



Fen Bilgisi Öğretmen Adaylarının Pedagojik Muhakeme Becerilerinin Öğrenci Yazılı Cevaplarını Yorumlamaları Yoluyla İncelenmesi *

Metin ŞARDAĞ* Kemal İZCİ**

• **Geliş Tarihi:** 10.10.2020 • **Kabul Tarihi:** 18.05.2021 • **Çevrimiçi Yayın Tarihi:** 24.05.2021

Öz

Bu çalışmanın amacı, fen bilgisi öğretmen adaylarının ortaokul öğrencilerinin yazılı sınav cevaplarında nelere dikkat ettiklerine, bu cevapları nasıl yorumladıklarına ve bu cevapları nasıl kullanmayı planladıklarına bakarak pedagojik muhakeme becerilerini incelemektir. Çalışma nitel araştırma desenlerinden olan durum çalışması ile gerçekleştirilmiştir. Çalışmaya ölçme ve değerlendirme dersini almış 76 fen bilgisi öğretmen adayı katılmıştır. Veri toplama aracı olarak ortaokul öğrencilerinin ışıklarda renk oluşumuyla ilgili iki açık uçlu kavramsal soruya verdikleri cevaplardan oluşan bir form kullanılmıştır. Toplanan veriler içerik analizi yöntemiyle analiz edilmiştir. Elde edilen bulgulara göre, öğretmen adaylarının öğrenci cevaplarını yüzeysel olarak inceledikleri, öğrencilerin problem yaşadıkları hususları tespit etmek yerine cevabın doğruluğu ve yanlışlığını ön plana çıkardıkları, yorumlama yaparken ortaya koydukları iddiaları destekleyecek detaylı deliller sunmadıkları ve öğrencinin farklı yanıtları arasında bağlantı kurmadan kavramsal öğrenmesi hakkında yorumlamalar gerçekleştirdikleri görülmektedir. Bu sonuçlar, öğretmen adaylarının eğitimlerinde ölçme değerlendirme hususunda öğrenci cevaplarının nasıl yorumlanıp kullanılabileceğine yönelik uygulamalı eğitimler verilmesinin önem arz ettiğini ve bu şekilde pedagojik muhakeme becerilerini geliştirebileceklerini ortaya koymaktadır.

Anahtar sözcükler: pedagojik muhakeme, ölçme ve değerlendirme, biçimlendirici değerlendirme, öğretmen eğitimi, öğretmen adayları

Atıf:

Şardağ, M. ve İzci, K. (2021). Fen bilgisi öğretmen adaylarının pedagojik muhakeme becerilerinin öğrenci yazılı cevaplarını yorumlamaları yoluyla incelenmesi. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 53, 495-520. doi: 10.9779/pauefd.808745.

* Bu çalışma 28-30 Eylül 2016 tarihlerinde Trabzon'da gerçekleştirilen 12. Ulusal Fen Bilimleri ve Matematik Eğitimi kongresinde sözlü bildiri olarak sunulmuştur.

* Arş. Gör. Dr., Van Yüzüncü Yıl Üniversitesi, <https://orcid.org/0000-0003-2162-8289>, metinsardag@yyu.edu.tr

** Doç. Dr., Necmettin Erbakan Üniversitesi, <https://orcid.org/0000-0002-4228-8845>, kizci@erbakan.edu.tr

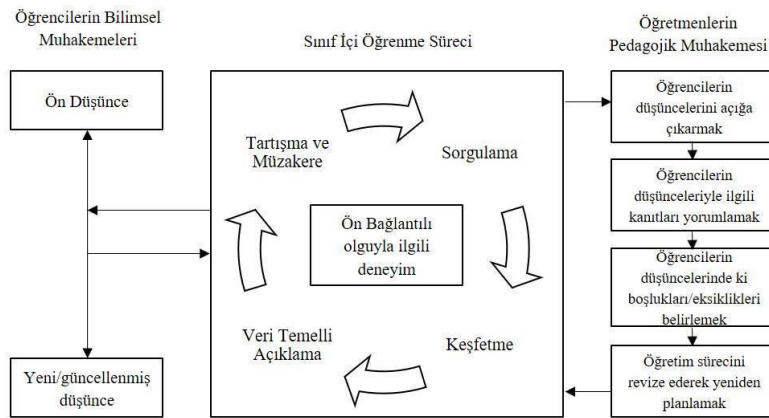
Giriş

Etkili bir fen öğretimi öğrencilerin öğrenme ortamlarına taşıdıkları ve ön yaşantıları sayesinde geliştirdikleri bilgi ve fikirleri temele alarak yeni öğrenmelerin bunlar üzerine yapılandırılmasını gerektirmektedir (Barnhart ve van Es, 2015; Ruiz-Primo ve Furtak, 2007; Talanquer, Tomanek ve Novodvorsky, 2015). Bu süreç oldukça karmaşık ve zorlayıcı olmasının yanı sıra öğretmenlere de oldukça önemli sorumluluklar yüklemektedir. Süreçte öğretmenlerden mevcut pedagojik yaklaşımlarını değiştirerek öğrencilerin sadece doğru ya da yanlış cevaplarına odaklanmak yerine “öğrencilerin okula getirdikleri ön düşüncelere ve öğrencilerin bu düşüncelerden en iyi şekilde nasıl anlayışlarını geliştireceklerine” odaklanmaları beklenmektedir (National Research Council-NRC, 2012, s. 256).

Fen bilgisi öğretmenleri ev ödevleri, sınavlar ve laboratuvar raporları gibi formel ve sınıf içi tartışmalar, soru-cevap ve gözlem gibi formel olmayan yollarla öğrencileriyle ilgili bilgilere ulaşabilmektedirler. Bir başka deyişle, fen bilgisi öğretmenleri pek çok yolla öğrencilerinden dönüt alma ve aldıkları bu dönütlere göre yürütecekleri öğretim sürecini planlama, değerlendirme ve revize etme şansına sahip olabilmektedirler (Harshman ve Yeziarski, 2015). Bu süreçte öğretmenlerden beklenen öğrenciler hakkında farklı yollarla elde ettikleri verileri öğretimle ilgili kararlar alırken kullanmalarıdır. Bu süreç alanyazında veri-güdümlü sorgulama (data-driven inquiry) veya biçimlendirici değerlendirme olarak da bilinmektedir (Furtak, 2012; Harshman ve Yeziarski, 2015).

Biçimlendirici değerlendirmenin öğrencilerin öğrenmeleri ve motivasyonları üzerine olumlu etkisinin olduğu pek çok araştırma ile (Black ve Wiliam, 1998; Vogelzanga ve Admiraal, 2017) ortaya konulmasının yanı sıra araştırmacıların biçimlendirici değerlendirmeyle ilgili öğretmenlerin inançları, algıları, bilgileri ve uygulamaları üzerine yoğunlaştıkları da görülmektedir (Furtak, 2012; Lyon, 2011; Siegel ve Wissner, 2011). Yapılan çalışmaların sonuçları öğretmenlerin öğrencilerin öğrenmelerini açığa çıkarma kabiliyetleri ve elde ettikleri bilgileri etkili bir şekilde yorumlayıp kullanarak dönüt verme ve dersi düzenleme becerilerinin öğrenci başarısını artıran oldukça önemli etkenler olduğunu göstermektedir (Furtak, 2012; Ruiz-Primo ve Furtak, 2007). Bu yüzden, öğretmenlerden öğrenci cevaplarını çok iyi analiz etmesi, cevaplarda ki sorunlu kısımların farkına varması ve bunları üretken bir şekilde yorumlayarak öğrencilerin zorlandıkları alanlarda onlara yardımcı olarak öğrenmeyi desteklemesi beklenmektedir (Bennett, 2011). Fakat yapılan araştırmalar maalesef öğretmenlerin öğrenci cevaplarına yüzeysel olarak baktıklarını, yaptıkları yorumların ise öğrencilerin kavramsal anlamasını engelleyen sorunlu kısımları

ortaya çıkarmaktan ziyade cevapların doğruluğu veya yanlışlığı üzerine odaklandıklarını dolayısıyla da öğrenmeyi destekleyemediklerini göstermektedir (Ateh, 2015; Furtak, 2012). Bu yüzden, öğrenci cevaplarındaki sorunlu kısımlara dikkat eden, bunların kavramsal anlamayı nasıl sınırlandırdığını algılayan ve bu verilerden yola çıkarak uygun yöntemlerle kavramsal öğrenmeyi destekleyebilen pedagojik muhakeme becerisine sahip öğretmenler yetiştirmek önemli bir husustur. Şekil 1’de görüldüğü gibi öğretmenlerin pedagojik muhakeme becerilerini kullanmasını gerektiren süreç kritik bir öneme sahiptir ve etkili bir öğretimin önemli bir parçasını oluşturmaktadır.



Şekil 1. *Duyarlı öğretim süreci ile öğrencilerin fen öğrenmeleri arasında ki ilişki (Sabel, Forbes ve Zangori, 2015, s. 422).*

Pedagojik muhakeme becerileri hem öğretim ortamını etkiler hem de öğretim ortamından etkilenir. Bir başka deyişle, öğretmenin öğrencilerin ön bilgilerine duyarlı bir öğretim yaklaşımı sergilemesi hem sunacağı öğrenme ortamını hem de öğrencilerin hedeflenen olguya ilgili anlayış gelişimini etkilemektedir (Sabel ve diğerleri, 2015).

Uluslararası alanda birçok araştırmacının öğretmen adaylarının öğrenci cevaplarına dikkat etme ve yorumlama becerilerini belirlemeye ve geliştirmeye yönelik pedagojik muhakeme becerilerini inceledikleri görülmektedir (Barnhart ve van Es, 2015; Luna ve Sherin, 2017; Talanquer ve diğerleri, 2015). Bu çalışmalar yaklaşımsal olarak incelenecek olursa pedagojik muhakeme becerileri üzerine gerçekleştirilen çalışmaların öncelikle matematik eğitimi alanında başladığı ve yoğunlaştığı görülmektedir (Luna ve Sherin, 2017; Schack, Fisher ve Wilhelm, 2017; Son 2013). Son zamanlarda fen eğitimi alanında da çalışmalar yürütülmektedir (Larkin 2012; Levin, Hammer ve Coffey 2009; Talanquer ve diğerleri, 2015). Çalışmalarda muhakeme becerileri öğretmenlik uygulaması dersleri, mikro öğretim uygulamaları, öğrenci yazılı cevapları kapsamında incelenmekle birlikte genelde

video kayıtları veya yazılı dokümanlar veri kaynağı olarak kullanılmaktadır (Lam ve Chan, 2020; Luna ve Sherin, 2017). Video kayıtları ve gözlemler genelde öğretimin gerçekleşmesi esnasında öğretmenin farklı yollar ile (ör., soru-cevap) öğrencilerden elde ettiği verileri anlık olarak analiz etmesi, yorumlaması ve öğretimde kullanma becerisini incelemek için kullanılmaktadır. Yazılı dokümanlar ise genelde öğrencilerden elde edilen verilerin ders dışı bir zaman diliminde kullanılarak öğretmenlerin inceleme, yorumlama ve kullanma becerilerinin araştırılması için tercih edilmektedir. Bu durumda (ertelenmiş fark etmede) öğretmenlerin öğrenci cevaplarını değerlendirmede daha fazla zamana sahip olması ve daha detaylı şekilde öğrenci cevaplarını analiz ederek etkili öğretimsel kararlar alabilmesi söz konusudur (Lam ve Chan, 2020). Bu yüzden genellikle öğretmen adayları üzerine yapılan çalışmalarda yansıtıcı öğretimi (reflective teaching) desteklemek amacıyla ertelenmiş fark etmeye odaklanılmaktadır. Araştırmacıların ertelenmiş fark etme için veri toplama şekillerine bakıldığında ise iki ayrı yaklaşım öne çıkmaktadır. Bu yaklaşımlardan ilkinde göre öğretmen adaylarının öğrenci cevaplarını bireysel olarak yazılı bir şekilde değerlendirmesi (Talanquer ve diğerleri, 2015) söz konusu iken ikincisinde öğretmen adaylarının işbirlikli olarak (3'lü-5'li gruplar) verilen yazılı öğrenci cevaplarını değerlendirmeleri sağlanmaktadır (Murray ve diğerleri, 2020).

Öğretmenlerin pedagojik muhakeme becerileri üzerine yapılan çalışmaların sonuçlarına bakıldığında karma bulgularla karşılaşmaktadır. Dini, Sevan, Caushi ve Picon (2020) bu karma sonuçları sergileyen öğretmenlerin kıyaslamasından yola çıkarak bir model önermektedir. Bu modele göre öğretmenler ya daha otoriter (more authoritative) ya da daha diyaloglu (more dialogic) bir bakış açısıyla pedagojik muhakeme becerilerini ortaya koymaktadırlar. Daha otoriter bir yaklaşım sergileyen öğretmenler, öğrenci düşüncelerini incelerken bazı özel bilgi ve becerilere bakmakta, bunların bilimsel olarak doğru olup olmadığını kontrol etmekte ve öğrencileri bilimsel doğruya direkt ulaştıracak öğretimsel yolları kullanmayı önermektedirler. Daha diyaloglu bakış açısında ise öğretmenler öğrencilerin mevcut bilgi ve deneyimlerini göz önünde bulundurmakta, verdikleri cevaplardaki mantıksal uyumsuzlukları belirlemekte, bu uyumsuzlukların nedenleriyle ilgili varsayımlar oluşturmakta, öğrencilerin bu uyumsuzlukları fark edebilecekleri, üzerine düşünebilecekleri ve kendi öğrenmelerini geliştirebilecekleri öğretimsel ortamlar sunmayı önermektedirler. Diğer araştırma sonuçları da Dini ve diğerlerinin (2020) önerdiği modeli desteklemektedir. Öğretmenlerin/öğretmen adaylarının çoğunluğu öğrenci cevaplarını değerlendirirken verilen cevapları betimlemekte, cevaplardaki bilgilerin doğruluğu-

yanlılığına odaklanmakta ve öğrencileri bilimsel doğruya ulaştıracak genel pedagojik yolları önermektedirler (Ateh, 2015; Aydeniz, Doğan, 2016; Chan ve Yau, 2021; Furtak, 2012; Gotwals ve Brimingham, 2016; Murray ve diğerleri, 2020; Talanquer ve diğerleri, 2015). Az sayıda da olsa bazı çalışmalarda ise öğretmenlerin/öğretmen adaylarının öğrenci cevaplarını değerlendirirken öğrencinin verdiği cevabının nedeniyle ilgili tahmini gerekçelere ve altında yatan mantığa odaklanmakta, öğrencinin sunduğu düşünceyi ve muhtemel gerekçeleri yorumlamakta ve öğrencinin merkezde olduğu ve kendi öğrenmesini geliştirebileceği aktif öğretimsel uygulamalar önermektedirler (Dini ve diğerleri, 2020; Murray ve diğerleri, 2020; Talanquer ve diğerleri, 2015).

