

Milli Eğitim Bakanlığı Tarafından Yayınlanan Test Maddelerinin Madde Yazım İlkelerine Uygunluğunun İncelenmesi

Eren Can Aybek^a, Seval Kula Kartal^b ve Özen Yıldırım^c

Öz

Araştırmada Milli Eğitim Bakanlığı (MEB) Ölçme, Değerlendirme ve Sınav Hizmetleri Müdürlüğü tarafından yedinci sınıf düzeyi için sunulan maddelerin ölçme ve değerlendirme ilkelerine uygunluğunun incelenmesi amaçlanmıştır. Bu doğrultuda fen bilgisi ve sosyal bilgiler derslerinden tesadüfi olarak seçilen üçer üniteden toplam 120 madde incelenmiştir. Maddelerin incelenmesinde alanyazın taraması sonucunda araştırmacılar tarafından geliştirilen madde inceleme formu kullanılmıştır. Verilerin analizinde betimsel istatistiklerden yararlanılmıştır. Araştırma sonucunda maddelerde doğru yanıtı yönlendirecek ipuçlarının bulunduğu, maddeyi yanıtlamak için gerekli olmayan metin ve görsellere yer verildiği, bileşik yanıt vermeyi gerektiren (I, II, III, IV gibi) maddelerin sıkça tercih edildiği belirlenmiştir. Bununla birlikte, maddelerde dilbilgisi kurallarına uyulduğu, ana fikrin seçeneklerde değil madde kökünde verildiği, maddelerin belirli bir görüşe ya da kanıya dayalı olmadığı, yönergelerin açık ve anlaşılır olduğu gözlemlenmiştir.

Anahtar Kelimeler: Çoktan seçmeli madde yazım ilkeleri, ölçme ve değerlendirme, test geliştirme

Makale Hakkında

Gönderim tarihi: 30.10.2021
Düzeltilme tarihi: 16.06.2022
Kabul tarihi: 08.09.2022
Elektronik Yayın Tarihi:
31.12.2022

Giriş

Başarılı ve nitelikli insan gücü potansiyeli bir toplumun kalkınmasında önemli bir etkidir. Bu nedenle öğrencilerin başarılı olmaları kendileri, aileleri ve içinde yaşadıkları toplum açısından büyük önem taşımaktadır (Bangchang, 2015). Güncel yaklaşımda başarı kavramı, bireyin bilişsel, duyuşsal ve psikomotor alandaki becerilerini geliştirmeyi içerecek biçimde genişlemiştir. Başarının bu tanımı gereği, okulların bireylerin bilgili, sorumlu ve içinde yaşadığı toplumla ilgilenen yetişkinler olarak yetişmesine katkı sağlamayı amaçlaması gerekmektedir (Elias vd., 1997).

Bireylerin bilişsel gelişimlerine katkı sağlamak için okullar öğrencilerin hem alt düzey hem de üst düzey zihinsel becerilerini geliştirmek durumundadır (Haladyna,

^aSorumlu yazar, Pamukkale Üniversitesi, Eğitimde Ölçme ve Değerlendirme Anabilim Dalı, erencan@aybek.net, ORCID: 0000-0003-3040-2337

^bPamukkale Üniversitesi, Eğitimde Ölçme ve Değerlendirme Anabilim Dalı, sevalk@pau.edu.tr, ORCID: 0000-0002-3018-6972

^cPamukkale Üniversitesi, Eğitimde Ölçme ve Değerlendirme Anabilim Dalı, ozenyildirim@pau.edu.tr, ORCID: 0000-0003-2098-285X

1997; Kutlu vd., 2014). Böyle bir gelişim, öğrencilere planlı ve sistematik bir öğretim sürecinin sunulmasını ve süreç içerisindeki gelişmelerinin izlenmesini gerekli kılmaktadır. Eğitimde, bir ders, ünite ya da dönem süresince öğrencilerin bilgi ve becerilerinde meydana gelen değişimleri gözlemek amacıyla başarı testleri yaygın biçimde kullanılmaktadır (Cohen ve Swerdlik, 2010). Anastasi (1961) başarı testlerinin, öğretim programlarının değerlendirilmesi ve iyileştirilmesi, eğitim sürecinin ve bu sürece ait çıktıların değerlendirilmesi gibi çok çeşitli amaçlarla kullanıldığını ifade etmektedir.

Başarı testlerinin psikometrik özelliklerinin yüksek olabilmesi için bireyler hakkında doğru ve yeterli bilgiler sağlanması gerekmektedir (Adkins, 1947). Ölçme araçlarının ölçülen özellik hakkında doğru bilgiler vermesi ise nasıl bir geliştirilme süreci sonucunda oluştuğuna bağlıdır. İyi bir test, test geliştirme ilke ve kurallarının dikkatli ve özenli bir biçimde uygulanmasıyla ortaya çıkmaktadır (Cohen ve Swerdlik, 2010). Psikometrik açıdan nitelikli bir test geliştirmek için önemli olan aşamalardan biri de maddelerin alan yazında yer alan madde yazım ilkelerine göre yazılmasıdır. Test maddelerinin psikometrik nitelikleri, başarı testinin psikometrik niteliği açısından belirleyici olduğundan, uzmanlar tarafından madde yazım ilkeleri üzerinde önemle durulmaktadır.

Hem sınıf içi hem de geniş ölçekli test uygulamalarında en sık kullanılan madde türlerinden biri çoktan seçmeli madde türüdür (Cheung, 2003). Diğer madde türlerine göre daha yaygın olarak kullanıldığından, bu madde türünün incelenmesi üzerinde daha fazla sayıda çalışma yapılmıştır. Örneğin, Haladyna ve Downing (1989) çoktan seçmeli madde geliştirme sürecine ilişkin bilgi içeren 46 kitabı incelemiş ve sonucunda 43 adet çoktan seçmeli madde yazım ilkesi belirlenmiştir. Bu araştırmanın devamı niteliğindeki bir başka çalışmada, Haladyna vd. (2002) ilkeleri yeniden gözden geçirmiş ve 31 ilkeden oluşan yeni bir liste oluşturmuştur. Yapılan çalışmada madde yazım ilkeleri maddenin içeriği, türü, düzeni, maddenin kökü ve seçenekleri başlıkları altında düzenlenmiştir. Maddenin içeriği başlığı altında, test maddesiyle önemli bir öğrenme kazanımını ölçmeyi hedefleme, görüşe dayalı maddelerden kaçınma, özgün materyaller kullanma gibi ilkeler yer almaktadır. Madde türü başlığı altında kullanılması önerilen çoktan seçmeli madde tiplerine yer verilirken, düzen başlığı altında genel düzen, yazım ve dilbilgisi kurallarına değinilmektedir. Madde kökünün yazılması başlığı altında maddenin ne sorduğunun madde kökünde açık biçimde ifade edilmesi, madde kökünde olumsuz ifadelerden kaçınılması gibi ilkeler bulunmaktadır. Madde seçeneklerini yazma başlığı altında ise seçeneklerin içerik, dilbilgisi, uzunluk açısından benzer olması, anlamlı ve mantıklı çeldiricilerin yazılması gibi ilkelere yer verilmiştir.

Madde yazım süreçleri için oluşturulan ilke ve öneriler çeşitlilik gösterse de uzmanlar tarafından ortak biçimde vurgulanan bazı ilkeler bulunmaktadır. Bu ilkeler içerisinde, test maddesiyle önemli öğrenme kazanımlarını ölçmeyi hedefleme, anlamı belirsiz hale getiren ifadelerden kaçınma, grubun düzeyine uygun bir dil kullanma, maddenin doğru cevabı için ipucu oluşturacak nitelikteki ifadelerden kaçınma gibi öneriler yer almaktadır (Haladyna, 1997; Gronlund ve Waugh, 2009; Turgut ve Baykul, 2010; Popham, 2011; Nitko ve Brookhart, 2011).

