

## Doküman İnceleme Yönteminde Geçerlik ve Güvenirlik: Eğitim Bilimleri Araştırmaları Bağlamında Kuramsal Bir İnceleme

### Validity and Reliability in Document Analysis Method: A Theoretical Review in the Context of Educational Science Research

Umut Birkan Özkan<sup>1</sup>

<sup>1</sup>Sorumlu Yazar, Doç. Dr., Milli Savunma Üniversitesi, uozkan@msu.edu.tr, (https://orcid.org/0000-0001-8978-3213)

**Geliş Tarihi:** 02.03.2023

**Kabul Tarihi:** 15.05.2023

#### ÖZ

Bu çalışma, doküman inceleme yönteminde geçerlik ve güvenilirlik kavramlarını eğitim bilimleri bağlamında incelemektedir. Kavramsal tarama modelinde desenlenmiş nitel bir araştırmadır. Kavramsal tarama, iki aşamada gerçekleştirilmiştir. İlk aşamada çalışmanın metodolojik bağlamı göz önünde bulundurularak gerçek problemin anlaşılması için mevcudiyetin keşfi Buca Eğitim Fakültesi Dergisi örnekleminde yapılmıştır. 2020-2022 yılları arasında yayımlanmış 161 makale betimsel analize tabi tutulmuştur. İkinci aşamada ise çalışmanın odağındaki kuramsal inceleme, birinci aşamada elde edilen bulguların ışığında ayrıntılı bir şekilde yapılmıştır. Bulgular incelendiğinde çalışmanın örneklemindeki doküman incelemesi kullanılan eğitim bilimleri makalelerinin %70'inde çalışmanın geçerliğine, %46'sında ise güvenilirliğine yönelik herhangi bir açıklamanın yapılmadığı görülmüştür. Geçerlik ve güvenilirlik konularının üzerinde tartışılması gerektiğini ortaya koyan bu bulgular doğrultusunda pozitivist, yorumlayıcı ve eleştirel paradigmların doküman incelemesine yaklaşımları çerçevesinde geçerlik ve güvenilirlik stratejileri önerilmiştir. Pozitivist paradigmda yerleşik olan geçerlik ve güvenilirlik kriterlerinin yanında yorumlayıcı ve eleştirel paradigma eksenindeki doküman incelemesi yöntemi kullanılan araştırmalar için geçerlik ve güvenilirlik stratejilerine ilk kez bu araştırmada yer verilmiştir. Bu çalışmada ortaya konulan geçerlik ve güvenilirlik stratejileri kural koyucu olarak alınmak yerine, bağlamı ve amacı göz önüne alındığında, bir eğitim bilimleri çalışmasının öğelerinin gözden geçirilmesine ve değerlendirilmesine rehberlik etmek için kullanılması önerilmektedir.

**Anahtar Kelimeler:** Doküman incelemesi, geçerlik, güvenilirlik, pozitivist paradigma, yorumlayıcı paradigma, eleştirel paradigma.

#### ABSTRACT

This study examines the concepts of validity and reliability in the context of educational sciences using the document analysis method. It is qualitative research designed in the conceptual survey model. First, the existence was explored to understand the real problem, considering the methodological context of the study, in the Journal of Buca Faculty of Education sample. 161 articles published between 2020 and 2022 were subjected to descriptive analysis. Second, the theoretical examination of the focus of the study was carried out in detail in light of the findings obtained. The study found that many educational science articles using document analysis did not explain validity and reliability. Thus, the study proposed validity and reliability strategies for research using the document analysis method within the approaches of positivist, interpretive, and critical paradigms. These proposed strategies should be used as a guide to review and evaluate the elements of an educational science study in its context and purpose, rather than as prescriptive rules. The validity and reliability strategies proposed in this study should be used not as prescriptive rules, but rather

to guide the review and evaluation of the elements of an educational science study, considering its context and purpose.

