

PSİKOLOJİK ÖLÇMELERİN KISA TARİHİ VE 21. YÜZYILDA SINIF İÇİ DURUM BELİRLEME ANLAYIŞI

A BRIEF HISTORY OF PSYCHOLOGICAL MEASUREMENTS AND AN APPROACH OF CLASSROOM ASSESSMENT IN THE 21ST CENTURY

Ömer KUTLU¹, Özge ALTINTAŞ²

Öz:Psikolojik özelliklerin ölçülmesinde 1800'lü yıllardan itibaren önemli değişimler yaşanmıştır. Bir yandan zekâ, tutum, algı, kişilik, yetenek, başarı gibi psikolojik yapıların ölçülmesinde, diğer yandan ölçme kuramlarında önemli gelişmeler olmuştur. Psikometri alanındaki bu ilerlemeler okul öğrenmelerinin ölçülmesine katkı sağlamış, ancak doğrudan okul öğrenmelerinin kendisiyle ilişkili olamamıştır. 1900'lü yıllardan itibaren güvenilirlik ve geçerlik üzerinde yapılan çalışmalar, sınıf içi başarının belirlenmesi süreçlerinde de etkili olmuştur. Öğretmen yapımı testlerin istenen güvenilirlik ve geçerlik değerlerine sahip olabilmeleri için çok sayıda maddeden oluşan testlerin kullanımı yaygınlaşmıştır. Bu durum öğretmenlerin, testlerinde konu boyutu açısından çoğunlukla öğrenme düzeyini gösteren kapsam temelli öğrenmelere; bilişsel boyutlar açısından ise çoğunlukla hatırlama, anlama ve uygulama düzeyindeki bilişsel süreçlere yönelmesine neden olmuştur. Bu değişimler olumlu yönde olsa da okulların yerine getirmesi gereken öğrenmelerin gerçek yaşam durumlarında kullanılması işlevinin gelişmesi önünde engel oluşturmuştur. Öğrenciler okullardan temel bilgileri öğrenen ancak bu bilgilerini doğal ve toplumsal yaşam durumlarında istenen düzeyde kullanamayan bireyler olarak mezun olmuşlardır. Eğitimde sözü edilen olumsuz koşulların değiştirilmesi isteği eğitim-öğretim süreçlerinde güncellemeleri kaçınılmaz kılmaktadır. Öğrencilerin başarısında etkili olduğu düşünülen ve 21. yüzyıl becerileri olarak adlandırılan bilişsel, içsel ve kişilerarası yeterliklerin de ölçme ve durum belirleme sürecine dahil olması bu açıdan önemli görülmektedir. Bu çalışma sınıf içi ölçmelerin dünü, bugünü ve geleceği ekseninde ele alınmış ve savunulan düşünce açıklayıcı bir ölçme ve durum belirleme örneğiyle desteklenmiştir.

Anahtar sözcükler: ölçme, değerlendirme, durum belirleme, 21. yüzyıl becerileri, bilişsel beceriler, içsel beceriler, kişilerarası beceriler, öğrenci başarısı, sınıf içi durum belirleme

Bu makaleye atıf vermek için:

Kutlu, Ö. ve Altıntaş, Ö. (2021). Psikolojik ölçmelerin kısa tarihi ve 21. yüzyılda sınıf içi durum belirleme anlayışı. *Trakya Eğitim Dergisi*, 11(3), 1599-1620.

Cite this article as:

Kutlu, O., & Altintas, O. (2021). A brief history of psychological measurements and an approach of classroom assessment in the 21st century. *Trakya Journal of Education*, 11(3), 1599-1620.

Abstract: The measurement of psychological constructs has changed significantly since the 1800s. On the one hand, psychological constructs such as intelligence, attitude, perception, personality, ability and success have been measured, and on the other hand, there have been important developments in measurement theories. These advances in psychometrics have contributed to the assessment of school learning, but have not been directly related to school learning itself. Since 1900, studies of reliability and validity have also been effectively used to determining classroom success. In order for teacher-made tests to achieve desirable reliability and validity values, the use of tests consisting of many items has become widespread. This has led teachers to turn towards mostly content-based learning that indicates learning levels by the content dimension of tests. It also oriented them towards cognitive processes at the remember, understand, and apply levels. Even if these changes were positive, they hindered the development of *the use of learning in real-life situations*, which should be a skill that schools teach their students. Graduates of schools learned basic information from them, but could not use their knowledge at the desired level in real-life situations. The desire to improve education inevitably requires it to be updated. In this regard, it is important that cognitive, intrapersonal, and interpersonal competencies that affect students' achievement and called as 21st century skills are also included in measurement and assessment processes. This study examines the past, present and future of classroom assessments, and its claims are supported by an explanatory example of measurement and assessment.

Keywords: measurement, evaluation, assessment, 21st c skills, cognitive skills, intrapersonal skills, interpersonal student achievement, classroom assessment

¹ Dr. Öğr. Üyesi, Ankara Üniversitesi, Eğitim Bilimleri Fakültesi, Eğitimde Ölçme ve Değerlendirme Anabilim Dalı, Ankara/Türkiye, e-posta: omerkutlu@ankara.edu.tr, ORCID: 0000-0003-4364-5629.

² Öğr. Gör. Dr., Ankara Üniversitesi, Eğitim Bilimleri Fakültesi, Eğitimde Ölçme ve Değerlendirme Anabilim Dalı, Ankara/Türkiye, e-posta: oaltintas@ankara.edu.tr, ORCID: 0000-0001-5779-855X.

EXTENDED ABSTRACT

Introduction

Throughout human history, people's reactions to stimulants based on their psychological constructs have been a matter of curiosity. This curiosity goes back to Plato, Aristotle and other Greek philosophers. They thought about many of the issues of human nature that concern psychologists today, including *memory, learning, motivation, perception, dreams and irrational behavior*. In the history of psychological measurements, especially since the 1800s, the measurement and evaluation of human behavior has changed greatly. On the one hand, there have been advances in the measurement of psychological constructs such as intelligence, attitude, perception, personality, ability and success, and on the other hand, measurement theories such as Classical Test Theory and Item Response Theory have influenced the evolution and development of the field.

Although advances in the field of psychometrics have made some contributions to the measurement of school learning, they are not directly related to school learning. Most approaches to the assessment of school learning were based on traditions from the past. Until the 1850s, this meant either oral and written examinations. After the 1850s, the use of written examinations became widespread. An understanding of assessment and evaluation as written examinations that assessed the details of acquired knowledge was dominant in schools because students were expected to memorize large amounts of information. Examinations required large amounts of information as a response to a small set of question that were intended to evaluate whether the students had learned the relevant information. Examination questions were designed to focus on the subjects students were expected to learn. Developments in the field of experimental psychology since the 1850s led to the understanding that a small number of measurements are not sufficient to measure the psychological construct of interest, and thus measurements of a large number of individuals were needed. Studies of reliability in mid-twentieth century found that when the number of items measuring the same construct and the number of individuals increased, the random errors decreased and reliability increased. This led to the widespread use of examinations consisting of many items in schools.

Constructs that represent knowledge schools intended to impart were assessed with different kinds of item such as short answer questions, fill-in-the-blank questions, matching, true-false questions and multiple-choice questions. It was deemed important to use large numbers of items representing the content to ensure that tests had content validity. Efforts to use too many items also affected cognitive levels, another dimension of content validity. Responding to test items that assess cognitive levels such as *applying, analyzing and evaluating*, the upper levels of cognition, takes too much time, thus reducing the numbers of items, so test developers had to write items at the lower cognitive levels of remembering and understanding. Questions such as: *How reliable is the reliability value? What is the deviation of the scores from the measurements? and What is the score range?* prompted concerns about the validity of reliable tests. Although the definition of validity has not changed over the years, discussions have led to a change in perspective and questions about whether tests are able to measure what they are intended to measure and whether tests adequately cover the content they are intended to evaluate. Studies of reliability and validity have provided important information to test developers about how to reduce errors in test scores. These two approaches also led to an increase in the number of items on tests based on the idea that more items would reduce errors in test scores.

In order to attain desirable reliability and validity values, teachers were directed to use tests consisting of many items. This led them to turn towards content-based learning, which assesses learning levels in the content dimension, and to avoid cognitive processes at the knowledge, comprehension and application levels of Bloom's taxonomy. So, they frequently used short-answer, fill-in-the-blank, matching, true-false and multiple-choice questions to monitor their students' success (formative) and determine their students' success levels (summative). Teachers tended to determine whether their students learned the topics covered in their classes and to structure classroom assessments to eliminate their *learning deficiencies and difficulties*. Even if those changes were positive, they hindered the development of the function of *using learning in real-life situations*. This negligence also failed to develop the basic life skills that enable students to use their learning in real-life situations. For these reasons, many societies have started to question the measurement and assessment approaches of schools, especially since the 1980s. These discussions recognized that schools were not adequately teaching the cognitive, affective

and social skills that prepare students for life. Today societies accept that change in education is inevitable and think that the systems currently in use may not be sufficient. In this study, brief information about the history of psychological measurements has been given and determinations have been made about how the understanding of in-classroom assessment will be in the twenty-first century. Besides, an exemplary assessment model is presented for the twenty-first century measurement approach.

Method

It is clear that the approach of determining student success with many questions, the starting point of this critical study, is inadequate for developing the skills students need, especially in the twenty-first century. This study found that the understanding of classroom assessment, which will inevitably change in the twenty-first century, should be based on real-life situations and feedback. Therefore, it proposes a skill-based classroom assessment model for the twenty-first century with examples of measurement and assessment based on open-ended items that allow students to gather information and skills, and use them to design their responses.

The model proposed in this study describes how a course assesses cognitive, intrapersonal and interpersonal skills. The *learning outcome* is defined first. Next, a case study (situation) for the learning outcome is given and associated with each of the three skill sets. In this study, learning outcome refers to the association of cognitive, intrapersonal and interpersonal skills with the course content. The proposed measurement and assessment approach ensures internal consistency (reliability) by assessing more than one construct related to the same content, and it ensures content validity by using a good example of the content to be assessed.

Findings

The desire to improve education inevitably requires it to be updated. The findings of international student achievement studies such as PIRLS, PISA and TIMSS, which have been implemented since the 2000s, provide important data to decision makers in education and indicate precautions they need to take. Findings in biology, medicine and psychology have enabled us to understand the nature of human behavior in more detail, and they indicate that school learning should be reorganized rapidly. The twenty-first century needs to educate people to live in peace with nature and society. This century should reveal new knowledge by making use of old knowledge produced for school learning. It is important that cognitive, intrapersonal and interpersonal competencies, which affect students' success and are referred to as twenty-first century skills, should be included in measurement and evaluation.

The concept of learning outcome defined in this study includes much more than "the expression" as it is currently used in today's curricula. The expressions often refer to the relationship of content to skills. Learning outcomes can be both the product of using cognitive, intrapersonal and interpersonal skills in the context of a task, and the use of cognitive, intrapersonal and interpersonal skills by associating them with a situation, as in this study.

Discussion and Conclusion

It is extremely important for schools to develop students' cognitive, intrapersonal and interpersonal skills. In the twenty-first century, schools are expected to develop these competencies using measurement and assessment practices in the classroom. Rather than exam papers or homework-based repeating what they learn in class, it should be ensured that students use their competencies in real-life situations. This calls for measurement and assessment approaches that prioritize feedback and require the use of more than one skill at a time.

Education is seen as the most important tool for solving many environmental and social problems, and improving educational practices has often been discussed, especially in the last 30 years. The OECD's *Future of Education and Skills 2030* project and report was intended to help answer two broad questions: *What knowledge, skills, attitudes and values will today's students need to thrive in and shape their world?* and *How can instructional systems develop these knowledge, skills, attitudes and values effectively?* The research concerning these questions makes the importance of the understanding addressed in this study stand out. This study is unique because it examines the history of assessment and evaluation and because of its new approach to measurement and assessment.

GİRİŞ

Her insani etkinlik aynı zamanda toplumsal ilişkilerin sonucudur. İnsanlık tarihi boyunca genç nüfusun eğitilmesi, toplumsal yaşamda gerekli olan görevlerin ve rollerin kazandırılması toplumların

öncelikleri arasında yer almıştır. Bu kazanımların biçimini, toplumun sosyoekonomik, siyasi ve kültürel yapısı yanında, doğadaki değişimler de etkilemiştir (Ercan, 2009; Harari, 2015, 2016). Eğitim toplumların varlıklarını sürdürebilmeleri için vazgeçilmezdir. Eğitimin başarısı ve bu başarının izlenmesi, eğitimin değerlendirilmesini de beraberinde getirmiştir. Planlı bir eğitim ve güçlü bir değerlendirme toplumların geleceğini de güvence altına almıştır. Bu çalışmada, ölçmelerin tarihi hakkında kısa bilgiler verilmiş 21. yüzyılda sınıf içi durum belirleme anlayışının nasıl olacağına ilişkin belirlemeler yapılmıştır. Ayrıca 21. yüzyıl ölçme anlayışı için örnek bir durum belirleme modeli de sunulmuştur.

İnsanların gerek doğada gerekse toplumda meydana gelen uyarıcılar karşısındaki psikolojik özelliklerine dayalı tepkileri, insanlık tarihi boyunca merak konusu olmuştur. Bilimsel anlayışın gelişmediği dönemlerde bu tepkiler çoğunlukla doğa üstü güçlerle açıklanmaya çalışılmıştır (Harrari, 2015). Ancak Sümer, Mezopotamya, Eski Yunan, Eski Çin, Eski Mısır ve Maya gibi uygarlıklar sayıları; astronomi, fizik, matematik, zaman ölçümü, tarım, hayvancılık gibi konularda, toplumsal yaşamın sürekliliğini sağlamak için kullanmışlardır. Bu çalışmalar aklın ve bilimin kullanıldığı en eski çalışmalar olarak bilinmektedir (Ronan, 2003).

Özellikle 1500'lü yıllardan itibaren Kopernik (1473-1543), Leonardo (1452-1519), Bruno (1548-1600), Galileo (1564-1642), Kepler (1571-1630) gibi düşünürlerin dünyanın evrendeki konumuyla, doğanın işleyiş kurallarıyla ve insan bedeniyle ilgili yaklaşımları ve çalışmaları insana ve insan davranışlarına bakış biçiminin değişmesinde büyük rol oynamıştır. Wootton (2019), Kopernik, Galileo, Newton, Darwin ve Heisenberg gibi bilimsel dönüşümlerde rol oynayan bilim insanlarının yaptıkları konusunda sağlam bir kavrayışa sahip olduklarını ve savundukları düşüncelerin benimsenmesi halinde ciddi sonuçlar doğacağına farkında olduklarını belirtmektedir. Nitekim gerçekleşen Rönesans ve Reform hareketleri hem fizik kurallarıyla hem de insanla ilgili bilgilerin gözlemlere ve kanıta dayalı olmasının önemini ortaya koymuştur (Freely, 2014).

