



DOI: 10.18039/ajesi.725475

Matematikte Kavram Karikatürlerinin Değerlendirilmesine Yönelik Puanlama Anahtarının Geliştirilmesi¹

Ahmet YILDIZ², Hasan ES³, Ali TÜRKDOĞAN⁴

Geliş Tarihi: 22.04.2020

Kabul Tarihi: 12.12.2020

Türü: Araştırma Makalesi

Öz

Bu araştırmanın amacı matematik dersi için hazırlanan kavram karikatürlerini değerlendirebilmek için geçerli ve güvenilir bir dereceli puanlama anahtarı (DPA) geliştirmektir. Araştırma verileri Sivas il merkezinde görev yapan matematik öğretmenlerinin oluşturduğu 36 kavram karikatürünün değerlendirilmesiyle elde edilmiştir. Değerlendirme işlemi matematik eğitimi alanında yüksek lisans yapmış iki matematik eğitimcisi gerçekleştirmiştir. Geçerlik çalışmalarında öncelikle uzmanların DPA hakkındaki görüşleri içerik, yapı ve ölçüt bağlamında tek tek incelenmiştir. Bunların DPA'ya uygunluğu ve yeterliği konusunda uzmanların görüşleri arasındaki uyum yüzdelerinin kabul edilebilir düzeyde ve %82 ile %95 arasında değiştiği görülmüştür. Geçerlik çalışması kapsamında ayrıca kapsam geçerlik indeksi de hesaplanmış ve her bir alt boyuta ilişkin kapsam geçerlik indeksleri 0.80 ile 1 arasında bulunmuştur. Hem uyum yüzdeleri hem de kapsam geçerlik indeksleri dikkate alındığında geliştirilen DPA'nın kavram karikatürlerini değerlendirmek için geçerli bir ölçme aracı olduğu görülmüştür. Güvenirlik çalışmalarında klasik test kuramı kapsamında iki puanlayıcı arasındaki uyum toplam puanlar için "Spearman Korelasyon Katsayısı" ile belirlenirken her bir alt ölçütteki uyum "Cohen's Kappa" ile belirlenmiştir. DPA'ya iki değerlendiricinin verdiği toplam puanlar arasında pozitif, anlamlı ve yüksek ($Rho=0.856$, $p<0.01$) bir ilişkinin olduğu görülürken, alt ölçütlere ilişkin de Cohen's Kappa sonuçları 0.706 ve 0.797 arasında yer aldığı görülmüştür. Güvenirlik çalışmalarında ayrıca madde tepki kuramı kapsamında Çok Yüzeysel Rasch Ölçme Modeli dikkate alınarak puanlayıcı, puanlanan (kavram karikatürü) ve madde yüzeyleri için güvenirlilik hesaplamaları yapılmıştır. Bu yüzeylere ait güvenirlilik indeksleri de sırasıyla 0.87, 0.80 ve 0.89 bulunmuştur. DPA'nın toplam puan için hesaplanan Spearman korelasyonu ve alt ölçüt puanları için hesaplanan Cohen's Kappa değerleri ile puanlayıcı, puanlanan(kavram karikatürü) ve madde yüzeylerine ait güvenilirlik indeksleri bir arada değerlendirildiğinde oluşturulan DPA'nın güvenilir bir puanlama anahtarı olduğu görülmüştür. Araştırmanın bulguları ilgili literatür dikkate alınarak tartışılmış ve çeşitli önerilerde bulunulmuştur.

Anahtar kelimeler: dereceli puanlama anahtarı, kavram karikatürü, matematik eğitimi, performans değerlendirme

Atf: Yıldız, A., Es, H. ve Türkdoğan, A. (2021). Matematikte kavram karikatürlerinin değerlendirilmesine yönelik puanlama anahtarının geliştirilmesi. *Anadolu Journal of Educational Sciences International*, 11(1), 250-267. DOI: 10.18039/ajesi.725475

¹ Bu çalışma, birinci yazarın ikinci ve üçüncü yazarların danışmanlığında Gazi Üniversitesi Eğitim Bilimleri Enstitüsü Matematik Eğitimi Anabilim Dalı'nda gerçekleştirdiği doktora tezinden türetilmiştir.

² (Sorumlu Yazar) Dr., Sivas Bilim ve Sanat Merkezi, Türkiye, ahmetyildiz58@gmail.com, <https://orcid.org/0000-0002-9149-5859>

³ Dr. Öğr. Üyesi, Gazi Üniversitesi, Eğitim Fakültesi, Matematik ve Fen Bilimleri Eğitimi Bölümü, Türkiye, hasanes@gazi.edu.tr, <https://orcid.org/0000-0003-0216-5426>

⁴ Dr. Öğr. Üyesi, Cumhuriyet Üniversitesi, Eğitim Fakültesi, Matematik ve Fen Bilimleri Eğitimi Bölümü, Türkiye, aliturkdogan@hotmail.com, <https://orcid.org/0000-0002-7732-8173>



DOI: 10.18039/ajesi.725475

Development of Rubric for the Evaluation of Concept Cartoons in Mathematics¹

Ahmet YILDIZ², Hasan ES³, Ali TÜRKDOĞAN⁴

Submitted by: 22.04.2020

Accepted by: 12.12.2020

Type: Research Article

Abstract

The aim of this study is to develop a valid and reliable rubric in order to evaluate the concept cartoons prepared for mathematics lesson. Research data were obtained by evaluating 36 concept cartoons created by mathematics teachers working in Sivas city center. Two mathematics educators who had a master's degree in mathematics education carried out the evaluation process. In the validity studies, experts' opinions about rubric were examined one by one in the context of content, structure and criteria. It has been observed that the percentages of agreement between the opinions of the experts about their suitability and competence with rubric are at an acceptable level and varied between 82% and 95%. Within the scope of the validity study, the content validity index was also calculated and the content validity indexes for each sub-dimension were found between 0.80 and 1. Considering both the concordance percentages and the content validity indexes, it has been seen that the developed rubric is a valid measurement tool for evaluating the concept characteristics. In reliability studies, within the scope of classical test theory, the harmony between two raters was determined by "Spearman Correlation Coefficient" for the total scores, while the fit in each sub-criterion was determined by "Cohen's Kappa". While it was observed that there was a positive, significant and high ($Rho = 0.856$, $p < 0.01$) relationship between the total scores given by the two evaluators to rubric, Cohen's Kappa results were found between 0.706 and 0.797 for the sub-criteria. In the reliability studies, reliability calculations were made for rater, scored (concept cartoons) and item surfaces, taking into account the Multi-Facial Rasch Measurement Model within the scope of item response theory. The reliability indices of these surfaces were found to be 0.87, 0.80 and 0.89, respectively. When rubric's Spearman correlation calculated for total score and Cohen's Kappa values calculated for sub-criterion scores were evaluated together with the reliable indexes of rater, scored (concept cartoons) and item surfaces, it was seen that the rubric created was a reliable scoring key. The findings of the study were discussed considering the relevant literature and various suggestions were made.

Keywords: concept cartoon, mathematics education, performance evaluation, rubric

Cite: Yıldız, A., Es, H. and Türkdoğan, A. (2021). Development of rubric for the evaluation of concept cartoons in mathematics. *Anadolu Journal of Educational Sciences International*, 11(1), 250-267. DOI: 10.18039/ajesi.725475

¹ This study is derived from the first author's doctoral dissertation, which he conducted under the supervision of the second and third authors in the Department of Mathematics Education at Gazi University Institute of Educational Sciences.

² (Corresponding author) Dr., Sivas Science and Art Center, Turkey, ahmetyildiz58@gmail.com, <https://orcid.org/0000-0002-9149-5859>

³ Assistant Professor, Gazi University, Faculty of Education, Department of Mathematics and Science Education, Turkey, hasanes@gazi.edu.tr, <https://orcid.org/0000-0003-0216-5426>

⁴ Assistant Professor, Sivas Cumhuriyet University, Faculty of Education, Department of Mathematics and Science Education, Turkey, aliturkdogan@hotmail.com, <https://orcid.org/0000-0002-7732-8173>

Giriş

Eđitim alanındaki çalışmaların nihai hedeflerinden biri de öğrenenlerin daha iyi ve kolay nasıl öğrenebileceklerini ortaya çıkarma olmuştur. Bu hedefe ulaşırken de yeni öğretim yöntem ve teknikleri geliştirilmiştir. Geliştirilen yöntem ve teknikler, öğrencilerin derse aktif katılımlarına, yaratıcı ve özgün fikirler üretip bunları paylaşımlarına önem vermektedir. Bu bağlamda öğrencinin aktif olarak sürece katan kavram karikatürleri ile gerçekleştirilen öğretimlerin etkili olduğu söylenebilir.