**Keywords:** Document analysis, validity, reliability, positivist paradigm, interpretivist paradigm, critical paradigm.

## GİRİŞ

Eğitim bilimleri alanındaki pek çok araştırmada dokümanlardan çeşitli şekillerde yararlanılmakta ve araştırma süreci dokümanlardan elde edilen veriler üzerine inşa edilebilmektedir. Dokümanların eğitim bilimlerinde kullanım şekillerinden biri diğer bilim alanlarında olduğu gibi literatür taraması amaçlı kullanımıdır. Literatür taraması, genel olarak, geçmiş araştırmaları toplamanın ve sentezlemenin sistematik bir yolu olarak tanımlanabilir (Snyder, 2019). Bir diğer kullanım şekli ise bu kuramsal incelemenin kapsamına giren bilimsel bir araştırma yöntemi olarak doküman incelemesidir. Doküman incelemesi yöntemine ilişkin yabancı alanyazında çeşitli yayınlar bulunmaktadır (Asdal & Reinertsen, 2022; McCulloch, 2004; Rapley, 2008; Silverman, 2014). Genel olarak bu yayınlarda, tüm çeşitlilikleriyle dokümanların zengin ve sistematik bir incelemesini yapabilmek için bir çerçeve sunulmaktadır. Doküman incelemesine, bilimsel araştırma yöntemlerini konu alan Türkçe yayımlanmış bazı kitapların belirli bölümlerinde değinilmektedir (Balaban Salı, 2018; Baş & Akturan, 2017; Yıldırım & Şimşek, 2018). Bununla birlikte, doküman incelemesini münferit olarak ve derinlemesine inceleyen az sayıda Türkçe çalışma vardır. Bu çalışmalardan biri olan Eğitim Bilimleri Araştırmaları İçin Doküman İnceleme Yöntemi (Özkan, 2021) adlı kitabın ilk baskısı 2019 yılında yayımlanmıştır. Özkan (2021), eğitim bilimleri alanında yapılacak çalışmaların yöntemsel boyutuna katkıda bulunmayı amaçlamış ve nitel araştırma yöntemlerinden biri olan doküman inceleme yöntemini, “araştırmanın veri setini oluşturan birincil veya ikincil kaynak olarak nitelendirilen çeşitli dokümanların elde edilmesi, gözden geçirilmesi, sorgulanması ve analizi” olarak tanımlamıştır (Özkan, 2021, ss. 2). Doküman incelemeyi bir araştırma yöntemi olarak gören başka yeni çalışmalar da bulunmaktadır. Fakat bağlam olarak eğitim bilimleriyle sınırlı değildirler. Örneğin, Kırıl'ın (2020) çalışması, doküman inceleme yöntemine dair daha genel bir incelemedir. Kırıl (2020, 185) çalışmasında doküman inceleme yöntemini “araştırılan konu ile ilgili dokümanların bilimsel esaslara uygun olarak incelenmesi” biçiminde ifade etmektedir. Sak ve diğerleri (2021), derleme bir çalışma üreterek doküman analizi ile ilgili çeşitli tartışmalara değinmiş ve doküman analizinde dikkat edilmesi gereken konuların altını çizmekle yetinmişlerdir. Sak ve diğerlerine (2021) göre eğitim bilimleri araştırmalarında araştırmacılar, kullandıkları belgelerin güvenilir ve geçerli kaynaklar olmasını sağlamalı ve araştırma konusuyla ilgili birincil kaynaklara ulaşmaya çalışmalıdırlar.

Yerli ve yabancı alan yazında doküman incelemesi ile ilgili yapılan yayınların da üzerinde durduğu gibi bilimsel bir araştırma yöntemi olarak doküman incelemesinin kullanımı giderek daha karmaşık bir sürece dönüşmektedir. Eğitim bilimleri alanında üretilmiş olan ve üretilmeye devam edilen dokümanların, bütünden ziyade parçalara ayrılmış veriler barındırması, birden fazla bilim alanıyla ilgili olması\* ve muazzam bir hızla yayılması karmaşıklığın temel nedenleri arasında sayılabilir. Ayrıca, teknolojik gelişmeler eğitim bilimleri alanındaki çok sayıda dokümana erişimi kolaylaştırmakta ancak bu durum çok sayıda dokümanı incelemeyi ve değerlendirmeyi zorlaştırmaktadır. Bu güçlüklerden biri eğitim bilimi araştırmacılarının araştırma konusuyla ilgili dokümanların tümüne hakim olma idealini gerçekleştirememesi, diğeri ise her bir dokümanın aynı zamanda çok dikkatli bir değerlendirme de gerektirmesidir (Tosh, 2002). Dolayısıyla, yirmi birinci yüzyılda doküman incelemesi yapmanın zorlukları ve

---

\* Benzer bir görüşe kitabında yer veren McCulloch (2004, 4), eğitim biliminin “sosyoloji, sosyal politika, antropoloji, siyaset, psikoloji ve hukuk gibi bir dizi sosyal bilim disiplini ve geleneği ile yakın bağlantılara sahip” olduğunu belirtmektedir.

fırsatları ancak bu yeni gelişmelerin önemi kabul edilerek tartışılabilir (McCulloch, 2004). Bununla birlikte, doküman inceleme yönteminde geçerli olan temel kurallar, yirmi birinci yüzyılın teknoloji ağırlıklı yeni gelişmelerinin etkisinin görüldüğü doküman incelemesine dayalı eğitim bilimleri araştırmaları için yeniden ele alınmalıdır. Bu bağlamda, üzerine tartışılması gereken temel kavramlardan ikisi geçerlik ve güvenilirliktir.