İnsanın psikolojik özellikleriyle ilgili merak Platon, Aristo ve diğer Yunan düşünürlerine kadar uzanmaktadır. Bu düşünürler *bellek, öğrenme, güdü, algı, rüyalar ve akıl dışı davranışlar* gibi insan doğasıyla ilgili günümüzdeki psikologların da ilgilendiği pek çok konu hakkında düşünmüşlerdir. 19. yüzyılın son çeyreğine kadar filozoflar insan doğasını birtakım kurgulara, sezgilere ve kendi sınırlı kişisel deneyimlerine dayalı genellemeler yoluyla incelemiştir. Daha sonra biyoloji ve fizik alanlarında, başarıları önceden kanıtlanmış bilimsel araç ve yöntemlerin insan doğasına ilişkin sorunlara uyarlanmasıyla büyük bir değişiklik meydana gelmiştir. Araştırmacılar insan zihnini araştırırken kontrollü gözlem ve deneyi güvenilir bir yol olarak benimsemiştir. Bu sayede psikoloji, felsefi köklerinden ayrılarak ayrı bir bilim kimliğini oluşturmaya başlamıştır (Schultz ve Schultz, 2011).

Psikolojik ölçmelerin özellikle 1800'lü yıllardan itibaren tarihi incelenecek olursa, son 200 yıl içinde insan davranışlarının ölçülmesi ve değerlendirilmesi süreçlerinde önemli değişimler yaşanmıştır. Bir yandan zekâ, tutum, algı, kişilik, yetenek, başarı gibi psikolojik yapıların ölçülmesi konularındaki ilerlemeler (Anastasi ve Urbina, 1997; Cronbach, 1990; Öner, 1997; Schultz ve Schultz, 2011) bir yandan da Klasik Test Kuramı (KTK) ve Madde Tepki Kuramı (MTK) gibi ölçme kuramlarındaki zenginleşmeler (Ackerman, Gierl ve Walker, 2003; Crocker ve Algina, 1986; Gulliksen, 1950; Hambleton ve Swaminathan, 1985; Kline, 1986; Lord ve Novick, 1968; Reckase, 2009) alanın değişmesinde ve gelişmesinde etkili olmuştur.

Arkeoloji biliminin bulguları, insanlığın yazılı tarihinin MÖ 8000'li yıllara kadar gittiğini göstermektedir. Yazıyı ilk olarak Sümerlerin MÖ 3000-3500 yılları arasında keşfettiği bilinmekte, tarihte okula benzer ilk yapıların Sümerler tarafından kurulduğu belirtilmektedir. Sümer kayıtlarına göre günlük yaşamın önemli ayrıntılarından biri eğitim ve okul faaliyetleridir (Gündüz, 2014). Günümüz okullarına benzeyen yapısıyla, Sümer okullarında tanrıbilim, bitkibilim, hayvanbilim, madenbilimi, coğrafya, matematik, dilbilim gibi alanlarda eğitim alan öğrencilerden, bu bilimlere katkıda bulunan bilginler yetişmiştir (Kramer, 1999, 2002; Çığ, 2002). MÖ 800'lü yıllardan itibaren Çin (matematik, konuşma, müzik, süsleme sanatları, dini ve felsefi öğretiler eğitimi), Mısır (hukuk, resim, musiki, tarih, coğrafya, tıp matematik ve mühendislik bilimleri eğitimi), Yunan (okuma, yazma, hesap, resim, müzik ve spor, edebiyat, askerlik hukuk eğitimi), Roma (hukuk, felsefe, tıp, dil bilgisi, retorik ve diyalektik, aritmetik, geometri, kozmografya eğitimi) toplumlarında okullar bireylerin yetiştirilmesinde önemli görevler üstlenmiştir (Köçer ve Koçoğlu, 2020). Orta Çağ, MS 3. yüzyıldan itibaren ortaya çıkmıştır. Eğitim çoğunlukla kilise ya da kilise mensupları tarafından denetlenmiştir. Özellikle, 800-1500 yılları arasında kilise öğrencilerin tutumları, davranışları ve inançları üzerinde oldukça büyük etkiye sahipti; dilbilgisi, astronomi, mantık, matematik, felsefe ve din, eğitim sırasında dikkate alınan temel alanlardı. Kilise kendi gündemini öğrencilere dayatmakta; iyi bir Hıristiyan yetiştirmek ve kilisenin geleceğini güvence altına almak istemektedir (Güven, 2019).

1800'lü yılların ortalarına kadar okullardaki ölçme ve değerlendirme uygulamaları yazılı ve sözlü sınavlar şeklindeydi. Bazı kaynaklar MÖ 2000'li yıllarda Çin'de askeri sınavların kullanıldığını belirtmektedir (Hogan, 2013). Bununla birlikte, 3000 yıllık bir süre boyunca Çin'de yeterli testlerinin kullanıldığı bilinmekteydi (Cohen ve Swerdlik, 2009). William Edward Burghardt Du Bois (1868-1963), yazılı sınavların ilk kez 1500'lü yılların sonlarında Cizvit okullarında ortaya çıktığının altını çizmiştir. Cizvit liselerinde sınavlar komisyon sınavı şeklinde yıl sonunda düzenlenmekteydi. Başarılı olanlar bir üst sınıfa geçerken başarısız olanlar sınıf tekrarı yaparlardı (İstek, 2019). Cizvit okullarının öğretim programlarının kılavuzlarında yazılı sınavların uygulanması konusunda katı kurallar söz konusudur.

1800'lü yıllara doğru bilimsel ilgi ve toplumun bilinci zihinsel hastalıklara yönelmiştir. Fransa'dan Philippe Pinel (1745-1826), İngiltere'den Samuel Tuke (1784-1857), Amerika'dan Benjamin Rush (1746-1783) tüm çabalarını zihinsel hastalıkların tanınması ve tedavisi üzerine yöneltmişlerdir. Örneğin, bilimsel açıdan zihinsel geriliği de içeren akıl hastalıklarının teşhisi için yöntemler ortaya çıkmaya başlamıştır. "Séguin Form Board Test (1866)" ilk kez bu dönemde ortaya çıkmıştır (Hogan, 2013). Édouard Séguin (1812-1880), zihinsel geriliği olan bireylerin çaresiz olduğu anlayışını reddetmiş; bu kişilerin eğitilebileceğini savunmuştur. Benzer şekilde, Jean-Étienne Dominique Esquirol (1772-1840) ve Alfred Binet (1857-1911), normal okul koşullarında eğitim alamayacak durumda olan zihinsel engelli çocukların okuldan tamamen dışlanması değil, bu çocukların eğer eğitilebilecek durumda iseler özel sınıflara atanmasını savunmuşlardır (Anastasi, 1976). Yine bu dönemde etkin bir biçimde zihin, zihinsel yetenekler ve zekânın kökenleri üzerine araştırmalar yapan Sir Francis Galton (1822-1911), yeteneğin nesiller boyunca nasıl aktarıldığını "Hereditary Genius (1869)" adlı kitabıyla ortaya koymuştur. Tepki süresi, görme keskinliği ve fiziksel güç gibi insanın duyuşal ve motor işlevinde bireysel farklılıkların var olduğunu göstermeye yoğunlaştığı çalışmalarla Galton, günümüzde bilimsel psikolojinin en önemli alanlarından biri olan, insanın bireysel farklılıklarıyla ilgili bir bilgi arayışını da başlatmıştır (Kaplan ve Sacuzzo, 2008).

Bu yıllarda ölçümlerin ya normları yoktu ya da güvenilir ölçümler bulunmamaktaydı; ancak yine de bu çabalar, çalışmalar sonrasında geliştirilecek olan ölçmelerin öncüsü niteliğinde olmuşlardır. Bu dönemin anlamlı gelişmelerinden biri, Boston Okul Komitesi tarafından geliştirilen yazılı sınavların resmi uyarlamasıdır. Bu uyarlama çalışması okul müdürü Horace Mann (1796-1859) gözetimindeydi. Mann, yazılı sınavların tüm öğrencileri bir düzene soktuğunu, daha geniş bir içerik kapsamına izin verdiğini, söz konusu şans ögesini azalttığını ve sınav görevlileri tarafından kayırmacılık olasılığını ortadan kaldırdığını düşünüyordu. Aynı zamanda tüm okulların öğrencilerini değerlendirmeleri gerektiğinin de şiddetli bir savunucusuydu (Domino ve Domino, 2006).

Psikolojik ölçmeler açısından bu dönemdeki önemli gelişmelerden biri de Wilhelm Wundt (1832-1920) tarafından ilk deneysel psikoloji laboratuvarının kurulması olmuştur (Bruno, 1996). Deneysel psikoloji ilk zamanlarda fizyoloji ve sonrasında psikofizikle eşdeğerdi. Bir takım felsefi görüşlere deneysel bir temel bulmak ve bu görüşlerin geçerliğini kanıtlamak amacıyla 1879 yılında Leipzig Üniversitesi'nde kurulan bu laboratuvar ile Wundt test dünyasına, ölçmelerin keskinliği ve test koşullarının standardizasyonu konusunda katkıda bulunmuştur (Hogan, 2013; Schunk, 2012).

Deneysel çalışmalar, insan davranışlarının anlaşılmasına ve bu davranışların altında yatan nedenlerin açıklanmasına katkı getiren önemli bulgulara imza atmış olsa da zekâ, ilgi, tutum gibi örtük psikolojik yapıların anlaşılmasına ve açıklanmasına yönelik fazla katkı sağlayamamıştır. Psikolojik yapılar üzerindeki kavramsal ve kuramsal çalışmaların derinleşmesi, bu yapıların ölçülebilir ve gözlenebilir biçimde işevuruk tanımlarının yapılması psikolojik ölçme araçlarının geliştirilmesi konusundaki mantığın ortaya çıkmasına yol açmıştır. Bu konudaki ilk ışık Galton tarafından yakılmıştır (Janda, 1992). Daha sonra Wundt'un da doktora öğrencisi olan James McKeen Cattell (1860-1944) 1890 yılında yazdığı "Mental Tests and Measurements" adlı eserinde psikolojinin bir deney ve ölçü temeline dayanmadıkça, fizik bilimlerinin kesinliğine ulaşamayacağını belirterek ancak çok sayıda bireye bir dizi zihinsel test ve ölçüm uygulanarak bu yönde bir adım atılabileceğini vurgulamıştır (Cattell, 1890). Bu tarihten önce (1883-1886 yılları arası) Cattell, insanın tepki süreleri ve bireysel farklılıkları üzerine dokuz makale yayınlamıştır (Thorne ve Henley, 2001). *Zihinsel Test* terimini ilk kez kullanan Cattell'in doktora tezi, Galton'un reaksiyon süresindeki bireysel farklılıklar üzerine yaptığı çalışmaya dayanmıştır. Böylece Cattell, modern testlerin geliştirilmesine yol açan güçleri sürdürmüş ve canlandırmıştır (Kaplan ve Sacuzzo 2008). Cattell zihinsel işleyişin anahtarının temel süreçler olduğuna inanıyordu. Duyusal keskinlik, reaksiyon süresi, bir çizginin görsel ikiye bölünmesi ve kısa zaman aralıkları gibi alanları kapsayan, 10'u dayanak noktası olarak kabul eden, 50 testten oluşan bir takım oluşturdu. Bunları üniversite öğrencilerine bugünün SAT'ın ve ACT'nin kavramsal kökenleri olarak akademik başarıyı yordamayı umarak uygulamıştır (Hogan, 2013).

1800'lü yılların sonlarına doğru, bireysel farklılıkların nedenleri ve bireylerin sahip oldukları psikolojik özellikler arasında ilişkiler olduğu düşüncesi, bir yandan da istatistikle ilgili çalışmaların hız kazanmasına neden olmuştur. Korelasyon katsayısı ve güvenilirlik kavramı henüz bulunmamıştı. Ancak dönemin sonunda ortaya çıkan bu kavramlar zihinsel testlere damgasını vurmuştur (Budak, 2008). Galton, bir ölçümler dizisine ait dağılımın ortalama değeri etrafındaki değişimin, iki rakamla anlamlı bir şekilde tanımlanıp özetlenebileceğini ileri sürmüştür. Böylelikle insanın psikolojik özellikleriyle ilgili bir dizi ölçüm, bu iki rakama indirgenebilmiştir. Galton'un istatistikle ilgili bu çalışmaları bilimin en önemli ölçümlerinden biri olan korelasyonu ortaya koymuştur. Korelasyonla ilgili ilk bilgi 1888'de ortaya çıkmıştır (Schultz ve Schultz, 2011).

Kökleri Galton'un çalışmalarına kadar uzanan korelasyonla ilgili çalışmalar, Pearson Momentler Çarpım Korelasyon Katsayısı adı altında Karl Pearson (1857-1936) tarafından matematiksel olarak formüle edilmiştir. Galton'un çalışmaları daha sonra Karl Pearson'un korelasyon üzerine yaptığı çalışmalarla birleştirildiğinde, bireye ilişkin özellikler ve kalıtımla ilgili ilk bilimsel çalışmaların temellerinin atılmasına öncülük etmiştir. Bunun yanında hem Galton'un hem de Wundt'un çalışmalarından etkilenerek modern anlamda zekânın ölçülmesi için kuramsal temelleri ortaya atan ilk bilim insanı Charles Spearman'dır (Bartholomew, 2004). Bu döneme damgasını vuran Spearman'ın zekânın ölçülmesine katkısı diğer bilim insanlarından farklı olmuştur. 1904 yılında yazdığı "General Intelligence, Objectively Determined and Measured" adlı makalede *g faktörü* adını verdiği *Genel Zekâ* kavramını açıklayarak zekâyâ ilişkin iki faktör kuramını ortaya koymuştur. Bunun yanı sıra faktör analizi tekniğini geliştirmiş, insan zekâsının deneysel anlamda incelenmesine ilişkin girişimlerde bulunmuştur. Önceki kuramlar insan zekâsının incelenmesini felsefî temellere dayandırırken, Spearman bunu test sonuçlarına, yani bir anlamda matematiksel bir temele dayandırmıştır. Spearman, Pearson korelasyonunun non-parametrik karşılığı olan Sıra Farkları Korelasyonu ile günümüzde kullanılan Faktör Analizi'nin ilk versiyonunu da geliştirmiştir (Hogan, 2013).