Kavram karikatürleri fen öğretmenlerinin yeni öğretim yöntemleri bulma ihtiyaçlarını gidermek için Brenda Keogh ve Stuart Naylor tarafından 1999' da geliştirilmiştir (Van der Mark, 2011). Çoktan seçmeli bir problem soran kavram karikatürü mizahi karikatür formatından farklıdır. Kavram karikatürü, aktarılmak istenen problemi görseller kullanılarak üç ya da daha fazla karakterin kendi arasında yaptığı tartışmalar şeklinde sunar. Kavram karikatüründe bulunan karakterler, bilimsel olarak kabul edilebilir olanlar da dahil olmak üzere alternatif fikirlerin konuşma balonları ile sunulduğu bir diyalog içinde yer alır (Keogh ve Naylor, 1999). Bu alternatif fikirler gerçek olaylara, ortak düşüncelere ya da yanlış anlamalara dayanmaktadır (Samkova ve Hospesova, 2016). Kavram karikatürlerinde kısa ve öz bir metin bulunduğundan dil becerileri zayıf olan öğrenciler de anlatılan görüşleri rahatlıkla anlayabilirler (Keogh and Naylor, 1998).

Kavram karikatürlerinde görsellerin kullanımı öğrencilerin dikkatlerini derse çekmekte ve ilgilerini artırmaktadır (Roesky ve Kennepoh, 2008). Böylece öğrencilerin aktif katılımı ile eğlenceli bir şekilde öğrenme gerçekleşmektedir (Balım, İnel ve Evrekli, 2008). Kavram karikatüründeki görsel öğeler ve karakterler arasındaki diyaloglar çocuklara fikirlerinin değerli olduğu hissini vermektedir. Böylelikle öğrenciler de soruları samimi bir şekilde cevaplandırabilmektedirler (Uzođlu, Yıldız, Demir ve Büyükkasap, 2013). Kavram karikatürleri öğrencilerde merak uyandırarak araştırma, sorgulama ve eleştirel düşünme gibi farklı düşünme becerilerinin gelişmesine de katkı sunar (Demirci ve Özyürek, 2017; Long ve Marson, 2003). Ayrıca öğrencilerin araştırma yaparken ortaya çıkan problemleri keşfetmelerine, çözüm üretmelerine ve bilimsel gerçeklere ulaşmalarına yardımcı olur (Jamal, Ibrahim ve Surif, 2019; Kabapınar, 2009).

Kavram karikatürü kullanımı öğrenciler üzerindeki yanlış verme konusundaki kaygıyı azaltarak öğretim faaliyetlerine daha rahat katılabilmelerini sağlar (Kabapınar, 2005). Öğrencilerin fikirlerini belirlemeye yardımcı olarak sınıf ortamında fikirlerini tartışmalarını, bilgilerini sorgulamalarını ve kavramsal öğrenmeyi gerçekleştirmelerini sağlar (Evrekli, 2010; Naylor, Keogh ve Downing, 2007). Ayrıca konuşma balonlarında yer alan bazı fikirler problem durumuna göre alternatif fikirler içerdiğinden öğrencilerin kavram yanlışlarının belirlemede de kavram karikatürü kullanımı etkili bir yöntemdir (Keogh ve Naylor, 1999). Kavram karikatürü kullanılarak bu kavram yanlışlarının nedenleri de sınıfta tartışılabilir (Balım vd., 2008).

Kavram karikatürlerinin yukarıda sıralanan faydaları sağlması için Keogh, Naylor ve Wilson (1998) tarafından açıklanan aşağıdaki özelliklere sahip olması gerekmektedir;

- ✓ Her yaştaki öğrencinin kolayca anlayabileceđi kısa metinlerden oluşmalıdır. Böylece okuryazarlık becerisi sınırlı öğrenciler için de dikkat çekici olur.
- ✓ Günlük olaylara uygulanabilen bilimsel fikirlere odaklanılmalıdır. Böylelikle öğrenenler günlük olaylar ile bilimsel fikirler arasındaki ilişkiyi görmüş olacaktırlar.
- ✓ Üç ya da daha fazla karakter yer almalıdır.
- ✓ Karakterlerin savunduđu alternatif fikirlerden en az biri bilimsel olarak kabul edilen görüş, diğerleri ise kavram yanlışsı içeren görüşler olmalıdır.
- ✓ Alternatif fikirler birbirine eşit statüde verilmelidir. Böylece öğrenen karikatüre bakıp doğru cevabı hemen bulamayacaktır.

Kavram karikatürlerinin taşınması gereken özelliklerin ne kadarına sahip olduğunu değerlendirmek için kullanılacak ölçme aracının belirlenmesi de gereklidir. Öğrenciler hakkında kararların alınması için çeşitli ölçme araçlarından faydalanılarak bilgi toplanabilir. Fakat yapılacak ölçmenin amacı, ölçülecek olan özelliklerin doğası (bilgi, beceri, tutum, algı vb.) ve öğrencilerin özellikleri gibi kriterler hangi ölçme aracının kullanılacağına belirlenmesinde etkilidir. Öğrencilerin bir konudaki başarısını ölçmek için yazılı, çoktan seçmeli, kısa yanıtı vb. testler kullanılabilir. Ancak bir işin yerine getirilmesi (örneğin kavram karikatürü oluşturma) için gerekli olan becerilerin bilgi testleri vasıtasıyla ölçümü zor olabilmektedir. Çünkü bir işi nasıl yapacağını bilmek ile bu işi belirlenen kriterlere uygun olarak yapmak arasında bir ilişki olmayabilir. Bu nedenle eğitimde öğrenci performansları genellikle söz konusu performansa uygun değerlendirme araçları kullanılarak dolaylı olarak ölçülmektedir. Bu ölçüm işlemine de performans değerlendirme denilmektedir (Tekin, 2009).

Performans değerlendirmede öğrenciden bilgi ve becerisini bir ürün oluşturarak göstermesi istenmektedir (Stiggins, 1994). Böylelikle bilgilerin ne derece edinildiğinin yanı sıra edinilen bilgilerin yeni durumlara ya da gerçek yaşam durumlarına ne derece aktarıldığı da ölçülebilmektedir (Biemer, 1993). Performans değerlendirmede kullanılacak yaygın ölçme araçlarından birisi de Dereceli Puanlama Anahtarıdır [DPA]. DPA, her bir performans için ölçütleri listeleyen ve performansta nelerin olması gerektiğini gösteren bir puanlama aracıdır (Popham, 1997). DPA'da, öğrenenden yerine getirmesi gereken performans ve bu performansın değerlendirilmesinde hangi davranışların daha önemli olduğu açıklanmaktadır (Arter ve McTighe, 2000).

DPA, sürece ve sonuca ilişkin değerlendirmeye imkan veren bir ölçme aracıdır. Süreç boyunca öğrenenler öğrenme hedefine ulaşmada neler yapmaları gerektiğini DPA'yı kullanarak belirleyebilirler. Diğer bir deyişle bu ölçümlerde öğretmen öğrencileri uygulama öncesinde haberdar etmektedir. Süreç sonunda ise öğretmen, ortaya çıkan ürünü belirlenen ölçütlere göre değerlendirebilmektedir (Mertler, 2001; Popham, 1997). Ayrıca, iyi yapılandırılmış DPA'lar, ölçme işleminin puanlayıcıdan ve zamandan bağımsız hale getirmesinden dolayı geçerli ve güvenilir ölçme araçlarıdır (Moskal ve Leydens, 2000). Bu nedenle öğrenenlerin performanslarını değerlendirme için DPA kullanımı puanlayıcı kaynaklı ölçme hatalarını asgari düzeye indirerek ölçme işlemini daha güvenilir kılacaktır (Dunbar, Brooks ve Miller, 2006).

DPA üç bölümden oluşmaktadır (Popham, 1997). Bunlar;

- ✓ *Değerlendirme Ölçütleri:* Öğrencilerin kabul edilebilir performansları ile kabul edilemez performanslarını birbirinden ayırmak için kullanılır.
- ✓ *Ölçüt Tanımlamaları:* Öğrencilerin performanslarındaki niteliksel farklılıkları belirtme biçimi ifade eder.
- ✓ *Puanlama Stratejisi:* Değerlendirmenin amacına bağlı olarak performansın değerlendirilmesinde bütüncül ya da analitik DPA kullanılabilir. Değerlendirilecek performans daha alt boyutlara ayrıştırılabildiğinde ve daha detaylı puanlama yapmak istenildiğinde analitik DPA kullanılabilir (Haladyna, 1997). Performans daha alt boyutlara ayrıştırılmıyorsa bütüncül DPA kullanılabilir (Brookhart, 1999).

