

Interdisciplinary Teaching Design Self-Efficacy Scale: The validity and reliability study

Hülya Sevim Kılınç^a  Gülsüm Gök^b  Nezaket Bilge Uzun^c 

^a Ministry of National Education, Mersin, Türkiye, hulyaaksahin@gmail.com

^b Mersin University, Mersin, Türkiye, gulsumgok@mersin.edu.tr

^c Mersin University, Mersin, Türkiye, buzun@mersin.edu.tr

ABSTRACT

The aim of this study is to develop a valid and reliable instrument to assess teachers' self-efficacy in designing interdisciplinary teaching. During the scale development process, the scaling by cumulative sums approach based on participant responses was employed. Validity evidence was gathered through content and construct validity procedures. An item pool was generated based on focus group interviews and a comprehensive literature review, followed by expert review for content validation. The study sample comprised 350 teachers for exploratory factor analysis (EFA) and 231 teachers for confirmatory factor analysis (CFA). The final version of the scale obtained through analyses consists of 21 items measured on an 11-point scale ranging from 0 to 100, structured under two factors, namely, Competencies and Challenges. The scale demonstrated high reliability, with Cronbach's alpha coefficients of 0.95 and 0.92 for the respective factors. The total variance explained by the two subscales collectively is 66.21%. The confirmatory factor analysis (CFA) yielded satisfactory model fit, with all goodness-of-fit indices meeting the recommended thresholds for acceptable to excellent fit, supporting the structural validity of the model $\chi^2 = 1413.84$, $df = 516$, ($\chi^2/df = 2.74$), $RMSEA = .09$, $CFI = .92$, $RFI = .86$, $NNFI = .92$, $SRMR = .05$. These findings provide robust evidence that the scale is a scientifically and psychometrically valid and reliable instrument for assessing teachers' self-efficacy in designing interdisciplinary instruction.

Article Type
Research

Article Background

Received:
July 14, 2025
Accepted:
September 8, 2025

Keywords

Teacher Self-efficacy,
Interdisciplinary
Teaching Design, Scale
Development, Scale
Validation

To cite this article: Kılınç H. S., Gök, G., & Uzun, N. B. (2026). Interdisciplinary Teaching Design Self-Efficacy Scale: The validity and reliability study. *International Journal of Turkish Education Sciences*, 14 (1), 85-108. <https://doi.org/10.46778/goputeb.1741954>

Corresponding Author: Gülsüm Gök, e-mail: gulsumgok@mersin.edu.tr

Introduction

Real-life challenges are increasingly complex and multifaceted, demanding solutions that transcend disciplinary boundaries. To adequately prepare learners for the demands of the 21st-century world, educational curricula must cultivate a foundational understanding of multiple disciplines and foster the ability to integrate diverse perspectives and methodologies. Some disciplines overlap, containing common interdisciplinary areas. While we can address each discipline separately, it is also possible to examine two or more disciplines together. Disciplines that are integrated through coherent relationships can be presented in conjunction. In an educational program, linking one subject to others is referred to as an interdisciplinary approach or interdisciplinary teaching. Interdisciplinary teaching can be defined as an educational approach that integrates knowledge from multiple academic disciplines around a common theme or purpose (Darbellay, 2022; Martinot, 2023).

The core of interdisciplinary teaching lies in the teacher's ability to integrate or connect different subjects. Rather than a random or disorganized approach to interdisciplinary relationships, there is a conscious effort involved. In interdisciplinary teaching, a problem, topic, or specific concept is taken as a starting point, and knowledge and skills from various disciplines are integrated to illuminate the subject from multiple perspectives (Aydın & Balım, 2005; Yıldırım, 1996). It fosters the need to consider different perspectives based on the exchange of information and develops the ability to see things from the perspective of other disciplines (Apostel, 1970; Jacobs, 1989).

While interdisciplinary teaching offers numerous benefits, its widespread adoption is often hindered by teachers' lack of confidence in their ability to design and implement effective interdisciplinary lessons (Wu, 2024). Teachers' knowledge, perspectives, perceptions, and self-efficacy regarding interdisciplinary education shape their teaching in the classroom as an integrated whole (You, 2017). At this point, identifying teachers' self-efficacy in designing interdisciplinary teaching environments provides an insight into their perceptions of interdisciplinary teaching abilities. To address this issue, it is crucial to develop reliable and valid instruments to measure teacher self-efficacy in this domain. Therefore, the aim of this study is to develop a valid and reliable measurement tool to assess teachers' self-efficacy in designing their teaching based on an interdisciplinary approach.

Interdisciplinary Teaching Approach

Interdisciplinary teaching, defined as the blending of two or more subjects (St Clair & Hough, 1992), has been implemented in K-12 education for more than a hundred years (Grossman et al., 2001). It enables the integration of various fields of knowledge into instruction, helping students learn by exploring relevant problems, issues, and questions. Creating meaningful learning experiences through projects or challenges rooted in real-life concerns has been shown to enhance student engagement and support holistic learning (Chen, 2020). The foundation of interdisciplinary teaching lies in the integration or association of knowledge, which is facilitated by the teacher. In interdisciplinary instruction, a problem, topic, or specific concept serves as the focal point, and knowledge and skills that can illuminate it from different perspectives are integrated from relevant disciplines (Aydın & Balım, 2005; Yıldırım, 1996).

In the interdisciplinary teaching approach, the aim is to facilitate learning by bringing together a central theme, topic, or concept through a disciplinary lens, establishing connections between disciplines to foster a holistic perspective. To achieve this goal, various integration methods are employed when combining disciplines. Choi and Park (2006) and Kezar and Elrod (2012) categorize the integration of disciplines under the headings of multidisciplinary, interdisciplinary, and transdisciplinary approaches. The multidisciplinary approach involves several disciplines independently contributing their own perspectives to a shared problem, without integrating their insights. However, the interdisciplinary approach aims to develop shared understandings and skills by integrating knowledge and methods from multiple disciplines through collaboration. The transdisciplinary approach transcends academic boundaries by combining insights from multiple fields and involving non-academic stakeholders to develop practical solutions for real-world issues. Similarly, Jacobs (1989) groups the methods of integrating disciplines under four different curriculum forms. In the parallel discipline design, teachers align lessons across different subjects but teach them separately. In the complementary discipline design, related subjects are brought together in a shared unit around a common theme. The interdisciplinary design involves combining multiple subjects into broader, integrated units taught periodically. Finally, the integrated day centers learning around themes and real-life issues from the child's perspective, shaping the entire school day.

The research highlights the effectiveness of interdisciplinary teaching, demonstrating its capacity to improve critical thinking, creativity, and adaptability among students—skills essential for addressing complex societal challenges (Rafiq et al., 2024; Scott & White, 2024). By integrating knowledge from varied fields, this approach promotes a holistic understanding of subjects, offering students a wider perspective and a more integrated learning experience (Rafiq et al., 2024; Martinot, 2023). Interdisciplinary teaching enhances the learning process by encouraging critical thinking, efficient communication, and the exploration of hybrid topics, which in turn promotes a deeper comprehension of content (Holley, 2024). Through collaborative inquiry and interpretation, it engages both students and teachers in the practical application of knowledge, facilitating connections across disciplines and enhancing long-term learning and academic achievement (Drake & Burns, 2004; Guercio, 1998; McDonald & Czerniak, 1994; Suraco, 2006; White & Carpenter, 2008). Moreover, this methodology leads individuals to develop self-awareness, understand their surroundings, and drive meaningful change. By encouraging the exchange of knowledge and the consideration of multiple perspectives, it cultivates the ability to approach issues from the standpoint of other disciplines, fostering critical and higher-order thinking skills (Apostel, 1970; Gür, 2003; Jacobs, 1989; Karacaoğlu, 2011; Yıldırım, 1996).