Haladyna vd. (2002) de bu bağlamda çoktan seçmeli madde yazım ilkelerini ortaya koymuş ve bu ilkeler maddelerin madde yazım ilkelerine göre incelendiği gelecek araştırmalar için bir temel oluşturmuştur. Bu ilkeler geliştirilirken, görgül çalışmaların bulgularına ve çoktan seçmeli madde yazımına ilişkin alan yazındaki kaynaklarda yer alan görüşler kullanılmış ve birleştirilmiştir. Örneğin, Downing (2005) tarafından yapılan çalışmada, dört farklı testte yer alan maddelerin Haladyna vd. (2002) tarafından oluşturulan 31 ilkeye uygunluğu üç farklı uzman tarafından incelenmiştir. Araştırmada, 219 test maddesi içerisinde toplam 100 maddenin bu ilkelerden en az birini ihlal ettiği bulunmuştur. En çok ihlal edilen ilkelerin madde kökünde soru cümlesi kullanmama ve olumsuz ifade kullanma olduğu görülmüştür. Sireci vd. (1998) tarafından yapılan çalışmada, Haladyna ve Downing (1989) tarafından belirlenen ilkelere göre maddeler incelenmiş ve en sık yapılan ihlallerin madde kökünde soru cümlesi yerine tamamlanmamış ve bileşik çoktan seçmeli madde tipinin kullanılması olduğunu belirtilmiştir. Tarrant vd. (2006)'nin çalışmasında incelenen 2770 çoktan seçmeli test maddesinde en sık gözlenen ihlaller arasında madde kökünde açık olmayan ya da gereksiz bilgi ve yönerge, olumsuz madde kökü, mantıksız çeldirici kullanımı yer almaktadır.

Bu araştırmanın amacıyla benzer bir amaca sahip bir araştırma Şengül-Avşar ve Barış-Pekmezci (2020) tarafından yapılmıştır. Bu araştırmada, Milli Eğitim Bakanlığı tarafından yayınlanan Türkçe dersi için geliştirilmiş beceri temelli madde örnekleri incelenmiştir. Bu inceleme sonucunda araştırmacılar, bileşik yapıdaki çoktan seçmeli madde tipinin ve olumsuz madde kökünün sıklıkla kullanıldığını bulmuştur. Ayrıca, araştırmacılar sorularda kullanılan görsel ve yazılı materyallerin soruları yanıtlamak için gerekli olduğunu, ancak bu materyallerin etkili biçimde kullanılmadığını belirtmiştir.

Bu konuyla ilgili bazı araştırmalarda, madde yazımında gözlenen ihlallerin madde istatistikleri üzerindeki etkilerine odaklanılmıştır. Örneğin, Martinez vd. (2009) maddede birbirine içerik olarak benzer olmayan çeldiricilerin kullanılmasının maddenin ayırt ediciliği üzerinde küçük ancak olumsuz bir etkisi olduğunu göstermiştir. Ayrıca, madde ayırt ediciliğinin “yukarıdakilerin hiçbiri” seçeneğinin doğru yanıt olması durumunda da olumsuz etkilendiği bulunmuştur. Diğer çalışmalardan farklı olarak, araştırmacılar seçeneklerinin uzunluklarının denk olmaması, seçeneklerde “her zaman, asla” gibi ifadelerin kullanılması ve “yukarıdakilerin hepsi” seçeneğinin kullanılmasının madde güçlüğü ve ayırt ediciliğini olumsuz etkilemediğini bulmuştur. Ancak, Rush vd. (2016) tarafından yapılan bir çalışmada, “yukarıdakilerin hepsi” seçeneğinin kullanılmasının maddeyi kolaylaştırdığı ve ayırt ediciliğini düşürdüğü belirtilmiştir. Ayrıca, bu çalışmada mantıksız çeldiricilerin kullanılmasının da madde güçlüğü ve ayırt ediciliğini olumsuz etkilediği görülmüştür.

İlgili araştırmalar incelendiğinde, çoktan seçmeli madde türünün yazımına ilişkin belirli ilke ve önerilerin geliştirildiği görülmektedir (Haladyna ve Downing, 1989; Haladyna vd., 2002). Buna ek olarak, uygulamada hangi ilkelerin en sık ihlal edildiğini belirlemeyi amaçlayan araştırmalar yürütülmüştür (Downing, 2005; Sireci vd., 1998). Ayrıca, araştırmacılar tarafından ilkelerin geçerliğini belirlemek amacıyla bu

ihlallerin madde istatistikleri üzerindeki etkisini inceleyen görgül çalışmalar yapılmıştır (Martinez vd., 2009; Rush vd., 2016). Ancak, yurt içi alan yazın incelendiğinde hem sınıf içi hem de ulusal geniş ölçekli test uygulamalarında çoktan seçmeli madde geliştirmede hangi ilkelerin göz önünde bulundurulduğunu, hangi ilkelerin sıklıkla ihlal edildiğini inceleyen araştırmaların çok sınırlı olduğu görülmüştür (Şengül-Avşar ve Barış-Pekmezci, 2020).

Madde yazımı üzerinde yürütülen araştırmalar (Anıl ve Acar, 2008; Gelbal ve Kelecioğlu, 2007), öğretmenlerin ölçme araçlarını yapılandırmadaki bilgi eksikleri nedeniyle desteklenmeye ihtiyaç duyduklarını ortaya koymaktadır. Bu nedenle, Millî Eğitim Bakanlığı (MEB) tarafından son yıllarda öğretmenlere ölçme araçlarının sahip olması gereken özellikler, test maddelerinin geliştirilmesi gibi konularda seminerler düzenlenmekte ve öğretmenlerin ölçme ve değerlendirme alanındaki gelişimi desteklenmektedir. Seminerlerin sonrasında bakanlık tarafından eğitime katılan öğretmenlerden kendi konu alanlarında madde yazmaları istenmektedir. Öğretmenler tarafından geliştirilen maddeler içerisinden uygun görülerek seçilen maddeler bakanlığın internet sitesinde (<https://odsgm.meb.gov.tr/>) ücretsiz bir biçimde öğrencilerin erişimine sunulmaktadır. Ancak, bu testlerde yer alan maddelerin geliştirilme sürecine ya da madde yazımında dikkat edilen noktalara ilişkin ayrıntılı bilgilere yer verilmemektedir. Oysa, öğretmenler tarafından geliştirilen ve ulusal düzeyde uygulamaya açılmış olan bu maddelerin, çoktan seçmeli madde yazımında sıklıkla yapılan ihlalleri ortaya koymak ve hangi ilkelere daha çok dikkat edildiğini belirlemek için iyi bir kaynak olduğu düşünülmektedir. Bu nedenle, araştırmadan elde edilen bulguların, geniş bir öğrenci grubunun kullanımına açılan soruların geliştirilmesi gereken yönlerine ilişkin bilgi sağlanması ve daha nitelikli soruların geliştirilmesinde yol gösterici olması beklenmektedir.

Bu araştırmada MEB Ölçme, Değerlendirme ve Sınav Hizmetleri Müdürlüğü tarafından her sınıf düzeyi için öğrencilerin erişimine sunulan maddeler içerisinden 7. sınıf düzeyinde Fen Bilgisi ve Sosyal Bilgiler konu alanlarından üçer ünite için hazırlanmış maddeler seçilmiştir. Bu maddeler, madde yazım süreçlerinde dikkat edilmesi gereken ilkelere uygunluğu açısından incelenmiştir. Bu sayede, öğretmenler tarafından madde yazım süreçlerinde uyulması gereken ilkelerin ne düzeyde göz önünde bulundurulduğu, maddelerde en çok hangi ilkelere uygunluk açısından sorunlar yaşandığı ortaya koyulmuştur. Araştırma sonucunda elde edilen bulguların öğretmenlerin madde yazımında desteklenmesi süreci açısından yol gösterici olacağı düşünülmektedir.

