örneklemde daha az etkilenen χ^2/sd oranı bunun yerine kullanılabilir bir ölçüt olarak kabul görmektedir (Şimşek 2007). Tablo 1'den maddelerin çıkarılmasıyla birlikte, bu değer mükemmel uyuma yaklaşıma başladığı görülmektedir. Öyle ki, modelde yapılan modifikasyon sonucu orijinal ölçekle tam olarak mükemmel bir uyum yakalanamasa da, elde edilen değer, uyumun mükemmel olduğuna yakın olduğunu kanıtlar niteliktedir. Bununla birlikte, GFI, CFI ve TLI uyum indeks değerlerinin ancak modelde yapılan modifikasyon sonucu iyi uyuma erişebildiği görülmektedir. Nitekim bu haliyle modelin uyum için belirlenen kesme noktalarını asgari ölçüde de olsa sağladığı ve orijinal ölçekle yeterli uyumu gösterdiği ifade edilebilmektedir. Benzer şekilde RMSEA uyum indeksinin de modifikasyon öncesindeki analizlerde iyi uyum gösterdiği ve modifikasyonla birlikte mükemmel uyuma erişebildiği görülmektedir. Yapılan analizlerdeki SRMR uyum indeks değerlerinin ise, maddelerin çıkarılmasıyla birlikte ancak mükemmel uyumu yakaladığı tespit edilmektedir. Öyle ki, maddeler çıkarılmadan önce bir uyumdan bahsetsek de, bu uyumun iyi düzeyde kaldığı ve mükemmel ulaşamadığı göze çarpmaktadır. Buna göre orijinal ölçekle uyum için belirlenen ölçüt değerlerinin tümünü sağlaması bakımından DFA III sonucunda, İÖÖ-OD Ölçeği'nin nihai şekli elde edilmiştir. Bu doğrultuda oluşturulan modele ilişkin path diyagramı aşağıda sunulmuştur.



Şekil 3. İki faktörlü modele ilişkin path diyagramı $n=421$, $\chi^2=537,987$ $sd=246$, $p < 0.001$

Model incelendiğinde, veri okuma faktörü altındaki maddelerin yük değerlerinin 0.54 ile 0.69 arasında, veriler arası okuma faktörünün altındaki maddelerin yük değerlerinin ise, 0.31 ile 0.64 arasında değişim gösterdiği ve tüm yük değerlerinin istatistiksel olarak anlamlı olduğu görülmektedir ($p < 0.01$). Maddelerin faktör yük değerlerinin genellikle 0.45 ve üzeri olması istenmekte, ancak faktör yükü 0.30 olan maddelerin de ölçekte yer alabileceği belirtilmektedir (Kline, 2011). Modelde yer alan maddelerin faktör yük değerlerinin 0.30 altında bulunmadığı görülmektedir. Bu durum maddelerin bu ölçütü uyumlu faktör yük değerlerine sahip olduğu ve bağlı olduğu faktörleri yeterli düzeyde yansıttığını ortaya koymaktadır. Ayrıca faktörler arasındaki korelasyon katsayısına bakıldığında, bu değer 0.95 olduğu tespit edilmektedir. Faktörler arasındaki korelasyonun çok uç değerler arasında olmaması ve istatistiksel olarak anlamlı olması geçerli bir ölçme aracı var olması istenen durumdur (Gündüz, & Coşkun, 2012). Buna göre ölçeğin faktörleri arasındaki ilişkinin ($r=0.95$, $p < 0.01$) çok uç değerler arasında bulunmadığı ve istatistiksel olarak anlamlı olduğu söylenebilmektedir.