Problem Durumu

Kavram karikatürü kullanımı öğrenme ve öğretme süreçlerindeki önemini kanıtlamıştır (Koutnikova, 2017). İlgili alan yazın incelendiğinde başta matematik eğitiminde olmak üzere diğer alanlarda da kavram karikatürü sıkça kullanılmaktadır. Samková (2020), kavram karikatürlerini ilköğretim öğretmenlerinin toplama işlemine ilişkin pedagojik alan bilgilerini belirlemede kullanmıştır. Yılmaz

(2020), ilkokul dördüncü sınıf öğrencileri ile yaptığı çalışmasında kavram karikatürü kullanımının öğrencilerin “Yerkabuđu ve yerin hareketleri” konusundaki akademik başarılarını artırmada olumlu katkılarının olduğunu görmüştür. Atasoy (2020) ortaokul 5. ve 8. sınıf öğrencilerinin “epistemolojik inançlarını” kavram karikatürleri kullanarak tespit edip karşılaştırmıştır. Serttaş ve Türkođlu (2020), ortaokul 7. sınıf öğrencilerinin “Astronomi” konusundaki kavram yanlışlarını kavram karikatürleri kullanarak tespit etmişlerdir. Aygün, Karadeniz ve Bütüner (2020), kavram karikatürü uygulamalarının 5. sınıf öğrencilerini matematik dersinde aktifleştirdiđini, fikir alışverişinde bulunabildikleri, matematiksel dil becerilerinin geliştirdiđi ve böylelikle kendi ifadeleriyle matematiksel cümleler kurabilmelerini sağladığını belirtmişlerdir. Karaca, Kuzu ve Çalışkan (2020), kavram karikatürü kullanımının 7. sınıf öğrencilerinin matematik dersine olan dikkat ve ilgilerini artırmanın yanı sıra çokgenler konusundaki başarılarına da olumlu etkilerinin olduğunu belirtmişlerdir. Balım, İnel-Ekici ve Özcan (2016), kavram karikatürü destekli öğretimin sorgulama becerilerini geliştirdiđini belirtmişlerdir. Bu araştırmaların bir kısmında, geçerlik ve güvenilirliğine dair açıklama yapılmadan araştırmacılar tarafından geliştirilen kavram karikatürlerinin kullanıldığı belirtilmiştir (Atasoy, 2020; Serttaş ve Türkođlu, 2020; Yılmaz, 2020). Aygün, Karadeniz ve Bütüner (2020), Balım, İnel-Ekici ve Özcan (2016), Karaca, Kuzu ve Çalışkan (2020) ile Samková (2020) ise başka araştırmacılar tarafından geliştirilen kavram karikatürlerini kullanmıştır. Bu nedenle geliştirilen kavram karikatürlerinin önyargılardan ve kişisel yargılardan bağımsız olarak yeterliğini değerlendirmenin çözülmesi gereken bir problem olduğu görülmektedir. Söz konusu bu araştırmada “Kavram karikatürlerinin değerlendirilmesi için oluşturulan dereceli puanlama anahtarı geçerli ve güvenilir midir?” sorusuna cevap aranacaktır.

Araştırmanın Amacı ve Önemi

Kavram karikatürlerinin değerlendirilmesinde kullanılacak ölçme aracının önyargıdan ve kişisel yargılamadan bağımsız derecelendirmeler verebilmesi önemlidir. Böylece kavram karikatürlerinde dair sağlam ve geçerli yorumlara ulaşılabilir. Kavram karikatürlerinin değerlendirilmesine yönelik yapılan literatür incelemesinde sadece Şaşmaz-Ören’in (2009) fen bilgisi öğretmen adaylarının kavram karikatürü oluşturma becerilerinin dereceli puanlama anahtarıyla değerlendirdiđi araştırması ile karşılaşılmıştır. Matematik eğitiminde araştırmacıların kullandıkları kavram karikatürlerinin yeterliğinin değerlendirilebilmesi için de geçerli ve güvenilir bir dereceli puanlama anahtarının oluşturulmasının gerekli olduğuna kanaat getirilmiştir. Bu bağlamda bu araştırmanın amacı; matematik eğitiminde hazırlanacak kavram karikatürlerinin değerlendirilmesinde kullanılabilecek bir dereceli puanlama anahtarının geliştirilmesidir. Literatürde önemli bir eksikliği giderecek olan araştırmanın bu bağlamda önemli olduğu düşünülmektedir.

Yöntem

Araştırma Deseni

Bu çalışma, kavram karikatürlerini değerlendirmeye yönelik bir dereceli puanlama anahtarının geliştirildiđi betimsel bir çalışmadır. Betimsel araştırmalar bir davranışın düzeyini tespit etmek amacıyla da kullanılabilir (McMillan ve Schumacher, 2014). Bu çalışmada da sonraki çalışmalarda kullanılabilecek bir dereceli puanlama anahtarı geliştirildiğinden betimsel araştırma modeli tercih edilmiştir.

Çalışma Grubu

Bu çalışma betimsel nitelikte olduğundan evren ve örneklem yerine araştırmanın desenine uygun olarak bir çalışma grubu belirlenmiştir. Erkuş (2012) ölçme araçlarının geliştirildiği çalışmalarda amaçlı örnekleme yönteminin kullanılmasını önermektedir. Bu çalışmada amaçlı örnekleme yöntemlerinden kolay ulaşılabilir örnekleme yöntemi tercih edilmiştir. Plano-Clark ve Creswell (2015), amaçlı örneklemede araştırmacıların çalışma için en uygun kişileri göz önüne alarak katılımcıları bir amaç doğrultusunda kasten seçtiklerini belirtmişlerdir. Kolay ulaşılabilir örnekleme yöntemi ise araştırmaya hız ve pratiklik kazandırır. Ayrıca araştırmacının erişilmesi kolay ve yakın olan durumları seçmesine imkan tanır (Yıldırım ve Şimşek, 2013). Bu bağlamda araştırmanın verileri araştırmanın amacına uygunluk, uygulanabilirlik ve kolay ulaşılabilirlik dikkate alınarak Sivas il merkezinde görev yapan ve kavram karikatürü hazırlama ile ilgili eğitim alan altı matematik öğretmeninin herhangi bir konu/öğrenme alanı kısıtlaması olmadan hazırladıkları 36 adet kavram karikatüründen elde edilmiştir. Puanlayıcılar ise matematik eğitimi alanında yüksek lisans yapmış ve kavram karikatürü hazırlama ile ilgili eğitim almış iki uzmandan oluşmaktadır. Çalışmaya katılan öğretmenler Öğrt1, Öğrt2, ... Öğrt6 şeklinde kodlanırken puanlayıcılar P1 ve P2 şeklinde kodlanmıştır. Çalışma grubuna ait demografik özellikleri Tablo 1’de yer almaktadır.

Tablo 1. Çalışma grubu demografik bilgileri

Öğretmenler	Cinsiyet	Yaş	Eğitim Durumu	Mesleki Deneyim
Öğrt1	K	36	Yüksek Lisans	12
Öğrt2	E	27	Lisans	5
Öğrt3	E	34	Yüksek Lisans	11
Öğrt4	E	28	Lisans	6
Öğrt5	E	29	Lisans	5
Öğrt6	E	34	Lisans	10
P1	E	33	Yüksek Lisans	9
P2	K	28	Yüksek Lisans	6

Tablo 1 incelendiğinde öğretmenlerin farklı eğitim durumlarına ve mesleki deneyimlere sahip kişilerden belirlendiği görülmektedir. Puanlayıcılardan ise biri 9 diğeri 6 yıllık mesleki deneyime sahiptir.

Veri Toplama Araçları

Araştırmada veri toplama aracı olarak; öğretmenlerin oluşturdukları kavram karikatürleri ve bu karikatürleri değerlendirmek için hazırlanan dereceli puanlama anahtarı kullanılmıştır.

Kavram Karikatürleri

Kavram karikatürü hazırlama ile ilgili eğitim alan öğretmenlerden, herhangi bir sınıf düzeyi ya da öğrenme alanı kısıtlaması olmadan kavram karikatürü çizmeleri istenmiştir. Öğretmenler toplam 36

adet kavram karikatürü çizmişlerdir. Çizilen bu kavram karikatürlerinin öğrenme alanlarına göre dağılımı Tablo 2’de yer almaktadır.

Tablo 2. Öğretmenlerce çizilen kavram karikatürlerinin öğrenme alanları

Öğrenme Alanları	f	%
Sayılar ve İşlemler	10	28
Cebir	8	22
Geometri ve Ölçme	7	19
Veri İşleme	6	17
Olasılık	5	14
Toplam	36	100

Tablo 2 incelendiğinde öğretmenlerin en çok sayılar ve işlemler öğrenme alanında olmak üzere tüm öğrenme alanlarında kavram karikatürü hazırladıkları görülmektedir.

Dereceli Puanlama Anahtarının Geliştirilmesi

Araştırma verilerini elde etmek için kavram karikatürlerinin değerlendirilmesinde kullanılacak bir dereceli puanlama anahtarı oluşturulmuştur. Bu dereceli puanlama anahtarının oluşturulmasında Goodrich (2001) tarafından önerilen adımlar takip edilmiştir. Bu adımlara ilişkin detaylı açıklamalar aşağıda yer almaktadır.

1. adım, beklenen performansa göre DPA’nın kullanım amacının belirlenmesi: Oluşturulan DPA’ların amacı matematik eğitimi alanında hazırlanan kavram karikatürlerini değerlendirmektir.