Yöntem

Araştırma Modeli

MEB tarafından yayınlanan örnek test maddelerinin madde yazım ilkelerine uygunluğunu incelemeyi amaçlayan bu araştırma tarama türündedir. Tarama araştırmaları, bir grubun özelliklerini belirlemeyi amaçlamaktadır (Fraenkel, vd., 2012).

Mevcut çalışmada da bir grup maddenin madde yazım ilkelerine uygunluk özelliği belirlendiğinden dolayı tarama araştırmasına uygun bir desen seçilmiştir.

Evren ve Örneklem

Araştırmanın evrenini MEB tarafından yayınlanan Fen Bilgisi dersinden 140 ve Sosyal Bilgiler dersine ait toplam 144 adet 7. sınıf düzeyinde madde oluşturmaktadır. Bu maddeler MEB Ölçme, Değerlendirme ve Sınav Hizmetleri Genel Müdürlüğü'nün sayfasında yer alan beceri temelli testler içerisinde seçilmiştir. Araştırmacılardan ikisi lisans eğitimini ortaöğretim fen ve matematik alanları eğitiminde tamamlaması dolayısıyla, yayınlanan testlerden ortaöğretime en yakın olan sınıf düzeyi tercih edilmiştir. İlgili testler içerisinde üçer üniteye ait test tesadüfi olarak seçilmiştir. Tesadüfi seçimi sağlayabilmek adına random.org sitesinde yer alan bir araçtan yararlanılmıştır (Haahr ve Haahr, 2021). Buna göre toplam yedi ünite içerisinde fen bilgisi dersinden 2, 5 ve 6. üniteler; sosyal bilgiler dersinden ise 1, 6 ve 7. üniteler araştırmaya dahil edilmiştir. Bu ünitelerin her birinden ilk 20 madde olmak üzere seçilen 120 test maddesi araştırmacılar tarafından incelenmiştir.

Veri Toplama Aracı

Araştırma kapsamında belirlenen maddeler, araştırmacılar tarafından geliştirilen madde inceleme formu kullanılarak incelenmiştir. Bu form araştırmacılar tarafından alanyazın taramasına dayalı olarak oluşturulmuştur (Gronlund ve Waugh, 2009; Haladyna ve Downing, 1989; Haladyna vd., 2002; Nitko ve Brookhart, 2011; Popham, 2011; Turgut ve Baykul, 2010). Formda çoğunlukla maddelerin teknik özelliklerini içeren ilkelere yer verilmiştir. Bunun en önemli nedeni, formun ölçme ve değerlendirme alan uzmanları tarafından kullanılacak temel ilkeleri içermesini sağlamaktır. Formun son halinde toplam 20 ilke bulunmaktadır (Ek-1). Bu ilkeler *maddenin geneli*, *madde kökü* ve *seçenekler* olmak üzere üç başlık altında düzenlenmiştir. Maddenin geneli başlığında maddenin bir bütün olarak dil bilgisi kurallarına uygun olup olmadığı, maddeyi yanıtlamada gerekli olmayan ifadelerin bulunup bulunmadığı gibi ilkeler incelenirken, madde kökü başlığında madde kökünün anlaşılabilirliği, yanlış ifadeler içerip içermediği gibi ilkeler incelenmiştir. Seçenekler başlığında ise seçeneklerin mantıksal bir sırada verilip verilmediği, tekrar eden ifade bulunup bulunmadığı gibi ilkeler yer almaktadır. İnceleme formundaki her bir madde ise *evet*, *hayır* ya da *uygun değil* kategorileriyle yanıtlanmaktadır.

Madde inceleme formuyla yapılan belirlemelerin tutarlılığını ortaya koymak için, incelenen her ünitenin ilk beş maddesi olmak üzere toplam 30 madde üç araştırmacı tarafından ortak olarak incelenmiştir. Araştırmacılar arasındaki tutarlılık Fleiss' kappası ve uyum yüzdesine dayalı olarak belirlenmiştir. Araştırmacıların tutarlılığı Fen Bilgisi dersine ait ünitelerde yer alan maddelerde 0,73, sosyal bilgiler dersi maddelerinde 0,85 olarak bulunmuştur. Üç araştırmacı arasındaki uyum yüzdesi fen bilgisi dersi için %86,67, sosyal bilgiler dersi için %91,33 olarak elde edilmiştir. Bu değerler farklı araştırmacılar tarafından madde inceleme formuna dayalı olarak tutarlı

ölçme yapılabildiğini göstermektedir. Daha sonra araştırmacılar bir araya gelerek uyumsuzluk bulunan durumlar için ortak bir karar vermiştir.

Verilerin Analizi

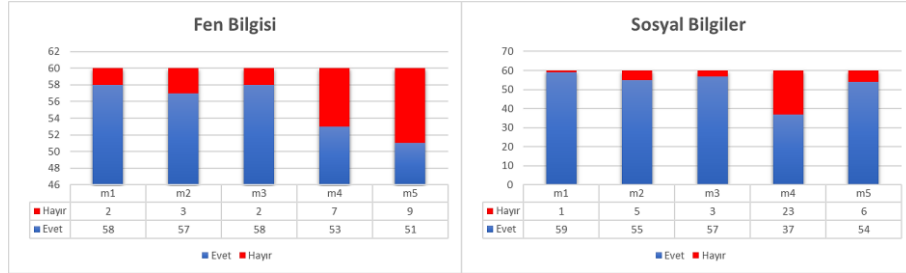
Toplanan veriler betimsel istatistiklerden frekans ve yüzde analizi ile çözümlenmiştir. Bu amaçla Jamovi 1.8.1 (The jamovi project, 2021) yazılımı kullanılmış ve çapraz tablolar aracılığıyla *Evet*, *Hayır* ve *NA* seçeneklerinin derslere göre dağılımı hesaplanmıştır. “Evet” seçeneği maddede ilgili ilkeye uyulduğunu, “hayır” seçeneği maddenin ilgili ilkeyi ihlal ettiğini göstermektedir. “NA” seçeneği ise ilgili ilkenin incelenen madde için uygulanabilir olmadığını ifade etmektedir. Bu seçeneklere dağılımları gösteren grafikler Microsoft Office Excel yazılımı kullanılarak oluşturulmuştur. Araştırmacılar arasındaki uyumu gösteren Fleiss’ kapp ve uyum yüzdeleri R (R Core Team, 2021) yazılımında “irr” paketi kullanılarak hesaplanmıştır (Gamer vd., 2018).

Bulgular

Araştırma bulguları, ölçme aracında yer aldığı sıra ile *maddenin geneli*, *madde kökü* ve *seçenekler* olmak üzere üç kısımda sunulmuş ve ilke ihlallerine ilişkin madde örnekleri ve yorumlarına yer verilmiştir. Bu bağlamda, **maddenin genel** olarak incelendiği *maddenin geneli* başlığındaki beş ilkeye ait yüzde grafikleri Şekil 1’de sunulmuştur.