Bilimsel bir çalışmada araştırma tasarımına, verilerin elde edilme biçimine ve verilerin analizine ilişkin geçerlik ve güvenilirliğin raporlanması genel olarak istenen bir unsurdur. Çünkü geçerlik ve güvenilirlik, pozitivist/postpozitivist<sup>†</sup> gelenekte araştırma kalitesinin temel bir göstergesi olarak merkezlenmektedir (Miller, 2008a). Hammersley (1987) eski tarihli bir makalesinde, geçerlik ve güvenilirliğe yapılan vurgunun nitel araştırmalardan ziyade nicel araştırmalarda daha sık görüldüğünü ancak nitel araştırmacılar tarafından giderek daha fazla kabul edildiğini, temel konuların her ikisi için de geçerli olduğunu ifade etmektedir. Hammersley'nin çalışmasının üzerinden uzun yıllar geçmiş olsa da altını çizdiği konunun geçerliğini koruduğu söylenebilir. Geçerliğin yaygın tanımı, belirli bir aracın ölçmeyi istediği şeyi gerçekten ölçebildiğinin kanıtlanması (Kabakçı Yurdakul, 2013; Miller, 2008a) olduğu görüşüne dayansa da nitel araştırma geleneğinin bilgi gelişimine sunduğu katkıyı güçlendirmek için çok sayıda dönüşüm geçirmiş ve birçok farklı biçimde ele alınmıştır (Whittemore vd., 2001). Cohen ve diğerlerine (2005) göre nicel araştırmalarda geçerlik dikkatli örnekleme, uygun ölçme araçları ve verilerin uygun istatistiksel analizi yoluyla geliştirilebilirken, nitel araştırmada geçerlik, elde edilen verilerin doğruluğu, derinliği, zenginliği, kapsamı, araştırmanın katılımcıları, çeşitlemenin kapsamı ve araştırmacının tarafsızlığı üzerinden ele alınabilmektedir. Esasen, nitel araştırmalarda geçerlik, bulguların doğruluğu ile ilgili olup (Kabakçı Yurdakul, 2013; Noble & Smith, 2015), araçların, süreçlerin ve verilerin uygunluğu anlamına gelmektedir (Leung, 2015). Nitel araştırmalarda geçerlik kavramına farklı bir pencereden bakan Baltacı (2019), araştırmacının sorunu mümkün olan en tarafsız şekilde çözme derecesi olarak geçerliği tanımlamakta ve elde edilen verilerin gerçek durumu ne ölçüde yansıttığının önemli olduğunu altını çizmektedir.

Geçerlik kavramıyla ayrılmaz biçimde bağlantılı olan bir kavram güvenilirliktir (Miller, 2008a). Bağlantının temelleri ise güvenilirliğin geçerlik hakkında kanıtlar sağlamasına ve kullanılan araştırma stratejilerinin yararlılığı hakkında bilgi vermesine dayanmaktadır (Hammersley, 2006). Güvenirlik, zaman içinde araçlar ve katılımcılar üzerinde tekrarlanabilirlik ile eş anlamlı kullanılmaya başlanmıştır (Baltacı, 2019; Cohen vd., 2005). Aslında güvenilirlik kavramına yönelik buna benzer tanımlamalar nicel bir araştırmaya daha uygundur. Çünkü nitel araştırma alanının paradigma ve metodolojik çeşitliliği nedeniyle güvenilirlik bu kadar tekdüzelikle tanımlanamamaktadır (Miller, 2008b). Bununla birlikte, nicel araştırmalarda güvenilirlik -geçerlik kavramında olduğu gibi- özelliklerin ölçülebileceği tekniklerle ilgilidir. Farklı paradigmalara sahip nitel araştırmalarda ise güvenilirliğin ölçme tekniklerine ve/veya tekrarlanabilirliğe indirgenerek tanımlanması zorlayıcıdır ve epistemolojik olarak sezgilere aykırıdır (Leung, 2015). Nitel araştırmalarda olgular kişiye ve bağlama göre sürekli değişebileceğinden aynı kitle üzerinde tekrarlanan çalışmalar aynı sonuçları vermeyebilir ve bunun yanında farklı araştırmacılar aynı verileri farklı algılayıp yorumlayabilir (Kabakçı Yurdakul, 2013). Dolayısıyla, doküman inceleme yönteminin içinde yer aldığı nitel araştırma alanında güvenilirlik için farklı bir tanımlama yapma ihtiyacı doğmaktadır. Basit bir tanımlamayla nitel araştırmalarda güvenilirlik, sonuçların zaman içindeki tutarlılığını değerlendirir (Hayashi Jr vd., 2019). Daha kapsamlı bir tanımlamayla güvenilirlik, bulguları etkilemiş olabilecek kişisel ve araştırma yöntemi kaynaklı önyargıların hesaba katılması da dahil olmak

---

<sup>†</sup> Doğanay (2018, 19) iki gelenek arasındaki ayrımı şu şekilde yapmaktadır: “Pozitivist paradigma; nesnellik, ölçülebilirlik, yordanabilirlik, kontrol, örüntü, yasa oluşturma ve davranışın kurallarını belirleme gibi ilkeleri gerçekleştirme hedefler. Pozitivist ötesi paradigma dünyayı, üzerinde yaşayanların bakış açılarına göre anlamayı ve yorumlamayı hedeflemektedir. Birincisinde gözlenen olay, ikincisinde de anlam ve yorum ön plandadır.”