2. adım, DPA türünün belirlenmesi: Analitik ve bütünsel olmak üzere iki çeşit DPA vardır. Analitik DPA daha detaylı incelemeye imkan sunmasının yanı sıra puanlayıcılar arasında güvenilirliği de sağlamaktadır (Knoch, 2009; Mertler, 2001). Ayrıca DPA’nın türüne karar verirken değerlendirmenin amacı da göz önünde bulundurulmuştur. Kavram karikatürünün içerisinde birçok öge bulundurması da dikkate alınarak matematik eğitimi alanında hazırlanan kavram karikatürlerini değerlendirebilmek için analitik DPA hazırlanmasına karar verilmiştir.

3. adım, beklenen performansa göre puanlanacak özelliklerin (ölçüt) belirlenmesi: Literatür incelenerek kavram karikatürlerinin değerlendirilmesinde kullanılan ölçme araçları incelenmiştir. Şaşmaz-Ören (2009) fen bilgisi öğretmenlerinin hazırladıkları kavram karikatürlerini değerlendirmek için geliştirdiği DPA’da ölçütleri içerik bilgisi, problemin sunumu, çözüm önerileri, özgünlük ve ilgi çekicilik, bilimsel dil kullanımı ve düzen şeklinde belirlemiştir. Kavram karikatürünün hazırlanması adımlarını açıklayan araştırmalar (Dabell, 2008; Duban, 2013; Duban, Aydođdu ve Evrekli, 2015; Göksu, 2014) da incelenmiştir. Sonuç olarak kavram karikatürünün yapısı da dikkate alınarak DPA’da yer alacak ölçütler; (i) İçerik, (ii) Problem Durumu, (iii) Alternatif Fikirler, (iv) Metin Yazımı ve (v) Tasarım olacak şekilde belirlenmiştir.

4. adım, ölçütler için başarı düzeylerinin belirlenmesi ve puanlandırılması: Haladyna (1997) güvenilir bir puanlama için dörtlü derecelendirme önermektedir. Şaşmaz-Ören (2009) “önemli eksiklikleri olan performansa” bir puan verirken “mükemmel performansa” da dört puan vermiştir.

Benzer şekilde Demir ve Yıldırım (2019) “başlangıç düzeyinde olan performans” bir puan verirken “tam olarak başarılı performans” dört puan vermişlerdir. Bu bağlamda, grafik materyallerin hazırlanmasında göz önünde bulundurulması gereken her bir ölçüt için performans düzeyleri ve puanları; Başlangıç düzeyinde performans (1 puan), Geliştirilmesi gereken performans (2 puan), Başarılı performans (3 puan) ve Mükemmel performans (4 puan) olacak şekilde düzenlenmiştir.

5. adım, ölçütlerin her bir başarı düzeyi için performans tanımlarının yapılması: Ölçütlerin özellikleri de dikkate alınarak tüm başarı düzeyleri için performans tanımlamaları yapılmıştır. Bu tanımlamalar, Durmaz (2007), Kabapınar (2005), Keogh ve Naylor (1999) ve Kılınç’ın (2008) çalışmalarında belirttikleri, kavram karikatürü hazırlanmasında dikkat edilmesi gereken hususlar göz önüne alınarak hazırlanmıştır. Başlangıç düzeyi önemli eksiklikleri olan performans olarak tanımlanırken eksiksiz ve örnek gösterilebilecek nitelikte olan performanslar mükemmel performans olarak tanımlanmıştır.

6. adım, uzman görüşlerinin alınması: Hazırlanan DPA’lar hakkında bir Türkçe ve iki matematik eğitimcisiinden uzman görüşü alınmıştır. Uzmanlardan DPA’ya ilişkin içerik, yapı ve ölçüt bakımından “yeterli, kısmen ve yeterli değil” kategorilerine göre incelemeleri ve varsa düzeltme önerilerini belirtmeleri istenmiştir.

7. adım, geliştirilen puanlama anahtarının pilot çalışması: Geliştirilen DPA’nın pilot uygulaması, matematik eğitimi literatüründe yer alan kavram karikatürlerinin değerlendirilmesiyle gerçekleştirilmiştir.

8. adım, puanlama anahtarının geçerlik ve güvenilirlik çalışmaları: Pilot çalışma sonucunda elde edilen veriler kullanılarak DPA’nın geçerlik ve güvenilirlik çalışmaları yapılmıştır. Geçerlik çalışmaları ve güvenilirlik çalışmasına dair detaylı bilgiler bulgular kısmında yer almaktadır.

Veri Toplama Süreci

Araştırma kapsamında öncelikle öğretmenler tarafından hazırlanan kavram karikatürleri birinci yazarın doktora tez çalışması kapsamında 2019-2020 eğitim öğretim yılı bahar döneminde toplanmıştır. Toplanan kavram karikatürleri yine 2019-2020 eğitim öğretim yılında puanlayıcılar tarafından değerlendirilmiştir.

Veri Analizi

Geçerlik çalışmaları kapsamında, DPA hakkında uzman değerlendiricilerden gelen görüşler arasındaki uyum Miles ve Huberman (1994) tarafından önerilen uyum yüzdesi formülü ile hesaplanmıştır.

$$\text{Uyum Yüzdesi} = \frac{\text{Uzlaşma sayısı}}{\text{Uzlaşma sayısı} + \text{Uzlaşmama sayısı}}$$

Geçerli bir uyum yüzdesinin %70’in üzerinde olması gerekmektedir (Miles and Huberman, 1994).

Geçerlik çalışması kapsamında değerlendiriciler arasındaki uyum yüzdesinin hesaplanmasından sonra daha detaylı bilgiler elde etmek için kapsam geçerlik indeksi (KGI) de hesaplanmıştır. DPA’nın kapsam geçerliği, en az üç uzmandan görüş alınmasını öneren “Davis tekniği” kullanılarak hesaplanmıştır. Bu teknikte uzman görüşleri ilgili madde tamamen uygun ise 4, bazı ufak değişiklikler

gerektiriyorsa 3, çok fazla deęişiklik gerektiriyorsa 2 ve uygun deęil ya da tamamen deęişiklik gerektiriyorsa 1 puan verilerek derecelendirilmektedir. KGİ, ilgili maddeye 3 ve 4 puan veren uzman sayısının toplam uzman sayısına bölünmesiyle bulunur ve bu deęerin 0.80'den büyük olması gerekmektedir (Yurdugöl, 2005).

DPA'nın güvenilirliğini denemek için birden fazla istatistiksel yöntem kullanılabilir. Bu istatistiksel yöntemler, uyum düzeyini toplam puan üzerinden ya da her ölçüt için ayrı ayrı incelemeleri; iki puanlayıcı arasındaki ya da ikiden fazla puanlayıcı arasındaki uyum düzeyine bakmaları; puanlayıcıların puanlarının normal dağılım gösterip göstermemeleri açısından birbirinden farklılık göstermektedir (Büyüköztürk, 2007; Şencan, 2005).

Bir performansın değerlendirilmesinde en önemli hata kaynağının puanlayıcılar olmasından hareketle güvenilirlik çalışması kapsamında puanlayıcılar arasındaki güvenirlığe bakılmıştır. Puanlayıcılar arası güvenilirlik, puanlayıcıların performansı değerlendirirken farklı kişilere ve maddelere verdikleri puanlar arasındaki tutarlılığı ifade etmektedir (Atılğan, Kan ve Dođan, 2011). Puanlayıcılar arasındaki güvenilirlik hem Klasik Test Kuramı (KTK) hem de Madde Tepki Kuramı (MTK) kapsamında değerlendirilmiştir. KTK kapsamında iki puanlayıcının DPA'ya verdikleri toplam puanlar arasındaki tutarlık düzeyi "Spearman Korelasyon Katsayısı" ile belirlenirken her bir alt ölçütteki tutarlılık düzeyi de "Cohen's Kappa" ile belirlenmiştir. Spearman korelasyonu (Rho) 0.70'ten büyükse burada yüksek düzeyde bir ilişkinin olduğu söylenebilir. Cohen's kappa deęeri için ise 0.00 - 0.20 arası önemsiz uyuşma; 0.21 - 0.40 arası orta derecede uyuşma; 0.41 - 0.60 arası kabul edilebilir uyuşma; 0.61 - 0.80 arası önemli derecede uyuşma ve 0.81 - 1.00 arası mükemmel uyuşma olarak ifade edilir (Büyüköztürk, 2007). KTK kapsamındaki analizler SPSS paket programından yararlanılarak gerçekleştirilmiştir.

MTK kapsamında ise Çok Yüzeysel Rasch Ölçme Modeli (ÇYRÖM) dikkate alınarak puanlayıcı, puanlanan (kavram karikatürü) ve madde yüzeysel için güvenilirlik hesaplamaları yapılmıştır. ÇYRÖM kapsamındaki analizler FACET (Linacre, 2014) paket programından yararlanılarak gerçekleştirilmiştir.

Geçerlik ve Güvenirlik / İnanırdıcılık Sağlama Amacıyla Alınan Önlemler

Çalışma grubunun farklı demografik özelliklere sahip kişilerden oluşmasına ve böylece araştırma probleminin yeterince incelenmesine çalışılmıştır. Veri toplama araçları uzman kişilerce incelenmiş ve onlardan gelen dönütler dikkate alınarak gerekli revizyonlar yapılmıştır. Araştırmanın problem durumu, amacı, önemi, yöntemi, veri toplama araçları ve veri analizi detaylı bir şekilde açıklanarak her birinin birbiriyle uyumlu ve tutarlı olmasına dikkat edilmiştir.