Şekil 1

İncelenen Maddelerin Maddenin Genelile İlgili İlkelere Uygunlukları



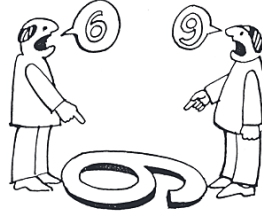
Şekil 1 incelendiğinde hem fen bilgisi hem de sosyal bilgiler derslerine ait örnek maddelerin maddenin geneli başlığındaki ilk üç ilkede %90’ın üzerinde bir uyuma sahip olduğu görülmektedir. Bu ilkeler sırasıyla şöyledir:

- *Maddede belirsiz ve hedef grubun düzeyine uygun olmayan kelimelerden (teknik terim, yabancı kelime, aşına olunmayan kelimeler) kaçınılmıştır.*
- *Dilbilgisi kurallarına uyulmuştur.*
- *Ana fikir seçeneklerde değil madde kökünde verilmiştir.*

Hem fen bilgisi hem de sosyal bilgiler dersleri için incelenen maddelerin bu ilkelere uygun olarak hazırlandığı anlaşılmaktadır. Öte yandan, maddeler 4 ve 5 numaralı ilkeler (*maddede doğru cevaba yönlendirecek ip uçlarından kaçınılmıştır, maddeyi yanıtlamada gerekli olmayan ifadelerden kaçınılmıştır*) açısından incelendiğinde, bu ilkelere uygun olan madde oranlarının düştüğü görülmüştür. Hatta, sosyal bilgiler dersi için 4 numaralı ilkenin karşılanma oranı %61,7 olarak bulunmuştur. Özellikle sosyal bilgiler dersine ait maddelerin, ölçülmek istenen özelliğe ek olarak okuduğunu anlama becerisini de ölçtüğü gözlenmiştir. İncelenen maddelerin yaklaşık olarak %38'inin, madde ile ölçülen özelliğe sahip olmayan ancak öncüldeki ipucunu kullanabilen öğrenciler tarafından doğru yanıtlanabilecek nitelikte olduğu görülmüştür. Araştırmacıların bu ilkelere ilişkin belirlemelerine bir örnek olması amacıyla sosyal bilgiler dersinden bir örnek madde Şekil 2'de verilmiştir:

Şekil 2

Sosyal Bilgiler Dersi İçin Örnek Madde – 1



İletişimle ilgili bu görsel aşağıdaki özdeyişlerden hangisi ile ilişkilendirilebilir?

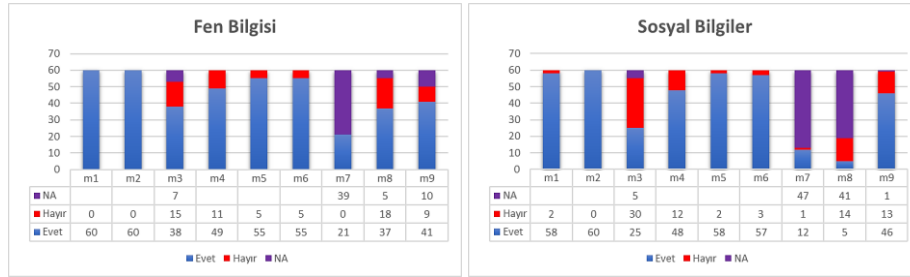
- A) İnsan, anlamadığı şeye sahip olamaz. (Goethe)
- B) Ne kadar bilirsen bil söylediklerin karşındakinin anladığı kadardır. (Mevlâna)
- C) Evinizin eşğini temizlemeden, komşunuzun damındaki karlardan şikâyet etmeyiniz. (Konfüçyüs)
- D) Her doğrunun iki yüzü vardır, birini kabullenmeden önce iki taraftan da bakmış olmak her zaman daha iyidir. (Aisopos)

Şekil 2'de verilen maddede bir görselin bulunduğu ve öğrencilerden özdeyişleri bu görselle ilişkilendirmelerinin beklendiği görülmektedir. Ancak, *D* seçeneğindeki özdeyişte *iki yüz, iki taraftan bakmak* gibi ifadeler kullanılmıştır. Görselde ise bir sayıya iki farklı taraftan bakan iki insan görülmektedir. Dolayısıyla, bu maddenin, öğrencileri doğru yanıtı ulaştıracak ipuçları içerdiği ifade edilebilir.

İnceleme formunun ikinci kısmında ise madde köküne ait ilkeler yer almaktadır. Maddelerin, madde köküyle ilgili ilkelere uygunluklarına ilişkin yüzdeler Şekil 3'te verilmiştir.

Şekil 3

İncelenen Maddelerin Madde Kökü ile İlgili İlkeler Uygunlukları



Şekil 3 incelendiğinde hem fen bilgisi hem de sosyal bilgiler maddelerinin, araştırmacılar tarafından **madde kökü** ana başlığına ait 1, 2, 5 ve 6 numaralı ilkeler açısından uygun bulunduğu görülmektedir. Bu ilkeler sırasıyla şöyledir:

- *Aşırı detay ya da aşırı genel içerikten kaçınılmıştır.*
- *Madde bir görüşe ya da kanıya dayalı değildir.*
- *Madde kökündeki yönerge açık ve anlaşılardır.*
- *Madde kökünde belli bir grup (cinsiyet, din, ırk, vb.) için yanlı olan bir ifade yoktur.*

İncelenen maddelerin, aşırı detay ya da genel içerikten uzak, görüşe ve kanıya dayanmayan, açık ve anlaşılır bir yönergeye sahip ve belli bir gruba yanlı ifadelerin içermeyen maddelerden oluştuğu anlaşılmaktadır. Bu nedenle, maddeler, 6 numaralı ilkede yer alan “yanlı ifade içermemesi” açısından araştırmacı görüşlerine dayalı olarak incelenmiştir.

Maddelerin 3, 4, 8 ve 9 numaralı ilkeler uygunluk oranlarının (%15 üzerinde hayır) daha düşük olduğu bulunmuştur. Bu ilkeler sırasıyla şöyledir:

- *Bileşik yanıt vermeyi gerektiren (I, II, III ve IV biçiminde öncül içeren) maddelerden kaçınılmıştır.*
- *Madde kökünün okuma yükü en aza indirilmiştir.*
- *Kullanılan görsel maddeyi yanıtlamak için gereklidir.*
- *Kullanılan metin maddeyi yanıtlamak için gereklidir.*

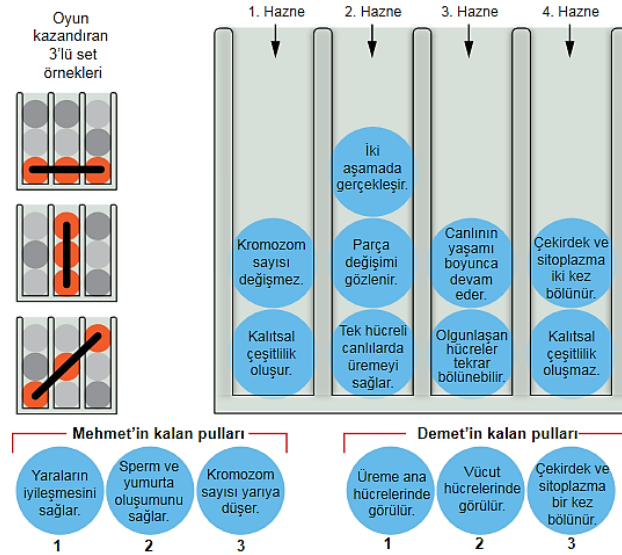
İlgili alanyazında bileşik yanıt vermeyi gerektiren maddelerin, özellikle ikili puanlama yapıldığı durumlarda kullanılması gerektiği ifade edilmektedir (Haladyna ve Downing, 1989; Haladyna vd., 2002). Ancak, bu araştırmada incelenen fen bilgisi maddelerinin yaklaşık olarak dörtte biri, sosyal bilgiler maddelerinin ise yaklaşık olarak yarısı bileşik yanıt vermeyi gerektiren maddelerden oluşmaktadır. Madde kökünün okuma yükünün en aza indirilmesi gerekliliği ise fen bilgisi maddelerinin dörtte birinde, sosyal bilgiler

maddelerinin ise beşte birinde sağlanamamıştır. Bu ilkenin ihlal edildiği örnek bir fen bilgisi maddesi Şekil 4'te verilmiştir:

Şekil 4

Fen Bilgisi Dersi İçin Örnek Madde – 1

ÜÇLÜ SET OYUNU
Mehmet ve Demet üzerinde mayoz ve mitoz bölünmenin özelliklerinin yazılı olduğu pulları eşit olarak paylaşıp aşağıdaki oyunu oynamaktadırlar. Oyunun amacı şekilde gösterilen örneklerde olduğu gibi aynı bölünme çeşidine ait özelliklerin yazılı olduğu pulları yatay, dikey ya da çapraz olarak aynı doğrultu oluşturmasını sağlayıp oyunu kazanmaktır. Aşağıda oyunun belli bir anındaki durumu gösterilmiştir.