Padagojik muhakeme becerilerinin geliştirilmesine yönelik yürütülen araştırmalar incelendiğinde olumlu katkıların meydana geldiği görülmektedir. Örneğin, Gotwals ve Brimingham (2016) ders/ünite planlarını, yazılı öğrenci cevapları ve öğretim videoları üzerine yansımalarını ve öğretim videolarını kullanarak fen bilimleri öğretmen adayının pedagojik muhakeme becerilerini ve gelişimlerini incelemişlerdir. Bu inceleme esnasında özellikle adayların öğrencilerin düşüncelerini açığa çıkarma (eliciting), önemli kısımları belirleme (identifying), yorumlama (interpreting) ve öğretimsel cevap verme (responding) becerileri üzerine yoğunlaşmışlardır. Sonuçlar katılımcıların öğrencilerin düşüncelerini açığa çıkarma becerilerinin zamanla geliştiğini ve daha kaliteli ve kapsamlı öğrenci düşüncelerini açığa çıkarabilecek sorular sormaya başladıklarını göstermiştir. Benzer bir şekilde Barnhart ve van Es (2015) fen bilimleri öğretmen adaylarının kendi öğretimlerinin videolarından seçtikleri çeşitli kesitler vasıtasıyla öğretmen adaylarının derslerinde ki önemli noktaları fark etmede başarılı olduklarını ortaya koymuşlardır. Araştırma sonuçları ayrıca başarılı bir analiz ve öğretimsel düzenleme için öğretimdeki önemli kısımları fark etmenin ön koşul olduğunu fakat ileri bir fark etme becerisinin başarılı bir analiz ve karşılık verme becerisiyle sonuçlanamayabileceğini de göstermiştir. Son olarak mentör öğretmen desteğiyle oluşturulan ölçme araçlarını kullanan öğretmen adaylarının zengin öğrenci cevaplarına erişime sahip olduğunu ve bunu öğretime yansıttıklarında öğrenci başarısını destekleyebilecekleri belirtilmektedir (Kang ve Anderson, 2015).

Ulusal düzeyde incelemeler gerçekleştirildiğinde ise fen bilgisi öğretmen adaylarına yönelik her ne kadar pedagojik muhakeme becerilerinin alt yapısını oluşturan ölçme değerlendirmeye yönelik algıları (İzci & Şardağ, 2016) veya uygulamalarına yönelik çalışmalar (İzci, 2018) dikkat çekse de Aydeniz ve Doğan'ın (2016) çalışması dışında öğretmenlerin pedagojik muhakeme becerilerini belirlemeye veya geliştirmeye odaklanan

neredeyse hiçbir çalışmanın gerçekleştirilmediği görülmektedir. Aydeniz ve Doğan (2016) araştırmalarında Newton'un hareket yasaları, kinetik enerji, seri ve paralel bağlı elektrik devreleri ve durgun elektrik kavramlarına yönelik 7 soru hakkında ortaokul 8 sınıf öğrencilerinin üretmiş olduğu yanıtlardan 4'er tanesini fen bilgisi öğretmen adaylarına inceleyerek pedagojik muhakeme becerilerini belirlemeye çalışmışlardır. Araştırmacılar çalışmalarında ağırlıklı olarak öğretmen adaylarının dönüt bakış açıları üzerine odaklanmışlardır. Fakat alanyazın incelendiğinde biçimlendirici değerlendirmenin etkililiğini etkileyen etmenler arasında alan bağımlılık konusunun önemli bir etken olduğu görülmektedir (Bennet, 2011). Dolayısıyla birden fazla konu alanına bağlı olarak yürütülen bir çalışmadan ziyade belirli bir konuya odaklanan, fen bilgisi öğretmen adaylarının pedagojik muhakeme becerilerini inceleyen çalışmaların gerçekleştirilmesi önemli bir husustur. Uluslararası alanyazında konuya özgü olarak yürütülen pedagojik muhakeme becerileri çalışmaları (Talanquer ve diğerleri, 2015) söz konusu iken ulusalda bir eksiklik olarak karşımıza çıkmaktadır. Bu sebepten dolayı bu çalışmada belirli bir konuya yönelik olarak pedagojik muhakeme becerileri araştırılmaktadır. Ayrıca yukarıda üzerinde durulan çalışmalar bir bütün olarak ele alındığında temelde pedagojik muhakeme becerilerinin geliştirilebilmesi için gerçekleştirilen faaliyetleri içerdikleri, pedagojik muhakeme becerilerin sergilenmesinde etkili faktörlerin neler olabileceğine ve etkilerine odaklandıkları görülmektedir. Bu çalışma ile de fen bilgisi öğretmen adaylarının öğrenci yazılı sınav cevaplarında nelere dikkat ettikleri, bu cevapları nasıl yorumladıkları ve bu cevapları nasıl kullanmayı planladıkları üzerine odaklanarak pedagojik muhakeme becerilerin gelişmesinde etkili olabilecek faktörlerin irdelenerek alanyazına katkı sunulabileceği düşünülmektedir. Bu bağlamda yürütülen bu çalışma ile fen bilgisi öğretmen adaylarının ortaokul öğrencilerinin yazılı sınav cevaplarında nelere dikkat ettikleri, bu cevapları nasıl yorumladıkları ve bu cevapları nasıl kullanmayı planladıkları üzerine odaklanarak öğretmen adaylarının pedagojik muhakeme becerilerinin incelenmesi amaçlanmıştır. Bu sayede hem yukarıda değinilen alanyazındaki eksikliğin giderilmesine hem de var olan araştırma bulguları üzerine katkı sağlanacağı düşünülmektedir. Bu doğrultuda aşağıdaki iki temel araştırma sorusuna cevap aranmıştır.

1. Fen bilgisi öğretmen adayları öğrencilerin yazılı cevaplarını nasıl değerlendirmektedirler?
2. Fen bilgisi öğretmen adaylarının öğrencilerin yazılı cevaplarından yola çıkarak önerdikleri öğretimsel düzenlemeler nelerdir?

Yöntem

Bu araştırma, nitel araştırma desenlerinden biri olan durum çalışması temelinde yürütülmüştür. Durum çalışması pek çok alanda sıklıkla kullanılan bir araştırma yöntemi olup küçük grup davranışları, yönetsel süreçler gibi gerçek yaşam olaylarının özelliklerini ortaya koymak için (Yin, 2009) ve araştırmaya konu olan olayların bütüncül yorumuna ulaşmak için imkânlar sağlamaktadır (Yıldırım ve Şimşek, 2016). Bu çalışmada belirli bir grubun değerlendirmeye yönelik yaklaşımlarının incelenmesi ve bundan hareketle bütüncül bir şekilde pedagojik muhakeme becerilerinin ortaya konulması amaçlandığından araştırmanın durum çalışması esasları gereği yürütülmesine karar verilmiştir. Bu doğrultuda çalışmada fen bilgisi öğretmen adaylarının, öğrenci yanıtlarını değerlendirme durumları, durum olarak incelenmekte ve pedagojik muhakeme becerileri ortaya konulmaya çalışılmaktadır.

Çalışma Grubu

Çalışma grubu amaçlı örnekleme yöntemlerinden biri olan ölçüt örnekleme yöntemiyle belirlenmiştir. Ölçüt örnekleme yönteminde araştırmacılar tarafından belirlenmiş bir dizi ölçüt çerçevesinde seçimlerin yapılması söz konusudur (Yıldırım ve Şimşek, 2016). Bu doğrultuda çalışma grubu Türkiye'nin doğusunda bulunan bir devlet üniversitesinde ölçme değerlendirme dersini almış 3. ve 4. sınıfta bulunan 76 fen bilgisi öğretmen adayı olarak belirlenmiştir ve çalışma verileri 2015-2016 eğitim-öğretim yılı bahar döneminde toplanmıştır.

Veri Toplama Araçları

Bu çalışmada veri toplama aracı olarak araştırmacılar tarafından geliştirilmiş bir form kullanılmıştır. Form oluşturulurken iki basamaktan oluşan bir süreç yürütülmüştür. İlk basamakta bir Fen Bilgisi öğretmeni ile iletişime geçilerek içerisinde açık uçlu kavramsal soruların yer aldığı fen bilgisi dersi öğrenci sınav kâğıtları temin edilmiştir. Bu sınav kâğıtları incelendikten sonra içerisinden iki soru seçilmiştir. Seçilen sorular “Farklı renkteki ışıklar nasıl elde edilir?” ve “Gökyüzü ve denizin mavi rengi nasıl oluşur?” şeklindedir. Daha sonra bu iki soruya cevap veren, ilgili sınavdan yüksekte düşüğe farklı başarı seviyelerde not almış beş öğrencinin yanıtı belirlenmiştir. Belirlenen öğrenci yanıtları aşağıda Tablo 1'de sunulmuştur.

Tablo 1. Ortaokul Öğrencilerinin Fen Bilgisi Dersi Sınavında Sorulara Vermiş Oldukları Yanıtlar

| Öğrenciler | Verilen Cevaplar |
|------------|--|
| Öğrenci-1 | <p>1) Işıkları birbirine karıştırarak elde ederiz.</p> <p>2)Uzay siyah güneş ışığı da sarı olduğundan bu iki ışık çarpıştığında mavi renk oluşur. Bu yüzden de gök ve deniz mavi renkli görünürler.</p> |
| Öğrenci-2 | <p>1)Farklı renkteki ışıklar beyaz ışığın soğurulması ile ortaya çıkarlar. Bu ışıklar kırmızı, turuncu, sarı, yeşil mavi ve mordur.</p> <p>2) Aslında gök ve deniz berraktır. Ama ışık vurup yansıttığı için biz onları mavi renkte görürüz.</p> |
| Öğrenci-3 | <p>1) Üç ana renk olan kırmızı, mavi ve yeşile gelen ışınlar sayesinde diğer renkler oluşur.</p> <p>2) Gökyüzü mavidir. Gökyüzündeki mavilik ışınlar aracılığıyla yansır ve beyaz olan suya değer ve denizde mavi olur.</p> |
| Öğrenci-4 | <p>1) Beyaz ışığın yansıtması ve soğurmasıyla diğer renkler oluşur. Mesela beyaz renk içerisinde aslında kırmızı renkte vardır. Mesela beyaz ışığa kırmızı ışık yansıtılırsa kırmızı görülür çünkü beyaz renk diğer renklerin birleşmesiyle oluşmuştur.</p> <p>2) Beyaz ışığın içinde mavi renkte vardır. Böylece beyaz ışık gök ve denize mavi rengini yansıtır. Gök ve deniz mavi olur.</p> |
| Öğrenci-5 | <p>1)Farklı renkteki ışıklar filtreleme yöntemiyle elde edilirler. Buna örnek olarak trafik ışıkları verilebilir. Aslında kırmızı, sarı ve yeşil renklerin altında beyaz ışık vardır. Fakat filtre kullanılarak diğer renkler soğrulur ve biz sadece kırmızı, sarı ve yeşil renkleri görürüz.</p> <p>2) Deniz ve gök aslında renksizdirler. Mavi görülmelerinin nedeni beyaz ışıktaki renklerden sadece maviyi yansıtır diğer renkleri ise soğururlar. Bu yüzden de gök ve deniz mavi olarak görünürler.</p> |

İkinci basamakta ise seçilen öğrenci cevapları kullanılarak katılımcıların bu cevapları değerlendirecekleri form düzenlenmiştir. Forma son hali verildikten sonra ilgili formdaki öğrenci cevaplarını dikkate alarak bir ders saati süresince öğretmen adaylarından aşağıda belirtilen sorulara yazılı olarak yanıt vermeleri istenmiştir.

Soru 1) Aşağıdaki öğrenci cevaplarını kullanarak her bir öğrencinin renk oluşumu ile ilgili temel kavramları anlayıp anlamadıklarına nasıl karar verirsiniz? Her bir öğrenci cevabını aşağıda tartışınız.

Soru 2) Sizce her bir öğrencinin cevabına göre konuyu yanlış ya da eksik anladığını gösteren sorunlu noktalar nelerdir?

Soru 3) Öğrencilerin vermiş oldukları cevaplara göre öğrencilerin konuyu anlama seviyelerini düşükten yükseğe doğru sıralayınız? Niçin böyle sıraladığınızı da delilleriyle açıklayınız?

Soru 4) Öğrencilerin vermiş oldukları bu cevaplar bir öğretmen olarak size neleri gösterebilir ve neler yapmanıza yardımcı olur?

Soru 5) Bu cevaplardan yola çıkarak konuyu anlamayan öğrencilerinizin renk oluşumu konusunu daha iyi öğrenmeleri için bir öğretmen adayı olarak neler yapabilirsiniz?

Verilerin Analizi

Veri toplama formu aracılığıyla elde edilen veriler, Talanquer ve diğerleri (2015) tarafından geliştirilen öğretmenlerin öğrenci cevaplarını değerlendirme boyut ve kriterleri kullanılarak analiz edilmiştir. Bu boyutlar temelde iki başlık altında ele alınmıştır. İlki alandan bağımsız boyut, ikincisi ise alan bağımlı boyuttur. Alandan bağımsız boyut kendi içerisinde beş farklı alt boyut içermektedir. Tablo 2’de alandan bağımsız boyutlar ve kriterleri sunulmuştur.

Tablo 2. Alandan Bağımsız Alt Boyutlar ve Kriterleri

| Boyut | Düzye | Açıklama |
|-------------------------|-----------|--|
| Değerlendirme Yaklaşımı | Başlangıç | Değerlendirme çoğunlukla öğrenci anlaması hakkında çıkarım yapmaktan ziyade öğrenci yanıtının betimlemesine dayanır. |
| | Gelişen | Değerlendirme betimleme ve çıkarım durumlarının bir karışımıdır. |

| | | |
|-------------------------------|---------------|--|
| | İleri Düzy | Değerlendirmenin merkezinde öğrencinin düşünmesi ve anlaması hakkındaki çıkarımlar bulunur. |
| Değerlendirme Yönelimi | Başlangıç | Öğrencinin mevcut cevaplarında anlam aramaksızın öğrencinin yapmakta veya söylemekte başarısız olduğu şeyler, yanlış veya doğru cevap üzerine odaklanır. |
| | Gelişen | Belirli doğru cevaplar ve yanıtlarda bazı anlamlar bulma üzerine odaklanır. |
| | İleri Düzy | Öğrenci yanıtlarının arkasında yatan mantıksal akıl yürütmeyi anlamak için onları analiz etmeye odaklanır. |
| Değerlendirmedeki Özgünlük | Başlangıç | Belirli cevaplara dayalı olarak öğrenci tutum veya bilgileri hakkında genel durumlar üretir. |
| | Gelişen | Öğrencinin bir konuyu anlaması veya yanlış anlaması hakkında daha fazla odaklanılmış durumlar üretir. |
| | İleri Düzy | Öğrenci yanıtlarına dayandırarak öğrenci anlaması hakkında spesifik iddialar üretir. |
| Değerlendirme tutarlılığı | Başlangıç | Yalıtılmış hatalar olarak yanlışlar veya kavram yanlışlarının belirlenmesi ve betimlenmesi üzerine odaklanır. |
| | Gelişen | Öğrenci yanıtlarındaki tutarlılığı veya tutarsızlığı fark etmek söz konusudur. |
| | İleri Düzy | Öğrenci anlaması hakkına yorumlar oluşturmak için öğrencilerin yanıtlarında ki tutarlılıkları ve tutarsızlıkları kullanma söz konusudur. |
| Delil kullanımı | Başlangıç | Çıkarımlar yapılması halinde, iddiaları desteklemek için net deliller sunulmaz. |
| | Gelişen | Öğrenci anlaması hakkında yapılan çıkarımların |

bazılarını desteklemek için delil sunulur.

İleri
Düzyer Öğrenci anlaması hakkında yapılan çıkarımlar öğrencilerin cevaplarından elde edilen delillerle her zaman ilişkilendirilir.

Alan bağımlı boyut ise kendi içerisinde dört farklı alt boyuttan ve bunların kriterlerinden oluşmaktadır. Bu boyut ve kriterleri Tablo 3'te sunulmuştur.