Implementing interdisciplinary teaching methods offers several advantages for teachers. It promotes a comprehensive understanding of knowledge and enhances curriculum integration, which helps teachers address students' literacy challenges and adapt to future workforce needs (Rauf & Bhatti, 2024). Additionally, these methods improve teaching efficiency by facilitating collaboration, sharing expertise, and creating opportunities for professional development (Sadaa, 2024). They also help reduce feelings of isolation among educators while fostering supportive professional relationships and encouraging innovative teaching strategies (D'Souza et al., 2021; Yuan, 2023). Overall, interdisciplinary teaching

enhances instructional practices and nurtures an environment conducive to professional growth and development.

Although interdisciplinary teaching methods are promising to improve educational outcomes, successfully implementing them require instructional support, experience, careful planning, and a strong commitment from the institution to address any obstacles (Lenka & Singh, 2024; Li, 2023; Rafiq et al., 2024; Scott & White, 2024).

Teacher Self-efficacy

Bandura (1994) defines self-efficacy as “people’s beliefs about their capabilities to produce designated levels of performance that exercise influence over events to affect their lives. Self-efficacy beliefs determine how people feel, think, motivate themselves, and behave.” (p. 71). The stronger belief in one’s own abilities contributes to higher achievement as well as overall well-being. People with high self-efficacy consider obstacles as opportunities for learning and development, tackling difficulties with persistence and a desire to develop new and better skills. On the contrary, those with low self-efficacy often struggle with self-doubt, prefer to avoid challenges, and reflect lower motivation and persistence when confronted with adversity. According to Bandura (1977, 1997), self-efficacy is derived from four primary sources: mastery experiences, vicarious experiences, social persuasion, and emotional arousal. Among these, mastery experiences serve as the most influential factor, as they involve individuals facing challenges, overcoming them, and achieving success. Vicarious experiences refer to observing others’ achievements and using them as role models. Social persuasion involves receiving feedback on one’s performance, while emotional arousal encompasses the feelings of joy, excitement, and satisfaction experienced during the task.

The Social Cognitive Theory of Bandura posits that self-efficacy plays a crucial role in shaping individuals’ behaviors and outcomes. Similarly, in the context of education, teacher self-efficacy has been identified as a key factor influencing instructional practices and student achievement. Tschannen-Moran and Hoy (2001, p. 783) define a teacher’s sense of efficacy as a teacher’s “judgment of his or her capabilities to bring about desired outcomes of student engagement and learning, even among those students who may be difficult or unmotivated”. A teacher with high self-efficacy tends to be more effective in managing their classrooms, achieving student engagement, addressing challenges, and implementing instructional methods that support student learning (Zimmerman, 2000).

A useful measurement for teachers’ self-efficacy needs to consider their competence across a wide range of activities and tasks that they are supposed to conduct (Tschannen-Moran & Hoy, 2001). In detail, in a valid measure both of the personal competence and the task itself, considering the available resources and limitations within specific teaching environments, needed to be assessed (Tschannen-Moran et al., 1998). A well-developed questionnaire can provide valuable data to support teacher development and inform educational policy. However, Tschannen-Moran and Hoy (2001) specify that teacher efficacy is considered to be context-specific, and its measurement should take this into account. As interdisciplinary education becomes increasingly prevalent, measuring teacher self-efficacy specifically in this area is essential to understanding the factors that contribute to effective integration of diverse subject areas.

Designing a teaching based on an interdisciplinary approach requires time, effort, experience, knowledge, and collaboration (Gryson, Strubbe, Valcke, & Vanderlinde, 2024). Additionally, this approach must be flexibly planned to account for individual differences in students' thinking processes, and these differences must be carefully considered (Kubat, 2018). For effective course design, teachers are expected to possess the necessary and sufficient qualifications. In this regard, teachers' self-efficacy specifically in designing interdisciplinary teaching has a significant importance.

Assessing Teacher Self-Efficacy on Interdisciplinary Teaching

The commonly used self-efficacy scale for teachers is the Teacher Sense of Self-Efficacy Scale (TSES) developed by Tschannen-Moran and Hoy (2001). It is a valid assessment tool, including a wide range of teacher tasks required for effective teaching. The version of the scale adapted in Turkish by Capa, Cakiroglu, and Sarikaya (2005) consists of 24 items and three sub-dimensions: efficacy for classroom management, efficacy for student engagement, and efficacy for instructional strategies. Another common scale developed based on social cognitive theory is The Teacher Self-Efficacy Scale (Schmitz & Schwarzer, 2000), consist of 10 items, and forming two sub-dimensions, namely, coping behavior and innovative behavior. The scale adapted in Turkish by Yılmaz et al. (2024) consist of 8 items. Although the scales were not domain-specific, as opposed to the current study, their items provided insight into how to measure teachers' self-efficacy.

Although there is no specific scale to measure teachers' self-efficacy regarding the interdisciplinary teaching approach, several assessment tools have been developed within the context of STEM education. Given that STEM is inherently interdisciplinary, the items of these scales were carefully examined. For instance, the STEM Efficacy Survey (Buechel et al., 2024) includes questions related to integrated STEM implementation, professional development experiences in integrated STEM education, and teachers' perceived confidence levels in teaching integrated STEM. Similarly, the T-STEM Survey, designed to evaluate various dimensions of teaching in STEM education, consists of seven constructs, one of which specifically assesses teacher self-efficacy (The Friday Institute for Educational Innovation, 2012a). Another self-report questionnaire, the Teacher Self-Efficacy in iSTEM Scale (Yang et al., 2025), not only measures self-efficacy but also evaluates teachers' commitment to being implementers, designers, and disseminators of interdisciplinary STEM practices. Additionally, a scale developed by Chai et al. (2019) within the Technological Pedagogical and Content Knowledge (TPACK) framework aims to measure teachers' self-efficacy in integrated STEM (iSTEM). The Turkish-adapted version of this scale (Aktaş & Özmen, 2024) includes 24 items and four dimensions, focusing on teachers' self-efficacy in integrated STEM (b-STEM) and their technological pedagogical knowledge in science, mathematics, and engineering. Among these, the b-STEM dimension consists of five items that assess teachers' ability to integrate interdisciplinary STEM content knowledge into student-centered learning. For example, one item states: "I can design lessons that appropriately integrate interdisciplinary STEM content knowledge for student-centered learning."

However, these scales are focusing on science, math, and engineering areas, which basically narrow the context related to interdisciplinary teaching. In essence, interdisciplinary teaching

encompasses a broader range of subjects, encouraging students to draw connections between diverse fields, such as arts and humanities alongside STEM (Seo et al., 2024). It includes instructional strategies drawn from multiple disciplines without a specific focus on STEM. Therefore, it becomes evident that there is a need for a scale that more accurately encompasses interdisciplinary teaching, without limiting its scope to STEM, but rather including it in a broader context. Moreover, the previously mentioned scales consisted of Likert-scale items. Regarding this, Bandura (2005) states that scales with a limited number of rating levels are less sensitive and reliable compared to those offering a broader range of gradations. Therefore, in the current study, the items were developed to be rated in increments of ten, ranging from 0 to 100. Through developing a structured and validated tool to assess teacher self-efficacy on interdisciplinary teaching, this study aims to provide an understanding of areas where teachers feel less confident, thereby guiding professional development and educational policy. Investigating teacher perception in this context and developing higher self-efficacy is crucial, as interdisciplinary teaching is increasingly recognized for its potential to enhance educational outcomes by integrating knowledge across different subjects.

Method

The current study is a scale development study aiming to develop a valid and reliable scale to assess teachers' self-efficacy in designing teaching based on an interdisciplinary approach. Ethical approval for the study was given by Mersin University Educational Science Ethics Board Committee by the decision number 1/25 dated 12.11.2024. A scaling approach based on ranked sums was employed during the scale development process, focusing on positioning individuals differently on the scale according to their responses to the items (Crocker & Algina, 1986; Tezbaşaran, 2004). The purpose was to enable inferences to be drawn based on participants' responses. Therefore, in the scale development process, scaling by the cumulative sums technique, which is one of the rater-centered approaches, was utilized. In this section, the study group, the development process of the self-efficacy scale for designing interdisciplinary teaching and the data analysis procedures are explained.

Study Group

In this study, two separate groups were utilized; one was formed for exploratory factor analysis (EFA), while the other was an independent sample designated for confirmatory factor analysis (CFA).