Buna göre hamle sırası gelen oyuncuyla ilgili aşağıdaki ifadelerden hangisi kesinlikle doğrudur?

- A) Mehmet kalan herhangi bir pulunu 1. hazneye bıraktığı zaman oyunu kazanır.
B) Demet kalan herhangi bir pulunu 2. hazneye bıraktığında oyunu kazanır.
C) Mehmet kalan herhangi bir pulunu 3. hazneye bıraktığı zaman oyunu kazanır.
D) Demet kalan herhangi bir pulunu 4. hazneye bıraktığında oyunu kazanır.

Şekil 4'te verilen maddeyle öğrencilerin mitoz ve mayoz bölünmenin özelliklerine ilişkin bilgilerinin ölçülmesi amaçlanmıştır. Maddede öğrencilerin bu bilgileri oyuna dayalı bir durum üzerinden ölçülmektedir. Öğrencinin bu maddeyi doğru yanıtlaması için hem mitoz ve mayoz bölünmenin özelliklerini bilmesi hem de oyunun kurallarını anlaması gerekmektedir. Dolayısıyla, bu durumun hem gereksiz bir okuma yükü oluşturduğu, hem de ölçülmek istenen özellikle ilgisiz bir beceri kullanmayı gerektirdiği için bir geçerlik sorunu yaratabileceği ifade edilebilir.

Kullanılan görselin maddeyi yanıtlamak için gerekli olmasına ilişkin ilkeyi ihlal eden örnek bir fen bilgisi maddesi Şekil 5'te verilmiştir:

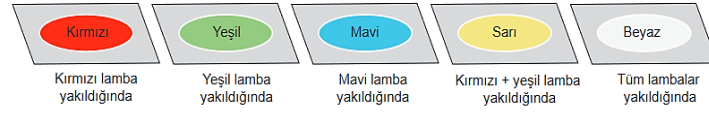
Şekil 5

Fen Bilgisi Dersi İçin Örnek Madde – 2

2. Işıkla ilgili bir etkinlik için şekildedeki düzenek hazırlanıyor.



Led lambalar tek tek veya birlikte yakıldığında karton bardağın üstten görünüşleri aşağıdakiler gibi oluyor.



Buna göre yapılan etkinlikle ilgili,

- I. Renkli ışıklar birleştirilerek farklı renkte ışıklar elde edilir.
- II. Beyaz ışık kendi rengindeki ışığı ve kendine yakın renkteki ışıkları yansıtır.
- III. Sarı renkli ışık mavi renkli ışıkla birleştirilirse beyaz renkli ışık görünür.

çıkarımlarından hangileri doğrudur?

- A) Yalnız I. B) Yalnız III. C) I ve II. D) I ve III.

Şekil 5'te verilen madde incelendiğinde, maddenin bileşik yanıt vermeyi gerektiren bir madde olduğu görülmektedir. Ayrıca, bu maddeye ait görselin ilk bölümünde anlatılan bardaklı düzenek maddeyi yanıtlamak için gerekli olmayan bilgiler içermektedir. Üstelik, bu maddeden tüm görsel çıkarıldığında, madde ışığın ve renklerin özelliklerini bilen bir öğrenci tarafından yanıtlanabilir niteliktedir.

Madde kökünde kullanılan metnin maddeyi yanıtlamak için gerekli olmadığı sosyal bilgiler dersine ait bir madde örneği Şekil 6'da verilmiştir:

Şekil 6

Sosyal Bilgiler Dersi İçin Örnek Madde-2

Bir ürünü ilk kez deneyecek olan tüketiciler öncelikle ürünün "dışsal" özelliklerine bakarak ürünü tercih edip etmeme kararı vermektedir. Bu özellikler ürün değerlendirmesinde önemli rol oynar ve tüketiciler ürünleri değerlendirme sürecinde ürünün fiyat, ambalaj, marka ismi, üretildiği ülke gibi dışsal özellikleri de bir kalite göstergesi olarak düşünerek karar verirler. Bu durum kalıp yargılarla ilgilidir.

Buna göre aşağıdakilerden hangisi ürüne yönelik bir kalıp yargı değildir?

- A) Bir ürün ne kadar pahalıysa o kadar kalitelidir.
- B) Japonlar teknolojik ürün üretiminde en iyi ülkedir.
- C) Cep telefonu alırken garanti belgesinin olmasına dikkat edilmelidir.
- D) Elektronik eşya alırken bilinen markalar tercih edilmelidir.

Şekil 6’da verilen madde, öncül bir metin içermektedir. Bu metinde kalıp yargılara ilişkin bilgi ve örnekler verilmiştir. Seçeneklerde ise birtakım önermeler bulunmaktadır. Öğrencilerden bu önermelerin kalıp yargı olup olmadığını belirlemeleri beklenmektedir. Metinde *fiyat, ambalaj, marka ismi, üretildiği ülke* gibi örneklere yer verilmiştir. Seçeneklerde ise *pahalı, Japonlar, bilinen marka* ifadeleri bulunmaktadır. Maddede kullanılan metnin hem maddeyi yanıtlamak için gerekli olmadığı hem de maddenin doğru yanıtına ilişkin ipuçları içerdiği ifade edilebilir. Bu haliyle, madde öğrencilerin sosyal bilgiler dersine ilişkin bilgi ve becerilerinden ziyade onların okuduğunu anlama becerilerini ölçmektedir. Öğrencinin bu maddeyi doğru yanıtlaması için okuduğu metni anlaması yeterlidir.

Madde kökü başlığı altında yer alan 7 numaralı ilke *madde kökündeki olumsuz ifadelerin vurgulanmasıyla* ilgilidir. Bu bağlamda hem fen bilgisi hem de sosyal bilgiler dersinden incelenen maddelerin büyük çoğunluğunun olumsuz madde köküne sahip olmadığı görülmüştür. Madde kökü olumsuz olan maddelerde ise olumsuz ifadelerin vurgulanması ilkesine uyulduğu belirlenmiştir.

İnceleme formunun üçüncü kısmında **maddenin seçeneklerine** ait ilkeler yer almıştır. Maddelerin, seçeneklerle ilgili ilkelere uygunluklarına ilişkin yüzdeler Şekil 7’de verilmiştir.

Şekil 7

İncelenen Maddelerin Seçenekler ile İlgili İkelere Uygunlukları



Grafik incelendiğinde, maddelerin seçeneklerle ilgili ilkeleri ihlal oranlarını genel olarak %20’nin altında olduğu görülmektedir. Karşılama oranı en yüksek olan 1, 4, 5 ve 6 numaralı sırasıyla şöyledir:

- Seçenekler mantıksal ya da anlamlı bir sırada düzenlenmiştir.
- Hepsi veya hiçbiri seçeneklerinden kaçınılmıştır.
- Seçeneklerde tekrar eden ifade yoktur.
- Anlamsız, maddeyle ilişkisiz (saçma) seçenekler yoktur.