üzere analitik prosedürlerin tutarlığı biçiminde ifade edilebilir (Noble & Smith, 2015). Benzer bir görüşü paylaşan Hammersley'e (2006) göre güvenilirlik, örneklerin farklı gözlemciler tarafından veya aynı gözlemci tarafından farklı durumlarda aynı kategoriye atanmasındaki tutarlık derecesini ifade eder. Yapılan tanımlamalarda açıkça görüldüğü gibi nitel araştırmalarda güvenilirliğin sağlanmasının altında yatan endişe, araştırma sürecinin verilerin doğası üzerinde bilmeden bir etki yaratan aşırı değişikliklerden bağımsız olarak tutarlı bir şekilde yürütülmesini sağlamaktır (Long & Johnson, 2000). Dolayısıyla, tutarlık, nitel araştırmalarda güvenilirliğin özüdür.

Nitel araştırmalardaki güvenilirlik ve geçerlik tartışmalarına bakıldığında, net bir tanımlar dizisinden ziyade kafa karıştırıcı bir fikir çeşitliliği olduğu görülmektedir. Farklı yazarların tanımları önemli ölçüde örtüşse de bazı farklılıkların olduğu göze çarpmaktadır. Çünkü, nicel bir araştırmanın geçerliğini ve güvenilirliğini belirlemenin çeşitli yolları bulunmaktadır (Cohen vd., 2005; Heale & Twycross, 2015). Ancak nitel araştırma yöntemlerinde geçerlik ve güvenilirliğin nasıl belirleneceğine ilişkin bir yol haritası ortaya koymak çok da mümkün görünmemektedir (Shenton, 2004; Silverman, 2014). Bunun en önemli sebeplerinden biri nitel araştırmaların güvenilirliğinin genellikle pozitivistler tarafından sorgulanması, geçerlik ve güvenilirlik kavramlarının diğer araştırma paradigmalarına dayalı çalışmalarda aynı şekilde ele alınmamasıdır (Shenton, 2004). Cohen ve diğerleri (2005), nicel ve nitel araştırmalarda geçerlik ve güvenilirliğin ele alınış şeklinin farklılık gösterse de, bu terimlerin her ikisinin de bu iki araştırma türüne uygulanabileceğini öne sürmektedir. Bu tartışmaya farklı bir pencereden bakan Endacott (2005) ise nicel araştırmalarda geçerlik ve güvenilirliği sağlamak için uygulanan kuralların genel olarak nitel araştırmalara uygun olmadığını ileri sürmekte ve başka araştırmacılar (Shenton, 2004; Yıldırım & Şimşek, 2018) gibi pozitivist araştırmacılar tarafından kullanılan kriterlere karşılık gelen Guba'nın (1981) ortaya koyduğu kriterlerin kullanılmasını önermektedir. Guba (1981) tarafından ortaya konulan kriterler nitel araştırmalarda yaygın olarak kullanılsa da doküman inceleme yönteminin kullanıldığı eğitim bilimleri araştırmaları özelinde geçerlik ve güvenilirliğin nasıl sağlanabileceğine yönelik bir fikir birliğinin olmamasının bu alandaki önemli bir kısıtlılık olduğu değerlendirilebilir<sup>‡</sup>. Bu kapsamda alanyazında yer alan sınırlı sayıda çalışmalardan (Kıral, 2020; Özkan, 2021; Sak vd., 2021; Yıldırım & Şimşek, 2018) farklı olarak, nitel bir araştırma yöntemi olan doküman incelemesinde tartışmalı bir konu olan geçerlik ve güvenilirliğin sağlanmasına yönelik çabaları değerlendirmek ve bu görüşlerden yola çıkarak bir senteze varmak ihtiyacı bu çalışmanın gerekçesini oluşturmaktadır. Bu gerekçe ile çalışmanın amacı doküman inceleme yönteminde geçerlik ve güvenilirlik ölçütlerine ilişkin tartışmaları eğitim bilimleri bağlamında değerlendirmektir. Bu kapsamda, “doküman inceleme yöntemi kullanılan bir araştırmada geçerlik ve güvenilirlik nasıl sağlanmalıdır?” sorunsalı üzerinden eğitim bilimleri alanında doküman incelemesine dayalı bir araştırmada geçerlik ve güvenilirlik stratejilerinin neler olması gerektiği tartışılacaktır.