Spearman'ın alana bir diğer ve önemli katkısı, testten elde edilen puanları *gözlenen puan* olarak düşünmesidir. Buna göre Spearman her gözlenen puanın içinde bir miktar gerçek bir miktar da hata puanı bulunduğunu vurgulayan eşitliği ($X_{gözlenen} = T_{gerçek} + E_{hata}$) temellendirmiştir (Spearman, 1904). KTK'nın güvenilirlik kavramına bakış açısı, Spearman'ın bu eşitliğine dayandırılmıştır. Spearman, test puanlarına karışan hatanın azaltıldığı ölçüde gözlenen puanların gerçek puana yaklaşacağını, bu eşitlikle ortaya koymuştur. Bu bilgi KTK'nın yapı taşı olmuştur (Crocker ve Algina, 1986).

Psikolojik özelliklerin ölçülmesi sürecinde, bireylerin testlerden elde ettikleri puanların birbiriyle olan tutarlılığı güvenilirliğin bir ölçüsü olarak düşünülmüş; güvenilirlik, KTK'daki gelişmelerle önce paralel testler arası korelasyon olarak tanımlanmıştır. Bu düşünceden hareketle güvenilirlik, aynı özellikleri ölçen eş değer (paralel) iki testi alan bireylerin puanları arasındaki tutarlılık (eş değer formlar güvenilirliği); aynı testi farklı zamanlarda alan aynı bireylerin puanları arasındaki tutarlılık (test-tekrar test güvenilirliği); bir testin iki eş değer yarıya bölünmesiyle oluşan aynı bireylere ait puanlar arasındaki tutarlılık (iki yarı güvenilirliği) ve bir testte yer alan maddelerin kovaryanslarına bağlı olarak elde edilen iç tutarlılık (Cronbach Alpha ve KR-20 güvenilirliği) olarak tanımlanmıştır (Gulliksen, 1967; Lord ve Novick, 1968; Nunnally, 1967; Thorndike, 1982).

Psikometri bilim alanında yapılan bu çalışmalar daha çok, insanların psikolojik özelliklerinin hatadan arınık biçimde nasıl ölçülebileceğiyle ilgiliydi. Bu çalışmalar okul öğrenmelerinin ölçülmesine bazı katkılar sağlasa da doğrudan okul öğrenmeleriyle ilişkili değildi. Okul öğrenmelerini ölçme yaklaşımları daha çok geçmişten gelen anlayışlara bağlı olarak yürütülüyordu. Oysa zekâ, tutum, kişilik gibi bireye ait özellikler henüz psikoloji bilimi içinde yeni ele alınıyor ve tanımlanmaya çalışılıyordu. Bu nedenle korelasyon ve testlerin güvenilirliği konusundaki çalışmalar bireyler arasındaki farklılıkların belirlenmesi amacıyla yapıyordu.

Eğitsel Ölçmelere Tarihsel Bir Bakış

Okul öğrenmelerinin ölçülmesinde 1850'li yıllara kadar yaygın olan anlayış, öğrenmelerin sözlü ve yazılı sınavlarla belirlenmesiydi. 1850'li yıllardan itibaren yazılı sınavların kullanımı yaygınlaşmıştır. O yıllarda okullarda daha çok yazdırmaya dayalı sorulardan oluşan ve öğrenmelerin ayrıntılarını ölçen yazılı sınavların kullanıldığı bir ölçme ve değerlendirme anlayışı egemendi. Çünkü öğrencilerden büyük miktarda bilgiyi ezberlemeleri bekleniyordu. Örneğin, belli bir ülkenin nehirleri, dağları, yöneticilerin isimleri, tarihin önemli olayları gibi birçok bilgiyi öğrencilerin öğrenmeleri gerekiyordu. Bu bilgilere sahip olan bireyin eğitilmiş bir birey olduğu düşünülüyordu (Raban, 2008). Yoğun bilgi içeren az sorulu ve yazdırma gerektiren bu sınavlar daha çok şu sorular etrafında değerlendirme yapıyordu: *Öğrenciler ilgili bilgileri öğrendi mi? Öğrencilerin belirli bir zamanda edindikleri bilgiler nelerdir?* Sınavlarda yer

alan soruların hazırlanmasında ise, öğrenciler için önemli olan ve öğrenilmesi gereken konular ön planda tutuluyordu.

Sözlü sınavlar öğrencinin derse her zaman hazır gelmesi gerektiği düşüncesinden yola çıkarak öğretmenin ön gördüğü bir zamanda ve onun karar vericiliğinde uygulanabilmekteydi. Yazılı sınavlar belirlenmiş tarihlerde ve öğretmen tarafından önceden hazırlanarak yapılıyordu. Sınavlar daha genel, ders alanı açısından önemli olduğu düşünülen konulara bağlı, çoğunlukla kısa ve uzun yanıtli sorulardan oluşuyordu. Bu durum 1900'li yılların başına kadar böyle devam etmiştir.

1850'li yıllardan itibaren deneysel psikoloji alanındaki gelişmeler ilgili psikolojik özelliğin ölçülmesi için az sayıda gözlemin yeterli olmadığına anlaşılmasına ve çok sayıda birey üzerinde gözlem yapma düşüncesine neden olmuştur (Cronbach, 1990). Bu süreçte bir psikolojik özellik hakkında karar verebilmek için yeterli sayıda gözlemin gerekli olduğu ve ilgili psikolojik özelliğin başka psikolojik özelliklerle olan ilişkisinin de ortaya konulması gerektiği düşüncesi doğmuştur (Anastasi ve Urbina, 1997). Bu durum hem çok sayıda birey üzerindeki gözlem ile tekrarlı ölçümlerin nasıl yapılacağı üzerindeki çalışmaların hem de psikolojik özellikler arasındaki ilişkilerin nasıl ortaya konulacağıyla ilgili çalışmaların artmasına yol açmıştır.

Karl Pearson'ın 1904 yılında, iki değişken arasındaki ilişkinin miktarını matematiksel olarak formüle etmesinden sonra, psikolojik özellikler üzerinde çalışma yapan bilim insanları bu bilgileri psikometri alanına nasıl transfer edecekleri konusunda düşünmeye başlamıştır. Korelasyon bilgisinden hareketle öncelikle güvenilirlik kavramına ulaşılmıştır. Brennan (2001) çalışmasında, Brown'un (1910) güvenilirlik katsayısını *aynı testin tekrarlı ölçümlerinden elde edilen korelasyon* olarak kullandığını söylemektedir. Brennan aynı yayında Traub (1997)'den alıntı yaparak, Kelley'nin (1923) güvenilirliği, *karşılaştırılabilir testler arasındaki korelasyon* katsayısı olarak tanımladığını belirtmektedir. Brown'un 1910 yılındaki güvenilirlik anlayışı, test tekrar test uygulamalarına işaret ederken Kelley'nin 1923 anlayışı ise, paralel testler uygulamasına işaret etmektedir. Guilford (1954) paralel testler kavramının Gullikson; eş değer formlar kavramının da Thorndike tarafından kullanıldığını belirtmektedir. Guilford güvenilirliğin kestirimine ilişkin standart yaklaşımları, üç başlık altında ifade etmektedir. Bunlar iki yarıya bölme, alternatif (paralel) formlar ve test tekrar test uygulamalarıdır. Bunların ortak amacı aynı testten türetilen ve aynı örneklemden elde edilen puanların korelasyonunu bulmaktır. Bu güvenilirlik hesaplama yolları, günümüzde birbirinden daha keskin biçimde ayrılmış olmasına rağmen aynı biçimde kullanılmaktadır. Bu süreçte, testlerin güvenilirliğine ilişkin ilk çalışmaların yoğunluklu biçimde 1960'lı yıllara kadar sürdüğü görülmektedir.

Sınıf İçi Ölçmelerde Güvenirlik ve Geçerlik Anlayışı

1900'li yılların ortalarında güvenilirlik üzerinde yapılan çalışmalar, aynı özelliği ölçen madde sayısı ve birey sayısı arttığında güvenilirliği etkileyen rastlantısal hatanın azaldığını ve güvenilirliğin arttığını ortaya koymuştur. Bu durum çok sayıda maddeden oluşan test ya da sınav uygulamalarının okul ortamlarında yaygınlaşmasına yol açmıştır. Öğrenci başarısının değerlendirilmesinde derslerin kazandırmayı hedeflediği davranışları temsil eden özellikler farklı madde türleriyle (kısa yanıtli, boşluk doldurmalı, eşleştirme gerektiren, doğru-yanlış, çoktan seçmeli gibi) ölçülmeye başlanmıştır. Dolayısıyla, testlerin kapsam geçerliğini sağlamada konu boyutunu temsil eden fazla sayıda maddenin kullanılması önemli görülmüştür. Fazla sayıda madde kullanma çabası aynı zamanda kapsam geçerliğinin diğer boyutu olan bilişsel düzeyleri de etkilemiştir. Bloom'un bilişsel sınıflamasının üst basamaklarını oluşturan *uygulama, analiz ve değerlendirme* gibi düşünme süreçlerine verilen yanıtların zaman alması ve madde sayısını azaltması test geliştiricileri çoğunlukla bilgi ve kavrama gibi bilişsel düzeylerde madde yazmaya zorlamıştır. Aynı zamanda; *Bulunan güvenilirlik değeri ne kadar güvenilir? Ölçmelerden elde edilen puanların sapması ne kadar? Puanlar hangi aralıkta yer alıyor?* gibi sorular güvenilir testlerin geçerliği konusunu gündeme getirmiştir (Cronbach ve Meehl, 1955; Jenkins, 1946).

Geçerlik, bir testle ölçülmesi istenilen özelliklerin başka özelliklere karıştırılmadan ölçülebilme derecesi olarak ele alınmaktadır (Anastasi, 1986; Cronbach ve Meehl, 1955; Ebel, 1956; Messick, 1989). Geçerlik tanımı yıllar içinde değişiklik göstermese de geçerlik kavramının kapsamı ve bu kavram üzerine yapılan tartışmalar geçerliğe bakış açısının da değişmesine yol açmıştır. Bu durum, testin ölçülmek istenen yapıyı ölçüp ölçemediğinden ya da testin bu yapıyı kapsayıp kapsamadığından daha farklı tartışmaları da gündeme getirmiştir. Bu tartışmalar çoğunlukla, *test puanları arasındaki ilişkilerin kuramsal boyuttaki ilişkilerle tutarlı olup olmadığı; test puanları kullanımının ve bu puanlara dayalı olarak verilen kararların etik, sosyal ve görgül açıdan doğrulanabilir olup olmadığı* konularında olmuştur (Borsboom, Mellenbergh ve van Heerden, 2004). Bunların yanında, *test puanlarının yordayıcı gücü, maddelerin yanlılık içerip içermediği, test puanlarının bireysel ya da kültürel etkenlerden etkilenip*

etiklenmediği gibi durumlar da geçerlik çalışmalarında ele alınmıştır. Güvenirlik ve geçerlik çalışmaları test puanlarına karışan hataları azaltma yolları hakkında test geliştiricilere önemli bilgiler sunmuştur. Aşağıda da ele alındığı gibi bu iki yaklaşım test puanlarına karışan hataları azaltacağı düşüncesiyle testlerde yer alacak madde sayısındaki artışı da ön plana çıkarmıştır.

Ölçme yapmak, gerçek miktarı bilinmeyen bir özelliğin miktarını, onun bir derece hatalı olabilecek ölçülerinden hareketle belirlemeye çalışmaktır (Özçelik, 1989). Buna göre, sınıf içi ölçmelerde kullanılacak bir ölçme aracının doğru ölçme yapabilmesi için gerekli olan koşullar KTK'nın güvenirlilik anlayışına dayandırılmıştır. Bu anlayışa göre, eğitsel başarı testlerinde öncelikle güvenirliliğin daha sonra da geçerliliğin sağlanması önemlidir. Okul başarı testleri geliştirilirken öğretmenlerin çoğunlukla akılcı ve mantıklı yolları kullandıkları bilinmektedir. Bir testin güvenirliliği, ölçme aracının kararlı ve tutarlı ölçme yapabilme gücüyle tanımlanmıştır. Ölçme aracının kararlı ve tutarlı ölçme yapabilme gücünün belirlenmesi; aynı araçla, aynı bireyler üzerinde, aynı ortamsal koşullarda davranış gözlemi gerektirmektedir. Ancak bu ideal durumun insan davranışlarının gözlemlenmesinde gerçekleşmesi olanaksızdır. Bu nedenle, eğitsel başarı testlerinde güvenirlilik hem aynı özelliği (davranışı, kazanımı) ölçen çok sayıda maddenin kullanılmasını hem de kısa yanıtı, boşluk tamamlamalı, eşleştirmeli gibi farklı türde maddelerin kullanılmasını gerektirmektedir (Thorndike ve Thorndike-Christ, 2014). Ebel ve Frisbie (1991), güvenirliliği etkileyen testle ilgili faktörler arasında *test uzunluğu* (madde sayısı, arzu edilen güvenirlilik değeri için gerekli madde sayısı) ve *test içeriği* (maddelerin aynı özelliği ölçme gücü - homojenlik/benzeşiklik) olduğunu belirtmektedirler.

Öğretmenler sınıf içi başarının belirlenmesi için, aynı kapsamı ölçen madde sayısını arttırmaya ve şansla elde edilen puanı azaltarak güvenirliliği büyütme odaklanmışlardır. Özçelik (1992-2), eğitimle kazandırılmak istenen yeni davranışların beklenen düzeyde öğrenilmiş olup olmadığına karar verebilmek için bu davranışların geçerliği ve güvenirliliği yeterli derecede yüksek ölçme araçlarıyla ölçülmesi gerektiğini belirtmektedir. Ayrıca, özellikle eğitimde kullanılan ve öğrenme düzeyini ölçmeyi amaçlayan testlerde birden fazla özelliğin ölçülmesi gerekliliğine vurgu yapmaktadır.

Geçerlik kavramı üzerinde yoğun çalışmaların yapıldığı 1940'lı yıllarda testin geçerliğini ortaya koyacak ölçütün testin dışından ya da testin içinden olup olamayacağı da tartışılmıştır. Sireci (1998) geçerlik üzerine ilk çalışmaları yapan bilim insanlarının, geçerlik değerlendirmelerinin yalnızca istatistiksel bakışıyla yapılmasının çok kısıtlayıcı olduğu yönündeki kaygılarını dile getirmiştir. Bununla ilgili olarak, Kelly'nin 1927 yılında testlerin geçerlik değerlendirmelerini desteklemek için uzman yargılarının da kullanılması önerisini dile getirmiştir. Rulon ise 1946 yılında, testin geçerliği hakkında bilgi verecek ölçütün testin içeriğiyle ve amacıyla ilgili olduğunu vurgulamıştır.