Study Group for the EFA

The study group for the EFA consisted of 350 in-service teachers in various subject areas. The participants volunteered to be part of this study. There were 226 females and 124 males in the group. In terms of educational background, most of the participants held a bachelor's degree and a master's degree, while fewer had a doctoral degree. Additionally, most of them were middle and high school teachers. Detailed information regarding their gender, educational background, and professional experience is presented in Table 1. Following the EFA based on the response of this group, additional data were collected using the final version of the scale to examine its structural validity through CFA.

Table 1

Demographic Information of the EFA Study Group

Variable	Categorical	f	%
Gender	Male	124	35.4
	Female	226	64.6
School Type	Preschool	3	.9
	Primary School	72	20.6
	Middle School	139	39.8
	High School	68	19.4
Educational Background	Bachelor's Degree	206	58.9
	Master's Degree	129	36.9
	PhD	8	2.3
Professional Experience	0-5 years	32	9.1
	6-10 years	34	9.7
	11-15 years	69	19.7
	16-20 years	84	24
	21 years and above	131	37.4

Study Group for the CFA

The study group for the CFA consisted of a total of 231 in-service teachers who voluntarily participated in the research, including 148 females and 83 males. Detailed explanations regarding their other demographic information are presented in Table 2.

Table 2

Demographic Information of the CFA Study Group

Variable	Categorical	f	%
Gender	Male	83	35.9
	Female	148	64.1
School Type	Preschool	2	.9
	Primary School	51	22.1
	Middle School	88	38.1
	High School	43	18.6
Educational Background	Bachelor's Degree	139	60.2
	Master's Degree	83	35.9
	PhD	5	2.2
Professional Experience	0-5 years	19	8.2
	6-10 years	23	10
	11-15 years	53	22.9
	16-20 years	52	22.5
	21 years and above	84	36.4

Scale Development

The Preliminary Scale Form was developed through a three-step process: focus group interviews, a literature review and item pool generation, and, lastly, the content validity assessment and data analysis procedures. During the first step, it includes the 'qualitative steps' such as focus group interviews and literature review in the scale development process. As a result of these qualitative steps, an item pool was developed. Then, according to expert's

opinions, a draft scale was generated. Lastly, the collected data from the implementation of this scale was analyzed.

Preliminary Scale Form

In this step, a focus group interview was conducted with three teachers at the doctoral level who had taken the course on Designing Interdisciplinary Teaching. Among the teachers, one specialized in curriculum design, another in instructional technology, and the third in pedagogical practice. The interview was scheduled at a time and day when their availability was least constrained. To form the participants for the focus group interview, a study group was established using the convenience sampling method. During the interview, questions were posed to define the self-efficacy of teachers in designing interdisciplinary teaching. These questions included:

- How should the learning environment be designed when planning the context?
- How should students be considered during the planning of the context?
- What should be considered when identifying a phenomenon (a problem worth solving and sufficiently complex, or a topic that integrates multiple learning areas)?
- What professional competencies should one possess?
- What cognitive competencies should one possess?
- If you have previously designed an Interdisciplinary Teaching course, please briefly describe it.

Literature Review and Item Pool Generation

In this step, articles and theses related to interdisciplinary teaching were reviewed (Choi & Pak, 2006; Jacobs, 1989; Kezar & Elrod, 2012; Morgil & Yilmaz, 1999; Yarımca, 2011). At this stage, statements defining interdisciplinary teaching were gathered. Additionally, previously developed scales in this context were examined to provide insights into item development regarding teacher self-efficacy and the interdisciplinary teaching profession. (Buechel et al., 2024; Chai et al., 2019; Schmitz & Schwarzer, 2000; The Friday Institute for Educational Innovation, 2012; Tschannen-Moran & Hoy, 2001; Yang et al., 2025). The expressions in the item pool, comprising a total of 60 items, were developed by the researchers and refined into statements that define interdisciplinary teaching.

Content Validity

For the content validity, 10 experts from various areas reviewed the items. The experts were consisting of professionals in the field of curriculum development, assessment and evaluation, science education (3 experts), a doctoral graduate in mathematics education, a doctoral graduate in curriculum development, and academics working in various disciplines. The expert evaluation form, which included the preliminary scale form, was sent via email for their evaluation. The experts were asked to evaluate each item in the preliminary scale form by indicating the degree to which the item was relevant to the construct (ranging from 4, highly

relevant to 1, not relevant at all), the clarity and comprehensibility of the item for the target audience (ranging from 4, very clear to 1, not clear at all), and the suitability of the item for the intended rating scale (ranging from 4, highly suitable to 1, not suitable at all). They were also requested to provide suggestions and corrections for the items, if they preferred. Following the expert evaluations, Davis' technique was utilized to calculate content validity. In this technique, the number of experts who selected the 'excellent match' option (4) for each item was divided by the total number of experts to calculate the content validity index (CVI). A CVI value of 0.80 was set as the criterion for content validity (Davis, 1992). As a result of the analysis, seven items (I10, I18, I41, I43, I51, I54, and I58) appeared not to meet the accepted CVI criterion of 0.80. In addition, after the expert evaluations and content validity analysis, 14 items (I14, I16, I22, I26, I34, I35, I45, I46, I48, I50, I53, I57, I59, and I60) were removed due to containing similar expressions that were deemed to convey the same meaning. Consequently, the 60-item form was reduced to 39 items based on the expert evaluations. Also, it was ensured that the positive and negative statements in the scale were distributed randomly within the measurement tool, avoiding any regular sequence or grouping (Karasar, 1995; Kırcaali İftar, 1999; Tezbaşaran, 1997), and a draft scale was generated accordingly.

Data Analysis Procedures

The construct validity of the scale was assessed through the exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). Additionally, convergent validity, which provides information about the relationship between factors and their underlying items, and Horn's Parallel Analysis, which offers supplementary evidence for determining the structure and number of factors, were employed. Factor analysis is a set of methods used to transform interrelated data structures into a smaller number of independent new data structures, in other words, to identify and, if necessary, name the variables (factors, dimensions, components) assumed to explain the underlying construct. The primary goal of EFA is to determine the factor loading matrix, which leads to identify the constructs and the items within these constructs (Alpar, 2018). Preliminary analyses, EFA, and CFA were conducted using IBM SPSS Statistics (Version 23) and Jamovi (Version 1.8.1) package programs.

Before conducting the EFA, preliminary analyses were conducted to test basic assumptions (missing data, univariate and multivariate outliers). Since the measured variable was an extreme trait, skewness in the distribution was one of the primary expectations, and no assumptions regarding normality were attempted to be met. No missing data or erroneous forms were observed in the dataset since the data was collected through a digital portal. Among the data from 450 participants, no univariate outliers were found (Z : 3.53 to -2.45, reference range: -4 to 4, Tabachnick & Fidell, 2007). Based on Mahalanobis distances, multivariate outlier analysis ($\chi^2 = 72.05466$, $p > 0.001$) identified 50 observations as multivariate outliers, which were subsequently excluded from the analysis. Following this stage, the presence of multicollinearity was assessed by examining whether tolerance values were above 0.20 and variance inflation factor (VIF) values were below 5 (Hair et al., 2010). Items I6, I8, I12, I14, and I17 were excluded from the analysis due to VIF values exceeding 5, while items I21, I22, I23, I26, I27, I28, I29, I30, I31, I34, I35, I36, and I37 were excluded due to tolerance values below 0.20. The tolerance and VIF values for the remaining 21 items occurred within the specified ranges, indicating no violation of multicollinearity.

The Durbin-Watson statistic was observed to be 1.96 indicating that the residuals are independent of one another since the value is very close to 2 (Kalaycı, 2005). In order to assess the suitability of the data for factor analysis, the Bartlett's Test of Sphericity and the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy were examined. A KMO value exceeding .80 is considered to indicate a very good level for factorability of data (Alpar, 2018), while a statistically significant result for Bartlett's Test of Sphericity indicates the meaningful correlations among variables in the dataset. This outcome also supports the assumption of multivariate normality (Çokluk et al., 2012). The inter-item correlations were assessed, and Varimax rotation was chosen for rotation (Field 2009; Saraçlı, 2011). The scree plot and Horn's parallel analysis were utilized to determine the factor structure. In the analysis of the data for the final structure, particular attention was paid to ensuring that each item's communality was at least .30, the minimum factor loading was .45, and the difference between cross-loadings on two factors exceeded .10 (Büyüköztürk, 2011; Tabachnick & Fidell, 2013).