3.2. Güvenilirlik

İÖÖ-OD Ölçeği'nin güvenilirliğini belirlemek için ölçeğin her bir faktörü ve geneli için Cronbach Alpha değerleri hesaplanmıştır. Buna göre ölçeğin on bir madde içeren birinci faktörüne (veri okuma) ait güvenilirlik katsayısı 0.87; on üç madde içeren ikinci faktörüne (veriler arası okuma) ait güvenilirlik katsayısı 0.82 olarak bulunmuştur. Bununla birlikte, ölçeğin geneli için hesaplanan güvenilirlik katsayısının 0.91 olduğu tespit edilmiştir. Bilimsel içerikli araştırmalarda ölçeğin iç tutarlılığı bağlamında tercih edilen Cronbach Alpha değerlerinin 0.70 ve üzeri olması yeterli olarak görülmektedir (Şencan, 2005). Bu doğrultuda İÖÖ-OD Ölçeği'nin güvenilir bir yapıya sahip olduğu ifade edilebilir. Diğer taraftan ölçekte yer alan maddelerin kişileri ne derece ayırt ettiğini incelemek amacıyla (i) %27 alt ve %27 üst grupların madde ortalamaları arasındaki farklılığa dayalı bağımsız gruplar için t testi ve (ii) madde toplam korelasyon değerleri hesaplanmıştır. Elde edilen sonuçlar Tablo 2'de yer almaktadır.

Tablo 2. %27'lik Üst ve Alt Gruplar için Bağımsız Gruplar t-testi ve Sonuçları ve Madde Toplam Korelasyonları

Faktör	Madde	Madde Toplam Korelasyonu	t (üst %27- alt %27)	Faktör	Madde	Madde Toplam Korelasyonu	t (üst %27- alt %27)
Veri Okuma	m1	0.57	12,289	Veriler arası Okuma	m12	0.58	13,383
	m2	0.60	13,201		m13	0.62	15,486
	m3	0.58	12,069		m14	0.63	12,585
	m4	0.65	15,217		m15	0.58	12,651
	m5	0.68	15,956		m16	0.60	14,264
	m6	0.66	13,796		m17	0.63	13,605
	m7	0.55	12,303		m18	0.59	12,752
	m8	0.55	11,848		m19	0.60	13,631
	m9	0.68	17,284		m20	0.39	7,850
	m10	0.65	16,444		m21	0.42	9,324
	m11	0.61	14,834		m22	0.42	8,149
				m23	0.48	11,775	
				m24	0.44	9,416	

p < 0.01

Tablo 2 incelendiğinde, ölçekte yer alan maddelerin madde toplam korelasyon değerlerinin 0.39 ile 0.68 arasında değişim gösterdiği görülmektedir. Madde toplam korelasyon değerlerinin 0.30 ve üzerinde olması ayırt ediciliğin iyi olduğunu göstermektedir (Büyükoztürk, 2011). Buna göre SETS-MS ölçeğinin kişileri ayırt etmeye elverişli olduğu ifade edilebilir. Bununla birlikte, %27'lik üst ve alt grubun puan ortalamalarına dayalı yapılan t testi sonuçlarına göre tüm maddeler için anlamlı bir fark olduğu tespit edilmiştir (p < 0.01). Bu durum her bir faktör altındaki maddenin ait olduğu faktörü yansıtacak davranışları içerdiğini göstermektedir. Buna göre yirmi dört maddelik İÖÖ-OD Ölçeği EK 1'de yer almaktadır.

4. Tartışma, Sonuç ve Öneriler

Bu çalışmada öğretmen adaylarının istatistiği öğretmeye yönelik öz yeterliklerini ölçmek amacıyla geliştirilmiş olan SETS Ölçeği'nin ortaokul seviyesindeki istatistik öğretimine odaklanan ölçeği için geçerlik ve güvenilirlik çalışmaları yapılarak Türk Eğitim Sisteminin yapısı çerçevesinde Türkçeye uyarlanması hedeflenmiştir. Bu doğrultuda ölçeğin yapı geçerliğine ilişkin bulgular DFA, güvenilirliğine ilişkin bulgular ise, Cronbach Alpha iç tutarlık katsayısı, %27'lik üst ve alt grup ortalamaları farkına dayalı bağımsız örneklemeler için t testi ve madde toplam korelasyon değerlerine başvurulmuş elde edilmiştir.