Sınırlılıklar

Bu araştırma kavram karikatürü hazırlama ile ilgili eğitim alan matematik öğretmenlerinin oluşturdukları kavram karikatürleri ve bu kavram karikatürlerinin matematik eğitimi alanında yüksek lisans yapmış ve kavram karikatürü hazırlama ile ilgili bilgisi olan puanlayıcıların değerlendirmeleri ile sınırlıdır.

Araştırmacının Rolü

Araştırmacı taslak DPA'nın geliştirilmesi, taslak metin hakkında uzman görüşlerinin alınıp değerlendirilmesi ve nihayetinde DPA'nın geçerlik ve güvenilirlik çalışmalarını yerine getirmiştir.

Etik Konular

Öğretmenlerin ve puanlayıcıların kurum ve kimlik bilgilerinin gizli olması için öğretmenler Öğrt1, Öğrt2, ..., Öğrt6 ve puanlayıcılar P1 ve P2 şeklinde kodlanmıştır. Katılımcıların çalışmaya gönüllü olarak katılabilecekleri ve istedikleri zaman çalışmadan ayrılacakları de belirtilmiştir.

Bulgular

Geçerlik ile İlgili Bulgular

Düzenlenen Analitik Dereceli Puanlama Anahtarı, biri Türkçe öğretmeni üçü matematik eğitimi uzmanı olmak üzere dört kişi tarafından incelenmiştir. Bu incelemede uzman değerlendirme formu kullanılmıştır. Uzmanların DPA hakkındaki görüşleri içerik, yapı ve ölçüt bağlamında tek tek incelenmiştir. Bunların DPA'ya uygunluğu ve yeterliği konusunda uzmanların görüşleri arasındaki uyum yüzdelerinin %82 ile %95 arasında değiştiđi belirlenmiştir. Hesaplanan bu uyum yüzdeleri, geliştirilen DPA'nın içerik, yapı ve ölçüt bakımından geçerli olduğunu kanıtlar değerdedir.

Geçerliğe dair daha detaylı bilgiler elde etmek için DPA'nın kapsam geçerlik indeksi (KGİ) de hesaplanmıştır. DPA için hesaplanan KGİ Tablo 3'te yer almaktadır.

Tablo 3. Puanlama anahtarı kapsam geçerlik indeksleri (KGİ) (sonra tamamla)

Boyutlar	Sorular	KGİ
İçerik	Değerlendirme ölçütleri performansın tüm yönlerini kapsıyor mudur?	1.00
	Değerlendirme ölçütleri performans dışı bir içeriđe sahip midir?	1.00
Yapı	Puanlama ölçütleriyle değerlendirilen performansın tüm önemli öğeleri mevcut mudur?	1.00
	İlgili performansla ilişkili olmayan herhangi bir değerlendirme ölçütü var mıdır?	1.00
Ölçüt	Değerlendirilecek ilgili performansın önemli öğelerine yer verilmiş midir?	1.00
	Puanlama ölçütleri, ilgili performansın öğelerini doğru olarak yansıtmakta mıdır?	1.00
	Puanlama ölçütleri ilgili performansın önemli öğelerini doğru ölçmekte midir?	1.00

Tablo 3'te görüldüğü üzere DPA'nın kapsam geçerlik indeksi 0.80'nin üzerindedir. KGİ'ler eşik değeri olan 0.80'den büyük oldukları için geliştirilen DPA'nın kavram karikatürlerini değerlendirmek için geçerli olduğuna kanaat getirilmiştir.

Uzmanlardan gelen öneriler dikkate alınarak kavram karikatürünün herhangi bir alt ölçütünde herhangi bir çizim/yazım/açıklama yapılmamışsa bu ölçüt için sıfır puan verileceđi DPA'ya eklenmiştir.

Güvenirlilik ile İlgili Bulgular

Geliştirilen DPA'nın güvenirliliđine dair çalışmalarda Klasik Test Kuramı (KTK) ve Madde Tepki Kuramı (MTK) kullanılmıştır. KTK bağlamında iki puanlayıcının DPA'ya verdikleri toplam puanlar arasındaki tutarlılığın belirlemesi için yapılan Spearman korelasyon katsayısı analizine ilişkin sonuçlar Tablo 4'te yer almaktadır.

Tablo 4. DPA'nın toplam puanlarına ilişkin spearman korelasyon sonuçları

Birinci değerlendiricinin puanları	İkinci değerlendiricinin puanları	
	Spearman's Rho	
		,856*
	P	,000
	N	36

Tablo 4 incelendiğine kavram karikatürlerinin değerlendirilmesinde kullanılacak olan DPA'ya iki değerlendiricinin verdiği toplam puanlar arasında pozitif, anlamlı ve yüksek ($Rho=0.856$, $p<0.01$) bir ilişkinin olduğu görülmektedir. Puanlayıcılar arasındaki bu tutarlılık, hazırlanan DPA'nın kavram karikatürlerinin değerlendirilmesinde yeterli düzeyde güvenilir olduğunu göstermektedir.

Ayrıca iki farklı puanlayıcının DPA'nın alt ölçütlerine verdikleri puanlar arasındaki uyum düzeyini belirlemek için de Cohen's Kappa analizi yapılmış ve elde edilen sonuçlara Tablo 5'te yer verilmiştir.

Tablo 5. DPA'nın alt ölçüt puanlarına ilişkin Cohen's Kappa sonuçları

Alt Ölçütler	Cohen's Kappa
İçerik	0.797
Problem Durumu	0.772
Alternatif Fikirler	0.774
Metin Yazımı	0.706
Tasarım	0.741

Tablo 5'te DPA'nın alt ölçütlerine ilişkin Cohen's Kappa sonuçları 0.706 ve 0.797 arasında yer almaktadır. Bu değerler de 0.61 ile 0.80 arasında yer aldığı için "önemli derecede uyuma" olarak yorumlanabilir. Bu da her bir alt ölçütte iki puanlayıcı arasındaki uyumun yüksek olduğunu ve dolayısıyla geliştirilen DPA'nın güvenilir olduğuna başka bir kanıttır.

ÇYRÖM dikkate alınarak puanlayıcı yüzeyine ait güvenilirlik indeksi 0.87, puanlanan (kavram karikatürü) yüzeyine ait güvenilirlik indeksi 0.80 ve madde yüzeyine ait güvenilirlik indeksi 0.89 olarak hesaplanmıştır. Bu değerler geliştirilen DPA'nın kavram karikatürlerinin değerlendirilmesi güvenilir bir ölçme aracı olduğunu göstermektedir.

DPA'nın toplam puan için hesaplanan Spearman korelasyonu (Rho) ve alt ölçüt puanları için hesaplanan Cohen's Kappa değerleri ile puanlayıcı, puanlanan (kavram karikatürü) ve madde yüzeylerine ait güvenilir indeksleri bir arada değerlendirildiğinde oluşturulan DPA'nın güvenilir bir puanlama anahtarı olduğu görülmüştür.

Sonuç, Tartışma ve Öneriler

Bu araştırmada matematik konularında hazırlanan kavram karikatürlerini değerlendirmek için analitik bir dereceli puanlama anahtarı geliştirilerek geçerlik ve güvenilirlik çalışmaları yapılmıştır. Söz konusu DPA Ek-1 de yer almaktadır.

Hazırlanan DPA, kavram karikatürü hazırlama becerisini değerlendirmek üzere beş ölçütten oluşmaktadır. Uzmanların görüşlerine göre DPA'nın içeriği, yapısı ve ölçütleri kavram karikatürü hazırlama becerisini değerlendirmede uygun ve yeterlidir. Ayrıca, uzman görüşleri arasındaki yüksek

tutarlılık hazırlanan DPA'nın geçerli olduğunu da göstermektedir. Geçerliđi daha detaylı incelemek için kapsam geçerlik indeksleri de hesaplanmıştır. Bu deđerler de kabul edilebilir düzeyde olup geliştirilen DPA'nın geçerliđine başka bir kanıttır.

İki deđerlendiricinin DPA'ya verdikleri puanlar arasında hem toplam hem de alt ölçütler bağlamında pozitif, anlamlı ve yüksek bir ilişkinin olduğu görülmüştür. Ayrıca Çok Yüzeyle Rasch Ölçme Modeli kapsamında hesaplanan puanlayıcı, puanlanan (kavram karikatürü) ve madde yüzeyle için güvenilirlik hesaplamaları da kabul edilebilir düzeydedir. Bu sonuçlar oluşturulan DPA'nın güvenilir bir puanlama anahtarı olduğu göstermekle birlikte puanlayıcıların tamamen uyumlu olmadığını da göstermektedir. Ancak DPA'nın puanlayıcılar arasındaki tutarlıđı artırarak objektifliđi artırdığı sonucuna varılabilir. Benzer çalışmalarda da DPA puanlayıcıları arasında yüzde yüz uyum olmasa da yüksek bir uyumun olduğu belirtilmiştir (Bilgen ve Dođan, 2017; Büyükkıdık ve Anıl, 2015; Knoch, 2009).