Makalenin çeşitli katkıları vardır. İlk olarak, bu makale doküman inceleme yöntemi özelinde kullanılan geçerlik ve güvenilirlik yaklaşımları ile Guba (1981) kriterleri gibi nitel araştırma yöntemlerinde genel kabul gören geçerlik ve güvenilirlik stratejileri arasında ayrım yapmaktadır. Nitel araştırmalarda kullanılan ve genel kabul gören geçerlik ve güvenilirlik stratejileri belirli bir fikir vermede etkilidirler ve ilgi duyanları tatmin edecek kanıtları sağlayabilirler. Ancak doküman inceleme yöntemini kullanan eğitim bilimleri araştırmaları için her zaman en iyi strateji olmayabilirler. Bunun yerine, doküman inceleme yöntemi özelinde bir geçerlik ve güvenilirlik yaklaşımı, bu alandaki boşlukları doldurmak için yararlı bir girişim olabilir. İkinci olarak, mevcut makale eğitim bilimleri araştırmalarında doküman incelemesi yöntemi kullanılırken karşılaşılabilecek pratik sorunlardan biri olan geçerlik ve güvenilirliği ele almaktadır.

<sup>‡</sup> Guba'nın (1981) bu kriteri ilk kez önermesinden bu yana 40 yılı aşkın bir süre geçmesine rağmen, hala önemli tartışmalara konu olmaktadır. Örneğin; Erlandson, Harris, Skipper ve Allen (1993, 30-35), Gomm, Hammersley ve Foster (2000, 100-101) gibi araştırmacılar Guba'nın kriterlerini gerçekleştiriminin uygulamada imkansız olduğunu savunmaktadır.

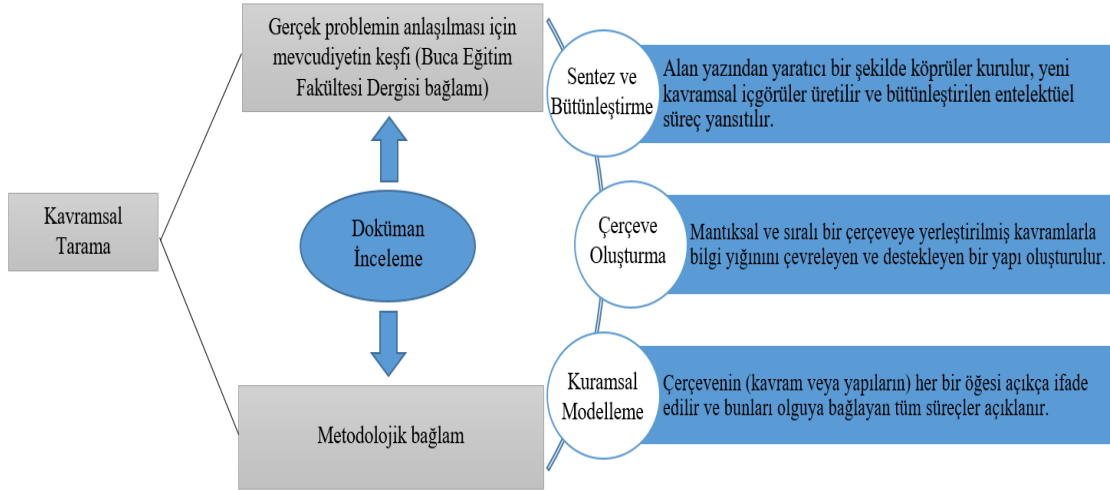
Doküman incelemesi yönteminde geçerlik ve güvenilirliğin ne olduğu, nasıl kullanılabilceği, geçerlik ve güvenilirliğin sağlanması için hangi kriterlerin göz önüne alınabileceği açıklanarak, eğitim bilimleri araştırma metodolojisine teorik ve pratik katkılarda bulunmaya çalışılmıştır. Son olarak, bu makale doküman inceleme yöntemi kullanılan makaleleri değerlendirirken geçerlik ve güvenilirliğin nasıl değerlendirileceğine dair hakemlere ve editörlere bazı yönergeler sağlayabilir. Bununla birlikte, bu çalışmanın, yazarlara, hakemlere ve editörlere olduğu kadar doküman inceleme yöntemi kullanılan makalelerin okuyucularına da yardımcı olması umulmaktadır.

## YÖNTEM

Bu çalışma, kavramsal tarama modelinde desenlenmiş nitel bir çalışmadır. Kavram taraması, farklı bilim ve disiplinlerdeki araştırma alanları çerçevesinde bir kavramın farklı yönlerini inceler (Baltacı, 2017). Bu bağlamda, kavramsal tarama modelinde desenlenen bu çalışmayla eğitim bilimleri araştırmacılarına doküman inceleme yönteminde geçerlik ve güvenilirlik konusunda bir yol haritası sunulmaya, alan yazında bu yönde bulunan eksiklik giderilmeye çalışılmıştır. Bununla birlikte, çalışmada temel alınan kuramsal incelemenin gerekliliğini ortaya koymak için Dokuz Eylül Üniversitesi Buca Eğitim Fakültesi Dergisinde 2020, 2021 ve 2022 yıllarında yayımlanan makalelerde betimsel analiz yapılmıştır. Betimsel analiz sonuçları bulgular bölümünde tablolaştırılarak mevcut durum açık bir şekilde ortaya koyulmuştur. Çalışmanın odağındaki kuramsal inceleme, bulguların ışığında Tartışma ve Sonuç bölümünde ayrıntılı bir şekilde yapılmış ve önerilerde bulunulmuştur. Bu çalışmada kullanılan araştırma yöntemi, araştırmanın doğası gereği doküman incelemesidir. Araştırmanın tasarımı Şekil 1’de belirtilmiştir.