Tablo 3. Alan Bağımlı Alt Boyutlar ve Kriterleri

| Boyut | Düzyer | Açıklama |
|-----------------------|--------------|--|
| Yorumlama kalitesi | Başlangıç | Öğrenci anlamasına yönelik yapılan yorumlamalar ulaşılabilir delil ile desteklenmemiştir. |
| | Gelişen | Yorumlamalar için sağlanan deliller sınırlı, yüzeysel veya basma kalıptır. |
| | İleri Düzyer | Öğrenci anlamasının mantıksal yorumlamaları oluşturulur. |
| Üretken düşünce | Başlangıç | Öğrencinin konuyu kavradığını değerlendirmek için konuyla ilgili temel kavramlara odaklanmak yerine önemsiz detaylara odaklanılır. |
| | Gelişen | Öğrencinin konuyu kavradığını değerlendirmek için konuyla ilgili temel kavramlarla birlikte önemsiz detaylara da odaklanılır. |
| | İleri Düzyer | Öğrencinin konuyu kavradığını değerlendirmek için konuyla ilgili temel kavramları önemsiz detaylardan ayırarak temel kavramlara odaklanılır. |
| Değerlendirme kapsamı | Başlangıç | Kavram ve fikirlerin değerlendirmesine sınırlı odaklanma söz konusudur. |
| | Gelişen | Değerlendirme aracılığıyla belirgin bir şekilde hedeflenenin |

| | | |
|----------------------|---------------|--|
| | | ötesinde ilgili alanlarda potansiyel bilgi eksikliğini tanıma söz konusudur. |
| | İleri Düzy | Değerlendirme aracılığıyla belirgin bir şekilde hedeflenenin ötesinde alanlarda ilgili yanlış anlamaları tanıma söz konusudur. |
| | Başlangıç | Temel bilgilerin yüzeysel ve güvenilmez anlamaları, mantıksal çıkarımların oluşturulmasını etkiler. |
| Bilimsel doğruluk | Gelişen | Bazı doğru olmayan yorumlamalar oluşturulur. |
| | İleri Düzy | Yorumlamalar bilimsel olarak doğrudur. |

Toplanan veriler Tablo 2 ve 3'te belirtilen kriterlere göre analiz edildikten sonra yüzde ve frekans değerleri hesaplanmıştır. Verilerin analizi sürecinde araştırmacılar tarafından bir Excel dosyası oluşturulmuş ve her bir öğretmen adayının her bir öğrenci yanıtı için yapmış olduğu değerlendirme alan bağımlı ve bağımsız boyutlarda incelenmiştir. İncelemeler gerçekleştirilirken 1 ve 2 olarak kodlamalar yürütülmüştür. Örneğin delil kullanımı boyutu için öğretmen adayı bir öğrenci yanıtını değerlendirirken delil kullanmamışsa 1, delil kullanmışsa 2 olarak kodlanmıştır. Daha sonra aynı öğretmen adayının bütün öğrenci yanıtları için elde edilen kodlamanın ortalaması alınmıştır. Bu ortalama 1.1 ve altında ise başlangıç, 1.2 -1.9 aralığında ise gelişen ve 2 ise ileri düzeyde olduğu kabul edilmiştir. Bu aralıklar belirlenirken öğretmen adaylarının iki soruya beş öğrencinin verdiği yanıtı değerlendirme durumları göz önünde bulundurulmuştur. Öğretmen adayı toplamda on değerlendirme gerçekleştirmektedir. On değerlendirme içerisinde en az aynı öğrencinin iki farklı soruya verdiği yanıtı kritere uygun değerlendirdiğinde 1.2 kritik değeri ortaya çıkmaktadır. 1.9 değeri ise on öğrenci yanıtının en az dokuz tanesinin ilgili kriterlere uygun şekilde değerlendirilmesiyle ortaya çıkmaktadır. Öğretmen adayı ilgili kriterlere göre tüm öğrenci yanıtlarını değerlendirdiğinde ise 2 ortalama puanı ortaya çıkmaktadır. Analizler gerçekleştirilmeden önce araştırmacılar arasında alan bağımsız ve bağımlı boyutların ne oldukları, nasıl analiz edilmesi gerektiği üzerine tartışmalar yürütülmüş ve fikir birliğine varılmıştır. Sonraki süreçte 20 öğretmen adayının yanıtı iki farklı araştırmacı arasında bağımsız olarak kodlanmış ve farklılık gösteren kısımlar üzerinde tartışılarak uzlaşmaya varılmıştır. Bu süreç içerisinde araştırmacılar arasındaki uyum, Miles ve

Huberman (1994) tarafından belirtilen güvenilirlik formülüyle hesaplanmıştır. Uyum indeksi .93 olarak tespit edilmiştir. Ayrıca daha sonraki bir zaman diliminde verinin tamamının analizini gerçekleştiren araştırmacı tüm verinin yaklaşık %15’lik kısmını tekrardan analiz ederek kodlamaların tutarlı olup olmadığını Miles ve Huberman (1994) güvenilirlik formülüyle araştırmıştır. Sonuç olarak %15’lik kısımda gerçekleştirilen kodlamaların aynı uyum indeksinin 1.00 olduğu bulunmuştur.

Bulgular

Araştırmaya katılan fen bilgisi öğretmen adaylarının öğrenci cevapları üzerinde yapmış oldukları değerlendirmeler analiz edildiğinde alandan bağımsız boyut için Tablo 4’te sunulan bulgulara ulaşılmıştır.

Tablo 4. Alan Bağımsız Alt Boyutlar ve Değerlendirmede Gerçekleştirilme Oranları

| Alt boyutlar | Başlangıç (1) | | Gelişen (2) | | İleri düzeyde (3) | | Ort. |
|----------------------------|---------------|------|-------------|------|-------------------|-----|------|
| | Sıklık | % | Sıklık | % | Sıklık | % | |
| Değerlendirme yaklaşımı | 42 | 61.8 | 26 | 38.2 | 0 | 0 | 1.38 |
| Değerlendirme yönelimi | 44 | 64.7 | 24 | 35.3 | 0 | 0 | 1.35 |
| Değerlendirmedeki özgünlük | 55 | 80.9 | 13 | 19.1 | 0 | 0 | 1.19 |
| Değerlendirme tutarlılığı | 41 | 60.3 | 23 | 33.8 | 2 | 2.9 | 1.41 |
| Delil kullanımı | 46 | 67.6 | 22 | 32.4 | 0 | 0 | 1.32 |

Alan bağımsız boyutunda, öğretmen adaylarının benimsemiş oldukları değerlendirme yaklaşımında %61.8 oranında başlangıç düzeyinde olduğu yani verilen cevapları değerlendirirken cevapların betimlemelerini kullandıkları tespit edilmiştir. Bu duruma örnek vermek gerekirse “Gök ve denizin mavi rengi nasıl oluşur?” sorusuna öğrenci “Aslında gök ve deniz berraktır. Ama ışık vurup yansıttığı için biz onları mavi renkte görürüz.” yanıtını vermiştir. Öğretmen adayı bu yanıtı “Bu öğrencimde denizin ve gök yüzünün parlak olduğunu gelen ışığın vurup yansıtması sonucunda mavi gördüğümüzü öğrenmiş doğru cevap vermiştir.” ifadesini kullanarak değerlendirmiştir. Yapılan değerlendirme incelendiğinde öğrencinin yanıtının yeniden ifade edildiği veya betimlendiği

görülmektedir. Bunun yanı sıra öğretmen adaylarının %38.2'lik kısmı ise değerlendirme yaparken betimsel ifadelerin haricinde çıkarımlar da yaptığı yani gelişen düzeyde olduğu görülmektedir. Bu duruma örnek olarak 3-26 kodlu öğretmen adayının değerlendirmesi verilebilir. Yukarıda verilen soruya öğrenciden “*Beyaz ışığın içinde mavi renkte vardır. Böylece beyaz ışık gök ve denize mavi rengini yansıtır.*” yanıtı gelmiştir. Öğretmen adayı ise “*Yansıtma olayını hem gök hem de deniz için geçerli olduğunu düşünerek öyle yanıtlamıştır. Bilgi yetersizliği vardır.*” ifadesini kullanarak değerlendirmesini yapmıştır. Yapılan değerlendirme incelendiğinde öğrencinin anlaması, bilgisi hakkında çıkarımlar yapıldığı görülmektedir. Bunun yanı sıra araştırmaya katılan öğretmen adaylarının hiçbirinin değerlendirme yaklaşımının ileri düzey olmadığı yani yapılan değerlendirmenin merkezinde öğrenci düşünme ve anlamasının çıkarımları olmadığı görülmektedir.

Bu durum ile paralellik gösteren değerlendirme yönelimlerinde ise öğretmen adaylarının %64.7'si sadece öğrencilerin cevaplarını doğru veya yanlış olarak değerlendirdikleri (ör., *Soruyu doğru cevaplamış.*), %35.3'lük kısmı ise öğrenci cevaplarının büyük bir kısmını analiz ederek cevabın arka planında yatan nedenler üzerine odaklandıkları görülmüştür. Örneğin 3-28 kodlu öğretmen adayı, gök ve denizin mavi renginin nasıl oluştuğu sorusuna “*Beyaz ışığın içinde mavi renkte vardır. Böylece beyaz ışık gök ve denize mavi rengini yansıtır. Gök ve deniz mavi olur.*” ifadesiyle yanıt veren bir öğrenciyi “*Gökte ışığın kırılması olayı meydana gelir. Denizde ise ışığın yansımaları. Öğrenci burada gök ve denizde oluşan renk olaylarını birbirine benzeterek yanlışlığa düşmüştür.*” ifadeleriyle değerlendirmektedir. Fakat değerlendirme yaklaşımı boyutunda olduğu gibi bu boyutta da ileri düzeyde katılımcı bulunmamaktadır. Yani tüm öğrencilerin cevaplarını analiz ederek sorunlu kısımların temelinde yatan nedenleri arayan, cevapların ardında yatan mantıksal akıl yürütmeyi anlamak için analiz eden herhangi bir katılımcı tespit edilememiştir.

Alan bağımsız boyut içerisinde yer alan değerlendirmedeki özgünlük alt boyutunda ise katılımcıların %80.9'lük kısmı öğrenci tarafından verilen yanıtlara dayalı olarak öğrencilerin tutum veya bilgileri hakkında genel değerlendirmeler yaptığı görülmektedir. Örneğin farklı renkteki ışıklar nasıl elde edilir sorusuna öğrenci “*Beyaz ışığın yansımaları ve soğrulmasıyla diğer renkler oluşur. Mesela beyaz renk içerisinde aslında kırmızı renkte vardır. Mesela beyaz ışığa kırmızı ışık yansıtılırsa kırmızı görülür çünkü beyaz renk diğerlerinin birleşmesiyle oluşmuştur.*” yanıtını vermiştir. Bu yanıt üzerine 3-38 kodlu öğretmen adayı, “*Doğru ve açıklayıcı bir cevap. Örnek vererek açıklamasıyla konu hakimiyeti varlığını anlarız.*” değerlendirmesini yapmıştır. Yapılan değerlendirme

incelendiğinde öğretmen adayının genel ifadeler kullandığı, öğrenci anlamasını değerlendirirken detaylar vermediği görülmektedir. Fakat katılımcıların %19.1’lik kısmı ise bazı öğrenci cevaplarında bilgi ve anlamayla ilgili genel ifadeler kullanırken bazı öğrenci cevaplarında ise kavramların anlaşılıp anlaşılmadığıyla ilgili detaylı açıklamalarda bulunmuştur. Bunların yanı sıra hiçbir adayın bir öğrencinin cevaplarına dayandırarak öğrenci anlaması hakkında spesifik iddialar ürettiği, bütün öğrencilerin cevaplarında ilgili kavramların öğrenilip öğrenilmediği ile ilgili detaylı bir açıklamada bulunduğu tespit edilememiştir.

Değerlendirme tutarlığı alt boyutunda yani öğrencinin kendi fikirleri içerisinde ve diğer öğrencilerin fikirleriyle olan tutarlılığın ele alındığı boyutta ise öğretmen adaylarının %60.3’ünün yapmış olduğu değerlendirmelerin tutarlığının başlangıç düzeyde olduğu görülmektedir. Bu durum öğretmen adaylarının yarısından fazlasının yanlışları ve kavram yanlışlarını izole bir şekilde belirlediği ve betimlediğini göstermektedir. %33.8’i ise gelişen düzeyde olup öğrencilerin kendi yanıtlarında veya diğer öğrencilerin yanıtlarıyla tutarlılık veya tutarsızlıkların farkına vardığını göstermektedir. %2.9’luk bir oranda ise öğretmen adaylarının öğrencilerin anlamaları hakkında yorumlar yapabilmek için belirledikleri tutarlılıkları ve tutarsızlıkları kullandıkları görülmektedir.

Son olarak öğretmen adaylarının delil kullanım durumları incelendiğinde %67.6’sının başlangıç düzeyinde olduğu tespit edilmiştir. Bu öğretmen adayları yaptıkları çıkarımlarda iddialarını desteklemek için net deliller sunmamaktadırlar. %32.4’ü ise gelişen düzeyde olup öğrenci anlamaları hakkında yaptıkları çıkarımları desteklemek için bazen delil kullanmaktadırlar. Ayrıca her zaman delil kullanan herhangi bir öğretmen adayı da tespit edilememiştir. Başlangıç düzeyi için gerçekleştirilen bir değerlendirmeye örnek olarak 4-12 kodlu öğretmen adayının “*Burada beyaz ışığın renklere ayrılması ve kırılması olayı hakkında bilgisi eksiktir.*” ifadesi örnek olarak verilebilir. Nitekim ifade incelendiğinde çıkarımın herhangi bir delile dayandırılmadığı görülmektedir. Gelişen düzeye örnek olarak 3-38 kodlu öğretmen adayının yanıtı verilebilir. “Gök ve denizin mavi rengi nasıl oluşur?” sorusuna öğrenci “*Aslında gök ve deniz berraktır. Ama ışık vurup yansıttığı için biz onları mavi renkte görürüz.*” yanıtını vermiştir. Öğretmen adayı bu yanıtı “*Burada genelleme yapmıştır. Deniz için yazdığı doğru ama gök için kırılmayı yazmamıştır. Eksik cevap vermiştir. Aynı zamanda denizin diğer renkleri soğurmasından bahsetmemiştir. Burada bir eksiklik vardır.*” şeklinde değerlendirmektedir. Bu değerlendirme incelendiğinde öğretmen

adayının yapmış olduğu çıkarımlar için deliller (ör., gök için kırılmayı yazmaması, soğrulmadan bahsetmemesi) kullandığı görülmektedir.

Alan bağımlı boyutlar incelendiğinde ise Tablo 5'te sunulan bulgulara ulaşılmaktadır.