Puanlayıcılar arasındaki farklılık puanlayıcıların kavram karikatürü hakkındaki bilgi düzeyleri, algıları ve yorumlamalarına göre de deđişiklik göstermiş olabilir. Buna puanlayıcı etkisi denilmektedir (Cronbach, 1990). Çeşitli araştırmalarda puanlayıcının katılıđı ya da cömertliđi gibi puanlayıcı etkilerini ortadan kaldırmada DPA'nın yetersiz olduğu belirtilmektedir (Güler ve Gelbal, 2011; İlhan, 2015). Puanlayıcıların ölçütleri farklı algılamış olmaları da puanlar arasındaki farklılıđın nedeni olabilir. Bilgen ve Dođan (2017) yaptıkları çalışmada geliştirdikleri DPA'nın güvenilirliđini incelerken puanlayıcılar ile ölçütler arasındaki etkileşimin puanlar arasındaki farklılıđa neden olabildiđini vurgulamışlardır.

Bu araştırmanın bulgularından hareketle uygulayıcılara ve diđer araştırmacılara bazı önerilerde bulunulmuştur. Geliştirilen DPA iki puanlayıcının verdikleri puanlar dikkate alınarak yapılmıştır. Daha fazla sayıda puanlayıcının yer aldığı ve bu nedenle farklı analiz tekniklerinin kullanılması gereken güvenilirlik çalışmaları yapılabilir. Geliştirilen DPA, matematik öğretmenlerinin oluşturdukları kavram karikatürlerinin deđerlendirilmesinde kullanıldığı için öğretmen adaylarının hazırladıkları kavram karikatürlerinin deđerlendirilmesinde de kullanılabilir olup olmadığı araştırılabilir. Matematik konularında hazırlanmış kavram karikatürlerinin analizinde kullanılan bu DPA revize edilerek farklı alanlardaki kavram karikatürlerinin deđerlendirilmesi için de kullanılabilir.

Kaynakça

- Arter, J. A. and McTighe, J. (2000). *Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance*. California: Corwin Press.
- Atasoy, S. (2020). Using concept cartoons to identify the epistemological beliefs of middle school students. *Journal of Science Learning*, 3(3), 165-173. <https://ejournal.upi.edu/index.php/jslearning/article/view/60> adresinden 10.11.2020 tarihinde erişilmiştir.
- Atılğan, H., Kan, A. ve Doğan, N. (2011). *Eğitimde ölçme ve değerlendirme*. Ankara: Anı Yayıncılık.
- Aygün, D., Karadeniz, M. H. ve Bütüner, S. Ö. (2020). Kavram karikatürü uygulamalarının 5. sınıf öğrencilerinin matematiksel sembol, terim/kavram kullanımına yansımaları. *International Journal of Educational Studies in Mathematics*, 7(3), 151-172. <https://dergipark.org.tr/tr/pub/ijesim/issue/56835/749497> adresinden 8.11.2020 tarihinde erişilmiştir.
- Balım, A.G., İnel, D. ve Evrekli, E. (2008). Fen öğretiminde kavram karikatürü kullanımının öğrencilerin akademik başarılarına ve sorgulayıcı öğrenme becerileri algılarına etkisi. *İlköğretim Online*, 7(1), 188-202. <http://ilkogretim-online.org//index.php?iid=2008-7-1.000&&jid=218&lng> adresinden 8.03.2020 tarihinde erişilmiştir.
- Balım, A. G., İnel-Ekici, D. and Özcan, E. (2016). Concept cartoons supported problem based learning method in middle school science classrooms. *Journal of Education and Learning*, 5(2), 272-284. <https://files.eric.ed.gov/fulltext/EJ1097432.pdf> adresinden 10.01.2020 tarihinde erişilmiştir.
- Biemer, L. (1993). Authentic assessment. *Educational Leadership*, 50(8), 81-82.
- Demir, S. B. ve Yıldırım, Ö. (2019). Yazılı anlatım becerilerinin değerlendirilmesi için dereceli puanlama anahtarı geliştirme çalışması. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 47, 457-473. <https://dergipark.org.tr/tr/pub/pauefd/issue/48574/588565> adresinden 18.11.2020 tarihinde erişilmiştir.
- Bilgen, B.Ö. ve Doğan, N. (2017). Puanlayıcılar arası güvenilirlik belirleme tekniklerinin karşılaştırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 8(1), 63-78. <https://dergipark.org.tr/tr/pub/epod/issue/28110/294847> adresinden 15.03.2020 tarihinde erişilmiştir.
- Brookhart, S. M. (1999). *The art and science of classroom assessment: The missing part of pedagogy*. ASHE-ERIC Higher Education Report 27(1). <https://eric.ed.gov/?id=ED432937> adresinden 12.03.2020 tarihinde erişilmiştir.
- Büyükkıdık, S. ve Anıl, D. (2015). Performansa dayalı durum belirlemede güvenilirliğin genellenebilirlik kuramında farklı desenlerde incelenmesi. *Eğitim ve Bilim*, 40(177), 285-296. <http://egitimvebilim.ted.org.tr/index.php/EB/article/view/2454> adresinden 11.03.2020 tarihinde erişilmiştir.
- Büyüköztürk, Ş. (2007). *Sosyal bilimler için veri analizi el kitabı*. Ankara: PegemA Yayıncılık.
- Cronbach, L.I. (1990). *Essentials of psychological testing*. New York: Harper and Row.
- Dabell, J. (2008). Using concept cartoons. *Mathematics Teaching Incorporating Micromath*, 209, 34-37. <https://eric.ed.gov/?id=EJ815105> adresinden 05.02.2020 tarihinde erişilmiştir.
- Demirci, F. and Özyürek, C. (2017). The effects of using concept cartoons in astronomy subjects on critical thinking skills among seventh grade student. *International Electronic Journal of Elementary Education*, 10(2), 243-254. <https://www.iejee.com/index.php/IEJEE/article/view/369> adresinden 25.02.2020 tarihinde erişilmiştir.
- Duban, N. Y. (2013). Sınıf öğretmenlerinin kavram karikatürlerini hazırlamaya ve kullanmaya yönelik görüşleri. *Akademik Araştırmalar Dergisi*, 56, 35-54.
- Duban, N.Y., Aydoğdu, B. ve Evrekli, E. (2015). Sınıf öğretmeni adaylarının fen ve teknoloji öğretimi-I dersinde hazırladıkları kavram karikatürlerinin değerlendirilmesi. *Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi Dergisi*, 26, 100-120. <https://dergipark.org.tr/tr/pub/zgefd/issue/47934/606371> adresinden 25.02.2020 tarihinde erişilmiştir.
- Dunbar, N.E., Brooks, C.F. and Miller, K.T. (2006). Oral communication skills in higher education: Using a performance-based evaluation rubric to assess communication skills. *Innovative Higher Education*, 31(2), 115-28. <https://link.springer.com/article/10.1007/s10755-006-9012-x> adresinden 24.02.2020 tarihinde erişilmiştir.