Şekil 7’de her iki derse ait uygunluk oranları incelendiğinde, 2 ve 3 numaralı ilkelerin (seçeneklerin içerik ve dilbilgisi olarak benzer olması, seçeneklerin yaklaşık olarak benzer uzunluğa sahip olması) diğer ilkelere göre daha fazla ihlal edildiği görülmektedir. 2 numaralı ilke fen bilgisi dersi için % 13,3, sosyal bilgiler dersi için

% 10 oranında; 3 numaralı ilke fen bilgisi dersi için % 11,7, sosyal bilgiler dersi için % 26,7 oranında ihlal edildiği belirlenmiştir. Üç numaralı ilkenin ihlaline ilişkin örnek bir madde Şekil 8’de verilmiştir:

Şekil 8

Fen Bilgisi Dersi için Örnek Madde – 3

İsmail’in geçirmiş olduğu bir hastalık sonrası bir bakteri türü vücuduna girmiş, ancak iki hafta sonunda kendiliğinden iyileşmişti. Bu süre içinde aylardır hazırlandığı bin metre yıldız sporcular koşu yarışmasına hiç antrenman yapamadan katılmak zorunda kalmıştı. Yarışın ilk üç yüz metresini rahat bir şekilde koşmuş ancak bacaklarındaki ani kasılmaları, son beş yüz metrede de enerjisinin tükendiğini hissetmişti. Terden sırsıklam olan İsmail nefes almakta güçlük çektiğini fark edince yarışı bıraktı.

Hayatından bir kesit verilen İsmail’in hücrelerindeki organellerle ilgili aşağıdaki yargılardan hangisine ulaşamaz?

- A) Golgi cisimciği İsmail’in terlemesi esnasında daha fazla görev almıştır.
- B) Vücuduna giren bakterilerin yok edilmesine lizozom organelleri yardımcı olmuştur.
- C) Nefes almakta güçlük çektiği sırada ribozom organelleri hücrelere gerektiği kadar oksijen taşıyamamıştır.
- D) Bacaklarındaki ani kasılmalar ve enerjisinin tükendiği hissi hücrelerindeki mitokondrilerin yeterli miktarda enerji üretememesi sonucudur.

Şekilde 8’de yer alan madde incelendiğinde seçeneklerin uzunluklarının benzer olmadığı görülmektedir. Maddeyi tahminle doğru yanıtlama davranışı gösterecek bir öğrencinin en kısa ya da en uzun seçeneğe yönelme eğiliminde olacağı göz önüne alındığında, bu maddenin doğru yanıtının bu iki seçenekten biri olmaması olumlu bulunmuştur. Bu maddede ihlal edilen önemli bir durum ise maddenin bilimsel hata içermesidir. Maddede öncül bir durum olarak verilen metnin ilk cümlesinde “*hastalık sonrası bakterinin vücuda girdiği*” biçiminde bir bilgi verilmiştir. Hastalık yapıcı bir bakterinin vücuda girmesi sonrasında hastalığın oluşacağı göz önüne alındığında, maddede yer alan bu bilginin bilimsel açıdan hatalı olabileceği ifade edilebilir.

Sonuç ve Tartışma

Bu araştırmada, MEB Ölçme, Değerlendirme ve Sınav Hizmetleri Müdürlüğü’nün web sitesi aracılığıyla öğrencilerin erişimine sunulan 7. sınıf düzeyindeki Fen Bilgisi ve Sosyal Bilgiler konu alanlarına dayalı farklı üniteler için hazırlanmış maddelerin, genel madde yazımı, madde kökü ve madde seçenekleri bakımından madde yazım ilkelerine uygunluğu incelenmiştir.

Maddeler genel madde yazım ilkelerine uygunluk açısından incelendiğinde hem fen bilgisi hem de sosyal bilgiler alanında maddelerin büyük bir çoğunluğunda hedef gruba uygun bir anlatım kullanıldığı, maddelerde öğrenciler için belirsiz olabilecek ifadelerden ya da aşına olmayan kelimelerden kaçınıldığı, ana fikrin madde kökünde açık ve anlaşılır bir biçimde ifade edildiği gözlenmiştir. İyi bir maddenin, öğrenciden beklenen görevi açık ve sade bir biçimde içermesi, ayrıca, öğrenciler için anlamlı ve önemli olan bir bilgiyi ve bilişsel süreci ölçmesi gerekmektedir (Haladyna, 2004). Maddenin içeriği ve dili öğrenci düzeyine uygun olmadığında, öğrenci

kendisinden nasıl bir yanıt beklendiğini ve madde ile neyin ölçülmek istendiğini anlamakta güçlük çekebilir. Bu durum, öğrencinin madde ile ölçülen kazanıma ulaşmış olmasına rağmen maddeye yanlış yanıt vermesine neden olabilir. Kullanılan madde türü ne olursa olsun, bir test maddesinin temel görevi ölçülen bilgi ya da beceriye sahip olan öğrencileri, bu bilgi ve beceriler açısından eksikleri olan öğrencilerden ayrılabilme (Nitko ve Brookhart, 2011).

Maddeler genel madde yazım ilkelerine uygunluk açısından incelendiğinde, maddelerde (öncül, kök veya seçenek) öğrencileri doğru cevaba yönlendirebilecek ipuçlarının bulunduğu görülmüştür. Madde yanıtlarını etkileyebilecek nitelikte ipuçları içeren madde sayısının fen bilgisine göre sosyal bilgiler dersinde daha fazla olduğu belirlenmiştir. Bu ipuçları özellikle maddelerde verilen öncüller (metin, resim, tablo, görsel vb.) içerisinde bulunmaktadır. Öğrencinin maddenin doğru yanıtına ilgili bilgi ve becerileri kullanarak mı, yoksa ipucunu fark ederek mi ulaştığı bilinemez. Bu durum, madde ile ölçülen psikolojik özellikten kaynaklı olmayan bir varyansın (construct irrelevant variance) oluşmasına neden olduğundan (Messick, 1989) madde puanlarının geçerliğini düşürebilir.

Araştırmada elde edilen bir diğer bulgu, maddelerde yanıt vermek için gerekli olmayan ifadelerin kullanılmasına ilişkindir. Madde yazımında uzun ve karmaşık cümle yapılarının, öğrenciler için zor olabilecek sözcüklerin ve maddeyi yanıtlamak için gerekli olmayan ifadelerin kullanılması, öğrencilerin maddedeki performanslarını olumsuz etkileyebilir. Maddenin yanıtlanması açısından gereksiz olan ifadeler öğrenci için önemli hale gelip onun dikkatini çekebilir. Bu durum, gerçekte zor olmayan bir maddeyi olduğundan daha zor ve daha az ayırt edici hale getirebilir (Nitko ve Brookhart, 2011; Yıldırım 2021).

Araştırmada madde kökü bağlamında inceleme yapıldığında genel olarak aşırı detay veya aşırı genel içerikten kaçınıldığı, madde kökünde kullanılan yönergelerin açık olduğu belirlenmiştir. İyi hazırlanmış bir çoktan seçmeli maddenin kökü, öğrencilerden beklenen görevi açık biçimde sunmalıdır. Ancak, madde kökü böyle değilse, öğrencinin madde kökündeki görevi anlamak için maddeyi tekrar tekrar okuması gerekebilir (Popham, 2011). Bu durum öğrencinin kendisinden beklenen asıl görevi anlamasını güçleştirebilir ve öğrencinin zaman kaybetmesine neden olabilir. Bu başlık altında yapılan incelemeler sonucunda, maddelerin çoğunluğunun belli bir görüşe veya kanıya dayalı olmadığı, yanlışlık içerebilecek ifadeler içermediği belirlenmiştir. Bir maddenin geçerli sonuçlar sağlayabilmesi, tüm yanıtlayıcılar için adil bir ölçme yapmasına bağlıdır. Madde yazımında herhangi bir katılımcı grubun duygu ve düşüncelerini rencide edecek bir dil, üslup ve bağlamın kullanımı veya belli bir grup için maddeyi yanıtlamanın avantaj veya dezavantaj oluşturması maddenin geçerliğini olumsuz etkiler (Yıldırım, 2021).