Araştırmada ilk etapta yapılan DFA I sonucunda hesaplanan uyum indeks değerlerinin iyi olmadığı tespit edilmiştir. Ayrıca ölçekte yer alan m25 ve m26 kodlu maddelerin faktör yük değerlerinin de anlamlı olmadığı görülmüştür. Nitekim faktör yüklerinin anlamlı çıkması, maddelerin faktörlere doğru yüklendiğinin önemli bir göstergesidir. Bu sebepten anlamlı olmayan maddelerin analizden çıkartılması gerekmektedir (Karagöz, 2017). Buradan hareketle bu maddelerin orijinal ölçekte yer alan faktörü (veriler arası okuma) yeterince yansıtmadığı sonucuna varılarak analizden çıkarılmasına karar verilmiştir. Böylece 24 madde üzerinden yapılan DFA II sonucunda, ölçekte yer alan tüm maddelerin faktör yüklerinin anlamlı olduğu; ancak ölçeğin Türkçe formunun uyum indeks değerlerinin orijinal ölçekle istenilen düzeyde uyum göstermediği tespit edilmiştir. Bu nedenle analiz sonucunda önerilen modifikasyonların yapılmasına karar verilmiştir. Nitekim modelin daha iyi uyum sağlaması açısından yapılacak düzeltmeler önemli olmaktadır. Lakin bu düzeltmelerin kuramsal yapıya uygun olarak yapılması gerekmektedir. Bu anlamda aynı faktöre ait maddelerin hata terimlerinin ilişkilendirmesi yapılmaktadır (Karagöz, 2017). Böylece kuramsal yapıya uygun olması şartıyla, önerilen ikililer arasında kovaryanslar oluşturularak analiz tekrarlanmıştır. Buna göre DFA III sonucu elde edilen uyum indeks değerleri ($\chi^2 = 537,987$, $p < 0.01$, $\chi^2/df = 2,18$, GFI=0,90, CFI=0,91, TLI=0,90, SRMR= 0,04, RMSEA=0,05) modele ait değerlerin kabul edilebilir sınırlar içinde olduğunu göstermiştir. Dolayısıyla 24 maddeden oluşan İÖÖ-OD Ölçeği'nin iki faktörlü yapısının doğrulandığı sonucuna varılmıştır. Ölçekteki maddelerin faktör yük değerleri 0.31 ile 0.69 arasında değişim göstermekte ve bağlı olduğu faktörü yansıtacak davranışları içermektedir. Ayrıca faktörler arasındaki korelasyon katsayısı da, 0.95 olarak hesaplanmıştır. Bu değer ölçeğin orijinal formu için hesaplanan değerle birlikte, her iki form için yüksek olduğu ifade edilebilir.

Araştırmada yapılan güvenilirlik analizi sonuçları, Cronbach Alpa değerlerinin ölçeğin faktörlerinden veri okuma için 0.87, veriler arası okuma için 0.82 olduğu görülmektedir. Bununla birlikte, ölçeğin tümü için hesaplanan değer 0.91 olduğu tespit edilmektedir. Bu değerlerin ölçeğin orijinal formu için hesaplanan

değerlere yakın olduğu anlaşılmaktadır. Ayrıca %27'lik üst ve alt grupların ortalama puanları arasındaki farka dayalı hesaplanan t testi sonuçları, puan ortalamaları arasında anlamlı bir farklılık olduğunu göstermiştir. Ayrıca ölçekteki maddelerin madde toplam korelasyon değerlerinin ise, 0.39 ile 0.68 arasında değiştiği gözlemlenmiştir. Bu doğrultuda her ne kadar bu değerlerin ölçeğin orijinal formu için hesaplanan değerlerin altında kaldığı tespit edilse de, Türkçe formda yer alan maddelerin kişileri ayırt edici özelliğe sahip olduğu söylenebilir.