Confirmatory factor analysis (CFA) examines the extent to which the empirical data fit the hypothesized model (DeVellis, 2017), and it is fundamentally based on structural equation modeling (Tabachnick & Fidell, 2013). Accordingly, CFA is used to test how well the structure identified in the EFA fits another independent sample, thereby allowing for the confirmation of the structure and the provision of evidence for construct validity. Basic assumptions were examined prior to CFA, and no violation was observed. Due to the digital collection of the data, no missing or erroneous entries were found. There was no univariate outliers in the dataset; however, based on Mahalanobis distance calculations for 350 participants, 50 multivariate outliers were excluded from the analysis, as their values exceeded the critical χ^2 value ($\chi^2(210) > 5903,55$ $p > .05$). To assess multicollinearity, tolerance values were found to be above 0.20, and variance inflation factor (VIF) values were below 5 for all 21 items. Multicollinearity was assumed, since both metrics fell within acceptable thresholds. Additionally, the Durbin-Watson statistic was calculated as 2, indicating that the residuals were independent.

In terms of reliability, Cronbach's alpha and stratified alpha coefficients were used for the EFA sample, while composite reliability (CR) was utilized for the CFA sample. Composite reliability, which evaluates the overall reliability of a set of latent constructs based on their indicators, is particularly useful for assessing internal consistency (Raykov, 1997). In this study, CR values were calculated based on CFA results and were employed as a verification tool alongside Cronbach's alpha coefficients.

Findings

In this section, findings on validity, EFA, and CFA results of the data, followed by reliability analyses, are presented, respectively.

Findings on Validity

Exploratory Factor Analysis

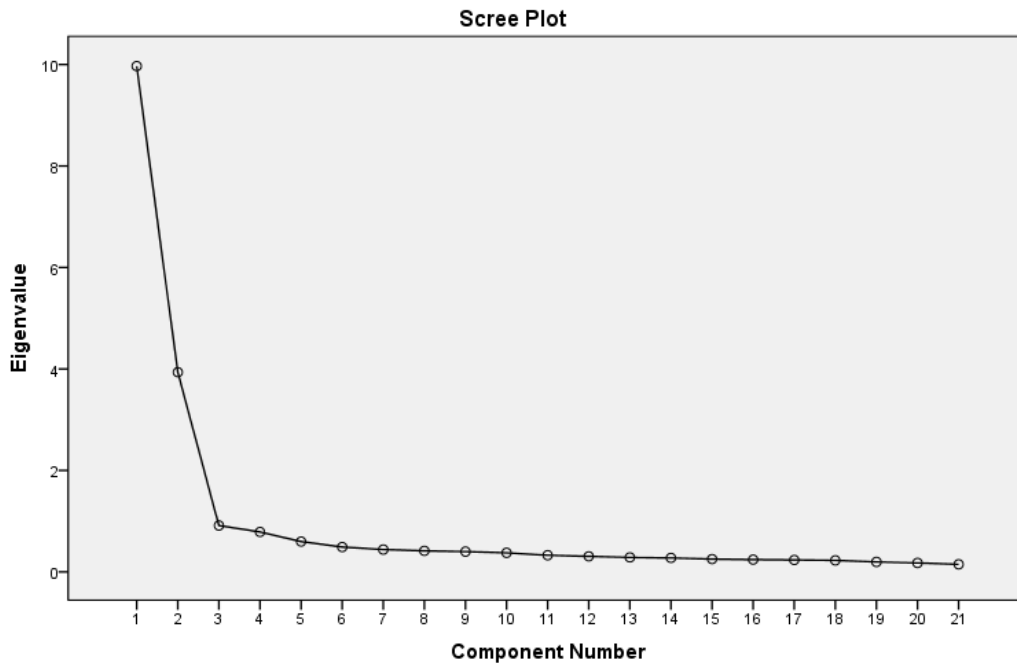
Preliminary analyses were conducted to assess the suitability of the data obtained from the 39-item pilot form for further analysis. The results of Bartlett's Test of Sphericity and the Kaiser-

Meyer-Olkin (KMO) Measure of Sampling Adequacy indicated no violation. Following the exploratory factor analysis (EFA) conducted on the 39 items, the number of factors was determined through a combination of the scree plot, Horn's parallel analysis, and the total variance explained. The findings regarding these analyses are presented below.

Scree plot presented in Figure 1 showed that the eigenvalues begin to gradually decline starting from the first factor. However, based on a subjective interpretation, the data appear to support the acceptance of a two- or three-factor structure.

Figure 1

Scree Plot



The total variance, as presented in Table 3, indicated a two-factor structure in which each factor had an eigenvalue greater than 1 and collectively accounted for 66.21% of the total variance. In determining the number of factors, more objective criteria were also considered—specifically, the results of Horn's Parallel Analysis, which is based on the Kaiser criterion. The items grouped under these two factors, along with their communalities, factor loadings, and corresponding factor assignments, are presented in Table 3. During the analyses, no items were identified with communalities below .30, factor loadings below .45, or cross-loadings on two factors with a loading difference of less than .10.

Table 3

Exploratory Factor Analysis for Scale

Item	Component		Extraction	Item-total correlation
	1	2		
Item33	.82		.72	.75
Item5	.82		.75	.72
Item11	.82		.75	.72
Item16	.81		.73	.71
Item2	.78		.71	.67
Item19	.77		.68	.67
Item38	.77		.66	.68
Item20	.76		.69	.66
Item9	.74		.65	.64
Item4	.74		.64	.63
Item1	.73		.67	.62
Item13	.73		.63	.63
Item39		.56	.76	.78
Item18		.57	.70	.73
Item25		.56	.68	.72
Item15		.54	.64	.71
Item10		.56	.66	.71
Item24		.60	.46	.46
Item32		.59	.63	.66
Item7		.56	.60	.66
Item3		.52	.50	.60
Empirical Eigenvalue	9.97	3.94	.91	.79
Generated Eigenvalue	1.46	1.39	1.32	1.27
TVE	47.47	18.74		
CPV	47.47	66.21		
Cronbach's Alpha	.95	.92		
Total Cronbach's Alpha		.93		

TVE: Total Variance Explained, CPV: Cumulative Percent Variance, KMO=.910; Barlett's test=6291.83, $p<.001$

The scale, consisting of 21 items, was found to form a two-factor structure. The eigenvalues of the two factors were 9.9 and 3.9, explaining a portion of the total variance as 47.46% and 18.74%, respectively. A variance rate of 40% or more is considered to be adequate or acceptable for multifactorial structures in the social sciences (Scherer et al., 1988), a total variance explained of 66.21% is considered to indicate a satisfactory model fit. The scale's first dimension comprises 12 items (1, 2, 4, 5, 9, 11, 13, 16, 19, 20, 33, 38) and the second dimension includes 9 items (3, 7, 10, 15, 18, 24, 25, 32, 39).