Madde köküne dayalı yapılan incelemelerde fen bilgisi ve sosyal bilgilerin her ikisinde de birleşik yanıt vermeyi gerektiren maddelere yer verildiği gözlenmiştir. Sireci ve diğerlerinin (1998), geniş ölçekli testlerde bulunan 285 madde üzerinde yaptıkları araştırmada en fazla ihlal edilen ilkenin birleşik yanıt gerektiren madde türü olduğu

belirlenmiştir. Bu kullanım maddenin anlaşılabilirliğini zorlaştırmakta ve madde ayırt ediciliğini düşürmektedir (Haladyna ve Downing, 1989; Kreiter ve Frisbie, 1989; Sireci vd. 1998).

Araştırmada, her iki alan için, madde köklerinde okuma yükünü artıran ve maddeyle ilişkili olmayan görsellere veya metinlere yer verildiği görülmüştür. Bu bulguyu destekler biçimde, Şengül-Avşar ve Barış-Pekmezci (2020) tarafından yapılan çalışmada, Milli Eğitim Bakanlığı Türkçe dersi maddelerindeki görsel ve metinlerin etkili bir biçimde kullanılmadığı ifade edilmiştir. Maddede uyarıcıların kullanımda yazarların dikkat etmesi gereken iki önemli soru bulunmaktadır. Bunlardan biri öncülün maddenin yanıtlanması için gerekli olup olmadığıyla ilgilidir. İkinci soru da öncülün öğrenci için anlamlı olup olmadığıyla ilgilidir. Eğer öncül kullanılmadan da madde yanıtlanabiliyorsa, bunlara maddede yer verilmemelidir. Gereksiz unsurlar maddenin yanıtlanmasını zorlaştırır ve öğrencide kafa karışıklığına neden olur (Bilican-Demir, 2021).

Madde kökünde olumsuz ifadenin kullanılması durumu incelendiğinde olumsuz madde kökünün az da olsa kullanıldığı ve olumsuzluk içeren kelimelerin vurgulandığı belirlenmiştir. Downing (2005)'in dört testteki 219 maddenin incelenmesi üzerine yaptığı araştırmasında testlerin % 35 ile % 65 oranında madde yazım ilkesini ihlal ettiği ve iki testte en çok görülen ihlalin maddenin olumsuz yazılması olduğu belirlenmiştir. Öğrencilerin olumlu maddelere göre olumsuz maddeleri anlaması daha zordur. Tamir (1993) öğrencinin olumsuz maddeleri anlamak için aynı maddenin olumlu haline göre iki kat daha fazla işleyen belleğini kullandığını belirtmiştir (Akt. Haladyna, 2004). Downing (2005) de olumsuz maddelerin madde güçlük düzeyinin diğer maddelere göre daha yüksek olduğunu bulmuştur.

Maddelerin seçenekleri incelendiğinde maddelerin büyük bir çoğunluğunda seçeneklerin mantıksal ve anlamlı bir sırada düzenlendiği, seçeneklerde tekrar eden ifadeler yer verilmediği belirlenmiştir. Seçeneklerde “yukarıdakilerin hepsi” veya “yukarıdakilerin hiçbiri” gibi ifadeler fen bilgisindeki maddelerde gözlenmezken, sosyal bilgilerde bazı sorularda bu ifadeler yer verilmiştir. Sireci vd. (1998) inceledikleri 285 madde içerisinde yukarıdakilerin hepsi veya hiçbiri ilkesinin ihlal edilmediğini, Downing (2005) ise 219 madde içerisinde bu ilkeyi ihlal eden madde sayısının çok az olduğunu belirtmiştir. Bu tür seçeneklerin kullanılması öğrenciye cevaplamada ipucu verebilir. Örneğin öğrenci, verilen seçenekler arasından ikisinin doğru olduğunu düşünürse bu durumda doğru yanıt hepsi seçeneği olacaktır (Nitko ve Brookhart, 2011). Haladyna (2004) bu tür seçeneklerin çoklu Doğru-Yanlıı türü maddeler dışında kullanılmamasını önermektedir.

Madde seçeneklerinin dil bilgisi ve içerik bakımından madde kökü ile ve kendi içerisindeki uyumu incelendiğinde diğer madde seçeneklerinde uyulması gereken ilkelere göre bu ilkenin daha fazla ihlal edildiği görülmüştür. Doğru yanıtla ilişkin ipucu içeren seçeneklere sahip maddelerin olduğu gözlenmiştir. Ayrıca, seçeneklerin uzunluk açısından birbirine benzer olmasına yeteri kadar dikkat edilmediği belirlenmiştir. Maddelerde doğru yanıt seçeneğinin çeldiricilere göre genellikle daha uzun olduğu

gözlendi. Bu bulguyu destekler biçimde, Tarrant vd. (2006)'nin çalışmasında, doğru yanıtın çeldiricilerden daha uzun olması en sık rastlanılan ihlaller arasında yer almıştır. Öğretmenlerin doğru seçeneği daha açık ve nitelikli yazma çabası bazı durumlarda doğru yanıt seçeneğinin çeldiricilere göre daha uzun olmasıyla sonuçlanabilir. Bu durum hem öğrenciye bir ipucu sağlar hem de maddedeki okuma yükünü artırır. Maddenin cevaplanmasında kolaylık sağlayan ipuçları ya da maddeyi yanıtlamayı zorlaştıran durumlar ölçülmek istenilen yapıyla ilgisiz bir varyansın oluşmasına neden olarak ölçme sonuçlarının güvenilirliğini ve geçerliğini düşürebilir (Downing, 2002; Messick, 1989).

Araştırmada, fen ve sosyal bilgiler dersleri için ulusal düzeyde öğrencilerin erişimine sunulmuş bazı maddelerin madde yazımında dikkat edilmesi gereken ilkelere uygunluğu incelenmiştir. Bu incelemeler, maddelerin çoktan seçmeli maddeler için istenen teknik niteliklere ne oranda sahip olduğunu belirlemekle sınırlandırılmıştır. Dolayısıyla, araştırmada maddelerin kazanım ve bilişsel boyutlara uygunluğu açısından bir inceleme yapılmamıştır. Ulusal düzeyde öğrencilerin erişimine sunulmuş olan maddelerin kazanım ve bilişsel süreçlere uygunluğuna ilişkin bir inceleme yapılabilir. Ayrıca, araştırmada yalnızca fen bilgisi ve sosyal bilgiler alanından tesadüfi olarak seçilen üç ünitadaki maddeler değerlendirilmiştir. Araştırmanın kapsamı geliştirilerek ilgili internet sitesinde sunulan diğer alanlardaki maddeler üzerinde de benzer incelemeler yapılabilir. Araştırmada, maddelerin yanlılığı yalnızca uzman görüşüne dayalı olarak değerlendirilmiştir. Maddelerin öğrencilere uygulanmasıyla elde edilecek veri üzerinde detaylı istatistiksel analizler yapılarak maddelerin yanlılığına ilişkin daha fazla bilgi elde edilebilir.

Kaynakça

- Adkins, D. C. (1947). *Construction and Analysis of achievement tests*. Government Printing Office.
- Anastasi, A. (1961). *Psychological Testing*. MacMillan Publishing Co., Inc.
- Anıl, D., ve Acar, M. (2008). Sınıf öğretmenlerinin ölçme ve değerlendirme sürecinde karşılaştıkları sorunlara ilişkin görüşleri. *Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 5(2), 44-61.
- Bangchang, K. (2015). Factors Affecting Academic Performance of Undergraduate Students. *International Journal of Multidisciplinary Approach & Studies*, 2(6), 205-215.
- Cheung, D. (2003). Guidelines for writing multiple-choice items. *Hong Kong Science Teachers' Journal*, 21(2), 1-11.
- Cohen, R. J., ve Swerdlik, M. E. (2010). *Psychological testing and assessment*. McGraw-Hill Companies, Inc.
- Bilican-Demir, S. (2021). Başarı testlerinin geliştirilmesi ve madde yazımı. Ö. Yıldırım, ve S. Kula-Kartal (Haz.), *Eğitimde ölçme ve değerlendirme* (s.124-164). Lisans Yayıncılık.