İlerleyen yıllarda, testlerin ölçmeyi hedeflediği *konu boyutu* ile konuların hangi *bilişsel düzeyde* testte yer alacağı düşüncesi ön plana çıkmıştır. Kapsam geçerliği kavramını Cronbach ve Meehl (1955), test maddelerinin ölçülmek istenen alanın bir örneklemini olarak tanımlamıştır. Turgut ve Baykul (2012), her bir bilişsel hedefin iki boyutlu bir belirtke tablosunda özetlenebileceğini ve tablonun bir boyutunda hedeflerin; ikinci boyutunda ise konuların sıralanacağını belirtmiştir. Ebel (1956), kapsam geçerliğinin test geliştirme sürecinde göz ardı edilmemesi gerektiğini vurgulamakta, test geçerliğinin incelenmesi için öncelikle kapsam geçerliği kanıtlarının toplanmasının gerekliliğine dikkat çekmektedir. Öğretmenlerin sınavlarında konuları ve bilişsel hedefleri saptamasının bunlar arasındaki oranları belirlemesinin, her konudan ve hedeften kaç soru sorulacağını ortaya konmasının önemini dile getirmektedir. Ayrıca, geçerliğin bir ön koşulu olan güvenirliliğin yüksek bir değer alabilmesi madde sayısı ile ilişkilendirilmiştir. Bu nedenle, testte yer alan madde sayısı ne kadar arttırıldığında güvenirlilik değerinin ne kadar artacağını gösteren bir formülasyon (test uzunluğu) hem test geliştiriciler tarafından hem de öğretmenler tarafından kullanılmaktadır (Thorndike ve Thorndike-Christ, 2014).

Bu açıklamalardan da görüleceği gibi, okul öğrenmelerinin ölçülmesinde öğretmen yapımı testlerin istenen güvenirlilik ve geçerlik değerlerine sahip olabilmesi için öğretmenler, çok sayıda maddeden oluşan testleri kullanmaya yönlendirilmiştir. Bu durum öğretmenlerin, testlerinde konu boyutu açısından çoğunlukla öğrenme düzeyini gösteren kapsam temelli öğrenmelere; bilişsel hedefler açısından ise Bloom sınıflamasının çoğunlukla bilgi, kavrama ve uygulama düzeyindeki bilişsel süreçlere yönelmesine neden olmuştur. Bu yaklaşım öğretmenlerin hem öğrencilerinin başarılarını izlemeye (formative) hem de onların başarı düzeylerini belirlemeye (summative) kısa yanıtı, boşluk doldurmalı, eşleştirme gerektiren, doğru-yanlış ve çoktan seçmeli madde türlerini sıklıkla kullanmalarına yol açmıştır. Bu anlayış öğretmenleri daha çok, öğrencilerinin derste ele alınan konuları "öğrenip-öğrenmediği" noktasına yönlendirmiştir. Sınıf içi ölçmelerin ise, öğrencilerin *öğrenme eksikliklerinin ve öğrenme güçlüklerinin* tamamlanmasına katkı getirici biçimde yapılandırılmasına yol açmıştır (Bloom, Hastings ve Madaus, 1971; Özçelik, 1992-2).

Bu deęişimler olumlu yönde olsa da okulların yerine getirmesi gereken *öęrenmelerin gerçek yaşam durumlarında kullanılması* işlevinin gelişmesi önünde engel olmuştur. Bu ihmal, öęrencilerin öęrenmelerini gerçekçi durumlarda kullanmalarını sağlayan temel yaşam becerilerinin gelişmesini de sağlayamamıştır. Öęrenciler okullardan, temel bilgileri öęrenen ancak bu bilgilerini doğal ve toplumsal yaşam durumlarında istenen düzeyde kullanamayan bireyler olarak mezun olmuşlardır. Bu gibi nedenlerle birçok toplum, özellikle 1980’li yıllardan itibaren okulların öęrenme yaklaşımlarını sorgulamaya başlamıştır. Bu tartışmalarda okulların yaşam için gerekli bilişsel, duyuşsal ve sosyal becerileri yeterince kazandıramadığı ele alınmıştır (Brookhart, 2009, 2013; Haladyna, 1997; Kutlu vd., 2015; Kutlu ve Kula-Kartal, 2018; Marzano, 1992; Marzano ve Kendall, 2007; NCLB, 2001; Popham, 2004, 2007; UNDP ve UNICEF, 1990).

Bloom ve dięerleri (1956) tarafından geliştirilen bilişsel sınıflama bir yandan öęretim programlarının doğasını deęiştirirken bir yandan da ölçme ve deęerlendirme biliminin test geliştirmeye olan bakış açısını deęiştirmiştir. Güvenirlięi artırma, kapsam ve yapı geçerlięini sağlama ve testlerin yordayıcı gücünü sağlama düşüncesi, Bloom sınıflamasının üst basamaklarına ilişkin madde sayısının artırılması önünde engel olmuştur. Ayrıca kısa yanıtı, boşluk doldurmalı, eşleştirme gerektiren, doğru-yanlıř, çoktan seçmeli gibi maddeler de bu düzeyleri ölçmede yetersiz kalmıştır. Çoktan seçmeli maddenin pek çok çeşidinin olması nedeniyle, bu madde türüyle her türlü bilişsel yeterlięin ölçülebileceęi düşüncesi de savunulmuştur (Özçelik, 1992-2; Turgut ve Baykul, 2012). Ancak çoktan seçmeli madde yazımının uzmanlık gerektirmesi, yazımının zaman alması, yanıtın seçenekler arasında verilmiş olması, bireye kendi yanıtını oluşturma fırsatı tanımaması ve üst düzey düşünme becerilerinin bu madde türüyle yeterince ölçülemeyeceęi düşüncesi bu savunuların uygulamada yerinin olmadığını ortaya koymuştur. Bloom’un bilişsel sınıflamasına ve Tam Öęrenme Modeli’ne dayanan okul uygulamaları, öęrencilerin öęrenmelerini gerçek yaşam durumlarında istenen düzeyde kullanamadıklarını ortaya koymuştur. Birçok ülke öęrencilerin çağın gerektirdięi temel yaşam becerilerini kazanamadıklarını ve bu durumdan okullardaki eğitim-öęretim anlayışının sorumlu olduğunu tartışmaya başlamıştır. 1948 yılında yayımlanan İnsan Hakları Evrensel Bildirgesi’nde dile getirilen, *her insanın eğitim hakkına sahip olduęu* maddesi dünya devletlerince kabul edilmiştir. Buna rağmen günümüzde birçok çocuk; temel eğitimi tamamlayamamış, okuma yazma öęrenememiş, toplumsal ve kültürel yaşam için gerekli temel bilgi ve becerileri yeterince kazanamamıştır. Bu durum Bildirge’nin eğitim konusundaki hedeflerine ulaşamadığını göstermektedir.

Eğitim dünya ülkeleri tarafından; savaş, açlık, çevre kirlilięi, çocuk ve kadına yönelik şiddet, gelir dağılımı adaletsizlięi, uyuşturucu kullanımı, terör, hakların ihlali gibi birçok sorunun çözümünde önemli bir araç olarak görülmektedir. Son yıllarda eğitim süreçleriyle ilgilenen bilim insanları yanıř eğitim politikalarının ve eğitim kararlarının da bu sorunlara yol açabileceęi üzerinde durmaktadır (Haladyna, 1997; Kutlu vd., 2015; Popham, 2000). Okul eğitimi hakkındaki çalışmalar günümüzde okullarda yürütülen eğitimin yetersizlięine vurgu yapmakta ve okul uygulamalarını eleştirmektedir (Gardner, 2019; Gatto, 2018; Harari, 2018; Illich, 2016; Khan, 2016; Wagner, 2015). Okul eğitiminin yetersizlięine vurgu yapan çalışmalar günümüz okullarının hem 21. yüzyılın gereklerine hem de içinde bulunan çağı yaşayacak genç kuşağın deęişen özelliklerine uymadığına vurgu yapmaktadırlar (Marzano ve Heflebower, 2012). 20. yüzyıldaki elektrik, elektronik, bilgisayar, makine, fizik, tıp, biyoloji gibi bilim alanlarındaki bilgilerin hızlı deęişmesi okulun öęrenci yetiştirme anlayışındaki beklentilerini de deęiştirmeye başlamıştır. 1850’li yıllarda yaşanan Endüstri Devrimi kendi içinde sık aralıklarla yeni buluşları beraberinde getirerek toplumsal ve doğal yaşamda köklü deęişikliklere neden olmuştur. Enerji, tekstil, otomotiv, demir-çelik, elektrik-elektronik, bilgisayar, kimya gibi alanlarda kısa aralıklarla devrim niteliğinde buluşlar yapılmış ve insanlığın kullanımına sunulmuştur. Bu durum teknolojiye dayalı bilgilerin, insanın sahip olduęu psikolojik özelliklerinden daha hızlı gelişmesine neden olmuştur. Okullar ise, hızla deęişen bu bilgi akışına yetişebilmek için öęrencilerde çoęunlukla bilişsel davranışların gelişimine odaklanmışlardır.

Birçok ülke 1980’li yılların sonlarında *Hiçbir Çocuğun Geride Kalmaması* gerektięi düşüncesini tartışmışlar ve “Herkes İçin Eğitim Dünya Beyannamesi”ni yayımlamışlardır. Beyanname genişletilmiş bir görüş çerçevesi ile yenilenmiş söz verme süreçlerini içermektedir. Bu görüşler řu noktaları kapsamaktadır (UNDP ve UNICEF, 1990):

1. Erişilebilirlięi evrenselleştirmek ve eşitlięi arttırmak,
2. Öęrenme başarısı üzerinde odaklanmak,
3. Temel eğitimin araçlarını ve kapsamını genişletmek,
4. Öęrenme ortamını zenginleştirmek,
5. Ortaklıkları güçlendirmek.

Herkes İçin Eğitim (Education for All) Küresel İzleme Raporu 2000’li yıllardan itibaren “Başarabilecek miyiz?” sorusunu sormakta ve 2015 yılı için yukarıda yenilenmiş ve verilmiş sözlerin gerçekleşme düzeyini araştırmaktadır (UNESCO, 2002). UNESCO (2013-4) raporu, 2015’e az bir sürenin kalmış olmasına rağmen, yenilenmiş ve verilmiş sözlerin başarılamayacağını açık bir biçimde görüldüğünü belirtmektedir. Rapor, 57 milyon çocuğun öğrenmede başarısız olduğunu, çünkü bu çocukların okul içinde olmadığını vurgulamaktadır. Ayrıca raporda eğitimdeki düşük kalitenin, öğrenmeyi okula gidenler için bile engellediği, ilkokul çağındaki çocukların üçte birinin okula gitmeler de gitmeseler de temel becerileri öğrenemedikleri dile getirilmektedir.

Sınıf İçi Durum Belirleme Anlayışında Değişim Kaçınılmaz

Eğitimde sözü edilen olumsuz koşulların değiştirilmesi isteği eğitim-öğretim süreçlerinde güncellemeler yapılmasını kaçınılmaz kılmaktadır. 2000’li yıllardan itibaren uygulanan PIRLS, PISA ve TIMSS gibi uluslararası öğrenci başarısını belirleme çalışmalarının bulguları eğitimde karar vericilere önemli veriler sunmakta ve eğitimde alınması gereken önlemleri işaret etmektedir. Ayrıca insan davranışlarının doğasını daha ayrıntılı anlamamızı sağlayan biyoloji, tıp ve psikoloji biliminin bulguları okulların öğrenmeleri hızla yeniden düzenlenmesi gerektiğinin de işaretini vermektedir.

Psikometri, ölçme ve değerlendirme bilim alanı uzun yıllar program geliştirmenin, öğrenme ve öğretme kuramlarının, psikoloji, biyoloji ve tıp biliminin zihin üzerindeki çalışmalarından etkilenmiştir ve etkilenmesini sürdürmektedir. Özellikle 1950’li yıllardan itibaren okul öğrenmelerinde etkili olan bilişsel, duyuşsal ve devimsel (psikomotor) alanların ne olduğu ve nasıl ölçüleceği hakkındaki çalışmalar (Anderson ve Krathwohl, 2001; Bloom vd., 1956; Carroll, 1963; Krathwohl vd., 1964; Marzano ve Kendall 2007; Thorndike, 1971), öğrenme ve öğretme kuramları çalışmaları (Bloom, 1976; Hergenhahn ve Olson, 2001; Hilgard ve Bower, 1975; Schunk, 2012) dikkat çekicidir. 1980’li yıllardan itibaren biyoloji, tıp, nörolojik psikoloji ve antropoloji biliminin beyin ve bellek üzerindeki çalışmaları eğitim bilimleri alanına bireyin öğrenmeleri hakkında daha fazla bilgi sunmuştur (Andreasen, 2005; Asbury ve Plomin, 2013; Carter, 2013; Demirsoy, 1998, 2018; Eagleman, 2013, 2016; Eagleman ve Brandt, 2019; Güvenç, 1997; Kılıç, 2019; Lewin, 1998). Beyinle ilgili çalışmaların bir uzantısı olarak bilişsel, duyuşsal ve devimsel alanları da içine alan *zihin* kavramıyla ilgili çalışmalar (Dehaene, 2014; Gardner, 1993; Kaku, 2014; Kandel, 2013; Kurzweil, 2015) ve bilişsel psikolojinin bulguları özellikle hem öğretim programlarının bilişsel kurama göre yeniden yapılandırılmasına hem de eğitimde ölçme ve değerlendirme yaklaşımlarının yenilenmesini sağlayacak anlayışların ortaya çıkmasına yol açmıştır. Yukarıda sözü edilen gelişmeler dikkate alındığında 21. yüzyılın, doğayla ve toplumla daha barışık yaşayacak bireylerin yetiştirilmesine gereksinim duyduğu görülmektedir. Bu yüzyıl, okul öğrenmeleri için üretilmiş eski bilgilerden de yararlanarak yeni bilgiler ortaya koyabilmelidir. Gregoryen takvime göre 21. yüzyıl, 1 Ocak 2001 - 31 Aralık 2100 tarihleri arasına karşılık gelen zaman dilimini kapsamaktadır.