The names of dimensions were assigned as Competences and Challenges based on the items loaded in factors. The items loaded in the Competences dimension highlight active competencies—skills teachers already possess or actively practice, such as interdisciplinary planning, research, and collaboration. These items measure the teacher's efficacy in their

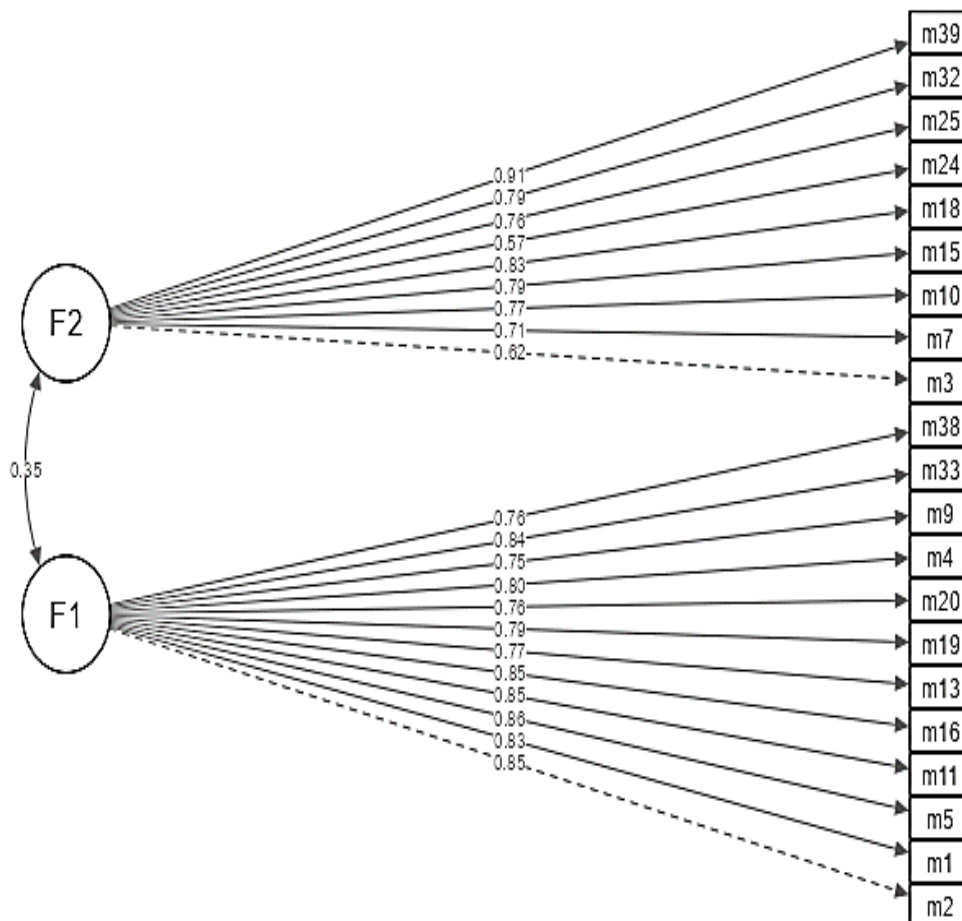
abilities, confidence, and proactive efforts in designing and implementing interdisciplinary learning experiences (e.g., I can research the same topic from the perspectives of different disciplines). In the items loading onto the Challenges dimension, teachers report difficulties, limitations, or negative perceptions regarding interdisciplinary approaches. This dimension reveals gaps in interdisciplinary teaching, framing the factor as areas needing support, training, or mindset shifts. (e.g., I struggle to define a topic, issue, or theme that integrates different disciplines.). The factor loadings of the items in the components ranged from .73 to .82 and .52 to .60, respectively (see Table 3).

Structural Model Validation

The confirmatory factor analysis (CFA) was conducted with data collected from a different sample group having similar characteristics, in order to verify the factor structure identified through the EFA. This analysis was carried out to provide additional evidence for the construct validity of the developed scale.

Figure 2

Standardized Path Coefficients of Interdisciplinary Teaching Design Self-Efficacy Scale



The model fit results indicated a good fit for the data since they were within the good and acceptable range, $\chi^2 = 1413.84$, $df = 516$, ($\chi^2 / df = 2.74$), RMSEA = .09, CFI = .92, RFI = .86; NNFI = .92,

SRMR =.05 (Çokluk, et al. 2010; Kline, 2011). Moreover, the path coefficients for items were found to be statistically significant ($p<.001$). In addition, the relationship between scale dimensions was indicating a moderate negative and significant correlation ($r=-.30$, $p<.01$).

Convergent and Discriminant Validity

In this study, for estimating internal consistency, composite reliability (CR) coefficients were calculated in addition to Cronbach's Alpha. Convergent validity refers to the assessment of the correlation levels among multiple indicators that are theoretically related and represent the same construct. In establishing convergent validity, factor loadings of the indicators were considered as well as composite reliability (CR) and average variance extracted (AVE) values (Yurdugül & Sırakaya, 2013). As presented in Table 4, factor loadings were between .57 and .91, meeting the suggested threshold of exceeding .50 (Hair et al., 2010). The AVE value ranges between 0 and 1. For convergent validity to be deemed adequate, the AVE value should exceed 0.50 (Ab Hamid et al., 2017; Shrestha, 2021). The CR values obtained for each factor were found to be .96 and .92. These values, which provide evidence for reliability by indicating the strength of the relationship between each factor and its underlying items, were found to be above .70, indicating that the measurements based on the items under each factor are reliable. By comparing the computed CR values with the AVE values calculated for each factor, findings related to convergent validity were obtained. The AVE values obtained for each factor were 0.66 and 0.59, respectively. Both values meet the criterion of being greater than 0.50 (Fornell & Larcker, 1981). The results indicate that the scale provides both convergent and discriminant validity criteria, and that the items, distributed under two distinct factors, are related both to each other and to the underlying constructs they are intended to measure.

Table 4

The Internal Consistency and Convergent Validity Results of the Scale

Factor	ITEM	Factor Load	CR	AVE
Competences	1	.85	.96	.66
	2	.84		
	4	.83		
	5	.82		
	7	.81		
	9	.79		
	10	.79		
	12	.78		
	14	.81		
	15	.80		
	19	.76		
Challenges	20	.96	.92	.59
	3	.83		
	6	.81		
	8	.81		
	11	.80		
	13	.83		
	16	.79		

Factor	ITEM	Factor Load	CR	AVE
	17	.78		
	18	.77		
	21	.67		
Total			.98	.65

Findings on Reliability

For the final version of the Interdisciplinary Teaching Design Self-Efficacy Scale, which consists of 21 items and two dimensions, internal consistency reliability coefficients—Cronbach’s Alpha and McDonald’s Omega—were calculated by data obtained from the EFA and CFA study groups using the Jamovi software (version 1.8.1). Cronbach’s Alpha, one of the most commonly used reliability coefficients in education and psychology, is noted to yield values close to the true reliability of a measurement tool, particularly when the instrument contains a sufficient number of items and is unidimensional (Cronbach, 1951). Although commonly attributed to Cronbach, the coefficient was originally introduced through the work of Guttman (1945) and is also referred to as a lower bound estimate of reliability (Sijtsma, 2009). As an alternative to Cronbach’s Alpha, McDonald’s Omega, estimates reliability based on item factor loadings (McDonald, 1999, Yurdugül, 2006). Omega typically produces reliability estimates that are equal to or higher than those yielded by Cronbach’s Alpha (Kula Kartal & Mor Dirlik, 2016). In this context, the reliability coefficients calculated using the data from both EFA and CFA study groups are presented in Table 5.

Table 5

Reliability Indices for Scale Dimensions in CFA

Factor	α	ω	AVE
Competences	0.957	0.957	0.649
Challenges	0.918	0.919	0.560

A Cronbach’s Alpha coefficient above 0.70 is generally expected to indicate adequate internal consistency reliability (DeVellis, 2012). In this context, for the Interdisciplinary Teaching Design Self-Efficacy Scale, the internal consistency coefficients were calculated as 0.957 for the Competences dimension and 0.918 for the Challenges dimension. These values are considered indicators of excellent internal consistency and suggest that the scale demonstrates high reliability both overall and at the subscale level. Additionally, McDonald’s Omega coefficients, calculated using the CFA data for each dimension, were found to be 0.957 and 0.919, respectively. These values, obtained through separate reliability analyses based on different datasets, indicate that the data gathered using this scale are likely to yield reliable results.