Tablo 5. Alan Bağımlı Alt Boyutlar ve Değerlendirmede Gerçekleştirilme Oranları

| Kriterler | Başlangıç (1) | | Gelişen (2) | | İleri düzeyde (3) | | Ort. |
|-----------------------|---------------|------|-------------|------|-------------------|------|------|
| | Sıklık | % | Sıklık | % | Sıklık | % | |
| Yorumlama kalitesi | 54 | 79.4 | 14 | 20.6 | 0 | 0 | 1.21 |
| Üretken düşünce | 50 | 73.5 | 17 | 25.0 | 1 | 1.5 | 1.28 |
| Değerlendirme kapsamı | 52 | 76.5 | 16 | 23.5 | 0 | 0 | 1.24 |
| Bilimsel doğruluk | 9 | 13.2 | 48 | 70.6 | 11 | 16.2 | 2.03 |

Öğretmen adaylarının yorumlama kalitesi öğrenci anlaması hakkında gerçekten ulaşılabilir olan delillere dayalı uygun ve makul çıkarımları ifade etmektedir. Yorumlama kalitesi her ne kadar alan bağımsız boyutta yer alan delil kullanımı alt boyutu ile aynı gibi düşünülse de delil kullanımında öğretmen adayların öğrenci anlaması hakkında çıkarımlarını desteklemek için ne ölçüde delil sunduğunu nitelendirmektedir. Alan bağımlı boyutta ise yorumlama kalitesiyle çıkarımların geçerliği nitelendirilmektedir (Talanquer ve diğerleri, 2015). Bu bakış açısıyla ele alındığında öğretmen adaylarının %79.4'ü başlangıç düzeyinde, %20.6'sı ise gelişen düzeyde olduğu tespit edilmiştir. Fakat hiçbir öğretmen adayı ileri düzeyde bulunmamaktadır. Delil kullanımı durumu ile yorumlama kalitesi karşılaştırıldığında öğrenci anlaması hakkında delil sunan bazı öğretmen adaylarının sunmuş oldukları delillerin, yapmış oldukları çıkarımlarının hatalı olduğu anlaşılmaktadır. Çünkü gelişen düzeyde bulunan öğrenci sayılarında farklılık bulunmaktadır. Başlangıç düzeyine örnek olarak 3-13 kodlu öğretmen adayının yanıtı incelenebilir. “Gök ve denizin mavi rengi nasıl oluşur?” sorusuna öğrenci, “Aslında gök ve deniz berraktır. Ama ışık vurup yansıttığı için biz onları mavi renkte görürüz” yanıtını vermiştir. 3-13 kodlu öğretmen adayı ise bu durumu “Burada beyaz ışığın renklere ayrılması ve kırılma olayı hakkında bilgisi eksiktir” şeklinde bir değerlendirme gerçekleştirmiştir. Öğretmen adayının yanıtı incelendiğinde öğrenci anlaması hakkında yapmış olduğu yorumlamaları ulaşılabilir olan herhangi bir

delille desteklemediği görülmektedir. Fakat aynı soru ve öğrenci cevabına yönelik 3-32 kodlu öğretmen adayı tarafından yapılan değerlendirme incelendiğinde yorum yapıldığı (“Eksik biliyor. Denizde ve gökte ayrı olaylar olacağını ayırt edemiyor. Genel bakıyor.”), delil sağlandığı fakat ortaya konulan delilin sınırlı ve yüzeysel olduğu (kırılma ve yansıma kavramlarına değinilmediği) görülmektedir.

Üretken düşünce boyutunda ise ölçülmesi hedeflenen kavramlar hakkında öğrencilerin anlama seviyeleri değerlendirilirken konuyla ilgili temel kavramlara odaklanılması ve detaylardan ayrılarak etkili ve öğrenmenin desteklenmesi için üretken olabilen bir değerlendirme yapılması söz konusudur. Bunun yanı sıra öğretmen adayının temel kavramlarla ilgili doğru bilimsel anlayışa sahip olması da gerekmektedir. Bu kapsamda incelendiğinde öğretmen adaylarının %73.5’inin başlangıç düzeyinde olduğu tespit edilmiştir. Bu bizlere öğretmen adaylarının öğrenci cevaplarını değerlendirirken hedef kavram için gerekli olan temel kavramları ya hiç göz önünde bulundurmadıkları ya temel kavramlarla ilgili bilimsel anlayışlara sahip olmadıkları ya da gereksiz veya alakasız kavramlara vurgu yaptıklarını göstermektedir. Örneğin 3-17 kodlu öğretmen adayı “Farklı renkteki ışıklar nasıl elde edilir?” sorusuna, Öğrenci-2’nin “Farklı renkteki ışıklar beyaz ışığın soğrulmasıyla ortaya çıkarlar. Bu ışıklar kırmızı, turuncu, sarı, yeşil, mavi ve mordur.” şeklindeki cevabını “Evet doğru. Beyaz ışık prizmadan geçerken renklere ayrılıyor ve renkleri KuTuSu YaMalıM kısaltmasına göre sıraladığı için konuya hâkim bir öğrenci.” açıklamasıyla değerlendirmektedir. Bu değerlendirme incelendiğinde öğretmen adayının temel kavramlardan olan soğrulma ve kırılmayı öğrenci yanıtında fark edemediği görülmektedir. Öğrenci yanıtını soğrulma kavramı temelinde vermiş olmasına rağmen öğretmen adayı değerlendirmesini kırılma temel kavramına göre gerçekleştirmiştir. Bu durumdan dolayı öğretmen adayının üretken bir değerlendirme gerçekleştiremediği ve başlangıç seviyesinde olduğu görülmektedir. Öğretmen adaylarının %25’nin ise hedef kavramı değerlendirirken temel kavramları göz önünde bulundurdıkları fakat öğrenci cevaplarındaki hatalı hususların kaynaklarının neler olabilecekleri üzerinde durmadıklarından dolayı istenilen düzeyde üretken bir değerlendirme yapamadıkları görülmektedir. Hedef kavramın anlaşılmasıyla ilgili temel kavramları kullanarak öğrencileri değerlendiren ve öğrenci yanıtlarındaki eksiklikleri, hataları ele alırken sebepleri üzerinde durarak üretken bir değerlendirme yapan öğretmen adayı sayısının (1) ise çok düşük olduğu görülmektedir.

Değerlendirme kapsamında ise öğretmen adayının öğrenci anlamasını değerlendirmede dikkate aldığı ilgili kavramlarının ve fikirlerin çeşitliliğine odaklanılmaktadır. Öğretmen adaylarının %76.5'lik kısmı öğrenci anlamasını değerlendirmede sınırlı düzeyde kavram ve fikirlerin değerlendirmesine, kavram ve fikirlerin ele alınmasına odaklandıkları görülmektedir. Öğretmen adaylarının %23.5'lik bir kısmı ise değerlendirme aracılığıyla belirgin bir şekilde hedeflenenin ötesinde potansiyel bilgi eksikliğine odaklanarak öğrenci anlamasını değerlendirdiği görülmektedir. Örneğin *“Farklı renkteki ışıklar nasıl elde edilir?”* sorusuna bir öğrenci *“Farklı renkteki ışıklar beyaz ışığın soğrulmasıyla ortaya çıkarlar. Bu ışıklar kırmızı, turuncu, sarı, yeşil, mavi ve mordur.”* yanıtını vermiştir. 3-10 kodlu öğretmen adayı bu yanıtı *“Öğrenci burada yansımayı soğrulmayla karıştırmış. Yine oluşan ana renkleri de karıştırmış. O yüzden soruyu yanlış cevaplamış”* şeklinde değerlendirmiştir. Öğretmen adayının değerlendirmesi ele alındığında sorunun temel kavramları olan kırılma ve soğrulmanın ötesinde ana renkler içinde değerlendirme yaparak hedeflenenin ötesinde yanlış anlamaları veya bilgi eksikliklerini değerlendirdiği görülmektedir. Bu alt boyutun ileri düzeyi için yani değerlendirme aracılığıyla belirgin bir şekilde hedeflenenin ötesinde ilgili yanlış anlamalara odaklanan öğretmen adayı bulunmadığı tespit edilmiştir.

Son olarak öğretmen adaylarının bilimsel kavramlara ve öğretilmesi hedeflenen düşüncelere yönelik kendi anlama seviyelerinin incelendiği bilimsel doğruluk boyutunda ise öğretmen adaylarının %13.2'si başlangıç düzeyinde olduğu belirlenmiştir. Yani öğretmen adaylarının uygun mantıksal çıkarımların oluşturulmasını engelleyen yüzeysel ve güvenilmez anlayışlara sahip oldukları görülmektedir. Bazı doğru olmayan yorumlamalar oluşturan ve kavramsal problemlerinin var olduğunu ortaya koyan öğretmen adayı oranı ise diğer bir deyişle gelişen düzeyde bulunan öğretmen adayı oranı ise %70.6'dır. Diğer taraftan ileri düzeyde bulunan öğretmen adayı oranı ise %16.2'dir. Bu katılımcıların öğrenci yanıtlarını değerlendirirken yapmış oldukları yorumlamaların bilimsel olarak doğru olduğu görülmektedir. Örneğin gök ve denizin renginin nasıl oluştuğu soruna verilen bir yanıtı değerlendiren 3-44 kodlu öğretmen adayı *“Çok iyi öğrenmemiş. Burada gök renginin oluşması yansıma değil ışığın kırılmasıyla oluşur.”* değerlendirmesini gerçekleştirmiştir. Öğretmen adayının değerlendirme ifadesi incelendiğinde bilimsel olarak herhangi bir hatanın olmadığı görülmektedir.

Veriler bir bütün olarak ele alındığında gerek alan bağımsız boyutta yer alan delil kullanımı gerekse alan bağımlı boyuttaki yorumlama kalitesinde yer alan gelişen düzey

oranları ile bilimsel doğruluk boyutunda yer alan ileri düzey oranlarının birbirlerine yakın çıkması çalışma için gerçekleştirilen analizlerin kendi içerisinde tutarlı olduğunu göstermektedir. Çünkü öğrenci anlamasını değerlendirirken bir öğretmen adayının delil kullanması, kullandığı delilin bilimsel olarak doğru olması öğretmen adayının kendi anlamasının bir yansımını ortaya koymaktadır.

İkinci alt problem kapsamında fen bilgisi öğretmen adaylarının öğrencilerin yazılı cevaplarından yola çıkarak önerdikleri öğretimsel düzenlemelerin ne olduklarına yönelik yapılan analizler neticesinde elde edilen bulgular Tablo 6'da sunulmuştur. Bu bulgular öğrencilerin yazılı cevaplarında öğretmen adaylarının fark ettikleri kavram yanlışları ve kavramsal boşlukları nasıl gidererek öğrencilerin öğrenmelerini desteklemeyi planladıklarını işaret etmektedir.

Tablo 6. Öğretmen Adaylarının Önerdikleri Öğretimsel Düzenlemeler

| Strateji | Sıklık | % | Strateji | Sıklık | % |
|--|--------|------|---|--------|------|
| Konuları tekrar etme | 48 | 70.6 | Gösteri deneyi yapmak | 17 | 25.0 |
| Deney yapma | 46 | 67.6 | Ödev verme | 16 | 23.5 |
| Yeni materyal ve gündelik hayattan örnekler kullanma | 43 | 63.2 | Hata yapılan yerlere göre dersi yeniden tasarlama | 13 | 19.1 |
| Uygulamalı ders yapma | 38 | 55.9 | Konuyla ilgili araştırma projesi yaptırmak | 13 | 19.1 |
| Daha fazla soru çözme | 34 | 50.0 | Önemli kavramlar üzerine daha fazla vakit ayırma | 11 | 16.2 |
| Kavram yanlışları düzeltme | 29 | 42.6 | Bireysel etkinlikler kullanma | 5 | 7.4 |
| Görsel materyaller/ diyagramlar kullanma | 24 | 35.3 | Kavram haritası kullanma | 3 | 4.4 |
| Dersi farklı yöntemlerle anlatma | 21 | 30.9 | Konuyla ilgili video izletme | 1 | 1.5 |
| Gözlem yaptırmak | 21 | 30.9 | | | |

Tablo 6 incelendiğinde fen bilgisi öğretmen adaylarının değerlendirdikleri öğrenci yanıtları üzerinden öğrencilerin gelişimini sağlamak için genellikle konu tekrarı, deney yapma, materyal ve gündelik hayattan yeni örnekler kullanma, uygulamalı ders yapma ve soru çözme yöntemleri gibi 17 farklı yöntemle planlama yaptıkları görülmektedir. Bunun içerisinde düşük oranda bireysel etkinlikler, kavram haritası kullanma ve konuyla ilgili video izletmeyi planlayan öğretmen adayları da bulunmaktadır. Öğretmen adaylarının tercih ettikleri öğretimsel düzenlemelere bakıldığında daha genel ve doğrudan öğrencilerin öğrenme eksiklerine odaklanmayan konuları tekrar etme, deney yapma ve yeni yöntemler ve materyaller kullanma gibi pedagojik düzenlemeleri daha çok tercih ettikleri göze çarpmaktadır. Bunun yanında konuyla ilgili kavram haritası kullanma, video izletme ve dersi kavram yanılgılarına göre ve önemli kavramlara daha fazla vakit ayırarak işleme gibi öğrencilerin zorlandıkları alanlara özgü pedagojik tercihlerin daha düşük oranda olduğu göze çarpmaktadır. Bir başka deyişle, katılımcıların çoğunluğu öğrencilerin öğrenme eksiklerini ve zorlandıkları konuları gidermek için konuya özgü ve etkili bir pedagojik düzenleme yapmada sınırlıdır.

Tartışma ve Sonuç

Bir sınıfta gerçekleştirilecek eğitimin kalitesi o sınıftaki öğretmenin öğretimsel kararları alırken öğrencilerin konu ile ilgili düşünce ve bilgilerini açığa çıkarıp kullanmasına bağlıdır (Bennet, 2011; Furtak, 2012). Bu durum öğretmenlerin öğrenci yanıtlarını nitelikli bir biçimde değerlendirmesini gerekli kılmaktadır. Bu düşünceye paralel olarak bu çalışmada öğretmen adaylarının öğrenci cevaplarında nelere dikkat ettikleri, bu cevaplarda ki zayıf ve kuvvetli kısımları ayırt edip edemedikleri ve bu cevapları nasıl yorumladıkları ve ne gibi öğretimsel faaliyetler planladıkları anlaşılmaya çalışılmıştır.