Öğrencilerin başarısında etkili olduğu düşünülen ve 21. yüzyıl becerileri olarak adlandırılan bilişsel, içsel* ve kişilerarası** yeterliklerin birlikte ölçme ve değerlendirme sürecine dahil edilmesi önemlidir. Bu yeterliklerin köklerinin yüzyıllar öncesine gittiği bilinmektedir. Ancak bu yeterliklerin ayrıntılarının araştırılması, ölçülebilir ve gözlenebilir bir yapıya oturtulması çoğunlukla 19. ve 20. yüzyıl içinde olmuştur. Günümüzde öğrencilerin gerçek yaşam için gerekli bu becerilerde yetkinleşmiş olarak okullardan mezun olmaları gerektiği düşüncesinin yaygınlaşması oldukça yenidir (Pellegrino, 1999; Greiff ve Kyllonen, 2016). Öğrencilerin kendi öğrenmelerini yönetebilmelerine olanak sağlayan; problem çözme, analitik düşünme, akıl yürütme, eleştirel düşünme, yaratıcılık gibi *bilişsel*; özgüven, özyeterlik, güdülenme, yılmazlık, problem çözmeye açıklık, kararlılık, esneklik, duyguları anlama gibi *içsel*; işbirliği yapma, iletişim kurma, uyum, bireysel farklılıklara değer verme, sorumluluk alma, ikna etme, hesap verebilir olma, üretkenlik, zaman yönetimi gibi *kişilerarası* özelliklerinin de ölçme sürecinin içine çekilmesi iyi bir ölçme ve durum belirleme anlayışının vazgeçilmez koşulu olarak kabul edilmektedir. Okul öğrenmelerinde, günlük yaşamda ve çalışma yaşamında ön plana çıkan bu beceriler 21. yüzyıl becerileri olarak adlandırılmaktadır (Collins, 2014; Kutlu ve Kula-Kartal, 2018; Kyllonen, 2012; Marzano ve Heflebower, 2012; Payne ve Kyllonen, 2012). Bu beceriler bu çalışmada, *bilişsel, içsel ve kişilerarası* olmak üzere üç grupta ele alınmıştır.

* Bu yeterlik alanı içinde yer alan beceriler, bireyin olaylar karşısında kendini anlamasını, yeni bilgiler öğrenmesini ve sahip olduğu duyuşsal özelliklerini bu durumlarla baş edebileceği biçimde yönetmesini içermektedir.

** Bu yeterlik alanı içinde yer alan beceriler, bireyin olaylar karşısında başkasını anlamasını, başka bireylerle etkileşim içinde olmasını ve karşılaştığı sorunları sosyal özelliklerini kullanarak yönetmesini içermektedir.

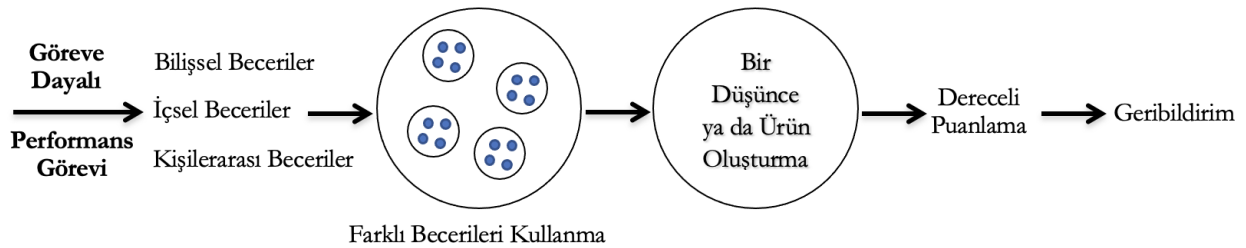
21. Yüzyılda Sınıf İçi Ölçmelerden Beklentiler

Beyin, sinir hücreleri, bellek ve öğrenme üzerindeki çalışmalar (Carter, 2013; Demirsoy, 1998, 2018; Eagleman, 2016; Eagleman ve Brandt, 2019; Kılıç, 2019) okul programları, dolayısıyla ölçme ve değerlendirmenin sınıf içi ölçmelere bakış açısını etkilemektedir. Sözü edilen gelişmeler yalnızca bilişsel davranışların ölçülmesine odaklı ölçme ve durum belirleme anlayışları yerine bireyin yaşamda kullandığı içsel ve kişilerarası becerilerinin de ölçülmesini ön plana çıkarmaktadır. Kılıç (2019), “Yeni Bilim ve Kültürün Kaynağı: Bağlantısal Bütünsellik” başlıklı YouTube konuşmasında, insan zihninin tek boyutlu değil birbiriyle ilişki içinde çalışan bağlantısal bir yapısının olduğunu belirtmekte ve bu durumu “çoklu bağlantısalılık” kavramıyla açıklamaktadır. Demirsoy (1998) ise bireyin, anne karnında başlayan sinaplaşma sürecinin ancak zengin öğrenme uyarıcılarıyla gelişeceğini; gerekli uyarıcılar sağlanmadığında kısa zaman diliminde çevreden gelen uyarıcıları algılamakta ve anlamakta zorlanacağını belirtmektedir.

Bu bilgilere dayanan bir anlayışla okulun yapılandırılması ve öğrencilerin yetiştirilmesi hem yeni bir okul düzenini hem de yeni bir ölçme ve durum belirleme anlayışını kaçınılmaz kılmaktadır. Okullarda yaygın kullanılan kısa yanıtı, boşluk doldurmalı, eşleştirme gerektiren, çoktan seçmeli gibi klasik ölçme yaklaşımlarına dayalı maddeler çoğunlukla öğrencileri ezberle yönlendirmekte ve öğrendiklerini yaşam durumlarında kullanmalarının önünde engel olmaktadır. Klasik ölçme yaklaşımları yapısı gereği temel bilgilerin ve alt düzey öğrenmelerin ölçülmesine daha fazla katkı sağlamakta, problem çözme, akıl yürütme, eleştirel düşünme, yaratıcılık gibi üst düzey düşünme becerilerinin ölçülmesinde yeterli olamamaktadır (Haladyna, 1997; Kutlu vd., 2015).

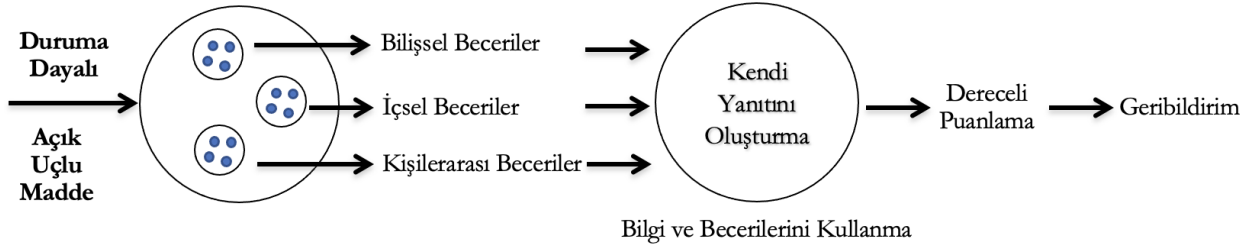
Günümüz toplumlarda eğitimde değişimin kaçınılmaz olduğunu kabul etmekte ve hali hazırda kullanılan öğrenme ve ölçme sistemlerinin yeterli olmayacağını düşünmektedirler. Stiggins (2002), sınıf içi ölçme ve durum belirleme uygulamalarının okul gelişimiyle bağlantısının kurulmasını ve bu uygulamalara yeni bir bakış açısıyla yaklaşılmasının gerekliliğini vurgulamaktadır. Bilim ve teknoloji anlayışındaki gelişmeler gelecekte insanların kendi öğrenmeleri ve öğrenme ortamları üzerinde daha fazla denetim sahibi olacaklarına işaret etmektedir. Yakın gelecekte bireylerin bir alanda yetkin olabilmeleri, *öğrenme, anlama, öğrenilenleri bir araya getirerek yaşamda kullanma* gibi vazgeçilmez üç özelliğe sahip olmalarını zorunlu kılmaktadır. Bu çalışmada öğrenci başarısı, öğrencilerin okulda kazandıkları bilişsel, içsel ve kişilerarası becerileri kendi bireysel özellikleriyle ilişkilendirerek gerçek yaşam durumlarında kullanma gücü olarak tanımlanmıştır. Alanyazında bu özellikler bir bütün olarak *üst düzey düşünme becerileri* olarak adlandırılmaktadır (Brookhart, 2010; Haladyna, 1997; Kutlu vd., 2015; Marzano, 1992; Popham, 2000).

Bu yüzyılda okulların öğrencilerde bu becerileri geliştirmesi son derece önemlidir. Sınıf içi ölçme ve durum belirleme uygulamalarında okulların, öğrencilerde sözü edilen üç yeterliği geliştirmesi beklenmektedir. Bir sınav kâğıdına ya da ders tekrarına dayalı ev ödevlerine puan verilmesinden daha çok öğrencilerin bu yeterlikleri yaşam durumlarında kullanması sağlanmalıdır. Bunun için geribildirim vermeyi önceleyen ve birden fazla beceriyi aynı anda kullanmayı gerektiren performans görevleri [Örnek çalışma için Kutlu ve Kula-Kartal’a (2018) bakılabilir] kullanılmalıdır. Göreve dayalı akış şeması Şekil 1’deki gibidir:



Şekil 1. Göreve dayalı akış şeması

Bunun yanında okullarda, öğrencinin birden fazla bilgiyi bir araya getirdiği bir ya da birkaç beceriye dayalı, kendi yanıtını yapılandırmasını sağlayan açık uçlu maddelere dayalı ölçme ve durum belirleme anlayışı da yaygınlaştırılmalıdır [Bu konuda yapılan bir çalışma için Kutlu’ya (2004) bakılabilir]. Bu anlayışın uygulanma sürecinin akış şeması ise Şekil 2’de verilmiştir:



Şekil 2. Açık uçlu madde uygulama süreci akış şeması

Şekil 1 ve 2’de sözü edilen durum belirleme süreçlerinin okul öğrenmelerinde ön plana çıkması öğrencilerin başarısını belirleme sürecinin yapısını değiştirmektedir. Bu tür maddelere dayalı bir ölçme süreci sonunda alınan ayrıntılı bir geribildirim ile öğrenciler temel bilgileri öğrenip öğrenemediklerini izleyebilecekler ve eğitim sürecinde gerekli önlemleri alabileceklerdir. Burada öğretmen gerçek bir yol gösterici olabilecek ve öğrenci kendi öğrenmelerini rahatlıkla yönlendirebilecektir.

Burada öğretmene düşen asıl görev, olgusal bilgilerin kullanılma yolları olan bilişsel, içsel ve kişilerarası becerileri gerçek yaşam durumlarına dayalı örnek olay ve durumlar içinde öğretmek, daha sonra bu öğrenmelerin gerçekçi durumlarda kullanımını sağlayacak açık uçlu ve performans görevleri göreve dayalı maddeleri vermektir. Üst düzey düşünme süreçleriyle ilgili olan görevlerin öğrenciler için yeni olması, onların kendi yeteneklerini kullanarak özgün yanıtlar vermeye çalışmalarını açısından önemlidir. Sınıf ve yaş düzeyine uygun olarak verilen görevlerin farklı kaynakları araştırmaya, okumaya, öğrencileri gözlem yapmaya, veri toplamaya özendirilmesi yerinde olacaktır. Görevlerde, öğrenci yanıtlarının onların görüşlerini yansıtacak yazılı anlatımı içermesi öğrencinin düşünsel gelişimi için gereklidir. Bu nedenle görevlerin açık uçlu maddelere dayanması, rapor, makale vb. yazmayı gerektirmesi, proje, plan, deney düzeneği gibi öneri oluşturmayı hedeflemesi kısaca derslere ilişkin öğrenmeleri kullandıracak durum belirleme yaklaşımlarının kullanılması önemlidir.

Durum belirleme maddelerinin hazırlanması ve maddelere verilen öğrenci yanıtlarının puanlanarak geribildirim verilmesi hem öğretmenlerde hem öğrencilerde bazı eğitsel donanımların olmasını gerektirmektedir. Öğretmenlerin göreve dayalı madde yazabilmeleri onların açık uçlu madde yazımında yetkin olmalarını gerektirmektedir. Açık uçlu ve göreve dayalı maddelerin puanlanmasında ise *dereceli puanlama anahtarları (scoring rubrics)* kullanılmaktadır. Dolayısıyla öğretmenlerin bu türden puanlama anahtarlarının geliştirilmesi ve geribildirim verme süreçlerini de iyi bilmeleri gerekmektedir. Ayrıca öğrencilerin kendilerini değerlendirmelerini sağlayacak değerlendirme formları hakkında bilgilendirilmeleri de yerinde olacaktır.

Önceki yüzyılların okul eğitimi, öğrenci başarısının geliştirilmesi açısından 21. yüzyıla çözülmesi gereken eğitsel sorunları da beraberinde getirmiştir. 2000’li yıllardan itibaren standart başarı testlerinden daha farklı yapıda geliştirilmiş olan PIRLS, PISA ve TIMSS gibi uluslararası öğrenci başarısını belirleme çalışmaları, katılımcı ülkelerin özellikle öğrencilerin okuma, fen ve matematik başarıları hakkında bilgiler ortaya koymuştur. Bu durum belirleme çalışmalarında yüksek puan alan ülkeler bile, yine bu çalışmalarda tanımlanan yeterlik düzeyleri açısından öğrencilerini istenen üst basamaklara tam olarak taşıyamamışlardır. Örneğin, 2003-2018 yılları arasında altı kez uygulanan PISA sonuçlarında bu durum değişmemiştir. Bu bulgu, öğrencileri üst yeterlik düzeyine taşımada bilişsel yeterliklere dayalı öğrenmelerin yeterli olmadığını; bilişsel yeterliklerin dışında sosyal ve duygusal yeterliklerin de bireylere kazandırılması gerektiğini ortaya koymuştur.

SONUÇ

Dünyamız 21. yüzyıla, 19. ve 20. yüzyıldan gelen birçok olumsuz etkiyle girmiştir. Bu etki 21. yüzyılda da devam ederek küresel iletişimin kolaylaştırdığı uluslararası karşılıklı bağımlılık, medyanın kontrolü, artan milliyetçilik, terörizm, büyük veri kullanımının bireylerin gizliliğini tehdit etmesi, dijitalleşmenin ve medyanın da yardımıyla sahte verilerin ve hikayelerin kolaylıkla yaratılıp, yönlendirilmesi gibi zorlukları da beraberinde getirmiştir. Eskiden doğal çevre insanın dışında yer alan bir parça olarak görülürken, bugün insanın yalnızca doğanın bir parçası olduğu daha büyük bir ekosistem olarak görülmektedir. Bu nedenle insanların doğayla uyum içinde yaşaması ve doğayla var olması beklenmektedir. Bu vizyonu gerçekleştirmek için herkesin harekete geçmesi; iş bölümünden paylaşılan sorumluluğa geçilmesi gerekmektedir. Bunu başarabilmek için herkes bilgiye, becerilere ve katkıda bulunma arzusuna sahip olmalıdır (OECD, 2019).