Interpretation of Scale Scores

The Interdisciplinary Teaching Design Self-Efficacy Scale (Appendix A) consists of 21 items each rated on an 11-point scale ranging from 0 to 100 in increments of 10 (i.e., 0, 10, 20, ..., 100). The scale assesses self-efficacy, with anchor points defined as follows: 0 = “not at all true of me” and 100 = “very true of me”. Participants select the numerical value that best reflects their level of assurance for each item. Scaling with 11 points was preferred, as it has been reported

to increase the sensitivity of measurement and the reliability of responses by allowing participants to express their perceptions in greater detail (Chyung, Roberts, Swanson & Hankinson, 2017; Özkan Bindak, 2021; Preston & Colman, 2000). In addition, Bandura (1977; 2001) states that sensitive intervals provide a more accurate measurement of the belief level while scales with fewer digits are less sensitive and reliable (Bandura, 2006).

The scale's structure includes two sub-dimensions, as detailed in Table 4. Items in the Competencies dimension are positively framed (e.g., When designing lessons, I can successfully incorporate both instructional approaches and subject matter from various disciplines.), while those in the Challenges dimension are negatively worded (e.g., I struggle to identify a topic, problem, or theme that adequately incorporates multiple disciplines.). The possible total scores range from 0 to 1200 (Competencies) and 0 to 900 (Challenges), while mean scores for items and both dimensions form a 0-100 metric for comparative analysis. A higher score on the Competencies dimension indicates greater self-efficacy, whereas a higher score on the Challenges dimension reflects lower self-efficacy. Therefore, when calculating the total self-efficacy score for the entire scale, it is recommended to reverse-code the Challenges items.

Discussion

The literature review within the scope of this study highlights the importance of interdisciplinary learning and the ability to design learning environments through this approach. Moreover, a need for a standardized and comprehensive measurement tool designed to assess teacher self-efficacy in this context is revealed. Acknowledging this gap, the current study was conducted with two main objectives. Firstly, it is aimed to develop a rigorously designed measurement instrument to assess the sub-dimensions of self-efficacy in designing teaching based on an interdisciplinary approach. Secondly, it is proposed to provide a valid and reliable tool capable of effectively measuring teachers' self-efficacy in designing interdisciplinary teaching within the Turkish educational context.

The validity studies for the development of the Interdisciplinary Teaching Design Self-Efficacy Scale focused on ensuring both content and construct validity. To establish content validity, focus group interviews and a comprehensive literature review were conducted, and an item pool was generated based on the findings. Subsequently, expert reviews were obtained for the items, and content validity ratios (CVR) and content validity indices (CVI) were calculated. Following comparison of these indices with critical values, the final items for the trial form were determined, resulting in 39 items selected from an initial pool of 60. For construct validity, exploratory factor analysis (EFA) was conducted. The analysis revealed that the final form consisted of 21 items, demonstrating a two-factor structure. The factor structure was further confirmed through Horn's parallel analysis (1965), supporting the retention of the two-factor model. Reliability analyses were performed at both the item and subscale levels. The calculated composite reliability (CR), along with convergent and discriminant validity indices, indicated that both the items and subscales represented the underlying structure adequately. The Cronbach's alpha and McDonald's omega coefficients computed for the confirmed structure and its factors exceeded critical thresholds, providing further evidence of reliability. Moreover, the ratio of the number of factors (Q) to the number of items (P), which is a

commonly recommended guideline for reliability (Haksitian et al, 1982; as cited in Stevens, 2002), was found to be less than 0.30, Q/P ratio ($2/21 = 0.095$), indicating support for reliability.

In this study, the primary objective was to develop a measurement designed to assess teachers' self-efficacy in designing interdisciplinary teaching and to test its validity and reliability. As a result of the research process, the intended scale was successfully developed. Following data analysis, the instrument was entitled the Interdisciplinary Teaching Design Self-Efficacy Scale.

The pilot form of the scale comprised 39 items. Based on the analyses conducted after data collection, 21 items were retained in the final version, while 18 items that did not meet the predetermined criteria were excluded. The number of factors was determined based on the scree plot, Kaiser's eigenvalues, and parallel analysis. Accordingly, it was concluded that the remaining 21 items formed into two factors, taking into account the scree plot and factor loading values obtained through exploratory factor analysis. These sub-dimensions were labeled Competences and Challenges. The first factor, Competences, comprised 12 items, all of which were semantically positive. The second factor, Challenges, consisted of 9 items, all of which were semantically negative in content. This emergent structure was subsequently evaluated through confirmatory factor analysis, and the standardized factor loadings, t-values, and significance levels were examined. Additionally, the calculated fit indices were found to be within excellent or acceptable thresholds.

The findings indicate that the Interdisciplinary Teaching Design Self-Efficacy Scale, which consists of 21 items across 2 factors (12 positive and 9 negative statements) rated on a 10-point scale, is a psychometrically valid and reliable measurement instrument. The scale is considered suitable for administration to teachers across various disciplines and is expected to validly and reliably reveal teachers' self-efficacy levels regarding the design of interdisciplinary teaching. The moderate correlations observed between the sub-dimensions of self-efficacy suggest that these components can be implemented independently as well as synergistically. However, considering these dimensions holistically rather than in isolation will further strengthen the utility and interpretive power of the scale.

Declarations

Ethics statement: This study was approved by the Ethics Committee of Mersin University, Educational Sciences. The ethics approval was granted on 12.11.2024 under the protocol number 2025/25. All procedures involving human participants were conducted in accordance with institutional and national ethical standards, the Helsinki Declaration, and applicable regulations. Informed consent was obtained from all participants.

Author Contributions (CRediT): H. S. Kılınç: Conceptualization, Investigation, Formal analysis, Resources, Writing – original draft. G. Gök: Conceptualization, Methodology, Visualization, Writing – review & editing, Supervision. N. B. Uzun: Formal analysis, Methodology

Funding: This research received no specific grant from any funding agency.

Conflict of Interest: The authors declare no conflict of interest.

AI Tools Disclosure: The authors used ChatGPT (OpenAI) solely for language refinement and grammar correction during the revision process. No AI tool was used to generate research ideas, theoretical content, data, analysis, or interpretations. All substantive intellectual contributions were made by the authors.

Data Availability Statement: Data cannot be shared due to ethical/legal restrictions.

References

- Ab Hamid, N. R., Alwi, S. F. S., & Mohd, R. (2017). An empirical analysis of the antecedents and performance consequences of using the Moodle platform. *International Journal of Information and Communication Technology*, 9(1), 1–16.
- Aktaş, İ., & Özmen, H. (2024). STEM-TPAB Öz-Yeterlik ölçeği: Türkçeye uyarlama çalışması [STEM-TPACK Self-Efficacy Scale: Adaptation Study into Turkish]. *Uludağ Üniversitesi Eğitim Fakültesi Dergisi*, 37(2), 798–829.
- Alpar, R. (2018). *Uygulamalı istatistik ve geçerlik güvenirlik [Applied Statistics: Validity and Reliability]*. Detay Yayıncılık.
- Apostel, L. (1970). *Interdisciplinarity: Problems of teaching and research in universities*. OECD Publishing.
- Aydın, G., & Balım, A. G. (2005). Yapılandırmacı yaklaşıma göre modellendirilmiş disiplinlerarası uygulama: Enerji konularının öğretimi [An interdisciplinary application based on constructivist approach: Teaching of energy topics]. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 38(2), 145–166.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215. <https://doi.org/10.1037/0033-295X.84.2.191>
- Bandura, A. (1994). Self-efficacy. In V. S. Ramachandran (Ed.), *Encyclopedia of human behavior* (Vol. 4, pp. 71–81). Academic Press.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52(1), 1–26. <https://doi.org/10.1146/annurev.psych.52.1.1>
- Bandura, A. (2005). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (Vol. 5, pp. 307–337). Information Age Publishing.
- Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307–337). Information Age Publishing.
- Buechel, C., Daugherty, M. K., Carter, V., & Topalcengiz, E. S. (2024). Elementary teachers' self-efficacy and its role in STEM implementation. *Turkish Journal of Education*, 13(3), 217–238. <https://doi.org/10.19128/turje.1267839>
- Büyükoztürk, Ş. (2011). *Sosyal bilimler için veri analizi el kitabı [Handbook of data analysis for social sciences]*. Pegem Akademi.
- Capa, Y., Çakıroğlu, J., & Sarıkaya, H. (2005). The development and validation of a Turkish version of teachers' sense of efficacy scale. *Eğitim ve Bilim*, 30(137), 74–81.