- Durmaz, B. (2007). *Yapılandırıcı fen öğretiminde kavram karikatürlerinin öğrencilerin başarıları ve duyuşsal özelliklerine etkisi (Muğla ili merkez ilçe örneği)*. (Yayımlanmamış yüksek lisans tezi). <https://tez.yok.gov.tr/UlusalTezMerkezi/tarama.jsp> adresinden 11.03.2020 tarihinde erişilmiştir.
- Erkuş, A. (2012). *Psikolojide ölçme ve ölçek geliştirme-I: Temel kavramlar ve işlemler*. Ankara: Pegem Akademi.
- Evrekli, E. (2010). *Fen ve teknoloji öğretiminde zihin haritası ve kavram karikatürü etkinliklerinin öğrencilerin akademik başarılarına ve sorgulayıcı öğrenme beceri algılarına etkisi*. (Yayımlanmamış doktora tezi). <https://tez.yok.gov.tr/UlusalTezMerkezi/tarama.jsp> adresinden 28.11.2020 tarihinde erişilmiştir.
- Goodrich, A. H. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, 57(5), 27-31. <https://eric.ed.gov/?id=EJ609600> adresinden 28.11.2020 tarihinde erişilmiştir.
- Göksu, F.C. (2014). *Doğrular, açılar ve çokgenler konularının kavram karikatür destekli yapılandırmacı öğrenme yaklaşımına göre işlenmesi*. (Yayımlanmamış Yüksek Lisans Tezi). <https://tez.yok.gov.tr/UlusalTezMerkezi/tarama.jsp> adresinden 11.03.2020 tarihinde erişilmiştir.
- Güler, N. ve Gelbal, S. (2010). Klasik test kuramı ve çok değişkenlik kaynaklı Rasch modeli üzerine bir çalışma. *Eğitim Araştırmaları-Eurasian Journal of Educational Research*, 38, 108-125. <http://ejer.com.tr/en/archives/2010-winter-issue-38> adresinden 14.03.2020 tarihinde erişilmiştir.
- Haladyna, T. M. (1997). *Writing test item to evaluate higher order thinking*. Michigan: Allyn & Bacon.
- İlhan, M. (2015). *Standart ve SOLO taksonomisine dayalı rubrikler ile puanlanan açık uçlu matematik sorularında puanlayıcı etkilerinin çok yüzeyli rasch modeli ile incelenmesi*. (Yayımlanmamış Doktora Tezi). <https://tez.yok.gov.tr/UlusalTezMerkezi/tarama.jsp> adresinden 11.03.2020 tarihinde erişilmiştir.
- Jamal, S.N.B., Ibrahim, N.H.B. and Surif, J.B. (2019). Concept cartoon in problem-based learning: A systematic literature review analysis. *Journal of Tecnology and Science Education*, 9(1), 51-58. <https://files.eric.ed.gov/fulltext/EJ1204838.pdf> adresinden 28.11.2020 tarihinde erişilmiştir.
- Kabapınar, F. (2005). Yapılandırmacı öğrenme sürecine katkıları açısından fen derslerinde kullanılabilecek bir öğretim yöntemi olarak kavram karikatürleri. *Kuram ve Uygulamada Eğitim Bilimleri Dergisi*, 5(1), 135-146.
- Kabapınar, F. (2009). What makes concept cartoons more effective?: Using research to inform practice. *Education and Science*, 34(154), 104-118. <http://eb.ted.org.tr/index.php/EB/article/download/552/39> adresinden 28.02.2020 tarihinde erişilmiştir.
- Karaca, Z., Kuzu, O. ve Çalışkan, N. (2020). Çokgenler konusunun öğretiminde kavram karikatürü kullanımının akademik başarıya etkisi. *Academia Eğitim Araştırmaları Dergisi*, 5(1), 110-125. <https://dergipark.org.tr/tr/pub/egitim/issue/53591/706531> adresinden 28.11.2020 tarihinde erişilmiştir.
- Keogh, B., Naylor, S. and Wilson, C. (1998). Concept cartoons: A newperspective on physics education. *Physics Education*, 33(4), 219-224. Adresinden <https://eric.ed.gov/?id=EJ57097> 15.01.2020 tarihinde erişilmiştir.
- Keogh, B. and Naylor, S. (1998). Teaching and learning in science using concept cartoons. *Primary Science Review*, 51, 14-16. adresinden <https://eric.ed.gov/?id=EJ565635> 15.11.2020 tarihinde erişilmiştir.
- Keogh, B. and Naylor, S. (1999). Concept cartoons, teaching and learning inscience: An evaluation. *International Journal of Science Education*, 21(4),431-446. <https://eric.ed.gov/?id=EJ584643> adresinden 18.01.2020 tarihinde erişilmiştir.
- Kılınç, A. (2008). *Öğretimde mizahi kavramaya dayalı bir materyal geliştirme çalışması: Bilim karikatürleri*. (Yayımlanmamış Doktora Tezi). <https://tez.yok.gov.tr/UlusalTezMerkezi/tarama.jsp> adresinden 11.03.2020 tarihinde erişilmiştir.
- Knoch, U. (2009). Diagnostic assessment of writing: A comparison of two rating scales. *Language Testing*, 26(20), 275-304. <https://eric.ed.gov/?id=EJ834136> adresinden 15.02.2020 tarihinde erişilmiştir.
- Koutnikova, M. (2017). The application of comics in science education. *Acta Educationis Generalis*, 7(3), 88-95. <https://content.sciendo.com/view/journals/atd/7/3/article-p88.xml> adresinden 28.11.2020 tarihinde erişilmiştir.
- Linacre, J.M. (2014). *A user's guide to FACETS Rasch-model computer programs*. <http://www.winsteps.com/a/facets-manual.pdf> adresinden 28.11.2020 tarihinde erişilmiştir.
- Long, S. and Marson, K. (2003). Concept cartoons. *Hand on Science*, 19(3), 22-23.
- McMillan, J. W. and Schumacher, S. (2014). *Research in education: Evidence-based inquiry*. Boston: Pearson.

- Mertler, C. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation*, 7(25). <https://eric.ed.gov/?id=EJ670687> adresinden 15.01.2020 tarihinde erişilmiştir.
- Miles, M. B. and Huberman, A. M. (1994). *Qualitative data analysis*. London: Sage Publications.
- Moskal, B. M. and Leydens, J. A. (2000). Scoring rubric development: Validity and reliability. *Practical Assessment, Research & Evaluation*, 7(10). <https://eric.ed.gov/?id=EJ638498> adresinden 18.02.2020 tarihinde erişilmiştir.
- Naylor, S., Keogh, B. and Downing, B. (2007). Argumentation and primary science. *Research in Science Education*, 37(1), 17-39. <https://link.springer.com/article/10.1007/s11165-005-9002-5> adresinden 28.02.2020 tarihinde erişilmiştir.
- Plano-Clark, V.L. and Creswell, J. W. (2015). *Understanding research: A consumer's guide*. Upper Saddle River, NJ: Pearson Education.
- Popham, W. J. (1997). What's wrong-and what's right-with rubrics. *Educational Leadership*, 55(2), 72-75. <https://eric.ed.gov/?id=EJ552014> adresinden 15.02.2020 tarihinde erişilmiştir.
- Roesky, H.W. and Kennepoh, L.D. (2008). Drawing attention with chemistry cartoons. *Journal of Chemical Education*, 85(10), 1355-1360. <https://pubs.acs.org/doi/abs/10.1021/ed085p1355> adresinden 28.11.2020 tarihinde erişilmiştir.
- Samkova, L. and Hospesova A (2015). *Using concept cartoons to investigate future teachers' knowledge*. Konrad Krainer; Nad'a Vondrová. Proceedings of CERME 9 - Ninth Congress of the European Society for Research in Mathematics Education, Prague, Czech Republic pp. 3241-324.
- Samková, L. (2020). Using Concept Cartoons to investigate future primary school teachers' pedagogical content knowledge on addition. *Quadrante*, 29(1), 36-51. <https://quadrante.apm.pt/index.php/quadrante/article/view/565> adresinden 28.11.2020 tarihinde erişilmiştir.
- Sertaş, S. and Türkoğlu, A. Y. (2020). Diagnosing students' misconceptions of astronomy through concept cartoons. *Participatory Educational Research*, 7(2), 164-182.
- Stiggins, R. J. (1994). *Student-centered classroom assessment*. New York: Macmillan Publishing Company.
- Şaşmaz-Ören, F. (2009). Öğretmen adaylarının kavram karikatürü oluşturma becerilerinin dereceli puanlama anahtarıyla değerlendirilmesi. *E-Journal of New World Sciences Academy*, 4(3), 994-1016. <https://dergipark.org.tr/tr/pub/nwsaedu/issue/19827/212429> adresinden 28.01.2020 tarihinde erişilmiştir.
- Şencan, H. (2005). *Sosyal ve davranışsal ölçümlerde güvenilirlik ve geçerlilik*. Ankara: Seçkin Yayıncılık.
- Tekin, H. (2009). *Eğitimde ölçme ve değerlendirme*. Ankara: Yargı Yayınevi.
- Uzoğlu, M., Yıldız, A., Demir, Y. ve Büyükkasap, E. (2013). Fen bilgisi öğretmen adaylarının ışıkla ilgili kavram yanılgılarının belirlenmesinde kavram karikatürlerinin ve açık uçlu soruların etkililiklerinin karşılaştırılması. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi*, 14(1), 367-388. <http://kefad.ahievran.edu.tr/Kefad/ArchiveIssues> adresinden 15.02.2020 tarihinde erişilmiştir.
- Van der Mark, M.H. (2011). *The use of narratives and concept cartoons in the professional development of teachers to achieve higher-order thinking skills and deep learning about the evolution of life and geological time*. (Yayınlanmamış Doktora Tezi). <https://www.uj.ac.za/library/research-support/Pages/theses-and-dissertations.aspx> adresinden 28.11.2020 tarihinde erişilmiştir.
- Yıldırım, A. ve Şimşek, H. (2013). *Sosyal bilimlerde nitel araştırma yöntemleri*. Ankara: Seçkin Yayıncılık.
- Yılmaz, M. (2020). Impact of instruction with concept cartoons on students' academic achievement in science lessons. *Educational Research and Reviews*, 15(3), 95-103. <https://eric.ed.gov/?id=EJ1250476> adresinden 28.11.2020 tarihinde erişilmiştir.
- Yurdugül, H. (2005). *Ölçek geliştirme çalışmalarında kapsam geçerliği için kapsam geçerlik indekslerinin kullanılması*. XIV. Ulusal Eğitim Bilimleri Kongresi'nde (28-30 Eylül, Denizli) sunulan bildiri. <http://yunus.hacettepe.edu.tr/~yurdugul/3/indir/PamukkaleBildiri.pdf> adresinden 28.11.2020 tarihinde erişilmiştir.