Sonuçlar öğretmen adaylarının öğrenci cevaplarını değerlendirme kabiliyetlerinin hem alandan bağımsız (Tablo 2) hem de alan bağımlı (Tablo 3) kriterler açısından başlangıç-gelişen seviyesi aralığında olduğu ve istenilen düzeyde bir kabiliyet sergileyemediklerini göstermektedir. Diğer bir ifadeyle öğretmen adaylarının öğrenci cevaplarına yüzeysel baktıkları, cevaplardan çıkarımlar yapmak yerine betimlemeler gerçekleştirdikleri, öğrencilerin problem yaşadıkları hususları tespit etmek yerine cevabın doğruluğu ve yanlışlığını ön plana çıkardıkları, öğrenci cevaplarını göz önünde bulundurarak öğrenci anlaması hakkında spesifik iddialar üretmedikleri, yorumlama yaparken ortaya koydukları iddiaları destekleyecek detaylı deliller sunmadıkları ve öğrencinin farklı yanıtları arasında bağlantı kurmadan kavramsal öğrenmesi hakkında yorumlamalar gerçekleştirdikleri

görülmektedir. Araştırmalar etkili bir öğretimin gerçekleşebilmesi için öğretmenin öğrencilerin mevcut durumları hakkında zengin veriler elde etmeleri ve bu verileri derinlemesine inceleyerek öğretimde kullanmalarının önemli olduğunu vurgulamaktadırlar (Furtak 2012; Harshman ve Yeziarski, 2015; NRC, 2012; Talanquer ve diğerleri, 2015). Fakat bu araştırmanın sonuçlarının da gösterdiği gibi öğretmen adayları genelde yüzeysel değerlendirmeler yaparak öğrenci cevaplarının doğru olup-olmadığına odaklanmaktadır. Bu yüzden de elde edilen verileri öğretimi destekleyecek şekilde kullanamamaktadırlar. Bu durumun temel nedenleri olarak öğretmen adaylarının değerlendirme yaptıkları konuya yönelik alan bilgisi eksikliği çekiyor olabilecekleri ve değerlendirmeyi öğretimi desteklemekten, şekillendirmekten ziyade not verme odaklı gerçekleştirmeleri gösterilebilir. Kang ve Anderson (2015) da paralel olarak bu durumun nedenini yeteri kadar gelişmemiş alan, öğrenme-öğretme ve değerlendirme bilgisi ve değerlendirme becerisi olabileceğini belirtmektedir. Nitekim bu araştırma sonucunda hem ortaya konulan alan bağımlı alt boyut sonuçlarda öğretmen adaylarının büyük bir oranının (%70.6) değerlendirme gerçekleştirirken alan bilgisine bağlı olarak bazı doğru olmayan yorumlamalar ürettiklerini hem de değerlendirme yönelimlerinin öğrenci yanıtlarının arkasında yatan mantıksal akıl yürütmeyi anlamak için onları analiz etmeye odaklanmadıklarını ortaya koymaktadır. Benzer sonuçların Aydeniz ve Doğan (2016), Gotwals ve Brimingham (2016), Otera (2006) ve Nathan ve Otera (2008) tarafından gerçekleştirilen çalışmalarda da ortaya konulduğu üzere öğretmen adaylarının öğrenci cevaplarına yüzeysel yaklaştıkları ve cevapları doğru-yanlış şeklinde yorumlamanın ötesine geçemedikleri görülmektedir. Ayrıca Aydeniz ve Doğan (2016) fen bilgisi öğretmen adaylarının biçimlendirici değerlendirmeye yönelik pedagojik kabiliyetlerini inceledikleri çalışmada öğretmen adaylarının öğrenme eksiklerine odaklanmadıklarını ve değerlendirmeyi öğretimsel kararlar almaya yönelik kullanımında yetersiz olduklarını ortaya koymaktadır. Bu kapsamda öğretmen adaylarının eğitimleri esnasında öğretim elemanları hem değerlendirme yaparken not vermeden ziyade biçimlendirici değerlendirmenin önemli olduğu konusunda örnek olmalı hem de farklı uygulamalar aracılığı ile etkili bir değerlendirmenin gerçekleşmesi için öğrenci cevaplarını değerlendirilirken alan bağımlı ve bağımsız kriterleri kullanmaları teşvik edilmelidir.

Diğer taraftan, bu çalışmada katılımcıların çoğunluğunun öğrenci cevaplarını yorumlarken delil kullanmadıkları bazı katılımcıların kullandıkları delillerde ise bilimsel olarak hataların olduğu belirlenmiştir. Bu sonuçlar alan yazın tarafından da desteklenmektedir. Örneğin Benzer şekilde Talanquer ve diğerleri (2015) öğretmen

adaylarının öğrenci cevaplarını kullanırken bilimsel olarak doğru olmayan bilgileri kullandıkları ve cevaplardan yaptıkları çıkarımların ise öğrenmeyi desteklemekten uzak olduğunu göstermektedir. Öğrenci verileri değerlendirilirken kaliteli ve çıkarımsal bir değerlendirmenin yapılabilmesi için delillerin kullanılması hem öğretimin değerlendirilmesinde hem de öğrenmenin desteklenmesinde önemlidir (NRC, 2012). Çünkü değerlendirilen kavramlar hakkında bilimsel bilgileri yetersiz olan bir öğretmenin öğrenmeyi değerlendirmesi güvenilir sonuçlar elde etmeyi tehlikeye attığı gibi biçimlendirici değerlendirme süreci kapsamında kavram yanlışlarının ve öğrenme boşluklarının belirlenerek öğretimi desteklemeyi de zorlaştırmaktadır (Coffey, Hammer, Levin ve Grant, 2011). Bundan dolayı değerlendirici kimliği taşıyacak olan öğretmen adaylarının hem kendi alanlarıyla ilgili alan bilgisine sahip olmaları hem de değerlendirme sürecinde deliller kullanmaları önem arz etmektedir. Bunun için ise öğretmenlik eğitimleri esnasında pedagojik bilginin yanında zengin bir alan bilgisiyle de öğretmen adayları desteklenmelidir.

Benzer şekilde bilimsel doğruluk ve delil kullanmaya paralel olarak öğretmen adaylarının üretken düşünce kapsamında da başlangıç-gelişen düzeyde olduğu görülmektedir. Öğrenci cevapları üzerine üretken bir yorumun yapılabilmesi için bilimsel doğruluk ve delil kullanımı önemli olmakla birlikte yeterli değildir. Üretken bir yorum yapılabilmesi için öğretmen adaylarının ilgili kavramlar kapsamında derinlemesine bir alan bilgisine sahip olmaları, kavramları önem ve öncelik sırasına göre ayırt edebilmeleri ve kavramlar arasındaki ilişkileri fark edebilmeleri gerekmektedir (Harshman ve Yeziarski, 2015; Talanquer ve diğerleri, 2015). Ancak bu sayede değerlendirme yaparken öğrencilerin nerelerde zorlandıklarını, hangi kavramsal ilişkileri kuramadıklarını ve kavramsal boşlukları belirleyerek düşük yeterlilikteki öğrencilerin öğrenmelerini destekleyebilecek üretken çıkarımlar yapabilir ve öğretimsel kararlar alabilirler (Gottheiner ve Siegel, 2012). Araştırmacılar, bu çalışmada da görüldüğü gibi, değerlendirme sürecinde en çok zorlanan alanın elde edilen verileri doğru ve üretken şekilde yorumlayıp öğretimi destekleyici şekilde kullanmak olduğunu belirtmişlerdir (Ateh, 2015; Kang ve Anderson, 2015; İzci & Siegel, 2019). Bu eksiklerini giderebilmek için öğretmenlik programlarında alana özgü kavram yanlışlarını belirleme ve giderme konusunda öğretmen adaylarının repertuarlarının zenginleştirilmesi önemlidir.

Ayrıca öğretmen adaylarının değerlendirme sürecinde dar kapsamda kavram ve düşünceler üzerine odaklandıkları görülmektedir. Bilimsel doğruluk açısından ele alındığında ise öğretmen adaylarının çoğunun gelişen düzeyde olduğu yani değerlendirme

sürecinde bazen doğru olmayan yorumlamalar oluşturdukları ve kavramsal problemlerinin var olduğu görülmektedir. Bu sonuçlar, öğretmen adaylarının eğitimlerinde ölçme değerlendirme hususunda öğrenci cevaplarının nasıl yorumlanıp kullanılabilmesine yönelik uygulamalı eğitimler verilmesinin önem arz ettiğini ve bu şekilde pedagojik muhakeme becerilerini geliştirebileceklerini ortaya koymaktadır.

Son olarak sonuçlar öğretmen adaylarının öğrenci cevaplarından yola çıkarak farklı yöntemlerle öğrencilerin öğrenme zorluklarını gidermeye çalıştıklarını göstermektedir (Tablo 6). Fakat katılımcıların büyük bir çoğunluğu, diğer araştırmacılar tarafından da belirtildiği şekilde (Aydeniz ve Doğan, 2016; Gottheiner ve Siegel, 2012) klasik yöntemler olan dersi tekrar anlatma, problem çözme ve deney yaptırma gibi genel ve öğrencilerin eksikliklerini ve zorluklarını birebir gidermede yetersiz yöntemlerle öğrenci başarısını artırmayı düşünmektedirler. Öğrencilerin zorlukları ve kavram yanılgıları giderilmediği müddetçe başarının artmasının sınırlı olacağı da bilinmektedir (Black ve Wiliam, 2006). Bu yüzden, öğretmen adayları farklı öğretim yöntemleri bilme ve kullanma hususunda desteklenmelidir.

Etik Kurul İzin Bilgisi: Bu araştırma, Van Yüzüncü Yıl Üniversitesi Rektörlüğü Sosyal ve Beşerî Bilimler Etik Kurulu'nun 13/05/2020 tarihli 2020/03-16. sayılı kararı ile alınan izinle yürütülmüştür.

Yazar Çıkar Çatışması Bilgisi: Yazarlar arasında herhangi bir çıkar çatışması bulunmamaktadır.

Yazar Katkısı: Yazar katkısı eşittir.

Kaynakça

- Ateh, C. M. (2015). Science teachers' elicitation practices: Insights for formative assessment. *Educational Assessment, 20*(2), 112-131.
- Aydeniz, M., & Dogan, A. (2016). Exploring pre-service science teachers' pedagogical capacity for formative assessment through analyses of student answers. *Research in Science & Technological Education, 34*(2), 125-141.

- Barnhart, T., & van Es, E. (2015). Studying teacher noticing: Examining the relationship among preservice science teachers' ability to attend, analyze and respond to student thinking. *Teaching and Teacher Education, 45*, 83–93.
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice, 18* (1), 5-25.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7-74.
- Black, P., & Wiliam, D. (2006). Developing a theory of formative assessment. In J. Gardner (Ed.), *Assessment and learning* (pp. 81-100). London: Sage
- Chan, K. K. H., & Yau, K. W. (2021). Using video-based interviews to investigate pre-service secondary science teachers' situation-specific skills for informal formative assessment. *International Journal of Science and Mathematics Education, 19*, 289-311.
- Coffey, J. E., Hammer, D., Levin, D. M., & Grant, T. (2011). The missing disciplinary substance of formative assessment. *Journal of Research in Science Teaching, 48*(10), 1109–1136.
- Dini, V., Sevian, H., Caushi, K., & Orduña Picón, R. (2020). Characterizing the formative assessment enactment of experienced science teachers. *Science Education, 104*(2), 290–325.
- Furtak, E. M. (2012). Linking a learning progression for natural selection to teachers' enactment of formative assessment. *Journal of Research in Science Teaching, 49*(9), 1181-1210.
- Gottheiner, D. G., & Siegel, M. A. (2012). Experienced middle school science teachers' assessment literacy: Investigating knowledge of students' conceptions in genetics and ways to shape instruction. *Journal of Science Teacher Education, 23*, 531–557.
- Gotwals, A. W., & Birmingham, D. (2016). Eliciting, identifying, interpreting, and responding to students' ideas: Teacher candidates' growth in formative assessment practices. *Research in Science Education, 46*(1), 365–388.
- Harshman, J., & Yeziarski, E. (2015). Guiding teaching with assessments: High school chemistry teachers' use of data-driven inquiry. *Chemistry Education Research and Practice, 16*(1), 93-103.

- İzci, K. (2018). Turkish science teacher candidates understandings of equitable assessment and their plans about it. *Journal of Education in Science, Environment and Health*, 4(2), 193-205.
- İzci, K. & Siegel, M. (2019). Investigation of an alternatively certified new high school chemistry teacher's assessment literacy. *International Journal of Education in Mathematics, Science and Technology*, 7(1), 1-19. DOI: 10.18404/ijemst.473605
- İzci, K. & Şardağ, M. (2016). Fen bilgisi öğretmen adaylarının sınıf içi ölçme-değerlendirme algıları. *Necatibey Faculty of Education Electronic Journal of Science & Mathematics Education*, 10(1), 439-471.
- Kang, H. & Anderson, C. W. (2015). Supporting preservice science teachers' ability to attend and respond to student thinking by design. *Science Education*, 99(5), 863–895.
- Lam, D. S. H. & Chan, K. K. H. (2020) Characterising preservice secondary science teachers' noticing of different forms of evidence of student thinking. *International Journal of Science Education*, 42(4), 576-597.
- Larkin, D. (2012). Misconceptions about misconceptions: Preservice Secondary science teachers' views on the value and role of student ideas. *Science Education* 96, 927–959.
- Levin, D., Hammer, D., & Coffey, J. (2009). Novice teachers' attention to student thinking. *Journal of Teacher Education* 60, 142–154.
- Luna, M. J., & Sherin, M. G. (2017). Using a video club design to promote teacher attention to students' ideas in science. *Teaching and Teacher Education*, 66, 282–294.
- Lyon, E. G. (2011). Beliefs, practices, and reflection: Exploring a science teacher's classroom assessment through the assessment triangle model. *Journal of Science Teacher Education*, 22(5), 417-435.
- Miles, M. B. & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. (2nd Edition). California: SAGE Publications.
- Murray, S. A.; Huie, R.; Lewis, R.; Balicki, S.; Clinchot, M.; Banks, G.; Talanquer, V.; Sevian, H. (2020). Teachers' noticing, interpreting, and acting on students' chemical ideas in written work. *Journal of Chemical Education*, 97(10), 3478–3489.

- National Research Council. (2012) *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington: The National Academies Press.
- Otero, V. (2006). Moving beyond the ‘Get it or don't’ conception of formative assessment. *Journal of Teacher Education*, 57(3), 247-255.
- Otero, V., & Nathan, M. J. (2008). Preservice elementary teachers' views of their students' prior knowledge of science. *Journal of Research in Science Teaching*, 45(4), 497-523.
- Ruiz-Primo, M. A. & Furtak E. M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching*, 44(1), 57-77.
- Sabel, J.L., Forbes, C.T., & Zangori, L. (2015). Promoting prospective elementary teachers' learning to use formative assessment for life science instruction. *Journal of Science Teacher Education*, 26(4), 419–445.
- Schack, E. O., Fisher, M. H., & Wilhelm, J. A. (2017). *Teacher noticing: Bridging and broadening perspectives, contexts and frameworks*. Springer.
- Siegel, M. A., & Wissehr, C. (2011). Preparing for the plunge: Preservice teachers' assessment literacy. *Journal of Science Teacher Education*, 22(4), 371-391.
- Son, J. (2013). How pre-service teachers interpret and respond to student errors: Ratio and proportion in similar rectangles. *Educational Studies in Mathematics* 84(1), 49–70.
- Talanquer, V., Bolger, M., & Tomanek, D. (2015). Exploring prospective teachers' Assessment practices: Noticing and interpreting student understanding in the assessment of written work. *Journal of Research in Science Teaching*, 52(5), 585-609.
- Vogelzang, J. & Admiraal, W. F. (2017). Classroom action research on formative assessment in a context-based chemistry course. *Educational Action Research*, 25(1), 155-166.
- Yin, R. (2009). *Case study research: Design and methods*. Thousand Oaks, CA: Sage.
- Yıldırım, A., & Şimşek, H. (2016). *Sosyal bilimlerde nitel araştırma yöntemleri*. (10. Baskı). Ankara: Seçkin Yayıncılık.



Exploring Prospective Science Teachers' Pedagogical Reasoning Skills Through Analyses of Student Written Responses*

Metin SARDAG* Kemal IZCI**

• **Received:** 10.10.2020 • **Accepted:** 18.05.2021 • **Online First:** 24.05.2021

Abstract

The purpose of this study was to explore prospective science teachers' pedagogical reasoning skills by analyzing what they look for in middle school student's written exam responses, how they evaluate these responses, and how they plan to use these responses. The study was conducted as a case study. Seventy-six prospective science teachers who have taken the assessment and evaluation course participated in the study. The data was collected using a form that contained the responses of middle school students to the two open-ended questions on color formation and light. It was analyzed through the content analysis method. According to the findings, prospective teachers evaluated examined the student responses superficially, prioritized whether the responses were correct or incorrect, over-diagnosing the subject matters students had problems with, did not provide evidence for the claims they put forward while evaluating, and developed evaluations on students' conceptual learning without establishing a connection among a student's responses to different questions. These results reveal the significance of having practical training on how to evaluate and use student responses as part of prospective teachers' training on assessment and evaluation and that, through this training, they can develop their pedagogical reasoning skills.