- Chai, C. S., Jong, M. S.-Y., Yin, H. B., Chen, M., & Zhou, W. (2019). Validating and modelling teachers' technological pedagogical content knowledge for integrative science, technology, engineering and mathematics education. *Educational Technology & Society*, 22(3), 61–73.
- Chen, X. (2020). Pre-service teachers' self-efficacy of interdisciplinary team teaching through the use of collaborative concept map. *International Journal of Technology in Teaching and Learning*, 15(2), 76–94.
- Choi, B. C. K., & Pak, A. W. P. (2006). Multidisciplinarity, interdisciplinarity and transdisciplinarity in health research, services, education and policy: 1. Definitions, objectives, and evidence of effectiveness. *Clinical and Investigative Medicine*, 29(6), 351–364.
- Chyung, S. Y., Roberts, K., Swanson, I., & Hankinson, A. (2017). Evidence-based survey design: The use of continuous rating scales in surveys. *Performance Improvement*, 56(7), 5–13. <https://doi.org/10.1002/pfi.21763>
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Harcourt.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334.
- Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2010). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları* [Multivariate statistics for the social sciences: Applications with spss and lisrel]. Pegem Akademi Yayıncılık.
- D'Souza, M. S., Nikku, B. R., & Field, C. (2021). Interdisciplinary teaching practices: Reflections from a teaching triangle. *Journal of Nursing Education and Practice*, 11(5), 10.
- Darbellay, F. (2022). When interdisciplinarity meets creativity: Exploring interdisciplinary creative teaching in the 21st century. In *Handbook of interdisciplinary teaching and administration*. Edward Elgar Publishing. <https://doi.org/10.4337/9781035309870.00029>
- Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research*, 5(4), 194–197. [https://doi.org/10.1016/s0897-1897\(05\)80008-4](https://doi.org/10.1016/s0897-1897(05)80008-4)
- DeVellis, R. F. (2012). *Scale development: Theory and applications* (3rd ed.). SAGE Publications.
- DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Sage.
- Drake, S. M., & Burns, R. C. (2004). *Meeting standards through integrated curriculum*. Association for Supervision and Curriculum Development.
- Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Sage.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50.
- Friday Institute for Educational Innovation. (2012a). *Teacher efficacy and attitudes toward STEM survey (T-STEM)*. North Carolina State University. <http://miso.ncsu.edu/articles/t-stem-survey-2>
- Grossman, P., Wineburg, S., & Woolworth, S. (2001). Toward a theory of teacher community. *Teachers College Record*, 103(6), 942–1012.
- Guercio, C. J. (1998). *An interdisciplinary curriculum and its positive effect on student motivation in the classroom* [Unpublished master's thesis]. Caldwell College.
- Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, 10(4), 255–282.
- Gür, T. M. (2003). Araştırma ve eğitiminde disiplinlerarasılık [Interdisciplinarity in research and education]. In O. N. Babüroğlu (Ed.), *Eğitimin geleceği* [The future of education] (pp. 182–203). Sabancı Üniversitesi Yayınları.

- Gryson, T., Strubbe, K., Valcke, T., & Vanderlinde, R. (2024). *Lifelong Learning through Teacher Design Teams for Interdisciplinary Teaching in Secondary Vocational Education: The Perspective of Different Stakeholders*. IntechOpen. <https://doi.org/10.5772/intechopen.114206>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson Education International.
- Holley, K. A. (2024). Interdisciplinary curricula. In *Interdisciplinary education* (pp. 300–303). Edward Elgar Publishing. <https://doi.org/10.4337/9781035317967.ch66>
- Jacobs, H. H. (1989). *Interdisciplinary curriculum: Design and implementation*. Association for Supervision and Curriculum Development.
- Kalaycı, Ş. (2005). *SPSS uygulamalı çok değişkenli istatistiksel analiz teknikleri* [Applied multivariate statistical analysis techniques using SPSS]. Asil Basım Yayın.
- Karacaoğlu, Ö. C. (2011). *Online eğitimde program geliştirme* [Curriculum development in online education]. İhtiyaç Yayıncılık.
- Karasar, N. (1995). *Bilimsel araştırma yöntemi* [Curriculum development in online education]. Sim Matbaası.
- Kezar, A., & Elrod, S. (2012). Facilitating interdisciplinary learning: Lessons from Project Kaleidoscope. *Change: The Magazine of Higher Learning*, 44(1), 16–25. <https://doi.org/10.1080/00091383.2012.635999>
- Kırcaali İftar, G. (1999). Ölçme [Measurement]. In *Sosyal bilimlerde araştırma yöntemleri* [Research methodology in social sciences]. Anadolu Üniversitesi Yayınları.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). The Guilford Press.
- Kubat, U. (2018). Identifying the Individual Differences Among Students During Learning and Teaching Process by Science Teachers. *International Journal of Research in Education and Science*, 4(1), 30–38. <https://doi.org/10.21890/IJRES.369746>
- Kula Kartal, S., & Mor Dirlik, E. (2016). Geçerlik kavramının tarihsel gelişimi ve güvenilirlikte en çok tercih edilen yöntem: Cronbach Alfa Katsayısı [Historical development of the concept of validity and the most preferred technique of reliability: Cronbach alpha coefficient]. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 16, 1865–1879.
- Lenka, S. K., & Singh, P. (2024). Implementing interdisciplinary teaching approaches: Challenges and strategies in teacher education under NEP 2020. *International Journal for Multidisciplinary Research*, 6(1). <https://doi.org/10.36948/ijfmr.2024.v06i01.13995>
- Li, Z. (2023). Analysis of “Teaching” and “Learning” in interdisciplinary learning. *Journal of Contemporary Educational Research*, 7(12), 248–252. <https://doi.org/10.26689/jcer.v7i12.5817>
- Martinot, A. (2023). *The foundations of an interdisciplinary pedagogy*. Oxford University Press eBooks. <https://doi.org/10.1093/oso/9780197511275.003.0005>
- McDonald, J., & Czerniak, C. (1994). Developing interdisciplinary units: Strategies and examples. *School Science and Mathematics*, 94(1), 5–10.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum Associates.
- Morgil, F. İ., & Yılmaz, A. (1999). Fen öğretmeninin görevleri ve nitelikleri, fen öğretmeni yetiştirilmesine yönelik öneriler [Roles and qualifications of science teachers, and suggestions for science teacher education]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 15(15), 181–186.