Extended Abstract

Introduction

Concept cartoons are drawings that present the messages that are intended to be conveyed by using visuals with discussions between three or more characters (Keogh and Naylor, 1999). Concept cartoon, in order to determine the relevant prior knowledge, possible misunderstandings and lack of knowledge of students before teaching a unit, subject or concept; In order to ensure meaningful learning during teaching; and post-teaching are drawings that can be used to evaluate teaching (Dabell, 2008).

The skills required to perform a job (for example, creating concept cartoons) can be difficult to measure through knowledge tests. Because there may be no relationship between knowing how to do a job and doing it according to the specified criteria. For this reason, student performances are generally measured indirectly using assessment tools suitable for the performance in question. This measurement process is also called performance evaluation (Tekin, 2009).

One of the common measurement tools that can be used in performance evaluation is the rubric. Rubric is a scoring tool that lists criteria for each performance and shows what should be in performance (Popham, 1997). Rubric consists of three parts (Popham 1997). These;

Assessment Criteria: It is used to distinguish between students' acceptable and unacceptable performances. *Criterion Definitions:* It refers to the way students express qualitative differences in their performances. *Scoring Strategy:* Depending on the purpose of the assessment, holistic or analytical GSK can be used to evaluate the performance.

In some of the studies in the literature, it was stated that the concept cartoons developed by the researchers without any explanation about their validity and reliability were used (Atasoy, 2020; Serttaş and Türkođlu, 2020; Yılmaz, 2020). Aygün, Karadeniz and Bütüner (2020), Balım, İnel-Ekici and Özcan (2016), Karaca, Kuzu and Çalışkan (2020) and Samková (2020) used concept cartoons developed by other researchers. Therefore, it is seen that evaluating the adequacy of the concept cartoons developed independently from prejudices and personal judgments is a problem that needs to be solved. In this study, "Is the rubric for evaluation of concept cartoons valid and reliable?" The answer to the question will be sought.

It was concluded that a valid and reliable rubric should be created in order to evaluate the adequacy of the concept cartoons used by researchers in mathematics education. In this context, the purpose of this research is; is the development of a rubric that can be used to evaluate the concept cartoons to be prepared in mathematics education. The research, which will fill an important deficiency in the literature, is thought to be important in this context.

Method

This study is a descriptive study in which a rubric for evaluating concept cartoons was developed. The data of the research were obtained from 36 concept cartoons prepared by six mathematics teachers who were trained in the preparation of concept cartoons in Sivas city center, without any subject / learning area limitation, taking into account the suitability to the purpose of the research, applicability and easy accessibility. As a data collection tool in research; the concept cartoons created by the teachers and the rubric prepared to evaluate these cartoons were used. Within the scope of validity studies, the consistency between the opinions of expert evaluators about rubric was calculated

with the formula for the percentage of agreement proposed by Miles and Huberman (1994). Within the scope of the validity study, the content validity index was also calculated to obtain more detailed information after calculating the percentage of agreement among the evaluators. Reliability among raters was evaluated within the scope of both Classical Test Theory and Item Response Theory. Analyzes within the scope of Classical Test Theory and analyzes within the scope of SPPS. Item Response Theory were carried out using the FACET package program.

Findings

Experts' views on rubric were examined one by one in terms of content, structure and criteria. It has been determined that the percentages of agreement between the opinions of the experts about their suitability and competence with rubric vary between 82% and 95%. These calculated compliance percentages prove that the developed rubric is valid in terms of content, structure and criteria. The content validity index of rubric is over 0.80. It has been concluded that this value is higher than the threshold value of 0.80 and is valid for evaluating the concept cartoons of the rubric developed.

It has been observed that there is a positive, significant and high relationship between the total points given by the two evaluators to the re rubric, which will be used in the evaluation of concept cartoons. In addition, it was observed that the agreement between the two raters in each sub-criterion was high and therefore the developed rubric was reliable. Considering the multi-faceted Rasch model, the reliability index of the rater surface was 0.87, the reliability index of the scored (concept cartoons) surface was 0.80, and the reliability index of the item surface was 0.89. These values show that the evaluation of concept cartoons of the developed rubric is a reliable measurement tool.

Conclusion, Discussion and Suggestions

According to the opinions of the experts, the content, structure and criteria of rubric are appropriate and sufficient to evaluate the concept caricature preparation skill. Content validity indices were also calculated to examine validity in more detail. These values are also at an acceptable level and are another proof of the validity of the developed rubric. It was observed that there was a positive, significant and high relationship between both the total scores and the sub-criteria given by the two evaluators to rubric. In addition, the reliability calculations for rater, scored (concept cartoons) and item surfaces calculated within the scope of the Multi-Surface Rasch Measurement Model are also at an acceptable level. In similar studies, it was stated that there was a high agreement among rubric raters, although there was no 100% agreement (Bilgen and Dođan, 2017; Büyükkıdık and Anıl, 2015; Knoch, 2009). The difference between raters may also have varied according to the level of knowledge, perceptions and interpretations of the raters about the concept cartoon. This is called the rater effect (Cronbach, 1990). Bilgen and Dođan (2017) emphasized that while examining the reliability of the rubric they developed in their study, the interaction between raters and criteria can cause differences between scores.

Since the developed rubric is used in the evaluation of the concept cartoons created by mathematics teachers, it can be investigated whether it can be used in the evaluation of the concept cartoons prepared by teacher candidates. This rubric, which is used in the analysis of concept cartoons prepared on mathematics subjects, can be revised and used for the evaluation of concept cartoons in different fields.

EK 1. Kavram Karikatürü İçin Analitik Dereceli Puanlama Anahtarı

ÖLÇÜTLER	Başlangıç Düzeyindeki Performans (1)	Geliştirilmesi Gereken Performans (2)	Başarılı performans (3)	Mükemmel performans (4)	PUAN
İçerik	İçerik tamamen matematik alanı dışındadır. İçerik birbiriyle tutarlı değildir. Gereksiz bilgilere yer verilmiştir.	İçerik tamamen matematik alanıyla ilgili dir. İçerik birbiriyle tutarlı değildir. Gereksiz bilgilere yer verilmiştir.	İçerik tamamen matematik alanıyla ilgilidir. İçerik birbiriyle tutarlıdır. Gereksiz bilgilere yer verilmiştir.	İçerik tamamen matematik alanıyla ilgilidir. İçerik birbiriyle tutarlıdır. Gereksiz bilgilere yer verilmiştir.	
Problem Durumu	Birden fazla problem vardır. Problem karmaşık bir şekilde belirtilmemiştir. Problem konuya uygun değildir.	Tek bir problem vardır. Problem karmaşık bir şekilde belirtilmemiştir. Problem konuya uygun değildir.	Tek bir problem vardır. Problem yalın ve anlaşılır bir şekilde belirtilmiştir. Problem konuya uygun değildir.	Tek bir problem vardır. Problem yalın ve anlaşılır bir şekilde belirtilmiştir. Problem konuya uygundur.	
Alternatif Fikirler	Problemin çözümüne ilişkin alternatif fikirler yoktur.	Problemin çözümüne ilişkin alternatif fikirler vardır. Alternatif fikirler birbirine eşit statüde verilmemiştir. Alternatif fikirlerden en az biri doğru değildir.	Problemin çözümüne ilişkin alternatif fikirler vardır. Alternatif fikirler birbirine eşit statüde verilmiştir. Alternatif fikirlerden en az biri doğru değildir.	Problemin çözümüne ilişkin alternatif fikirler vardır. Alternatif fikirler birbirine eşit statüde verilmiştir. Alternatif fikirlerden en az biri doğrudur.	
Metin Yazımı	Bilimsel terimlere yer verilmemiştir. Terimler açık, net, anlaşılır ve tutarlı değildir. Dil yazım kurallarına uyulmamıştır.	Bilimsel terimlere yer verilmiştir. Terimler açık, net, anlaşılır ve tutarlı değildir. Dil yazım kurallarına uyulmamıştır.	Bilimsel terimlere yer verilmiştir. Terimler açık, net, anlaşılır ve tutarlıdır. Dil yazım kurallarına uyulmamıştır.	Bilimsel terimlere yer verilmiştir. Terimler açık, net, anlaşılır ve tutarlıdır. Dil yazım kurallarına uyulmamıştır.	
Tasarım	Karakterler isimlendirilmemiştir. Kavram karikatürü ilgi çekici değildir. Bir düzen yoktur ve okunaksızdır.	Karakterler isimlendirilmiştir. Kavram karikatürü ilgi çekici değildir. Bir düzen yoktur ve okunaksızdır.	Karakterler isimlendirilmiştir. Kavram karikatürü ilgi çekicidir. Bir düzen yoktur ve okunaksızdır.	Karakterler isimlendirilmiştir. Kavram karikatürü ilgi çekicidir. Düzenli ve okunaklıdır.	
					TOPLAM

*Herhangi bir ölçütte hiçbir şey yapılmamışsa bu ölçüt için sıfır puan verilir.