### Şekil 1

#### Araştırmanın Tasarımı

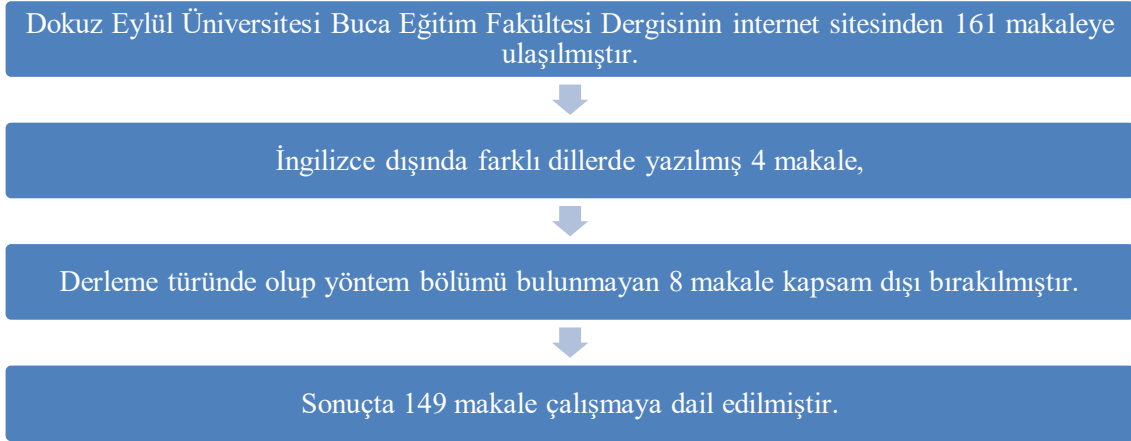


Örnekleme alınan dergilerin 2020, 2021 ve 2022 yıllarında yayımlanmasına yönelik sınırlamanın arka planındaki mantıksal gerekçe COVID-19 pandemisiyle ilişkilidir. 2019 yılı sonunda küresel bir salgına dönüşen COVID-19’un neden olduğu kapanmalardan ve izolasyon tedbirlerinden dolayı ölçek ve/veya anketler yoluyla nicel verilerin, görüşme ve/veya gözlem yoluyla nitel verilerin toplanmasında sıkıntılar yaşanmıştı. 2020, 2021 ve 2022 yıllarını kapsayan bu süreçte, araştırmacıların doküman inceleme yöntemiyle araştırma yapmaya ağırlık

vermiş olabilecekleri düşünülerek bu yıllar ile sınırlanmıştır. Bu kapsamda, 161 adet makalenin açık erişimine derginin <https://dergipark.org.tr/tr/pub/deubefd/archive> adresli internet sitesinden erişilmiş ve betimsel analiz yapılmıştır. Bu makaleler içinde doküman inceleme yöntemi kullanılan makaleler belirlenmiştir. Makalelerin belirlenmesine ilişkin süreç adımları Şekil 2’de sunulmuştur.

## Şekil 2

### Makale Belirleme Süreci



Şekil 2’de görüldüğü gibi 161 makaleden İngilizce dışında farklı dillerde yazılmış olan 4 makale ve derleme türünde olup yöntem bölümüne yer verilmeyen 8 makale kapsam dışı bırakılmış ve 149 makale araştırmaya dahil edilerek belirlenen kriterlere göre analiz edilmiştir. Makalelerin yöntem bölümü dikkatli bir biçimde okunmuş ve elde edilen veriler makale künyesine göre not edilmiştir. Makalelerin hangi araştırma yöntemiyle yapıldığı, kullanılan geçerlik ve güvenilirlik stratejileri bilgisayar ortamına kaydedilmiştir. Makaleler, araştırmada kullanılan araştırma yöntemine göre tekrar tekrar incelenmiş, geçerlik ve güvenilirliğe yönelik açıklamaların olup olmadığı ve varsa neler olduğu kısaca not edilmiştir. Makalelerin tekrar tekrar incelenmesinin sebebi, bazı çalışmalarda doküman inceleme yönteminin kullanıldığı belirtilmesine rağmen yöntem bölümünde farklı ifadeler kullanılmasıdır. Örneğin incelenen bir çalışmada, uluslararası alanda kabul gören alan indekslerinde yayımlanan makalelerin incelendiği dolayısıyla doküman inceleme yönteminin kullanılmasına rağmen çalışmanın yöntem bölümünde araştırma yönteminin içerik analizi yöntemi olduğu belirtilmiştir. Dolayısıyla, herhangi bir gözden kaçırmanın önüne geçmek için makalelerin araştırma yöntemleri tekrar tekrar incelenmiştir.