- Özkan Bindak, H. (2021). *Eğitim araştırmalarında ölçme ve değerlendirme* [Assessment and evaluation in educational research]. Pegem Akademi.
- Pajares, F., Hartley, J., & Valiante, G. (2001). Response format in writing self-efficacy assessment: Greater discrimination increases prediction. *Measurement and Evaluation in Counseling and Development*, 33(4), 214–221. <https://doi.org/10.1080/07481756.2001.12069027>
- Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104(1), 1–15. [https://doi.org/10.1016/S0001-6918\(99\)00050-5](https://doi.org/10.1016/S0001-6918(99)00050-5)
- Rafiq, S., Kamran, F., & Afzal, A. (2024). Investigating the benefits and challenges of interdisciplinary education in higher education settings. *Journal of Social Research Development*, 5(1), 87–100. <https://doi.org/10.53664/JSRD/05-01-2024-08-87-100>
- Rauf, A., & Bhatti, N. (2024). Interdisciplinary approaches to teaching social sciences in the 21st century. *Multidisciplinary Journal of Emerging Needs of Curriculum*, 1(2).
- Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement*, 21(2), 173–184.
- Saada, M. (2024). Effectiveness of interdisciplinary teacher teams among English teachers in public secondary schools in Jolo. *Journal of Education and Academic Settings*, 1(1), 1–16. <https://doi.org/10.62596/mcyw5g50>
- Saraçlı, S. (2011). Faktör analizinde yer alan döndürme metotlarının karşılaştırmalı incelenmesi üzerine bir uygulama [An application on comparison the extracting methods in factor analysis]. *Düzce Üniversitesi Sağlık Bilimleri Enstitüsü Dergisi*, 1(3), 22–26.
- Scherer, R. F., Luther, D. C., Wiebe, F. A., & Adams, J. S. (1988). Dimensionality of coping: Factor stability using the Ways of Coping Questionnaire. *Psychological Reports*, 62(3), 763–770. <https://doi.org/10.2466/pr0.1988.62.3.763>
- Schmitz, G. S., & Schwarzer, R. (2000). Selbstwirksamkeitserwartung von Lehrern: Längsschnittbefunde mit einem neuen Instrument. *Zeitschrift für Pädagogische Psychologie*, 14(1), 12–25.
- Scott, E. K., & White, B. R. (2024). An empirical study of cultivating innovative practice abilities in an interdisciplinary education environment in Australia. *Research and Advances in Education*, 3(5), 53–63. <https://doi.org/10.56397/rae.2024.05.06>
- Seo, S., Van Orman, D., Beattie, M., Paxson, L., & Murray, J. (2024). Transforming learning orientations through STEM interdisciplinary project-based learning. *Education Sciences*, 14(11), 1154. <https://doi.org/10.3390/educsci14111154>
- Shrestha, N. (2021). Factor analysis as a tool for survey analysis. *American Journal of Applied Mathematics and Statistics*, 9(1), 4–11.
- Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, 74(3), 421–438.
- Stevens, J. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Lawrence Erlbaum Associates.
- Suraco, T. L. (2006). An interdisciplinary approach in the art education curriculum [Unpublished master's thesis]. Georgia State University.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Experimental designs using ANOVA* (Vol. 724). Thomson/Brooks/Cole.

- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson.
- Tezbaşaran, A. (1997). *Likert tipi ölçek geliştirme kılavuzu [Manual for Developing Likert-Type Scales]*. Türk Psikologlar Derneği Yayınları.
- Tezbaşaran, A. (2004). Likert tipi ölçeklere madde seçmede geleneksel madde analizi tekniklerinin karşılaştırılması [A comparison of conventional item analysis techniques to construct likert type scales]. *Türk Psikoloji Dergisi*, 19(54), 77–90.
- Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, 17(7), 783–805. [https://doi.org/10.1016/S0742-051X\(01\)00036-1](https://doi.org/10.1016/S0742-051X(01)00036-1)
- Tschannen-Moran, M., Woolfolk Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research*, 68, 202–248 p.783.
- White, D. J., & Carpenter, J. P. (2008). Integrating mathematics into the introductory biology laboratory course. *Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 18(1), 22–38. <https://doi.org/10.1080/10511970701753415>
- Wu, P. (2024). Practice and Reflection on the Interdisciplinary Teaching in Primary and Secondary Physical Education Programs. *International Journal of Social Sciences and Public Administration*, 3(2), 442–447. <https://doi.org/10.62051/ijsspa.v3n2.53>
- Yang, K.-L., Wu, H.-K., Wu, J.-Y., Lin, K.-Y., & Hsu, Y.-S. (2025). Identity and experience matter: Differences between secondary STEM teachers' self-efficacy and commitment in integrated STEM teaching activities. *Eurasia Journal of Mathematics, Science and Technology Education*, 21(4), em2609. <https://doi.org/10.29333/ejmste/16082>
- Yarımca, Ö. (2011). Disiplinler arası yaklaşıma dayalı bir durum çalışması [A case study in interdisciplinary approach]. *Akademik Bakış Dergisi*, 25, 2–22.
- Yıldırım, A. (1996). Disiplinlerarası öğretim kavramı ve programlar açısından doğurduğu sonuçlar [The concept of interdisciplinary instruction and its implications for curricula]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 12, 89–94.
- Yılmaz, M., Köseoğlu, P., Gerçek, C., & Soran, H. (2024). Yabancı dilde hazırlanan bir öğretmen öz-yeterlik ölçeğinin Türkçeye uyarlanması [Adaptation of a teacher self-efficacy scale to Turkish]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 27, 260–267.
- You, H. S. (2017). Why Teach Science with an Interdisciplinary Approach: History, Trends, and Conceptual Frameworks. *Journal of Education and Learning*, 6(4), 66–77. <https://doi.org/10.5539/JEL.V6N4P66>
- Yuan, C. (2023). From knowledge transmission to thinking teaching: The integrated design of interdisciplinary curriculum for empowering minds. *Journal of Education, Teaching and Social Studies*, 5(3), 1–11. <https://doi.org/10.22158/jetss.v5n3p1>
- Yurdugül, H. (2006). The comparison of reliability coefficients in parallel, tau-equivalent, and congeneric measurements. *Ankara University Journal of Faculty of Educational Sciences*, 39(1), 15–37.
- Yurdugül, H., & Sırakaya, D. A. (2013). Çevrimiçi öğrenme hazır bulunuşluluk ölçeği: Geçerlik ve güvenilirlik çalışması [The scale of online learning readiness: a study of validity and reliability]. *Education and Science*, 38(169), 391–406.
- Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, 25(1), 82–91. <https://doi.org/10.1006/ceps.1999.1016>

13	10	Disiplinler arası bir ders planlarken problem durumu, konu ya da tema ile ilgili akademik bilgileri (makale, bildiri vb.) araştırırım.	When planning an interdisciplinary lesson, I research academic resources (such as articles or conference papers) related to the topic, problem, or theme.
15	11	Disiplinler arası öğrenme deneyimi tasarımına uygun öğretim teknikleri belirleyemem.	I cannot determine appropriate teaching techniques for designing interdisciplinary learning experiences.
16	12	Öğrencilerin sevmediği dersi sevdikleri derslerle birleştirerek sevmelerini sağlayabilirim.	I can help students develop an interest in subjects they dislike by integrating them with subjects they enjoy.
18	13	Aynı konuda farklı disiplinlerin bakış açılarını anlamakta zorlanırım.	I have difficulty understanding the perspectives of different disciplines on the same topic.
19	14	Okul dışı öğrenme ortamlarını, disiplinler arası öğrenme deneyimi tasarımına uygun olarak kullanabilirim.	I can use out-of-school learning environments in ways that are suitable for designing interdisciplinary learning experiences.
20	15	Disiplinler arası hazırlanmış örnek ders tasarımlarını incelerim.	I review sample lesson designs that are developed from an interdisciplinary perspective.
24	16	Kendi disiplinim dışında başka bir disiplinin dersime katkısının olmadığını düşünürüm.	I believe that disciplines other than my own do not contribute to my teaching.
25	17	Kendi disiplinimin başka disiplinlerle olan ilişkisini içeren etkinlikleri derslerimde kullanmakta zorlanırım.	I find it difficult to use activities that involve relationships between my own discipline and other disciplines in my lessons.
32	18	Dersimi işlerken konuda farklı disiplinlerden örnekler vermekte zorlanırım.	I find it difficult to provide examples from different disciplines while teaching my subject.
33	19	Farklı disiplinlerin kendi alanımla olan ilişkisini araştırabilirim.	I can investigate how different disciplines relate to my own field.
38	20	Disiplinler arası yaklaşımda öğrenme deneyimi tasarım sürecinde öğrencilerin fikirlerine başvurabilirim.	In the process of designing interdisciplinary learning experiences, I can seek students' opinions and input.
39	21	Disiplinler arası yaklaşımda öğrenme deneyimi tasarımına uygun öğretim yöntemleri belirlemekte zorlanırım.	I struggle to determine teaching methods that are appropriate for designing interdisciplinary learning experiences.

*The items have been translated into English to enhance readability for international audiences. The original Turkish items were employed in the actual data collection and analysis processes.

Item Distributions:

Factor 1 Competences Item Numbers: 1, 2, 4, 5, 7, 9, 10, 12, 14, 15, 19, 20

Factor 2: Challenges Item Numbers: 3, 6, 8, 11, 13, 16, 17, 18, 21

Reverse Item Numbers: 3, 6, 8, 11, 13, 16, 17, 18, 21

The final version of the scale includes 9 reverse items.