Keywords: pedagogical reasoning, assessment, and evaluation, formative assessment, teacher training, prospective teachers

Cited:

Şardağ, M. & İzci, K. (2021). Exploring prospective science teachers' pedagogical reasoning skills through analyses of student-written responses. *Pamukkale University Journal of Education*, 53, 495-520. doi: 10.9779/pauefd.808745

* This study was presented at the 12th National Science and Mathematics Education Congress held in Trabzon between 28th and 30th September 2016.

* Res. Asst. Dr., Van Yuzuncu Yil University, <https://orcid.org/0000-0003-2162-8289>, metinsardag@yyu.edu.tr

** Assoc. Prof. Dr., Necmettin Erbakan University <https://orcid.org/0000-0002-4228-8845>, kizci@erbakan.edu.tr

Introduction

Effective science education must be grounded on the knowledge and ideas students carry to the learning environment and have developed through their previous life experiences, and new learning must be built on this knowledge and ideas (Barnhart and van Es, 2015; Ruiz-Primo and Furtak, 2007; Talanquer, Tomanek and Novodvorsky, 2015). This process is highly complex and challenging, but it also entrusts a considerably significant responsibility to teachers. As part of the process, the teachers are expected to change their pedagogical approach and focus on "initial ideas students bring to school and how they may best develop an understanding" of scientific practices rather than focusing only on whether students' responses are correct or incorrect (National Research Council-NRC, 2012, p. 256).

Science teachers can obtain information on their students through formal methods such as homework, exams, and lab reports and informal methods such as in-class discussions, oral assessments, and observation. In other words, science teachers have the opportunity to receive feedback from their students through many methods and plan, evaluate, and revise their teaching process according to this feedback (Harshman and Yeziarski, 2015). Teachers are expected to incorporate the information they gathered about their students through various methods while making decisions about their teaching in this process. This process is also known as data-driven inquiry or formative assessment in literature (Furtak, 2012; Harshman and Yeziarski, 2015).

Many studies reveal that formative assessment has a positive effect on students' learning and motivation (Black and Wiliam, 1998; Vogelzanga and Admiraal, 2017), and researchers concentrate on teachers' beliefs, perceptions, knowledge, and practice in studying formative assessment (Furtak, 2012; Lyon, 2011; Siegel and Wissner, 2011). The results of studies indicate that the teachers' abilities to reveal students' learning, effectively interpret the information they have received, provide feedback, and revise their lectures are critical factors that increase student success (Furtak, 2012; Ruiz-Primo and Furtak, 2007). Therefore, teachers expect teachers to analyze student responses well, detect problematic parts in the responses, interpret these productively, and support student learning by helping them in areas they experience difficulty (Bennett, 2011). Nonetheless, existing studies, unfortunately, show that teachers evaluate student responses superficially, and their interpretation of these responses cannot support conceptual learning because rather than revealing problematic areas that prevent students' conceptual understanding, they focus on the correctness or incorrectness of the responses (Ateh, 2015; Furtak, 2012). Thus, training

teachers with pedagogical reasoning skills who pay attention to the problematic areas in student responses, understand how these limit conceptual understanding, and support conceptual learning through appropriate methods using this information is a significant issue. As shown in Figure 1, the process requiring teachers to use pedagogical reasoning skills has critical importance and forms an important part of effective learning.

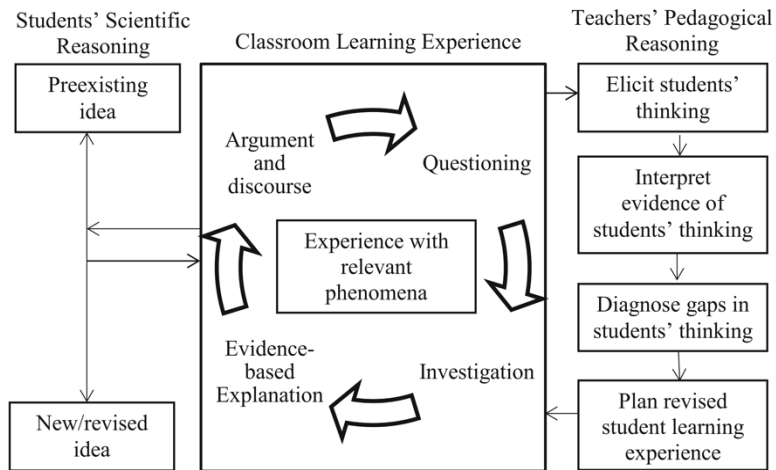


Figure 1. *The relationship between responsive teaching process and students' science learning (Sabel, Forbes and Zangori, 2015, p. 422).*

Pedagogical reasoning skills both affect and are affected by the teaching environment. In other words, for teachers to exhibit a teaching approach that is responsive to the students' preliminary knowledge affects both the teaching environment they will offer and the development of students' understanding of the target phenomenon (Sabel et al., 2015).

It is observed that, globally, many researchers have analyzed prospective teachers' pedagogical reasoning skills in order to determine and improve their ability to pay attention to and evaluate student responses (Barnhart and van Es, 2015; Luna and Sherin, 2017; Talanquer et al., 2015). If these studies are examined in terms of their approach, it is observed that studies conducted on pedagogical reasoning skills primarily started and concentrated on math education (Luna and Sherin, 2017; Schack, Fisher and Wilhelm, 2017; Son, 2013). Studies have recently been conducted on science education (Larkin, 2012; Levin, Hammer, and Coffey, 2009; Talanquer et al., 2015). In these studies, reasoning skills are analyzed within teaching practice lectures, micro-teaching practices, and students' written responses. Besides, video recordings and written documents are usually used as data sources (Lam and Chan, 2020; Luna and Sherin, 2017). Video recordings and observations are generally used to examine a teacher's skills in momentarily (in-the-moment noticing)

analyzing, interpreting, and using the data he/she obtained from students through various methods (e.g., asking questions) in his/her teaching. On the other hand, written documents are preferred to analyze teachers' skills in analyzing, interpreting, and using the data collected from students outside of the classroom time. In this case (delayed noticing), teachers have more time to analyze student responses, and they can make effective educational decisions by analyzing student responses in more detail (Lam and Chan, 2020).

For this reason, studies conducted on prospective teachers focus on delayed noticing in order to support reflective teaching. Two distinct approaches become prominent when researchers' data acquisition methods for belated perception are examined. While the first of these approaches involve prospective teachers individually evaluating student responses in writing (Talanquer et al., 2015), the second approach involves prospective teachers cooperatively (in groups of 3-5) evaluating written student responses (Murray et al., 2020).

The results of studies conducted on teachers' pedagogical reasoning skills indicate mixed findings. Dini, Sevian, Caushi, and Picon (2020) suggest a new model based on comparing teachers who exhibit these mixed findings. According to this model, teachers exhibit their pedagogical reasoning skills either more authoritative or more dialogic. Teachers who exhibit a more authoritative approach privilege certain special knowledge and skills while examining students' thoughts, checking whether they are scientifically correct, and suggesting educational methods that will supply the students directly with the scientific truth. In a more dialogic approach, teachers show regard to students' existing knowledge and experiences, determine the logical inconsistencies in their responses, develop hypotheses for the causes of these inconsistencies, and suggest providing educational environments in which students can realize these inconsistencies, contemplate them, and improve their learning. Results of other studies support the model suggested by Dini et al. (2020). While evaluating student responses, the majority of teachers/prospective teachers describe the responses, focus on the correct-incorrectness of the information in the responses, and suggest common pedagogical methods to supply the students with the scientific truth (Ateh, 2015; Aydeniz and Doğan, 2016; Chan and Yau, 2021; Furtak, 2012; Gotwals and Brimingham, 2016; Murray et al., 2020; Talanquer et al., 2015). Even though their number is limited, teachers/prospective teachers in certain studies focus on hypothetical justifications of the student responses and their underlying logic, interpret the thoughts students have presented and their hypothetical justifications, and suggest educational

methods in which students are at the center and can improve their learning while evaluating student responses (Dini et al., 2020; Murray et al., 2020; Talanquer et al., 2015).

When studies on improving pedagogical reasoning skills are examined, it can be seen that they make positive contributions. For example, Gotwals and Brimingham (2016) have analyzed prospective science teachers' pedagogical reasoning skills and their improvement using the projections of lecture/unit plans on written student responses and lecture videos and lecture videos. This analysis has especially focused on prospective teachers' skills in eliciting student thoughts, identifying significant areas, interpreting, and educational responding. Results show that participants' skills in eliciting student thoughts have developed in time and that they have started asking higher quality and comprehensive questions that will help elicit student thoughts. Similarly, Barnhart and van Es (2015) have revealed that prospective teachers successfully distinguished significant parts of their lectures through clips chosen by prospective teachers from videos of their lectures. The study results have also shown that distinguishing significant parts of teaching is a prerequisite of successful analysis and educational organization. However, a highly developed skill in distinguishing may not result in a successful analysis and a highly developed skill in responding. Lastly, it is indicated that prospective teachers who use assessment tools created with the support of mentor teachers have access to a rich pool of student responses, and they can support student success when they incorporate it in their teaching (Kang and Anderson, 2015).

When the analyses are brought down to the national level, despite studies on science teachers' perceptions on assessment and evaluation (Izci & Sardag, 2016), which constitutes the foundation of pedagogical reasoning skills, or their implementation (Izci, 2018), except Aydeniz and Doğan's (2016) study, studies that focus on determining and improving teachers' pedagogical reasoning skills are almost nonexistent. Aydeniz and Doğan (2016), in their study, aimed at determining prospective science teachers' pedagogical reasoning skills by having them examine responses of 8th-grade students to 7 questions on Newton's laws of motion, kinetic energy, series, and parallel circuits, and static electricity. Researchers have focused heavily on prospective teachers' feedback approaches in their study. However, the literature demonstrates that field dependency is significant among factors that influence the effectiveness of formative assessment (Bennet, 2011). Thus, it is significant to carry out studies that analyze prospective science teachers' pedagogical reasoning skills based on a specific subject matter rather than conducting studies based on multiple subjects. While

there are subject-specific studies in the international literature on pedagogical reasoning skills (Talanquer et al., 2015), the national literature lacks such studies. For this reason, this study analyzes pedagogical reasoning skills based on a specific subject matter. In addition, when the studies discussed above are considered as a whole, it is observed that they include activities conducted in order to improve pedagogical reasoning skills and focus on the factors affecting the exhibition of pedagogical reasoning skills and their effects. This study contributes to the literature by examining factors that may be effective in developing pedagogical reasoning skills by focusing on what prospective science teachers look for in students' written exam responses, how they interpret these responses, and how they plan to use these responses. Concordantly, the purpose of this study was to explore prospective science teachers' pedagogical reasoning skills by analyzing what they look for in middle school students' written exam responses, how they evaluate these responses, and how they plan to use these responses. Hence, it is expected to both fill the deficiency in the literature discussed above and the existing research findings. Accordingly, it seeks answers to the following main research questions.

1. How do the prospective science teachers evaluate students' written exam responses?
2. What educational regulations are suggested by prospective science teachers based on students' written exam responses?

Method

This study was conducted as a case study, one of the qualitative research designs. A case study is a research method used often in many fields and provides an opportunity to reveal characteristics of real-life events such as small group behavior and administrative processes (Yin, 2009) and to reach a holistic interpretation of events that are subject to study (Yıldırım and Şimşek, 2016). Since the purpose of this study is to examine a specific group's approach to evaluation and to reveal their pedagogical reasonings holistically based on this examination, the study was decided to be conducted as a case study. Accordingly, prospective science teachers' evaluation of student responses is assessed as a case, and the research aims to reveal their pedagogical reasoning skills.

This study was conducted with the approval granted by Van Yüzüncü Yıl University Presidency Social Sciences and Humanities Ethics Committee on 13/05/2020 with the serial 2020/03-16.

Participants

The study group was determined using criterion sampling among purposive sampling methods. The selections are made based on a set of criteria determined by the researchers in criterion sampling method (Yıldırım and Şimşek, 2016). Accordingly, the study group was determined as 76 prospective science teachers who study in their third or fourth years at a state university in eastern Turkey and completed the assessment and evaluation course. The data was collected during the spring semester of the 2015-2016 academic year.

Data Collection Tools

A form developed by the researchers was used as the data collection tool in this study. A 2-step process was conducted while developing the form. The first step was to contact a science teacher and obtain written exam papers for a science class with open-ended conceptual questions. These exam papers were examined, and two questions were selected. The questions are "How is light obtained in different colors?" and "How do the blue colors of the sky and the seas emerge?" The responses of five students who answered these questions and received varying grades in different success levels ranging from high to low in the exam were then determined. These student responses are provided below in Table 1.

Table 1. *Responses of Middle-School Students to Questions in the Science Exam*

| Students | Responses |
|-----------|---|
| Student-1 | <p>1) We obtain by mixing lights together.</p> <p>2) Since space is black and the sunlight is yellow, the color blue is created when these two lights hit one another. This is why the sky and the sea seem blue.</p> |
| Student-2 | <p>1) Lights in different colors emerge through the absorption of white light. These lights are red, orange, yellow, blue, and purple.</p> <p>2) The sky and the sea are actually transparent. However, we see them blue because light hits and is reflected.</p> |
| Student-3 | <p>1) The beams that hit the three main colors, red, blue, and yellow, make the other colors emerge.</p> <p>2) The sky is blue. The blueness of the sky is reflected through beams that</p> |

touch white water, causing the sea to turn blue as well.

Student-4 1) Other colors emerge through the reflection and absorption of white light. For example, inside the white light, there is the color red as well. For instance, if a red light is reflected on white light, it looks red, because the color white emerges through the combination of other colors.

2) There is the color blue inside the white light, as well. So, the white light reflects the color blue onto the sky and the sea. The sky and the sea become blue.

Student-5 2) Lights in different colors are obtained through the filtering method. Traffic lights can be given as an example of this. Underneath red, yellow, and red lights there is actually white light. But other colors are absorbed through filtering and we only see red, yellow, and green colors.

2) The sea and the sky are actually colorless. They are perceived as blue because they only reflect blue from white light colors and absorb other colors. This is why the sky and the sea appear blue.

In the second step, a questionnaire in which the participants will evaluate the student responses was created using these responses. Following the completion of the questionnaire, prospective teachers were asked to answer in writing to the following questions based on student responses from the questionnaire within a class hour

Question 1) How would you decide whether each student has understood the basic concepts in relation with the color formation or not using the student responses below? Please discuss each student's response below.

Question 2) What do you think are the problematic instances that mean that he/she misunderstood or only partially understood the topic, based on each student's responses?

Question 3) List student's levels of understanding of the subject from low to high according to their responses. Explain why you listed in this order with evidence.

Question 4) What may these responses by the students show you as a teacher and what do they help you to do?

Question 5) What can you do as a prospective teacher based on these responses in order for your students who did not understand the subject to understand color formation?

Data Analysis

The data collected through the data collection form were analyzed using teachers' evaluation of student responses dimension and criteria developed by Talanquer et al. (2015). These dimensions are discussed under two headings. The first is the domain-neutral dimension, and the second is the domain-dependent dimension. Domain-neutral dimension consists of 5 different sub-dimensions. Domain-neutral dimensions and their criteria are given in Table 2.