Araştırmanın geçerlik ve güvenilirliğini artırmak amacıyla nitel araştırma geleneğine ve doküman inceleme yöntemine uygun çeşitli çalışmalar yapılmıştır. Öncelikle, 161 makale dergi yönetim sisteminden indirilerek dijital ortamda depolanmıştır. Betimsel analizlerin sonuçları Microsoft tarafından Microsoft Windows işletim sistemleri tabanında çalışmak üzere yazılan bir tablolama programı olan Microsoft Excel ile tablolaştırılmıştır. Depolanan makaleler ve betimsel analiz sonuçlarının olduğu tablolar eğitim bilimleri alanında çalışmalar yapan iki farklı araştırmacıya gönderilerek kontrol ettirilmiştir. Yapılan kontroller sonucunda üç araştırmacının tüm sonuçlar üzerinde görüşbirliği içerisinde olduğu görülmüştür. Betimsel analiz tabloları ve makalelerin dijital nüshaları saklanmıştır. Çalışmanın geçerliğine yönelik olarak, çalışmanın yönteminin tüm basamakları detaylı bir biçimde sunulmuştur. Ayrıntılı örneklem, veri toplama ve analiz süreçlerinin sunulması çalışmanın geçerliğine katkıda bulunmuştur. Aynı zamanda, araştırma süreci aşamalarının detaylı betimlenmesi, araştırmanın tekrar edilebilmesinin yanında farklı bağlamlarda gerçekleştirilmesini de sağlayabilir. Araştırmada bağımsız iki araştırmacının

betimsel analiz süreçlerinde yer alması, çalışmanın dış denetçiler tarafından izlenmesini ve kontrol edilmesini sağladığından çalışmanın geçerliğinin artmasına önemli derecede katkıda bulunmuştur.

Bu makale etik kurul izni gerektiren çalışmalar (anket ya da ölçek uygulamayı gerektiren, görüşme ve gözlem içeren; doküman, resim, anket vb. diğerleri tarafından geliştirilen ve kullanım izni gerektiren çalışmalar) kapsamında girmediğinden etik kurul izni alınmamıştır.

## BULGULAR

Bu bölümde, araştırma tasarımında (Şekil 1) belirtilen mevcudiyetin keşfi için Buca Eğitim Fakültesi Dergisinde doküman inceleme yöntemi kullanılan eğitim bilimleri araştırmaları üzerinde yapılan betimsel analizlerden elde edilen bulgulara yer verilmiştir. Buca Eğitim Fakültesi Dergisinde 2020, 2021 ve 2022 yıllarında yayımlanan makalelerin toplam sayısı, doküman inceleme yöntemi kullanılan makalelerin sayısı ve doküman inceleme yöntemi kullanılan makalelerde geçerlik ve güvenilirliğe yönelik açıklamalar yapılan makalelerin sayısı Tablo 1’de sunulmuştur.

**Tablo 1**

*Betimsel Analiz Sonuçları*

Yıl	Dergi Numarası	Toplam Makale Sayısı	Doküman İnceleme Yöntemi Kullanılan Makale Sayısı	Çalışmanın Geçerliğine Yönelik Açıklama Yapılan Makale Sayısı	Çalışmanın Güvenirliğine Yönelik Açıklama Yapılan Makale Sayısı
2022	54	42	13	4	8
	53	37	10	2	6
2021	52	39	6	2	3
	51	25	3	2	2
2020	50	13	5	1	1
	49	5	-	-	-
<b>Toplam</b>		161	37	11	20

Tablo 1’de verilen değerler incelendiğinde doküman incelemesi kullanılan eğitim bilimleri makalelerinin %70’inde çalışmanın geçerliğine yönelik herhangi bir açıklamanın yapılmadığı görülmektedir. Bununla birlikte, doküman inceleme yöntemi kullanılan 37 eğitim bilimleri makalesinin yaklaşık yarısında (%46) güvenilirlikten hiç bahsedilmemiştir. Tablo 1’de sunulan betimsel analiz sonuçlarına göre mevcut durumda, doküman inceleme yönteminin kullanıldığı eğitim bilimleri araştırmalarının büyük bir bölümünde geçerlik ve güvenilirlik konularından bahsedilmediği söylenebilir. Dolayısıyla mevcut durum, bu konuların üzerinde tartışılması gerektiğini ortaya koymaktadır. Çalışmanın bundan sonraki bölümünde doküman inceleme yöntemi kullanılan eğitim bilimleri araştırmalarında geçerlik ve güvenilirlik konuları metodolojik bağlamda tartışılarak önerilerde bulunulmaya çalışılacaktır.