Sonuç olarak, Türkçe'ye uyarlanan İÖÖ-OD Ölçeği'nin orijinal ölçekte olduğu gibi iki faktörden oluştuğu, ölçeğin araştırmanın katılımcılarından elde edilen verilerle uyum içerisinde olduğu ve ölçeğin iç tutarlık katsayılarının yeterli seviyede olduğu görülmüştür. Böylece bu ölçeğin öğretmen adaylarının ortaokul düzeyindeki istatistiği öğretmeye yönelik öz yeterliğini ölçen çalışmalarda kullanılabileceği anlaşılmaktadır. Bu çalışmaların ölçeğin geçerlik ve güvenilirliğine ait sonuçlarının desteklenmesi açısından da önemli olacağı düşünülmektedir. Ayrıca orijinal ölçekte olduğu gibi ölçeğin Türkçe formunun da farklı öğrenim düzeylerinde kullanımı için geçerlik ve güvenilirlik çalışmalarının yapılabilir. Diğer taraftan bu araştırmanın ülkemiz ortaokul düzeyinde verilen istatistik öğretimiyle uyumlu olarak özgün ölçek geliştirme çalışmalarının yapılması açısından da bu alanda çalışan araştırmacılara fikir vereceği düşünülmektedir.

Kaynaklar / References

- Akoğlu, K. (2018). *Blending online coursework and small learning communities to examine professional growth in teaching statistics: A phenomenological case study* (Unpublished doctoral dissertation). North Carolina State University, North Carolina.
- Arseven, A. (2016). Öz yeterlilik: Bir kavram analizi. *Turkish Studies*, 11(19), 63-80.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman and Company.
- Bandura, A. (2006). Guide for constructing self-efficacy scales. In T. U. F. Pajares (Ed.), *Self-efficacy beliefs of adolescents* (pp. 307-337). Greenwich, CT: Information Age Publishing.
- Batur, A., Yiğit, N., & Baki, A. (2019, Ekim). İstatistiksel Okuryazarlık Öz yeterlik Ölçek Geliştirme Çalışması. *Uluslararası Fen, Matematik, Girişimcilik ve Teknoloji Eğitimi Kongresi Tam Metin Kitabı* içinde (ss.740-750). Denizli: OrEgDa Yayıncılık.
- Begg, A., & Edwards, R. (1999, December). *Teachers' ideas about teaching statistics*. Paper presented at the Combined Annual Meeting of the Australian Association for Research in Education and the New Zealand Association for Research in Education. Melbourne, Australia. Retrieved from <https://files.eric.ed.gov/fulltext/ED455195.pdf>.
- Büyüköztürk, Ş. (2011). *Veri analizi el kitabı (15. bs.)*. Ankara: Pegem A Yayıncılık.
- Callingham, R., & Burgess, T. (2014, July). How the curriculum shapes teachers' thinking: A comparison of New Zealand and Australian teachers' thinking about statistics. In K. Makar, B. deSousa, & R. Gould (Eds.), *Sustainability in statistics education* (pp. 13-18). Voorburg, The Netherlands: International Statistical Institute. Retrieved from http://iase-web.org/icots/9/proceedings/pdfs/ICOTS9_10C3_CALLINGHAM.pdf
- Callingham, R., & Watson, J. M. (2017). The development of statistical literacy at school. *Statistics Education Research Journal*, 17(1), 181-201.
- Carmichael, C. S., & Hay, I. (2009). The development and validation of the Students' Self- Efficacy for Statistical Literacy Scale. In R. Hunter, B. Bicknell & T. Burgess (Eds.), *Proceedings of the 32nd Annual Conference of the Mathematics Education Research Group of Australasia* (Vol. 1, pp. 97-104). Wellington: MERGA.
- Czerniak, C. M. (1990, April). *A study of self-efficacy, anxiety, and science knowledge in pre-service elementary teachers*. Paper presented at the Annual Meeting of the National Association of Research in Science Teaching, Atlanta, GA.
- Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2016). *Sosyal bilimler için çok değişkenli istatistik SPSS ve LISREL uygulamaları (4. bs.)*. Ankara: Pegem Akademi.
- De Mesquita, P. B., & Drake, J. C. (1994). Educational reform and self-efficacy beliefs of teachers implementing nongraded primary school programs. *Teaching and Teacher Education*, 70(3), 291-302.
- Dede, Y. (2008). Matematik öğretmenlerinin öğretimlerine yönelik öz-yeterlik inançları. *Türk Eğitim Bilimleri Dergisi*, 6(4), 741-757.
- delMas, R. (2004). A comparison of mathematical and statistical reasoning. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 79-95). Netherlands: Kluwer Academic Publishers.
- Enochs, L., Smith, P. L., & Huinker, D. (2000). Establishing factorial validity of the mathematics teaching efficacy eliefs instrument. *School Science and Mathematics*, 100(4), 194-202.
- Finney, S., & Schraw, G. (2003). Self-efficacy beliefs in college statistics courses. *Contemporary Educational Psychology*, 28(2), 161-186.
- Fitzmaurice, O., Leavy, A., & Hannigan, A. (2014). Why is statistics perceived as difficult and can practice during training change perceptions? Insights from a prospective mathematics teacher. *Teaching Mathematics and Its Applications*, 33, 230-248.

- Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: A pre-K-12 curriculum framework*. Alexandria, VA: American Statistical Association.
- Friel, S. N., Curcio, F. R., ve Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, 32(2), 124-158.
- Gal, I. (2004). Statistical literacy: Meanings, components, responsibilities. In J. B. Garfield and D. Ben-Zvi (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 47–78). Dordrecht: Kluwer Academic Publishers.
- Garfield, J., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching Practice*. Dordrecht, The Netherlands: Springer
- Gerez Cantimer, G., Şengül, S., & Akçin, N. (2020). Özel eğitim öğretmenlerinin matematik öğretimine yönelik öz-yeterlilik algıları. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 35(2), 306-319. doi: 10.16986/HUJE.2019052312
- Göloğlu-Demir, C. ve Çetin, Ş. (2010). İlköğretim matematik öğretmen adaylarının matematik öğretimine yönelik öz-yeterlilik inançları ölçeğinin geliştirilmesi geçerlik ve güvenirlik çalışması. *Ticaret ve Turizm Eğitim Fakültesi Dergisi*, (1), 164-175.
- Groth, R. E. (2007). Toward a conceptualization of statistical knowledge for teaching. *Journal for Research in Mathematics Education*, 38(5), 427-437.
- Gündüz, Y., & Coşkun, Z. S. (2012). Öğrenci algısına göre öğretmen etik değerler ölçeğinin geliştirilmesi: Geçerlik ve güvenirlik çalışması. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi*, 13(1), 111-131.
- Harrell-Williams, L. M., Lovett, J. N., Lee, H. S., Pierce, R. L., Lesser, L. M., & Sorto, M. A. (2017). Validation of scores from the high school version of the Self-Efficacy to Teach Statistics instrument using preservice mathematics teachers. *Journal of Psychoeducational Assessment*, 37(2), 194-208. <https://doi.org/10.1177/0734282917735151>
- Harrell-Williams, L. M., Lovett, J. N., Pierce, R. L., Sorto, M. A., Lee, H. L., & Lesser, L. M. (2017). The middle grades SETS instrument: Psychometric comparison of middle and high school pre-service mathematics teachers. In E. Galindo and J. Newton, (Eds.), *Proceedings of the 39th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 1064-1067). Indianapolis, IN: Hoosier Association of Mathematics Teacher Educators.
- Harrell-Williams, L. M., Sorto, M. A., Pierce, R. L., Lesser, L. M., & Murphy, T. J. (2014). Validation of scores from a new measure of preservice teachers' self-efficacy to teach statistics in the middle grades. *Journal of Psychoeducational Assessment*, 32, 40-50. doi:10.1177/0734282913486256.
- Harrell-Williams, L.M., Sorto, M.A., Pierce, R.L., Lesser, L.M., & Murphy, T.J. (2015). Identifying statistical concepts associated with high and low levels of self-efficacy to teach statistics to middle grades. *Journal of Statistics Education*, 23(1), 1–20.
- Irakleous, P., & Panaoura, A. (2015). Personal teaching efficacy in teaching statistics: A case study, *Mathitudes 1*(1), 1-18.
- Kahle, D. K. B. (2008). *How elementary school teachers' mathematical self-efficacy and mathematics teaching self-efficacy relate to conceptually and procedurally oriented teaching practices* (Doctoral dissertation). The Ohio State University, Columbus, OH.
- Karagöz, Y. (2017). *SPSS ve AMOS Uygulamalı Nitel-Nicel-Karma Bilimsel Araştırma Yöntemleri ve Yayın Etiği (1 bs)*. Ankara: Nobel Akademik Yayıncılık.
- Kline, R. B. (2011), *Principles and Practice of Structural Equation Modeling (3rd Edition ed.)*. New York: The Guilford Press.
- Lane, A. M., Hall, R., & Lane, J. (2004). Self-efficacy and statistics performance among Sport Studies students. *Teaching in Higher Education*, 9(4), 435-448.
- Lin, S., & Huang, C. (2013). The Construction and Validity of Scale of Self-Efficacy for Statistical Literacy, *International Conference on Advanced Information and Communication Technology for Education*. pp. 272-275, DOI: 10.2991/icaicte.2013.54.
- Lovett, J. N. (2016). *The preparedness of preservice secondary mathematics teachers to teach statistics: A cross-institutional mixed method study* (Unpublished dissertation). North Carolina State University.
- März, V., & Kelchtermans, G. (2013). Sense-making and structure in teachers' reception of educational reform. A case study on statistics in the mathematics curriculum. *Teaching and Teacher Education*, 29, 13–24.
- McGee, J. R., & Wang, C. (2014). Validity-supporting evidence of the self-efficacy for teaching mathematics instrument. *Journal of Psychoeducational Assessment*, 32(5), 390-403.
- Mercimek, O., & Pektaş, M. (2013). İstatistiğe yönelik öz-yeterlilik ölçeğinin geliştirilmesi. *Kastamonu Eğitim Dergisi*, 21(2), 759-776.
- Pajares, F. (2002). *Overview of social cognitive theory and of self-efficacy*. Retrieved from <http://www.emory.edu/EDUCATION/mfp/eff.html>