Sözü edilen sorunlarla ilgili önlemler alınmazsa ne yazık ki 21. yüzyılın genç kuşakları bu sorunları en ağır biçimde yaşayacaklardır. Doğal ve toplumsal birçok sorunun çözümünde eğitim en önemli araç olarak görülmekte ve eğitim uygulamaları kapsamında neler yapılabileceği özellikle son 30 yıldır tartışılmaktadır. OECD'nin 2018 yılında yayımladığı *Eğitimin ve Becerilerin Geleceği 2030* raporu ile 2019 yılında yayımladığı *Öğrenme Pusulası 2030* raporlarında, ülkelerin iki geniş kapsamlı soruya (ne ve nasıl) yanıt bulmalarına katkı sağlamak amaçlanmıştır. Bu sorulardan ilki “Günümüz öğrencilerinin bireysel ve toplumsal refahlarını sağlamada, kendilerini geliştirmek ve dünyalarını biçimlendirmek için hangi bilgi, beceri, tutum ve değerlere (yetkinliklere) gereksinimi olacak?” İkinci soru ise “Öğretim sistemleri bilgi, beceri, tutum ve değerleri etkili biçimde nasıl geliştirecek?” Bu iki kapsamlı soru etrafındaki çalışmalar incelendiğinde, yukarıda anlatılan ölçme ve durum belirleme yaklaşımlarının önemi ortaya çıkmaktadır. Bu çalışma hem ölçme ve değerlendirme anlayışının geçmişten günümüze tarihsel süreçlerini ele alması hem de ortaya koyduğu yeni ölçme ve durum belirleme anlayışı açısından özgünlük taşımaktadır.

KAYNAKÇA

- Ackerman, T. A., Gierl, M. J., & Walker, C. M. (2003). Using multidimensional item response theory to evaluate educational and psychological tests. *Educational Measurement: Issues and Practice*, 22(3), 37-53. <https://doi.org/10.1111/j.1745-3992.2003.tb00136.x>
- Anastasi, A. (1976). *Psychological testing* (4th Ed.). Macmillan Publishing.
- Anastasi, A. (1986). Evolving concepts of test validation. *Annual Review of Psychology*, 37, 1-15. <https://doi.org/10.1146/annurev.ps.37.020186.000245>
- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th Ed.). Prentice-Hall/Pearson Education.
- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives* (Complete Ed.). Longman.
- Andreasen, N. C. (2005). *The creating brain: The neuroscience of genius*. Dana Press.
- Asbury, K., & Plomin, R. (2013). *G is for genes: The impact of genetics on education and achievement*. John Wiley & Sons, Inc.
- Bartholomew, D. J. (2004). *Measuring intelligence: Facts and fallacies*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511490019>
- Bloom, B. S., Englehart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *The Taxonomy of educational objectives, handbook I: The Cognitive domain*. David McKay Co., Inc.
- Bloom, B. S., Hastings, J. T., & Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning*. Mc Graw-Hill.
- Bloom, B. S. (1976). *Human characteristics and school learning*. McGraw-Hill.
- Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review*, 111(4), 1061-1071. <https://doi.org/10.1037/0033-295X.111.4.1061>
- Bransford, J. D., & Stein, B. S. (1984). *The IDEAL problem solver*. W. H. Freeman.
- Brennan, R. L. (2001). An essay on the history and future of reliability from the perspective of replications. *Journal of Educational Measurement*, 38(4), 295-317. <https://www.jstor.org/stable/1435452>
- Brookhart, S. M. (2009). *Exploring formative assessment*. ASCD Publishers.
- Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. Copyright Clearance Center.
- Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. ASCD Publishers.
- Brown, A. L. (1978). Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 1, pp. 77-165). Hillsdale: Erlbaum.
- Bruno, F. J. (1996). *Psikoloji tarihi* (Çev. G. Sevdiren.). Kibele Yayınevi.
- Budak, İ. (2008). Üstün yeteneklilik kavramı ve tarihsel gelişim süreci. *Journal of Qafqaz University*, 22, 164-173. <http://eds.a.ebscohost.com/eds/pdfviewer/pdfviewer?vid=0&sid=3599e458-6976-4d89-8663-8de1e28e9a1b%40sessionmgr4007>
- Carroll, J. B. (1963). A model for school learning. *Teachers College Record*, 64, 723-733. <http://garfield.library.upenn.edu/classics1982/A1982NF37300001.pdf>

- Carter, R. (2013). *Beyin*. Alfa Yayıncılık.
- Cattell, J. M. (1890). Mental tests and measurements. *Mind*, 15(59), 373-381. <http://www.jstor.org/stable/2247264>
- Cohen, R. J., & Swerdlik, M. (2009). *Psychological testing and assessment. An introduction to tests & measurement* (7th Ed.). McGraw-Hill.
- Collins, R. (2014). Skills for the 21st Century: Teaching higher-order thinking. *Curriculum & Leadership*, 12(14), front page. http://www.curriculum.edu.au/leader/teaching_higher_order_thinking,37431.html?issueID=12910
- Crocker, L. M., & Algina, L. (1986). *Introduction to classical and modern test theory*. Holt, Rinehart and Winston.
- Cronbach, L. J. (1990). *Essentials of psychological testing*. Harper & Row.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281-302. <https://doi.org/10.1037/h0040957>
- Çığ, M. İ. (2002). *Ortadoğu uygarlık mirası*. Kaynak Yayınları.
- Dehaene, S. (2014). *Beyin nasıl okur?* Alfa Basım Yayım Dağıtım.
- Demirsoy, A. (1998). *Son imparatora öğretmenler: Bilim toplumu*. Meteksan A.Ş.
- Demirsoy, A. (2018). *Biyolojik saat: Belleğin ve davranışların evrimi*. Asi Kitap.
- Domino, G., & Domino, M. L. (2006). *Psychological testing: An introduction* (2nd Ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511813757>
- Eagleman, D. (2013). *Incognito: Beynin gizli hayatı*. Domingo Yayıncılık.
- Eagleman, D. (2016). *Beyin: Senin hikayen*. Domingo Yayıncılık.
- Eagleman, D., & Brant, A. (2019). *Yaratıcı tür: Fikirler dünyayı nasıl yeniden yaratıyor*. Domingo Yayıncılık.
- Ebel, R. L. (1956). *Essentials of educational measurement* (1st Ed.). Prentice-Hall.
- Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of educational measurement* (5th Ed.). Prentice-Hall of India Private Limited.
- Ercan, F. (2009). *Toplumlar ve ekonomiler*. Bağlam Yayıncılık.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906-911. <https://doi.org/10.1037/0003-066X.34.10.906>
- Freely, J. (2014). *Galileo'dan önce: Orta Çağ Avrupa'sında modern bilimin doğuşu*. Kolektif Kitap.
- Gardner, H. (1993). *Multiple intelligences: The theory in practice* (10th Ed.). Basic Books.
- Gardner, H. (2019). *Eğitilmemiş zihin*. Alfa Basım Yayım Dağıtım.
- Gatto, J. T. (2018). *Eğitim: Bir kitle imha silahı*. Edam Yayıncılık.
- Greiff, S., & Kyllonen, P. C. (2016). Contemporary assessment challenges: The measurement of 21st century skills. *Applied Measurement in Education*, 29(4), 243-244. <https://doi.org/10.1080/08957347.2016.1209209>
- Guilford, J. P. (1954). *Psychometric methods* (2nd Ed.). McGraw-Hill.
- Gulliksen, H. (1950). *Theory of mental tests* (1st Ed.). John Willey & Sons.
- Gulliksen, H. (1967). *Theory of mental tests* (6th Ed.). John Willey & Sons.
- Gündüz, M. (2014). Ünite 1: Eski çağda eğitim. *Eğitim Tarihi, Anadolu Üniversitesi* (1-29), Şişman, M., Editör. Anadolu Üniversitesi, AÖF Yayınları.
- Güven, İ. (2019). *Türk eğitim tarihi* (7. Baskı). Pegem Akademi.
- Güvenç, B. (1997). *İnsan ve kültür*. Remzi Kitabevi.
- Haladyna, T. M. (1997). *Writing test items to evaluate higher order thinking*. Viacom Company.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Kluwer Nijhoff Publishing.
- Harari, Y. N. (2015). *Hayvanlardan tanrılara SAPIENS: İnsan türünün kısa bir tarihi*. Kolektif Kitap.
- Harari, Y. N. (2016). *Homo Deus: Yarının kısa bir tarihi*. Kolektif Kitap.
- Harari, Y. N. (2018). *21. yüzyıl için 21 ders*. Kolektif Kitap.

- Hergenhahn, B. R., & Olson, M. H. (2001). *An introduction to theories of learning* (6th Ed.). Phoenix Color Corp.
- Hilgard, E. R., & Bower, G. H. (1975). *Theories of learning* (5th Ed.). Prentice-Hall Inc.
- Hogan, T. P. (2013). *Psychological testing - A practical introduction* (3rd Ed.). John Wiley & Sons, Inc.
- Illich, I. (2016). *Okulsuz toplum*. Şule Yayınları.
- İstek, G. (2019). Cizvit tarikatının liselerde verdiği eğitim ve buna karşı oluşan tepkiler: Avusturya örneği. *Akademik İncelemeler Dergisi*, 14(2), 163-198. <https://dergipark.org.tr/en/download/article-file/867609>
- Janda, L. H. (1992). *Psychological testing: Theory and applications*. Boston, MA: Allyn and Bacon.
- Jenkins, J. G. (1946). Validity for what? *Journal of Consulting Psychology*, 10(2), 93-98. <https://doi.org/10.1037/h0059212>
- Kaku, M. (2014). *Zihnin geleceği*. ODTÜ Yayıncılık.
- Kandel, E. R. (2013). *Belleğin peşinde: Yeni bir zihin biliminin doğuşu*. Boğaziçi Üniversitesi Yayınevi.
- Kaplan, R. M., & Saccuzzo, D. P. (2008). *Psychological testing: Principles, applications, and issues* (7th Ed.). Wadsworth, Cengage Learning.
- Khan, S. (2016). *Dünya okulu: Eğitimi yeniden düşünmek*. Yapı Kredi Yayınları.
- Kılıç, T. (2019, 24 Aralık). Prof. Dr. Türker Kılıç'ın yeni bilim ve kültürün kaynağı: Bağlantısız bütünsellik üzerine konuşması [Video]. YouTube. <https://www.youtube.com/watch?v=rLu2zXWvAX8&feature=youtu.be>
- Kline, P. (1986). *A handbook of test construction*. Methuen & Co. Ltd.
- Köçer, M., & Koçoğlu, E. (2020). *Eğitim tarihi*. Pegem Akademi.
- Kramer, S. N. (1999). *Tarih Sümer'de başlar: Yazılı tarihteki otuz dokuz ilk*. Kabalıcı Yayınevi.
- Kramer, S. N. (2002). *Sümerler*. Kabalıcı Yayınevi.
- Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1964). *Taxonomy of educational objectives: The classification of educational goals, handbook II: Affective domain*. David McKay Co., Inc.
- Kurzweil, R. (2015). *Bir zihin yaratmak: İnsan düşüncesinin esrarı*. İstanbul Üniversitesi Yayınları.
- Kutlu, Ö. (2004, 17-18 Ocak). *Tek soruyla öğrenci performansının belirlenmesi* [Bildiri Sunumu]. Eğitimde İyi Örnekler Konferansı, Sabancı Üniversitesi, İstanbul, Türkiye.
- Kutlu, Ö., Yalçın, S., & Pehlivan, E. B. (2010). İlköğretim programında yer alan kazanımlara dayalı soru yazma ve puanlama çalışması. *İlköğretim Online*, 9(3), 1201-1215. <http://ilkogretim-online.org.tr/index.php/io/issue/view/119>
- Kutlu, Ö., Doğan, C. D., & Karakaya, İ. (2015). *Ölçme ve değerlendirme: Performansa ve portfolyoya dayalı durum belirleme* (5. Baskı). Pegem Akademi.
- Kutlu Ö., & Kula-Kartal, S. (2018). The prominent student competences of the 21st century education and the transformation of classroom assessment. *International Journal of Progressive Education* 14(6). 70-82. <https://doi:10.29329/ijpe.2018.179.6>
- Kyllonen, P. C. (2012, May 7-8). *Measurement of 21st century skills within the common core state standards* [Paper presentation]. Invitational Research Symposium on Technology Enhanced Assessment. K12 Centre at ETS. <https://www.ets.org/Media/Research/pdf/session5-kyllonen-paper-tea2012.pdf>
- Lindenfield, G. (2014a). *Emotional confidence: Simple steps to build your confidence*. Harper Thorsons.
- Lindenfield, G. (2014b). *Assert yourself: Simple steps to build your confidence*. Harper Thorsons.
- Lord F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Addison-Wesley Publishing Company.
- Lewin, R. (1998). *Modern insanın kökeni*. TÜBİTAK Yayınları.
- Marzano, R. J. (1992). *A different kind of classroom: Teaching with dimensions of learning*. Alexandria: Association for Supervision and Curriculum Development.
- Marzano, R. J., & Kendall, J. S. (2007). *The new taxonomy of educational objectives*. Corwin Press.
- Marzano, R. J., & Heflebower, T. (2012). *Teaching and assessing 21st century skills*. Marzano Research.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.). *Educational measurement* (3rd Ed.). American Council on Education and Macmillan Publishing Company.