Table 2. *Domain-Neutral Dimensions and Their Criteria*

| Dimension | Level | Description |
|------------------------|----------|--|
| Evaluation Approach | Novice | The evaluation mostly depends on student response's description rather than making inferences about student understanding. |
| | Emerging | The evaluation is a mix of description and inference statements. |
| | Advanced | At the center of the evaluation are the inferences on student thinking and understanding. |
| Evaluation Orientation | Novice | Focuses on things that students fail to do or say, correct or incorrect responses without looking for meaning in the student's existing responses. |
| | Emerging | Focuses on certain correct responses and finding some meaning in responses. |
| | Advanced | Focuses on analyzing student responses in order to understand the logical reasoning that lay behind student responses. |
| Specificity in the | Novice | Produces general situations about student's attitudes and knowledge based on certain responses. |

| | | |
|-----------------------|----------|--|
| Evaluation | Emerging | Produces situations focused more on whether the student has understood or misunderstood a subject. |
| | Advanced | Produces specific claims regarding student's understanding based on student responses. |
| Analysis of Coherence | Novice | Focuses on determining and describing mistakes or misconceptions as isolated errors. |
| | Emerging | Notices consistency or inconsistency in student responses. |
| | Advanced | Uses consistencies and inconsistencies in student responses in order to build interpretations about student understanding. |
| Use of Evidence | Novice | Clear evidence is not presented to support claims if inferences are made. |
| | Emerging | Evidence is presented in order to support some of the inferences made about student understanding. |
| | Advanced | Inferences made about student understanding are always connected to evidence gathered from student responses. |

Whereas domain-dependent dimension consists of separate dimensions and their criteria. These dimensions and their criteria are given in Table 3.

Table 3. *Domain-Dependent Dimensions and Their Criteria*

| Dimension | Level | Description |
|-------------------------------|----------|--|
| Quality of the Interpretation | Novice | Interpretations of student understanding are not supported by accessible evidence. |
| | Emerging | Evidence provided for interpretations is limited, superficial, or stereotypical |
| | Advanced | Logical interpretations of student understanding are produced. |

| | | |
|-------------------------|----------|--|
| | Novice | Focus is on insignificant details instead of on basic concepts of the subject while evaluating student's comprehension of the subject. |
| Productive Thinking | Emerging | Focus is on both the basic concepts of the subject and insignificant details while evaluating student's comprehension of the subject |
| | Advanced | Basic concepts of the subject are distinguished from insignificant details, and the focus is on basic concepts of the subject while evaluating student's comprehension of the subject. |
| Scope of the Evaluation | Novice | There is limited focus on the assessment of concepts and ideas. |
| | Emerging | It recognizes a potential lack of knowledge in related areas prominently beyond the target of the evaluation. |
| | Advanced | It recognizes a potential lack of knowledge in areas prominently beyond the target of the evaluation. |
| Scientific Accuracy | Novice | Shallow or unreliable understanding of basic knowledge hinders the production of logical deductions. |
| | Emerging | Some incorrect interpretations are made. |
| | Advanced | Interpretations are scientifically accurate. |

Percent and frequency values of the data are calculated after the data were analyzed according to the criteria indicated in Tables 2 and 3. The researchers created an Excel file during data analysis, and each prospective teacher's evaluation of each student response was analyzed in domain-dependent and domain-neutral dimensions. The data were coded as 1 or 2 during analysis. For example, for the use of evidence dimension, if the prospective teacher did not use evidence while evaluating a student response, it was coded as 1; if he/she did use evidence, it was coded as 2. Then, the arithmetic mean of the coding of the same prospective teacher for all student responses was calculated. If this arithmetic means are 1.1 or below, it

is designated as a novice, if it is between 1.2-1.9, it is designated as emerging, and if it is 2 it is designated as advanced. In determining these ranges, prospective teachers' evaluation states of responses to the two questions by five students are considered. A prospective teacher makes 10 evaluations in total. Among these ten evaluations, if he/she evaluates at least two responses by the same student according to the criteria, the critical 1.2 value is obtained. The value 1.9 is obtained if he/she evaluates at least nine out of ten student responses according to the relevant criteria. If the prospective teacher evaluates all student responses according to the relevant criteria, the value 2 is obtained. Before the analyses were conducted, researchers discussed and reached a consensus about the domain-dependent and domain-neutral dimensions and how they should be analyzed. Then, the responses of 20 prospective teachers were independently coded by two different researchers, and then the differences were discussed and reconciled. The coherence between researchers in this process was calculated through the reliability formula developed by Miles and Huberman (1994). The index of fit was calculated as .93. Moreover, the researcher who analyzed the entire data has checked the consistency of the coding according to the Miles and Huberman (1994) reliability formula by reanalyzing around 15% of the entire data at another time. As a result, the coding completed in this 15% part was the same, and the fit index was calculated as 1.00.

Findings (Results)

Findings presented in Table 4 were obtained for the domain-neutral dimension when the evaluations made by prospective science teachers on student responses were analyzed.

Table 4. *Domain-Neutral Dimensions and Their Rates of Implementation in the Evaluations*

| Dimensions | Novice (1) | | Emerging (2) | | Advanced (3) | | Mean |
|-------------------------------|------------|------|--------------|------|--------------|-----|------|
| | f | % | f | % | f | % | |
| Evaluation Approach | 42 | 61.8 | 26 | 38.2 | 0 | 0 | 1.38 |
| Evaluation Orientation | 44 | 64.7 | 24 | 35.3 | 0 | 0 | 1.35 |
| Specificity in the Evaluation | 55 | 80.9 | 13 | 19.1 | 0 | 0 | 1.19 |
| Analysis of Coherence | 41 | 60.3 | 23 | 33.8 | 2 | 2.9 | 1.41 |
| Use of Evidence | 46 | 67.6 | 22 | 32.4 | 0 | 0 | 1.32 |

The evaluation approach adopted by prospective teachers was determined as novice level with a 61.8% ratio, which means that they used the descriptions of the responses while evaluating them. To illustrate with an example, the student responded to the question "How do the blue colors of the sky and the seas emerge?" with the statement, "*The sky and the sea are actually transparent. However, we see them blue because light hits and is reflected.*" A prospective teacher has evaluated this response with the statement, "*This student also has learned **that the sea and the sky are bright and that we see them as blue as a result of the reflection** of the light that hits and has responded correctly.*" When the evaluation is analyzed, it can be observed that the student response was restated or described. On the other hand, it is observed that 38.2% of the prospective teachers have also engaged in deduction besides descriptive statements while evaluating, which means they are at the emerging level. The evaluation of the prospective teacher with the code 3-36 can be given as an example. One student responded to the question above with the statement, "*There is the color blue inside the white light, as well. Therefore, the white light reflects the color blue onto the sky and the sea.*" The prospective teacher has evaluated this response with the statement, "*He/she thinks that reflection occurs both in the sky and in the sea, so he/she responded in this way. There is a lack of knowledge.*" When the evaluation is analyzed, it can be observed that inferences about student understanding and knowledge were made. However, none of the prospective teachers' evaluation approaches were determined as advanced, which means that inferences about student thinking and understanding were not at the center of the evaluations.

For evaluation orientations, which are in parallel with the case of evaluation approaches, 64.7% of the prospective teachers evaluated student responses only according to their correctness or incorrectness (e.g., *He/She responded to the question correctly.*), whereas 35.3% percent of them focused on underlying reasons behind responses through analyzing a large portion of student responses. For instance, the prospective teacher with the code 3-28 evaluated a student who responded to the question about how the blue color of the sky and the seas emerge with the statement. *So, the white light reflects the color blue onto the sky and the sea. The sky and the sea become blue.*" through the statement "*Refraction of the light occurs at the sky. In contrast, light reflects on the sea. The student mistakenly connected the phenomena related to color that occurs in the sky and the sea.*" However, there are no advanced-level participants in this dimension, as with the evaluation approach dimension. In other words, no participant was identified who seeks for the underlying

reasons that lay at the foundation of the problematic parts of the responses and tries to understand logical reasoning behind the responses through their analysis.

Whereas in the specificity in the evaluation dimension within the domain-neutral dimension, 80.9% of the participants made general evaluations on students' attitudes or knowledge based on the student responses. For example, a student has responded to the question on how to obtain light in different colors with the statement, "*Other colors emerge through the reflection and absorption of white light. For example, inside the white light, there is the color red as well. For instance, if a red light is reflected on white light, it looks red because the color white emerges through the combination of other colors.*" The prospective teacher with the code 3-38 has evaluated this response with the statement, "*A correct and explanatory answer. We understand that he/she has mastered the subject since he/she explains it through examples.*" When this evaluation is analyzed, it is observed that the prospective teacher uses general statements and does not provide details while evaluating student understanding. However, 19.1% of the participants, while making general statements about knowledge and understanding in some student responses, they have made detailed explanations about whether or not concepts were understood while evaluating other responses. On the other hand, no participant has produced specific claims on student understanding based on student responses and made detailed explanations about whether or not relevant concepts were understood or not for all student responses.

In the analysis of coherence dimension, meaning the dimension that addresses the coherence of a student's ideas with his/her other ideas and other students' ideas, evaluations made by 60.3% of prospective teachers were determined to be at the novice level. This indicates that more than half of the prospective teachers determined and described mistakes and misconceptions isolated. However, 33.8 percent of them are at the emerging level, and they are shown to be conscious of consistency or inconsistency in their answers or the answers of other students. 2.9% of the prospective teachers used the coherence and incoherence they identified to evaluate students' understanding.

Lastly, when the use of evidence of prospective teachers was analyzed, it was observed that 67.6% of them were at the novice level. These prospective teachers did not provide clear evidence to support their claims for the inferences they made. In contrast, 32.4% of the participants were at the emerging level and sometimes used evidence to support the inferences they made about students' understanding. There were no prospective teachers who provided evidence for all instances. The statement made by the prospective

teacher with the code 4-12 can be given as an example to a novice-level evaluation: "*His/Her knowledge on separation and refraction of white light into different colors is inadequate.*" When the statement is examined, it can be observed that the inference is not grounded on any evidence. A response by the prospective teacher with the code 3-38 can be given as an example to an emerging-level evaluation. A student responded to the question, "How do the blue color of the sky and the seas emerge?" "How do the blue colors of the sky and the seas emerge?" with the statement "*The sky and the sea are transparent. However, we see them blue because light hits and is reflected.*" The prospective teacher evaluated this response with the following statement: "*The student overgeneralized. What he/she wrote about the sea is true, but he/she did not mention refraction for the sky. His/Her answer is partial. He/She also did not mention the absorption of other colors by the sea. There is missing information.*" When this evaluation is examined, it is observed that the prospective teacher used evidence (e.g., not talking about refraction about the sky, not mentioning absorption) for the inferences he/she has made.

Findings of the analysis of domain-dependent dimensions are given in Table 5.

Table 5. *Domain-Dependent Dimensions and Their Rates of Implementation in the Evaluations*

| Criteria | Novice (1) | | Emerging (2) | | Advanced (3) | | Mean |
|-------------------------------|------------|------|--------------|------|--------------|------|------|
| | f | % | f | % | f | % | |
| Quality of the Interpretation | 54 | 79.4 | 14 | 20.6 | 0 | 0 | 1.21 |
| Productive Thinking | 50 | 73.5 | 17 | 25.0 | 1 | 1.5 | 1.28 |
| Scope of the Evaluation | 52 | 76.5 | 16 | 23.5 | 0 | 0 | 1.24 |
| Scientific Accuracy | 9 | 13.2 | 48 | 70.6 | 11 | 16.2 | 2.03 |

Quality of the interpretation of the prospective teachers refers to the consistent and reasonable inferences based on genuinely accessible evidence about student understanding. As much as the quality of the interpretation is thought to be similar to the use of evidence dimension among the domain-neutral dimensions, the use of evidence qualifies to what extent the prospective teachers provide evidence to support their inferences on student

understanding. Whereas in the domain-dependent dimension, the quality of the interpretation qualifies the validity of inferences (Talanquer et al., 2015). With this in mind, 79.4% of prospective students were determined to be at the novice level, and 20.6% were determined to be at the emerging level. None of the prospective teachers were at the advanced level. When the use of evidence and quality of the interpretation is compared, it is observed that the evidence provided by some prospective teachers who have provided evidence on student understanding and the inferences they made were inaccurate. The reason for that is that there is a difference in the number of students among emerging levels. To give an example to the novice level, a response by the prospective teacher with the code 3-13 can be examined. A student responded to the question "How do the blue color of the sky and the seas emerge?" with the following statement: *"The sky and the sea are actually transparent. However, we see them blue because light hits and is reflected."* The prospective teacher with code 3-13 evaluated this response with the following statement: *"He/She lacks knowledge of separating light into colors and refraction."* When the statement of the prospective teacher is examined, it is observed that any evidence does not support the inferences he/she has made about students' understanding. When the evaluation made by the prospective teacher with the code 3-32 is examined, it is observed that he/she has made inferences (*"Her knowledge is inadequate. He/She cannot distinguish the phenomena that occur in the sea and the sky. He/She generalizes."*) and provided evidence. However, the evidence is limited and superficial (refraction and reflection concepts were not mentioned).

Whereas the productive thinking dimension refers to making an evaluation that focuses on the basic concepts of a subject, which can be effective by not focusing on details and productive to support learning in the analysis of student understanding of the targeted concepts. In addition, the prospective teacher must have an accurate scientific understanding concerning the basic concepts. When the analysis is conducted within this framework, 73.5% of the prospective teachers were determined to be a novice. This indicates that prospective teachers either do not consider required basic concepts for the target concept or do not have a relevant scientific understanding of the basic concepts or highlight unnecessary or unrelated concepts while evaluating student responses. To give an example, a prospective teacher with the code 3-17 evaluates the following response to the question "How is light obtained in different colors?" by Student-2: *"Lights in different colors emerge through the absorption of white light. These lights are red, orange, yellow, blue, and purple."* The prospective teacher's evaluation is as follows: *"Yes, correct. White light*

separates into different colors when passing through a prism and this student has mastered the subject because he/she has listed the colors according to an abbreviation made for easy remembering (In Turkish: KuTuSu YaMalıM, underlined letters corresponding to the initials of colors)." When this evaluation is analyzed, it is observed that the prospective teacher could not detect the basic concepts of absorption and refraction in the student response. Although the student's response is based on the concept of absorption, the prospective teacher evaluated the student's response based on the concept of refraction. Thus, the prospective teacher could not make a product evaluation and was at the novice level. On the other hand, 25% of the prospective teachers did consider basic concepts while evaluating the target concept; however, because they did not underscore the incorrect parts in student responses, they could not make an evaluation that was as productive as desired. The number of prospective teachers who assess students using basic concepts related to the comprehension of the target concept and making a productive assessment by highlighting the causes of mistakes is fairly low (1).

The scope of the evaluation focuses on the variety of concepts and ideas the prospective teacher considers while evaluating student understanding. 76.5% of the prospective teachers focus on analyzing concepts and ideas at a limited level while evaluating student understanding. In contrast, 23.5% of the prospective teachers evaluate student understanding by focusing on a potential lack of knowledge beyond the target through their analysis. For instance, a student responded to the question "*How is light obtained in different colors?*" with the following statement: "*Lights in different colors emerge through the absorption of white light. These lights are red, orange, yellow, blue, and purple.*" The prospective teacher with code 3-10 evaluated this response with the following statement: "*The student confused reflection with absorption here. She also listed the resulting colors incorrectly. So, she responded to the question incorrectly.*" It can be observed that the prospective teacher goes beyond commenting only on refraction and absorption, which are the basic concepts of the subject, and makes an evaluation about primary colors, as well. In other words, he/she evaluates misconceptions and lack of knowledge beyond the targeted degree. No prospective teachers were determined who prominently focus on relevant misconceptions beyond what is targeted through the evaluation.