- Pierce, R., & Chick, H. (2011). Teachers' beliefs about statistics education. In C. Batanero, G. Burrill & C. Reading (Eds.), *Teaching statistics in school mathematics—challenges for teaching and teacher education: A joint ICMI/IASE study* (pp. 151–162). Dordrecht, The Netherlands: Springer
- Shaughnessy, J.M. (2007). Research on statistics learning and reasoning. In F.K. Lester (Ed.), *The Second Handbook of Research on Mathematics* (pp. 957–1010). Information Age Pub Inc.
- Smith, J. (1996). Efficacy and teaching mathematics by telling: A challenge for reform. *Journal for Research in Mathematics Education*, 27(4), 387-402.
- Steinberger, P. (2020). Assessing the Statistical Anxiety Rating Scale as applied to prospective teachers in an Israeli Teacher-Training College. *Studies in Educational Evaluation*, 64, 100829.
- Şencan, H. (2005). *Sosyal ve Davranışsal Ölçümlerde Güvenirlilik ve Geçerlilik*. Ankara: Seçkin Yayıncılık.
- Şimşek Ö. F. (2007). *Yapısal Eşitlik Modellemesine Giriş, Temel İlkeler ve LISREL Uygulamaları*. Ankara: Ekinoks.
- Tschannen-Moran, M. ve Woolfolk Hoy, A. (2001). Teacher Efficacy: Capturing an Elusive Construct. *Teaching and Teacher Education*, 17(7), 783-805.
- URL-1 (2020). Self Efficacy to Teach Statistics (SETS), <https://www.memphis.edu/sets/> adresinden 19 Nisan 2020 tarihinde erişilmiştir
- Watson, J. M. (2001). Profiling Teachers' Competence and Confidence to Teach Particular Mathematics Topics: The case of chance and data. *Journal of Mathematics Teacher Education*, 4(4), 305-337.
- Watson, J. M. (2013). Statistical literacy, a statistics curriculum for school students, the pedagogical content needs of teachers and the Australian Curriculum. *Curriculum Perspectives*, 33(3), 58-69.
- Wessels, H. (2014, July). Developing statistical knowledge for teaching of variability through professional development. In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education* (pp. 1-6). Flagstaff, Arizona, USA. Voorburg, the Netherlands: International Statistical Institute. www.iase-web.org
- Wheatley, K. F. (2005). The case for reconceptualizing teacher efficacy research. *Teaching and Teacher Education*, 21, 747–766. doi:10.1016/j.tate.2005.05.009
- Woolfolk Hoy, A. W., ve Burke-Spero, R. (2005). Changes in teacher efficacy during the early years of teaching: A comparison of four measures. *Teaching & Teacher Education*, 21(4), 343-356.
- Zee, M., & Kooman, H. M. Y. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustments, and teacher well-being: a synthesis of 40 years of research. *Review of Educational Research*. 86(4), 981-1015. doi:10.3102/0034654315626801.