- Downing, S. M. (2002). Threats to the validity of locally developed multiple-choice tests in medical education: Construct-irrelevant variance and construct underrepresentation. *Advances in Health Sciences Education*, 7(3), 235-241.
- Downing, S. M. (2005). The effects of violating standard item writing principles on tests and students: The consequences of using flawed test items on achievement examinations in medical education. *Advances in Health Sciences Education*, 10,133-143.
- Elias, M. J., Zins, J. E., Weissberg, R. P., Frey, K. S., Haynes, N. M., Kessler, R., Shwab-Stone, M. E., ve Shriver, Y. (1997). *Promoting social and emotional learning: Guideline for educators*. Association for Supervision and Curriculum Development.
- Fraenkel, J. R., Wallen, N. E. ve Hyun, H. H. (2012). *How to design and evaluate research in education*. McGraw Hill.
- Gamer, M., Lemon, J., ve Puspendra Singh, I. F. (2019). *irr: Various Coefficients of Interrater Reliability and Agreement. R package version 0.84.1*. <http://cran.r-project.org/web/packages/irr/irr.pdf>.
- Gelbal, S. ve Kelecioğlu, H. (2007). Öğretmenlerin ölçme ve değerlendirme yöntemleri hakkındaki yeterlik algıları ve karşılaştıkları sorunlar. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 33(33), 135-145.
- Gronlund, N. E., ve Waugh, C. K. (2009). *Assessment of student achievement*. Pearson Education, Inc.
- Haahr, M. ve Haahr, S. (2021). *True random number service*. Random.org. <https://www.random.org/>
- Haladyna, T. M. (1997). *Writing test items to evaluate higher order thinking*. Allyn & Bacon.
- Haladyna, T. M. (2004). *Developing and validating multiple-choice test items* (3. baskı). Lawrence Erlbaum Associates Publishers.
- Haladyna, T. M., ve Downing, S. M. (1989). A taxonomy of multiple-choice item writing rules. *Applied Measurement in Education*, 1, 37-50.
- Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple choice item-writing guidelines for classroom assessment, *Applied Measurement in Education*, 15(3), 309-333.
- Kreiter, C. D. ve Frisbie, D. A. (1989). Effectiveness of multiple true-false items, *Applied Measurement in Education*, 2(3), 207-216, https://doi.org/10.1207/s15324818ame0203_2
- Kutlu Ö., Doğan, C. D., ve Karakaya, İ. (2014). *Ölçme ve değerlendirme: Performansa ve portfolyoya dayalı durum belirleme*. Pegem Akademi.

- Martínez, R. J., Moreno, R., Martín, I., ve Trigo, M. E. (2009). Evaluation of five guidelines for option development in multiple-choice item-writing. *Psicothema*, 21(2), 326-330.
- Messick S. (1989). Validity. R.L. Linn (Haz.) *Educational Measurement* (3. baskı.) (s.13-104). American Council on Education and Macmillan.
- Nitko, A., ve Brookhart, S. M. (2011). Educational assessment of students. Pearson Education, Inc.
- Popham, W. J. (2011). *Classroom assessment: What teachers need to know*. Pearson Education, Inc.
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. R project. <https://www.R-project.org/>.
- Rush, B. R., Rankin, D. C., & White, B. J. (2016). The impact of item-writing flaws and item complexity on examination item difficulty and discrimination value. *BMC Medical Education*, 16(1), 1-10.
- Sireci, S. G., Wiley, A., ve Keller, L. A. (1998). *An empirical evaluation of selected multiple-choice item writing guidelines*. Annual Meeting of the Northeastern Educational Research Association.
- Şengül-Avşar, A. ve Barış-Pekmezci, F. (2020). Investigation of skill-based test items and sample items in the entrance examination for high schools for Turkish lesson. *International Journal of Education Technology and Scientific Researches*, 5(12), 677-718.
- Tarrant, M., Knierim, A., Hayes, S. K., ve Ware, J. (2006). The frequency of item writing flaws in multiple-choice questions used in high stakes nursing assessments. *Nurse Education Today*, 26, 662-671
- The jamovi project (2021). *jamovi*. (Version 1.8). <https://www.jamovi.org>.
- Turgut M. F. ve Baykul Y. (2010). *Eğitimde ölçme ve değerlendirme*. Pegem Akademi.
- Yıldırım, Ö. (2021). Geniş ölçekli testlerde madde yapıları ve hazırlanmasında dikkat edilmesi gerekenler (PISA ve TIMSS Örnekleri). *İstanbul Ölçme ve Değerlendirme Dergisi*. 1(2). 50-59.

Investigation of the Test Items Published by Ministry of National Education According to the Item Writing Guidelines

Abstract

The present study aims to investigate the consistency of the sample test items published by The Ministry of National Education of Turkey according to the item writing principles. The 120 items randomly selected from the items published for the seventh-grade science and social sciences courses. The selected items were examined by the researchers based on an item review form developed by the researchers, and the collected data were analyzed based on the descriptive statistics. The results of the study revealed that the items include clues referring to the correct answer, textual or visual materials unnecessary to answer the items. In addition, it was brought out that the complex multiple choice item format (e.g., I, II, II, IV etc.) are frequently used.

When the items are analyzed to see if they meet the style concerns or item stem writing principles (such as the use of correct grammar or clear directions in the item stem), it was found that the items meet the grammatical principles, the main idea are generally presented in the item stem instead of the choices, items have clear instructions, and they are not opinion-based.

Keywords: multiple-choice item writing principles, measurement and evaluation, test development

Ek-1: Madde İnceleme Formu

	Madde İnceleme Formu	Evet	Hayır	Uygulanamaz (NA)
	Maddenin Geneli			
1	Maddede belirsiz ve hedef grubun düzeyine uygun olmayan kelimelerden (teknik terim, yabancı kelime, aşına olunmayan kelimeler) kaçınılmıştır.			
2	Doğru dilbilgisi (noktalama işaretleri ve yazım kuralları) kullanılmıştır.			
3	Ana fikir seçeneklerde değil madde kökünde verilmiştir.			
4	Maddede doğru cevaba yönlendirecek ip uçlarından kaçınılmıştır.			
5	Maddeyi yanıtlamada gerekli olmayan ifadelerden kaçınılmıştır			
	Madde Kökü			
1	Aşırı detay ya da aşırı genel içerikten kaçınılmıştır.			
2	Madde bir görüşe ya da kanıya dayalı değildir.			
3	Bileşik yanıt vermeyi gerektiren maddelerden (I, II, III ve IV öncül maddeler) kaçınılmıştır.			
4	Madde kökünün okuma yükü en aza indirilmiştir.			
5	Madde kökündeki yönerge açık ve anlaşılırdır.			
6	Madde kökü belli bir gruba (cinsiyet, din, ırk, vb.) yanlı ifade yoktur.			
7	Madde kökündeki olumsuz ifadeler vurgulanmıştır.			
8	Kullanılan görsel maddeyi yanıtlamak için gereklidir.			
9	Kullanılan metin maddeyi yanıtlamak için gereklidir.			
	Seçenekler			
1	Seçenekler mantıksal ya da anlamlı bir sırada düzenlenmiştir.			
2	Seçenekler içerik ve dilbilgisi olarak benzeşiktir.			
3	Seçenek uzunlukları yaklaşık olarak eşittir.			
4	Hepsi veya hiçbiri seçeneklerinden kaçınılmıştır.			
5	Seçeneklerde tekrar eden ifade yoktur.			
6	Anlamsız, maddeyle ilişkisiz (saçma) seçenekler yoktur.			