- Murphy, K. R., & Davidshofer, C. O. (1998). *Psychological testing* (4th Ed.). Prentice-Hall.
- NRC. (2011). *Assessing 21st century skills: Summary of a workshop*. The National Academies Press. <https://doi.org/10.17226/13215>
- Nitko, A. J., & Brookhart, S. M. (2007). *Educational assessment of students* (5th Ed.). Pearson Education.
- NCLB. (2001). *The no child left behind act of 2001*. <https://www2.ed.gov/nclb/overview/intro/execsumm.pdf>
- Nunnally, J. C. (1967). *Psychometric theory*. McGraw-Hill.
- OECD. (2013). *PISA 2012 assessment and analytical framework: Mathematics, reading, science, problem solving and financial literacy*. OECD Publishing. <http://dx.doi.org/10.1787/9789264190511-en>
- OECD. (2018). *The future of education and skills: Education 2030*. OECD Publishing. <https://observatorioeducacion.org/sites/default/files/oecd-education-2030-position-paper.pdf>
- OECD. (2019). *OECD future of education and skills 2030: OECD learning compass 2030*. A Series of Concept Notes. OECD Publishing. https://www.oecd.org/education/2030-project/contact/OECD_Learning_Compass_2030_Concept_Note_Series.pdf
- Öner, N. (1997). *Türkiye’de kullanılan psikolojik testler*. Boğaziçi Üniversitesi Yayınları.
- Özçelik, D. A. (1989). *Test hazırlama kılavuzu* (2. Baskı). ÖSYM Eğitim Yayınları 5.
- Özçelik, D. A. (1992-2). *Ölçme ve değerlendirme*. ÖSYM Yayınları.
- Payne, D., & Kyllonen, P. C. (2012, January 11-13). *The role of noncognitive skills in academic success*. [Paper presentation]. 21st Century Knowledge and Skills: The New Curriculum and the Future of Assessment Center for Enrollment Research, Policy, and Practice (CERPP). Marriott Downtown, Los Angeles, United States. <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.418.8555&rep=rep1&type=pdf>
- Pearson, K. (1904). *On the theory contingency and its relation to association and normal correlation*. University of London.
- Pellegrino, J. W. (1999). *The evolution of educational assessment: Considering the past and imagining the future*. 6th William H. Angoff Memorial Lecture Series. Educational Testing Service.
- Popham, J. W. (2000). *Modern educational measurement: Practical guidelines for educational leaders* (3rd Ed.). Allyn and Bacon.
- Popham J. W. (2004). *America’s “failing” schools: How parents and teachers can cope with no child left behind*. Routledge.
- Popham, J. W. (2007). Formative assessment: False pathway to proficiency? Properly implemented formative assessments can improve student learning, even if they don’t measure proficiency. *Leadership Compass*, 4(3), 1-2.
- Raban, S. (2008). Introduction. In S. Raban (Ed.). *Examining the world: A history of the university of Cambridge local examinations syndicate* (pp.1-11). Cambridge University Press.
- Reckase, M. D. (2009). *Multidimensional item response theory*. Springer.
- Ronan, C. A. (2003). *Bilim tarihi: Dünya kültürlerinde bilimin tarihi ve gelişmesi*. TÜBİTAK Yayınları.
- Schultz, D. P., & Schultz, S. E. (2011). *A history of modern psychology* (10th Ed.). Wadsworth, Cengage Learning.
- Schunk, D. H. (2012). *Learning theories: An educational perspective* (6th Ed.). Pearson Education.
- Sireci, S. G. (1998). The construct of content validity. *Social Indicators Research*, 45, 83-117. <https://doi.org/10.1023/A:1006985528729>
- Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1), 72-101. <https://doi.org/10.2307/1412159>
- Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, 83(10), 758-765. <https://doi.org/10.1177/003172170208301010>
- Thorne, B. M., & Henley, T. (2001). *Connections in the history and systems of psychology*. Houghton Mifflin.
- Thorndike, R. L. (1971). *Educational measurement* (2nd Ed.). American Council on Education.
- Thorndike, R. L. (1982). *Applied psychometrics*. Houghton Mifflin Company.
- Thorndike R. M., & Thorndike-Christ T. (2014). *Measurement and evaluation in psychology and education* (8th Ed.). Pearson Education Limited.

- Turgut, M. F., & Baykul Y. (2012). *Eğitimde ölçme ve değerlendirme* (4. Baskı). Pegem Akademi.
- UNDP., & UNICEF. (1990, 5-9 March). *World declaration on education for all and framework for action to meet basic learning needs*. World Conference on Education for All: Meeting Basic Learning Needs. Jomtien, Thailand. <https://unesdoc.unesco.org/ark:/48223/pf0000097551>
- UNESCO. (2002). *Dünya rayında mı? Herkes için eğitim (EFA) küresel izleme raporu*. https://www.unesco.org.tr/Content_Files/Content/Sektor/Egitim/EFA%202002.pdf
- UNESCO. (2013-4). *Öğretme ve öğrenme: Herkes için kaliteli eğitimi gerçekleştirmek*. Herkes için eğitim (EFA) küresel izleme raporu. https://www.unesco.org.tr/Content_Files/Content/Sektor/Egitim/EFA2013-4.pdf
- Wagner, T. (2015). *21st century schools*. <https://www.21stcenturyschools.com/tony-wagner.html>
- Wootton, D. (2019). *Bilimin icadı: Bilim Devrimi'nin yeni bir tarihi*. Yapı Kredi Yayınları.

EKLER

Ölçme ve Durum Belirleme Örneği için Açıklama

Aşağıda bir ders kapsamının; bilişsel, içsel ve kişilerarası yeterlik alanlarına ait becerilerle nasıl ilişkilendirileceği ve nasıl ölçüleceğine ilişkin bir ölçme ve durum belirleme örneği verilmiştir. Bunun için öncelikle *öğrenme çıktısı* tanımlanmıştır. Daha sonra bu öğrenme çıktısını kapsayacak bir örnek olay (durum) sunulmuş ve her üç yeterlik alanıyla ilişkili olan beceriler tanımlanarak Tablo 1’de verilmiştir. Bu örneğin, öğretim programlarının yeniden düzenlenmesi sürecine de ışık tutacağı düşünülmektedir. Öğretim programlarının kazanımlar yerine öğrenme çıktıları biçiminde düzenlenmesi, ders kapsamıyla ilgili öğretim etkinliklerinin bu örnekte belirtildiği gibi çoklu etkileşime sahip bir yapıda işlenmesine ve öğrencilerin bilişsel, içsel ve kişilerarası becerilerini aynı anda kullanarak gelişmelerine katkı sağlayacaktır.

Bu çalışmada öğrenme çıktısı, bir dersin kapsamıyla ilişkili olan bilişsel, içsel ve kişilerarası becerilerin, kapsamla ilişkilendirilerek öğrencilere kazandırılmasını ve öğrenmelerin farklı becerilerle zenginleştirilmesini ifade etmektedir. Bu ifade günümüz öğretim programlarında yer alan “kazanım” ifadesinden çok daha fazlasını içermektedir. Kazanımlar çoğunlukla bir içeriğin (kapsam) bir beceriyle olan ilişkisini belirtmektedir (Kutlu vd., 2010). Öğrenme çıktısı, öğrencinin bir kapsama bağlı bilişsel, içsel ve kişilerarası becerileri aynı görevde kullanarak bir ürün ortaya koymasına biçiminde olabileceği gibi (Kutlu ve Kula-Kartal, 2018), bu çalışmada olduğu gibi bilişsel, içsel ve kişilerarası becerileri bir durumla ilişkilendirerek birbirinden bağımsız olarak kullanması biçiminde de olacaktır.

Bu ölçme ve durum belirleme yaklaşımıyla hem aynı kapsama ilişkin birden fazla özellik ölçülerek içtutarlılık (güvenirlilik), hem de maddelerin ölçülmek istenen alanın iyi bir örneklemini vermesiyle kapsam geçerliği sağlanmaya çalışılmıştır.

Ölçme ve Durum Belirleme Örneği

Ders : Sosyal Bilgiler
Düzey : İlkokul 4. Sınıf
Ünite : Haklarımız ve Sorumluluklarımız

Öğrenme Çıktısı:

Karşılaştığı sorunlarda haklarını ve sorumluluklarını bilen davranışlar gösterir.

Tablo 1.

Öğrenme çıktısına ulaşmak için 21. yüzyıl yeterlik alanlarına ilişkin ölçülmesi gereken beceriler

21. Yüzyıl Yeterlik Alanları		
Bilişsel Beceriler	İçsel Beceriler	Kişilerarası Beceriler
<p>Hak ve sorumluluk kavramlarını ayırt eder.</p> <p><i>Ayırt etme:</i> Bir olay ya da bir durum karşısında gösterilen tepkinin bunun dışındaki hiçbir olay ya da durum karşısında gösterilmemesidir. Bireyin iki uyarıcının benzer olmayan özelliklerini fark etmesidir.</p>	<p>Beklenmedik durum karşısında, hak ve sorumluluklarını kullanarak etkili bir davranış gösterir.</p> <p><i>Uyumlu olma:</i> Belirsiz durumlar ya da problem durumları karşısında etkili biçimde tepki vermesidir.</p>	<p>Haklarına ve sorumluluklarına sahip çıkarken iletişim diline dikkat eder (saygılı olur).</p> <p><i>Savunmacı iletişim:</i> Duyularını, tercihlerini, gereksinimlerini ve görüşlerini tehdit edici ya da diğer kişileri cezalandırıcı olmayacak bir biçimde ifade etmesidir.</p> <p><i>Saygılı olma:</i> Saygı, karşısındaki insanın varlığını kabullenmesi ve onun düşüncelerini kabul etmese bile onun düşüncelerinin dinlenmeye değer olduğunu ona hissettirmesidir.</p>
<p>Hak ve sorumlulukla ilgili bilgilerini karşılaştığı sorunların çözümünde kullanır.</p> <p><i>Problem çözme (problemi tanımlama):</i> Bireyin bir hedefe ulaşmak için topladığı gücün karşısına çıkan engellerle baş edecek uygun yöntemleri geliştirmesi ve bu yöntemleri kullanarak başarıya ulaşmasıdır.</p>	<p>Hak ve sorumluluğunu kullanırken kendine güvenir.</p> <p><i>Özgüvenli olma:</i> Bireyin korku, çekinme ve kuşku duymadan kendisine inanmasıdır.</p>	<p>Görüşlerini savunurken ikna edici (etkileyici) bir dil kullanır.</p> <p><i>Sosyal etki:</i> Kişileri belirli davranışları, inançları ya da tutumları benimsemeye yönlendirmesidir.</p> <p><i>İkna etme:</i> Birisini bir konuya her yönüyle inandırması ya da konuya inanmaya uygun duruma getirmesidir.</p>

Örnek Duruma Dayalı Maddeler

Öğrenme çıktısına ulaşmak için üç yeterlik alanında sekiz beceri tanımlanarak Tablo 1’de sunulmuştur. Bu becerilerden ikisi *bilişsel*, ikisi *içsel* ve dördü *kişilerarası* yeterlik alanına aittir. Öğretmenlerin, öğretim sürecinde bu becerilerin tümünü ölçen maddelere yer vermeleri yerinde olacaktır. Bu örnek duruma dayanan 10 madde hazırlanmış olmasına rağmen çalışma kapsamında, yalnızca dört madde örneğine yer verilmiştir.

Örnek Durum

Ege, marketten paketli olarak aldığı bir kutu beyaz peyniri eve geldiğinde açmış ve peynirin tadında, kokusunda ve renginde bozulmalar olduğunu görmüştür. Ürünü zaman kaybetmeden markete geri götürmüştür. Market yetkilisine ürünü göstererek durumu anlatmıştır. Ancak market yetkilisi paketin açılmış olduğunu ve paketi açılan ürünü geri alamayacaklarını ifade etmiştir.

1. madde

Her bireyin bir tüketici olarak bilmesi gereken hakları ve sorumlulukları vardır.

Bir tüketici olarak Ege’nin hakları ve sorumlulukları nelerdir? Bu olayla ilişkilendirerek yazınız.

Birinci madde Tablo 1'deki öğrenme çıktılarında verilen *bilişsel* yeterliklerden *ayırt etme* becerisiyle ilişkilidir. Bu maddeyle öğrenciden bir tüketici olarak hak ve sorumluluk kavramlarını bilmesi, bu iki kavramı aynı olayda (tüketim, alışveriş yapma) ilişkilendirerek birbirinden ayırt etmesi beklenmektedir.

Maddeye verilecek öğrenci yanıtları puanlanırken öncelikle, kavramların doğru anlamda kullanılıp kullanılmadığının belirlenmesi önemlidir. Bunun yanında, bu maddede öğrenciden yalnızca kavramları anlam düzeyinde öğrenmesi değil; bu kavramları karşılaşılan bir olayda doğru ve yerinde kullanması beklenmektedir. Anderson ve Krathwol (2001) tarafından geliştirilen Bloom'un güncellenmiş bilişsel sınıflamasında ayırt etme becerisi, öğrencinin verilen kavramları çözümlenerek bu kavramların farklılıklarını ortaya koymasını gerektirmektedir. Bu nedenle öğrenci yanıtlarında, kavramların anlamca bilinmesinin dışında; farklılıklarının doğru ve yerinde kullanılacak biçimde ayırt edilmiş olması da önemlidir. Bu noktada geribildirim ilişkisel düşünmeyi gerektiren bir ayırt etme becerisine odaklanması ve bu davranışı istenen düzeyde gösteremeyen öğrenciye bu becerinin nasıl kullanılacağıyla ilgili ipuçlarının sunulması gerektirmektedir.

2. madde

Bu olayda Ege yaptığı tüketimle ilgili bir sorunla karşılaşmıştır.

Bu sorunu tanımlayınız. Ege'nin bu sorunla ilgili hakları ve sorumlulukları nelerdir? Olayla ilişkilendirerek yazınız.

İkinci madde Tablo 1'deki öğrenme çıktılarında verilen *bilişsel* yeterliklerden problem çözmenin bir alt basmağı olan *problemi tanımlama* becerisiyle ilişkilidir. Bu maddede öğrenciden öncelikle sorunu (problemi) belirleyerek tanımlaması istenmektedir. Daha sonra öğrenciden bu soruna bilinçli bir tüketici olarak yaklaşması, hakları ve sorumlulukları çerçevesinde bir çözüm ileri sürmesi beklenmektedir.

Nitko ve Brookhart'a (2007) göre, problem çözmeye önemli olan, istenen hedefe nasıl ulaşılabileceğidir. Bir öğrenci istenen hedefe ulaşmanın doğru yolunu otomatik olarak bilmediği için öğrenciden bir ya da daha fazla düşünme becerisini kullanması beklenmektedir. Bu süreç problem çözme olarak adlandırılmaktadır.

Problem çözme süreci öncelikle, problemin çözümü için gerekli bilgilerin öğrenilmiş ve anlamlandırılmış olmasını gerektirir (Bransford ve Stein, 1984). Bu maddeye verilecek öğrenci yanıtları puanlanırken öncelikle, olayda sözü edilen sorunun tanımlanması gerekmektedir. Bir sorunun tanımlanması problem çözme adı verilen üst düzey düşünme beceri alanının öncelikli basamaklarından birini oluşturmaktadır. Problemi doğru tanımlamak, problemin doğru çözümü için önemlidir. Bir birey problemi doğru tanımlayamadığı zaman çözüm yolu ve bu çözümle ulaşılan sonuç yanlış olacaktır.

Haladyna (1997) problem çözmenin yalnızca zihinsel adımlardan oluşmadığını, aynı zamanda bireyin eyleme geçmesini gerektiren fiziksel adımları da içerdiğini belirtmektedir. Bu anlamda problem çözme düzeyinde bir maddenin süreçle ilgili bilgileri içermesi ve çözüme katkı sağlayacak fikirler hakkında ipuçları vermesi gerekir.