EK 1 / APPENDIX 1 Ortaokul Düzeyinde İstatistiği Öğretmeye yönelik Öz yeterlik Ölçeği

MADDELER		Hiç güvenmiyorum (1)	Çok az güveniyorum (2)	Biraz güveniyorum (3)	Güveniyorum (4)	Çok güveniyorum (5)	Tamamen güveniyorum(6)
Veri Okuma							
1.	Ortaokul öğrencilerini ilgilendiren istatistiksel bir soruya yanıt vermek için veri toplayabilirim.						
2.	Bireyler üzerinde yapılan gözlemler arasında doğal bir değişkenlik olacağını fark edebilirim.						
3.	Bireyleri birbirleriyle ve bir grupla karşılaştırmak için uygun grafik gösterimlerini ve sayısal özetleri seçebilirim.						
4.	Dağılımları özetlemek için grafikler (sütun, çizgi vs.) ve tablolar oluşturabilirim.						
5.	Dağılımları tasvir etmek için grafikler (sütun, çizgi vs.) ve tablolar oluşturabilirim.						
6.	Dağılımları özetlemek için kutu grafikleri oluşturabilirim.						
7.	Dağılımları tasvir etmek için kutu grafikleri, medyan ve açıklığı kullanabilirim.						
8.	Serpilme diyagramını kullanarak iki değişken arasındaki ilişkiyi belirleyebilirim.						
9.	Bir istatistiksel sonucu küçük bir gruptan büyük bir gruba genelleyeabilirim.						
10.	İstatistiksel sonuçların başka bir sınıf veya grupta farklı olabileceğinin farkındayım.						
11.	Bir sınıftan elde edilen veri setinden hareketle sınıf dışındaki herhangi bir popülasyona genelleme yapmanın sınırlılıklarını fark edebilirim.						
Veriler arası Okuma							
12.	Verilerde değişkenliğin olduğu bir soru ile kesinlik içeren modele dayalı bir soru (örneğin; hız ve zamanın belli olduğu $x=v \times t$ modelinde, uzaklığın değerini bulmak) arasında ayırım yapabilirim.						
13.	Bir istatistiksel soruyu açıklarken ne tür değişkenlerin ölçüleceğini ve bu değişkenlerin nasıl ölçüleceğini belirleyebilirim.						
14.	Aynı grupta yer alan bireyler arasındaki değişkenliği sayısal olarak ifade edebilirim.						
15.	Dağılımları özetlemek için histogramlar oluşturabilirim.						
16.	Dağılımları karşılaştırmak için histogramları kullanabilirim.						
17.	Dağılımları özetlemek için çeyrekler açıklığını ve 5-sayı özetlerini (minimum ve maksimum değer, alt ve üst çeyrek, medyan) hesaplayabilirim.						
18.	Dağılımları karşılaştırmak için çeyrekler açıklığı, 5-sayı özetlerini (minimum ve maksimum değer, alt ve üst çeyrek, medyan) ve kutu grafiklerini kullanabilirim.						
19.	Bir popülasyondan rastgele alınan örnekleme dayalı bir sonuca varırken örnekleme hatasının rolünü fark edebilirim.						
20.	Lineer modelleri kullanarak iki değişken arasındaki ilişkinin gücünü sayısal olarak ifade edebilirim.						
21.	İki ya da daha fazla grup arasındaki farklılıkları dağılımların merkezi, yayılımı ve şekli açısından açıklayabilirim.						
22.	Bir örneklemin geniş bir popülasyonu temsil edip etmeyeceğini fark edebilirim.						
23.	İlişki ölçümlerini yorumlayabilirim.						
24.	Gözleme ve deneye dayalı çalışmalar arasında ayırım yapabilirim.						