Lastly, in the scientific accuracy dimension, which refers to examining prospective teachers' understanding levels of scientific concepts and the thoughts targeted to be taught,

13.2% of the prospective teachers were determined to be at the novice level. In other words, prospective teachers had a superficial and unreliable understanding that would prevent them from making appropriate logical inferences. At the same time, the ratio of prospective teachers who made some incorrect inferences and put forward conceptual problems; in other words, prospective teachers who were at the emerging level is 70.6%. On the other hand, the advanced level prospective teacher ratio is 16.2%. Inferences of these participants while evaluating students' responses were observed to be scientifically accurate. For instance, the prospective teacher with the code 3-44 evaluated response to the question about how the blue color of the sky and the seas emerge with the following statement "*He/She did not learn very well. The color of the sky emerges through the refraction of the light, not its reflection.*" No scientific mistakes are observed when the evaluation statement of the prospective teacher is examined.

When the data were considered a whole, emerging level ratios of both the use of evidence in domain-neutral dimension and quality of the interpretation in domain-dependent dimension and advanced level ratios of the scientific accuracy dimension show that the analyses conducted are internally coherent. The reason for this is that the use of evidence by a prospective teacher while evaluating student understanding and its correctness puts forward the reflection of the prospective teacher's understanding.

Findings obtained from the analyses on the educational regulations suggested by the prospective science teachers based on students' written exam responses within the scope of the second sub-problem are presented in Table 6. These findings signal how the prospective teachers plan to support student learning by eliminating misconceptions and conceptual gaps detected in students' written exam responses.

Table 6. *Educational Regulations Suggested by Prospective Teachers*

| Strategy | f | % | Strategy | f | % |
|---|----|------|---|----|------|
| Review the material | 48 | 70.6 | Experiment Demonstration | 17 | 25.0 |
| Experimenting | 46 | 67.6 | Giving homework | 16 | 23.5 |
| Using new materials and examples from everyday life | 43 | 63.2 | Redesigning the lecture according to parts where students make mistakes | 13 | 19.1 |

| | | | | | |
|-----------------------------------|----|------|--|----|------|
| Conducting hands-on lessons | 38 | 55.9 | Having students do a research project on the subject | 13 | 19.1 |
| Providing more practice questions | 34 | 50.0 | Spending more time on important concepts | 11 | 16.2 |
| Correcting misconceptions | 29 | 42.6 | Using individual activities | 5 | 7.4 |
| Using visual materials/diagrams | 24 | 35.3 | Using concept maps | 3 | 4.4 |
| Lecturing with different methods | 21 | 30.9 | Using videos about the subject | 1 | 1.5 |
| Having students make observations | 21 | 30.9 | | | |

When Table 6 is examined, it is observed that the prospective science teachers did planning using 17 different methods such as reviewing the materials, experimenting, using new materials and examples from daily life, conducting hands-on lessons, and providing more example questions in order to ensure student improvement based on student responses. Even though their percentage is small, some prospective teachers plan to use individual activities, concept maps, and videos to teach the subject. When the educational regulations preferred by the prospective teachers are examined, it is striking that they prefer pedagogical regulations that are more general and do not focus directly on learning deficiencies of the students, such as repeating the subjects, experimenting, and using new methods and materials. On the other hand, the percentage of pedagogical choices specific to the areas students have difficulties with, such as using conceptual maps, watching videos, and spending more time on misconceptions and significant concepts, is low. In other words, the majority of the participants are limited in making subject-specific and effective pedagogical regulations that focus on eliminating students' learning deficiencies and difficulties.

Discussion and Conclusion

The quality of education in a classroom depends upon the ability of a teacher to reveal and use the thoughts and knowledge of students on the subject while making educational decisions (Bennet, 2011; Furtak, 2012). This necessitates the teachers to be able to evaluate

student responses sufficiently. Concordantly, this study aimed at exploring what prospective science teachers looked for in student responses, whether or not they could distinguish weak and strong parts of these responses, how they evaluated these responses and what kinds of educational activities they were planning.

Results show that prospective teachers' ability to evaluate student responses is in the novice-emerging range in terms of domain-neutral and domain-dependent criteria, and their ability to evaluate student responses is not at the desired level. In other words, it is observed that prospective teachers regarded student responses superficially; they described the responses rather than making inferences from the responses, underlined the correctness and incorrectness of the responses rather than detecting areas where students experience problems in, could not provide specific claims about student understanding using student responses, did not provide detailed evidence supporting their claims while evaluating, and made evaluations about a student's conceptual learning without making connections between different responses of the student. Researchers underline the importance of obtaining rich data about students' existing conditions, thoroughly examining these data, and incorporating them in teaching for effecting teaching (Furtak, 2012; Harshman and Yeziarski, 2015; NRC, 2012; Talanquer et al., 2015). However, as the results of this study illustrate, prospective teachers generally make external evaluations and focus on whether the student responses are correct or not. Therefore, they are unable to make use of the data they obtained in a way that supports teaching. Fundamental reasons for this may be that the prospective teachers experience a lack of knowledge in the subject that they are making evaluations and carry out the evaluation process with a focus on grading rather than supporting and shaping teaching. In parallel, Kang and Anderson (2015) indicate that the reason for this situation may be rudimentary areas, learning-teaching and evaluation knowledge, and evaluation skills. The results of this study's domain-dependent dimension indicate that a large majority of the prospective teachers (70.6%) both produced some incorrect interpretations that depend on area knowledge while making evaluations, and their evaluation orientations did not focus on analyzing student responses in order to understand underlying logical reasoning. Similar results are observed in studies conducted by Aydeniz and Doğan (2016), Otera (2006), and Nathan and Otera (2008), and these studies show that prospective teachers approach student responses superficially and do not go beyond interpreting the responses as correct or incorrect.

Moreover, Aydeniz and Doğan (2016), in their study analyzing the pedagogical skills of prospective teachers on formative assessment, showed that prospective teachers did not focus on learning deficiencies and were inefficient in using evaluation while making educational decisions. Teaching staff should be role models in demonstrating the importance of formative assessment rather than grading in evaluating student responses during prospective teacher training. Their use of domain-dependent and domain-neutral criteria in evaluating student responses should be encouraged to ensure effective evaluation through various exercises.

On the other hand, this study has established that most participants did not use evidence while evaluating student responses, and evidence provided by some participants was scientifically inaccurate. This finding is supported by the literature, as well. For instance, Talanquer et al. (2015) similarly demonstrate that prospective teachers use scientifically inaccurate knowledge while using student responses, and inferences they made from the responses were far from supporting learning. The use of evidence to ensure high quality and inferential evaluation while evaluating student data is significant in evaluating teaching and supporting learning (NRC, 2012). While the evaluation of a teacher whose scientific knowledge of the evaluated concepts is inaccurate to endanger obtaining reliable results, it also makes supporting teaching by determining misconceptions and learning gaps in the formative assessment process more difficult (Coffey, Hammer, Levin, and Grant, 2011) Thus, it is of utmost importance for a prospective teacher who will be carrying an evaluator title to both possess area knowledge in their field and use evidence in the evaluation process. In order to ensure this, prospective teachers should be supported with a rich area of knowledge, along with pedagogical knowledge, during their training.

In parallel with scientific accuracy and the use of evidence, prospective teachers are observed to be at a novice-emerging level in the context of productive thinking. Although scientific accuracy and evidence are significant in productive thinking over student responses, they are not sufficient alone. In order for a product evaluation to be produced, prospective teachers must have deep area knowledge on the relevant concepts, must be able to distinguish concepts according to their significance and priority levels, and determine the connections between concepts (Harshman and Yeziarski, 2015; Talanquer et al., 2015). Only in this manner can they make productive inferences that will help support students with low competence by determining where students experience difficulties, which conceptual relations they cannot establish, and conceptual gaps, and make educational decisions

(Gottheiner and Siegel, 2012). Researchers, as illustrated in this study, have indicated that the area in which the most difficulties are experienced in the evaluation process is interpreting the data obtained correctly and productively and using it in a way that supports teaching (Ateh, 2015; Kang and Adnerson, 2015; Izci & Siegel, 2019). In order to fill this gap, enriching the repertoires of prospective teachers in detecting field-specific misconceptions and eliminating them in teaching programs is crucial.

On the other hand, it is observed that prospective teachers focus on concepts and thoughts with a narrow scope in the evaluation process. Most of the prospective teachers are at an advanced level in terms of scientific accuracy, which means that they sometimes make incorrect interpretations and have conceptual problems. These results reveal the significance of having practical training on how to evaluate and use student responses as part of prospective teachers' training on assessment and evaluation and that, through this training, they can develop their pedagogical reasoning skills.

Lastly, the results show that prospective teachers eliminate students' learning difficulties through different methods based on student responses (Table 6). However, a large majority of the participants, as stated by other researchers (Aydeniz and Doğan, 2016; Gottheiner and Siegel, 2012), plan on increasing student success through general methods insufficient indirectly eliminating deficiencies and difficulties experienced by students such as repeating the lecture, providing example problems, and experimenting. It is known that as long as difficulties experienced by students and misconceptions are not eliminated, the increase in success will be limited (Black and William, 2006). Therefore, prospective teachers should be supported in learning and using different teaching methods.

Ethical Approval: This research was conducted with the permission of the Van Yüzüncü Yıl University Presidency Social Sciences and Humanities ethics committee with the decision no 2020/03-16 and dated 13/05/2020.

Conflict Interest: The authors declare that they have no conflict of interest.

Authors Contributions: Authors' contribution is equal.

References

- Ateh, C. M. (2015). Science teachers' elicitation practices: Insights for formative assessment. *Educational Assessment*, 20(2), 112-131.
- Aydeniz, M., & Dogan, A. (2016). Exploring pre-service science teachers' pedagogical capacity for formative assessment through analyses of student answers. *Research in Science & Technological Education*, 34(2), 125–141.
- Barnhart, T., & van Es, E. (2015). Studying teacher noticing: Examining the relationship among preservice science teachers' ability to attend, analyze and respond to student thinking. *Teaching and Teacher Education*, 45, 83–93.
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18 (1), 5-25.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-74.
- Black, P., & Wiliam, D. (2006). Developing a theory of formative assessment. In J. Gardner (Ed.), *Assessment and learning* (pp. 81-100). London: Sage
- Chan, K. K. H., & Yau, K. W. (2021). Using video-based interviews to investigate pre-service secondary science teachers' situation-specific skills for informal formative assessment. *International Journal of Science and Mathematics Education*, 19, 289-311.
- Coffey, J. E., Hammer, D., Levin, D. M., & Grant, T. (2011). The missing disciplinary substance of formative assessment. *Journal of Research in Science Teaching*, 48(10), 1109–1136.
- Dini, V., Sevian, H., Caushi, K., & Orduña Picón, R. (2020). Characterizing the formative assessment enactment of experienced science teachers. *Science Education*, 104(2), 290–325.
- Furtak, E. M. (2012). Linking a learning progression for natural selection to teachers' enactment of formative assessment. *Journal of Research in Science Teaching*, 49(9), 1181-1210.

- Gottheiner, D. G., & Siegel, M. A. (2012). Experienced middle school science teachers' assessment literacy: Investigating knowledge of students' conceptions in genetics and ways to shape instruction. *Journal of Science Teacher Education*, 23, 531–557.
- Gotwals, A. W., & Birmingham, D. (2016). Eliciting, identifying, interpreting, and responding to students' ideas: Teacher candidates' growth in formative assessment practices. *Research in Science Education*, 46(1), 365–388.
- Harshman, J., & Yeziarski, E. (2015). Guiding teaching with assessments: High school chemistry teachers' use of data-driven inquiry. *Chemistry Education Research and Practice*, 16(1), 93-103.
- Izci, K. (2018). Turkish science teacher candidates understandings of equitable assessment and their plans about it. *Journal of Education in Science, Environment and Health*, 4(2), 193-205.
- Izci, K. & Siegel, M. (2019). Investigation of an alternatively certified new high school chemistry teacher's assessment literacy. *International Journal of Education in Mathematics, Science and Technology*, 7(1), 1-19. DOI: 10.18404/ijemst.473605
- Izci, K. & Şardağ, M. (2016). Prospective Science Teachers' Perceptions of Classroom Assessment. *Necatibey Faculty of Education Electronic Journal of Science & Mathematics Education*, 10(1), 439-471.
- Kang, H. & Anderson, C. W. (2015). Supporting preservice science teachers' ability to attend and respond to student thinking by design. *Science Education*, 99(5), 863–895.
- Lam, D. S. H. & Chan, K. K. H. (2020) Characterising preservice secondary science teachers' noticing of different forms of evidence of student thinking. *International Journal of Science Education*, 42(4), 576-597.
- Larkin, D. (2012). Misconceptions about misconceptions: Preservice Secondary science teachers' views on the value and role of student ideas. *Science Education* 96, 927–959.
- Levin, D., Hammer, D., & Coffey, J. (2009). Novice teachers' attention to student thinking. *Journal of Teacher Education* 60, 142–154.
- Luna, M. J., & Sherin, M. G. (2017). Using a video club design to promote teacher attention to students' ideas in science. *Teaching and Teacher Education*, 66, 282–294.

- Lyon, E. G. (2011). Beliefs, practices, and reflection: Exploring a science teacher's classroom assessment through the assessment triangle model. *Journal of Science Teacher Education*, 22(5), 417-435.
- Miles, M. B. & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. (2nd Edition). California: SAGE Publications.
- Murray, S. A.; Huie, R.; Lewis, R.; Balicki, S.; Clinchot, M.; Banks, G.; Talanquer, V.; Sevian, H. (2020). Teachers' noticing, interpreting, and acting on students' chemical ideas in written work. *Journal of Chemical Education*, 97(10), 3478–3489.
- National Research Council. (2012) *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington: The National Academies Press.
- Otero, V. (2006). Moving beyond the 'Get it or don't' conception of formative assessment. *Journal of Teacher Education*, 57(3), 247-255.
- Otero, V., & Nathan, M. J. (2008). Preservice elementary teachers' views of their students' prior knowledge of science. *Journal of Research in Science Teaching*, 45(4), 497-523.
- Ruiz-Primo, M. A. & Furtak E. M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching*, 44(1), 57-77.
- Sabel, J.L., Forbes, C.T., & Zangori, L. (2015). Promoting prospective elementary teachers' learning to use formative assessment for life science instruction. *Journal of Science Teacher Education*, 26(4), 419–445.
- Schack, E. O., Fisher, M. H., & Wilhelm, J. A. (2017). *Teacher noticing: Bridging and broadening perspectives, contexts and frameworks*. Springer.
- Siegel, M. A., & Wissehr, C. (2011). Preparing for the plunge: Preservice teachers' assessment literacy. *Journal of Science Teacher Education*, 22(4), 371-391.
- Son, J. (2013). How pre-service teachers interpret and respond to student errors: Ratio and proportion in similar rectangles. *Educational Studies in Mathematics* 84(1), 49–70.
- Talanquer, V., Bolger, M., & Tomanek, D. (2015). Exploring prospective teachers' Assessment practices: Noticing and interpreting student understanding in the

assessment of written work. *Journal of Research in Science Teaching*, 52(5), 585-609.

Vogelzang, J. & Admiraal, W. F. (2017). Classroom action research on formative assessment in a context-based chemistry course. *Educational Action Research*, 25(1),155-166.

Yin, R. (2009). *Case study research: Design and methods*. Thousand Oaks, CA: Sage.

Yıldırım, A., & Şimşek, H. (2016). *Sosyal bilimlerde nitel araştırma yöntemleri*. (10. Baskı). Ankara: Seçkin Yayıncılık.