Anderson ve Krathwol (2001) uygulama basamağını, bir problemin çözümünde kullanılan işlem adımlarını içerdiğini, probleme ilişkin çözüm sürecinin belirlenebilmesi için problemin ilgili duruma uyarlanması gerektiğini belirtmektedir. PISA 2012 çerçevesinde, karşılaşılan bir problemin çözülebilmesi için ilk aşama problemin içeriğiyle ilgili bilgilerin bellekten çağrılmasıdır. Bu bilgiler bireyin uzun süreli belleğinde yoksa hem problemin anlaşılması hem de problemin tanımlanması ve çözümlenmesi olanaklı olmayacaktır. Bu durum öğrencinin karşılaştığı problem durumunda hızlı bir biçimde problemle ilgili değişkenlerin birbiriyle olan ilişkilerinin keşfedilmesini gerektirmektedir (OECD, 2013). Arka planda öğrencinin bu donanıma yeterince sahip olmaması karşılaşılan sorunun anlaşılmasının önünde bir engel oluşturacaktır. Öğrenci bu donanıma sahipse ancak o zaman bundan sonra sorunu tanımlayabilecek ve çözüme odaklanabilecektir. Bu nedenle çözümün doğru ve başarılı olabilmesi öncelikle sorunun doğru tanımlanmasına bağlıdır.

Bu kapsamda öğretmenlerin öğrenci yanıtlarında öncelikle, temel bilgilerin var olup olmadığına bakmaları, sonra problemin doğru tanımlanıp tanımlanmadığını belirlemeleri, daha sonra tanım ile çözüm arasındaki ilişkilerin doğruluğunu kontrol etmeleri yerinde olacaktır. Geribildirim bu ilişkilerden çıkan sonuçlara bağlı olarak oluşturulmalıdır.

3. madde

Market yetkilisi paketin açılmış olduğunu ve ürünü geri alamayacaklarını ifade etmiştir.

Ege, market yetkilisiyle başlangıçta nasıl bir konuşma yapsaydı sorunun çözümüne katkı sağlamış olurdu?

Üçüncü madde, Tablo 1’de verilen *kişilerarası yeterliklerden savunmacı iletişim ve ikna etme* becerilerinin birlikte kullanımıyla ilgilidir. Savunmacı iletişim bireyin duygularını, tercihlerini, gereksinimlerini ve görüşlerini diğer bireyleri tehdit edici ya da cezalandırıcı olmayacak biçimde ifade etmesidir [National Research Council (NRC), 2011]. Bu beceri, bireyin farklı görüşler öne sürmesini, kendine güvenmesini, girişken olmasını, haklarını ve gereksinimlerini savunmasını gerektirir. Marzano ve Heflebower (2012) savunmacı davranışı, bir ilişkinin her iki tarafın da kazanmasıyla sonuçlandığı (kazan-kazan) bir davranış türü olarak tanımlamaktadır. Bu davranışa sahip bireyler, başkalarına zarar vermeden ya da saygısızlık etmeden, kendilerini ve gereksinimlerini açıkça ifade ederler. Savunmacı davranış, karşılıklı saygıyı ve birbirini önemsemeyi özendirir. İkna etme becerisi ise bireylerin görüşlerini savunurken etkileyici bir dil kullanmasıdır. İkna etme karşındaki kişiyi bir konuya her yönüyle inandırmak ya da inanmaya uygun duruma getirmek, kendi düşüncemizin tarafına çekmektir. NRC (2011) raporunda bu süreçte bireyin konumunu savunması, açık fikirli olması, akılcı düşünceler ileri sürmesi, olumsuz durumları bireysel olarak algılamaması, yargılardan kaçınması ve mantıklı bakış açılarını bütünleştiren bir yaklaşım sergilemesinin önemi vurgulanmaktadır.

Bu maddede öğrenciden olumsuz bir durum karşısında ve haklı olduğu bir konuda karşısındaki bireyi inandıracak davranışı gösteren bir yanıt yazması beklenmektedir. Yukarıdaki açıklamaların ışığında öğrenci yanıtında olumlu vurguların, inandırıcı fikirlerin, etkileyici sözcük kullanımının gerçekleşip gerçekleşmediğine bakılmalıdır.

4. madde

Beklemediği bir durumla karşılaşan Ege, market yetkilisiyle tartışmak istememiş ve evine geri dönmüştür. Annesi, sorunun çözümü için Ege’nin kendine güvenen (özgüvenli) bir davranış sergilemesi gerektiğini belirtmiştir.

Buna göre Ege, markette hangi davranışı sergilemiş olsaydı özgüvenli bir birey davranışı göstermiş olurdu? Yanıtınızı bu olayla ilişkilendirerek yazınız.

Dördüncü madde, Tablo 1’deki öğrenme çıktılarında verilen *içsel yeterliklerden özgüvenli olma* becerisiyle ilgilidir. İçsel yeterlikler, bireylerin bilişsel süreçlerini denetleyen bir mekanizma oluştururlar. Flavell (1979), bu süreci bilişötesi (metacognition) olarak adlandırmıştır. Ona göre bilişötesi, bireyin kendi bilişsel süreçleri hakkındaki bilgilerini anlamasına, denetlemesine ve bilişsel süreçlerini harekete geçirmesine katkı sağlar. Brown (1978) ise, bilişötesi kavramını öğrencilerin öğrenme sırasında ya da problem çözme durumunda kullandıkları kendi düşünme süreçlerine ilişkin farkındalık düzeyi olarak tanımlamıştır. Tanımlardan da görüldüğü gibi, karşılaşılan bir problem durumunda bireyin, bilişötesi süreçlerini devreye sokabilmesi için hem problemin içeriğiyle ilgili bilgilere (hak ve sorumluluk kavramları hakkındaki bilgiler) hem de problemin çözülmesiyle ilgili becerilere (problemi tanımlama, probleme çözüm önerme) sahip olması gerekmektedir. Bu iki ön donanımın olmadığı durumda birey, kendine güvenmeyerek hakkını savunmaktan vazgeçmekte ya da daha saldırganca bir davranış gösterebilmektedir.

Özgüven bireyin kendini tanıması, değerinin bilincinde olması, başarabilme ve yapabilme duygusunu içerir. Özgüvene sahip bireyler kendi haklarını bilirler ve savunurlar; aynı zamanda karşısındaki kişinin de haklarını bilirler, suçluluk duygusuna kapılmazlar, zayıf ve güçlü yönlerini bilerek başarısızlığa karşı bir yaklaşım izlerler (Lindenfield, 2014a, 2014b).

Bu maddede öğrenciden bir tüketici olarak haklarından vazgeçmek yerine haklarının bilincinde olarak özgüvenli bir davranış içeren yanıt yazması beklenmektedir. Bunun yanında öğrenciden hak ve sorumluluk kavramlarını bildiğini gösteren cümleler kurması, problemin farkında olduğuna işaret eden anlatımlara yer vermesi ve kendine güvenen bir yaklaşım içinde haklarını savunan ifadeler kullanması gerekmektedir. Öğretmen geribildiriminin bu üç duruma dayalı olarak yapılandırılması önemlidir.

Öğretmenler öğrencilerden özellikle, bilişsel yeterlik alanıyla ilgili bilgileri ve becerileri yerinde ve doğru kullandıklarını görmeleri ve bir içsel yeterlik becerisi olan özgüven kavramıyla doğru ilişkilerin kurulup kurulmadığını belirlemeleri gerekmektedir.

Maddelerin Puanlanması

Dereceli puanlama anahtarları öğrencilere yaptıkları çalışmaların hangi ölçütlere göre değerlendirileceğini ve performanslarını hani düzeydeki puana denk geleceğini gösteren puanlama araçlarıdır. Bu sayede öğrenciler puanlama sonunda buldukları başarı durumunu ve ulaşmaları gereken başarı düzeyini görebilmektedirler (Kutlu vd., 2015). Bu çalışmada örnek olması açısından yalnızca 2. ve 3. maddeye ilişkin dereceli puanlama anahtarları sırasıyla Tablo 2 ve 3'te verilmiştir.

Her madde 0-10 arası puanlanmış; 10 puan öğrencinin ölçülen özellik açısından istenen yanıt eksiksiz verdiğini, 9-2 arası puanlar yanıtların kısmi doğruluk içerdiğini ve 10 puandan uzaklaştıkça temel bilgi ve becerilerde eksikliklerin arttığını göstermektedir. 0 puan olan boş yanıt öğrencinin maddenin ölçtüğü özellik konusunda bilgisinin olmadığını, 1 puan olan yanlış ve ilişkisiz yanıtları ifade etmektedir. Yanlış yanıtlar kendi içinde doğru, ancak ölçülen özelliğin doğrusu olmayan ifadeyi; ilişkisiz yanıtlar ise ölçülen özellik ile ilgili olmayan uydurma, saçma ifadeleri temsil etmektedir.

2. madde

Bu olayda Ege'nin yaptığı tüketimle ilgili karşılaştığı sorunu tanımlayınız.

Ege'nin bu sorunla ilgili hakları ve sorumlulukları nelerdir? Bu olayla ilişkilendirerek yazınız.

Tablo 2.

İkinci maddeye ait dereceli puanlama anahtarı

Yanıtlar	Başarı Puanı
En Doğru Yanıt	
Öğrenci, Ege'nin karşılaştığı sorunu tanımlar. Bu soruna haklar ve sorumluluklar çerçevesinde bulunduğu çözümü olayla ilişkilendirerek yazar. Örnek Yanıt: Ege'nin marketten aldığı peynir bozuk çıkmıştır. Ege bir tüketici olarak aldığı bozuk ürünü değiştirme ve iade etme hakkına sahiptir. Eğer satıcı ürünü değiştirmezse yasal yolları kullanarak tüketici hakları mahkemesine başvurma hakkı da vardır. Ege'nin aldığı ürünün son kullanma tarihine bakması, ambalajının sağlamlığını kontrol etmesi, fiş ya da fatura alması sorumluluğudur.	10
Uzak Doğru Yanıtlar	
1. Yanıt: Öğrenci, Ege'nin karşılaştığı sorunu tanımlar. Yaşadığı olaya çözüm önerirken haklarının tamamını sorumluluklarından yalnızca birini yazar.	
2. Yanıt: Öğrenci Ege'nin karşılaştığı sorunu tanımlar. Yaşadığı olaya çözüm önerirken sorumluluklarının tamamını haklarından yalnızca birini yazar. Örnek Yanıt: Öğrenci, Ege'nin yaşadığı olaya çözüm önerirken sorumluluklarının tamamını haklarından yalnızca birini yazar.	8
Örnek Yanıt: Öğrenci, Ege'nin yaşadığı olaya çözüm önerirken haklarından ve sorumluluklarından birer tanesini yazar.	6
Örnek Yanıt: Öğrenci, Ege'nin yaşadığı olaya çözüm önerirken haklarından ya da sorumluluklarından birini yazar.	4
1. Yanıt: Öğrenci, Ege'nin yaşadığı olaya çözüm önerirken haklarından ya da sorumluluklarından birini yazar.	
2. Yanıt: Öğrenci yalnızca, Ege'nin karşılaştığı sorunu tanımlar. Örnek Yanıt:	2
Boş Yanıt	0
Yanlış Yanıtlar	
1. Yanıt: Öğrenci, Ege'nin karşılaştığı sorunla ilgili, kendi içinde doğru ancak bu maddenin ölçtüğü özellik ile ilgili olmayan bir yanıt yazar.	
2. Yanıt: Öğrenci, Ege'nin haklarını ve sorumluluklarını ters yazar. Örnek Yanıt: Ege market yetkilisiyle tartışmıştır. Ege'nin bozuk ürünü değiştirmesi, değişimi yapamadığı durumda yasal yollara başvurması sorumluluğudur. Ürünün son kullanma tarihine bakması, fiş alması hakkıdır.	1
İlişkisiz Yanıtlar	
Öğrenci, Ege'nin yaşadığı olaya çözüm önerirken haklarını ve sorumluluklarıyla ilişkili olmayan	1

uydurma bir yanıt yazar.

Örnek Yanıt:

Ben olsam bir daha o markete gitmem.

3. madde

Market yetkilisi paketin açılmış olduğunu ve ürünü geri alamayacaklarını ifade etmiştir.

Ege, market yetkilisiyle başlangıçta nasıl bir konuşma yapsaydı, sorunun çözümüne katkı sağlamış olurdu?

Tablo 3.

Üçüncü maddeye ait dereceli puanlama anahtarı

Yanıtlar		Başarı
En Doğru Yanıt		Puanı
Öğrenci, sorunun çözümüne katkı getirecek (1) selamlaşma ifadesi kullanan, (2) sorunu ayrıntılı anlatan, (3) ürünü alırken kontrol ettiğini belirten, (4) etkileyici, inandırıcı ve nezaket sözcükleri kullanan olumlu bir konuşma örneği yazar.		
Örnek Yanıt: İyi günler, kolay gelsin. Dün sizden aldığım bir paket peynirle ilgili bir sorunun var, uygunsanız konuşabilir miyiz? Peyniri az önce açtım ancak peynir bozulmuş. Alırken paketi ve son kullanma tarihini kontrol etmişim. Sizin tüketici haklarına değer veren bir market olduğunuzu biliyorum. Eğer ürünü değiştirirseniz çok mutlu olurum. Konunun uzamadan, tüketici haklarına gidilmeden çözülmesini umut ediyorum. Size şimdiden teşekkür ederim.		10
Uzak Doğru Yanıtlar		
Öğrenci, sorunun çözümüne katkı getirecek 4 durumdan 3'ünü kullanan olumlu bir konuşma örneği yazar.		8
Örnek Yanıt: Öğrenci, sorunun çözümüne katkı getirecek 4 durumdan 2'sini kullanan olumlu bir konuşma örneği yazar.		5
Örnek Yanıt: Öğrenci, sorunun çözümüne katkı getirecek 4 durumdan 1'ini kullanan olumlu bir konuşma örneği yazar.		3
Boş Yanıt		0
Yanlış Yanıtlar		
Öğrenci, sorunun çözümüne katkı getirmeyen olumsuz bir konuşma örneği yazar.		
Örnek Yanıt: Bir yetkiliyle görüşebilir miyim? Marketiniz ne yazık ki bozulmuş ürünleri satıyor. Aldığım peynir bozuk çıktı, değiştirmezseniz sizi zor durumda bırakırım, tüketici hakları mahkemesine giderim.		1
İlişkisiz Yanıtlar		
Öğrenci, sorunun çözümüyle ilişkili konuşma örneği yazar.		
Örnek Yanıt: Peki siz bu ürünü kaç senedir satıyorsunuz?		1