## TARTIŞMA, SONUÇ VE ÖNERİLER

Doküman inceleme yönteminde geçerlik ve güvenilirlik tartışmalarını anlamak, yorumlamak ve kuramsal bir temele oturtmak için doküman inceleme yönteminin teorik çerçevesinin dayandığı araştırma paradigmlarına bakmak gerekir. Farklı araştırma paradigmları incelendiğinde, geçerlik ve güvenilirlik yaklaşımlarındaki farklılıklar ve ortaklıklar daha açık bir şekilde görülebilir. “Araştırma metodu, teorik yaklaşım, metodolojik yaklaşım kavramları paradigmlar tarafından belirlendiğinden” (Taş, 2018, 62), hangi anlam yüklenirse

yüklensin tek tip bir geçerlik veya güvenilirlik kavramından söz edilemez. Bu kavramlar, araştırma paradigmalarının temel varsayımlarıyla ilişkili olan metodolojik kabullere dayanmaktadır ve araştırma paradigmasından ayrı adımlar değildirler. Pozitivist epistemoloji, yorumlayıcı gelenek ve eleştirel söylem analizi, doküman inceleme yönteminde teorik çerçevenin yerleştirilebileceği üç geniş araştırma paradigmasıdır (Jupp & Norris, 1993; Neuman, 2014). Pozitivist, yorumlayıcı ve eleştirel yaklaşımları birbirini dışlayan yaklaşımlar olarak görmekten sakınmak gerekse de aralarındaki farklılıkların vurgulanması geçerlik ve güvenilirlik kavramlarının kuramsal incelemesi için önemlidir.

Pozitivist paradigma, farklılıklardan ziyade benzerliklere odaklanarak evrensel olanı yakalamaya çalışır (Taş, 2018). Pozitivist yaklaşım, çalışmanın nesnel, rasyonel, sistematik ve nicel doğasını vurgular (McCulloch, 2004). Bilimsel ifadelerin ve sonuçların doğrulanabilir olması pozitivism için önemlidir (Kaya & Büyükekemci, 2022). Pozitivist paradigmaya göre eğitim bilimleri de dahil olmak üzere tüm bilimlerin amacı, deneysel gözlem yoluyla ispat edilebilen ilkelere, yasalara ve genellemelere ulaşmaktır (Şişman, 1998). Neuman (2014), pozitivist paradigmaya göre bilimsel bir araştırmanın nihai amacının, insan davranışının evrensel nedensel yasalarını keşfetmek ve belgelemek için bilimsel açıklama elde etmek olduğunu belirtmektedir. Pozitivist yaklaşımın ifade edilen nitelikleri düşünüldüğünde toplumun bir meta, nesne olduğu, deneysel bir inceleme neticesinde sosyal olgunun incelenebileceği ve sorunsalların bu yöntemlerle düzelebileceği sonucu çıkarılabilir. Pozitivist yaklaşımın doğası gereği, toplum bir meta ve bir nesne olup ampirik analizler sonucunda toplumsal olgular incelenebilir ve sorunlar bu yöntemlerle düzelebilir (Akyıldırım, 2020). Pozitivist paradigmada oluşturulan bilginin kaynağı, var olan gerçekliğin bağımsız gözlemleri ve ölçümleridir (Çıvık & Sezerel, 2018).

Pozitivist anlayış neden sorusuna odaklanırken, yorumlayıcı araştırma nasıl sorusuna odaklanır (Çıvık & Sezerel, 2018). Dolayısıyla yorumlayıcı araştırma nesnel olmaktan çok öznedir (Thanh & Thanh, 2015). Yorumlayıcı araştırmacılar için sosyal araştırmanın amacı, sosyal hayata dair bir anlayış geliştirmek ve insanların doğal ortamlarda anlamı nasıl inşa ettiğini keşfetmektir (Neuman, 2014). Pozitivist yaklaşım yerine anlama, yorumlama ve açıklama yöntemlerini kullanan bir yaklaşımla toplumsal yapının doğasını anlamaya çalışmanın daha tutarlı bir adım olarak görülmesi yorumlayıcı anlayışın temelini oluşturmuştur (Akyıldırım, 2020). Yorumlayıcı anlayışa göre toplumsal olgular nesnel, dışsal ve önceden belirlenmiş değil, bireyler tarafından toplumsal olarak inşa edilmiştir (Jupp & Norris, 1993). Namlı Altıntaş ve Erdem (2021, 40), yorumlayıcı anlayışa çağdaş bir bakış açısı getiren Schleiermacher'ın<sup>§</sup> çalışmalarında yorumlayıcı anlayışı, “tüm metinlere uygulanabilecek bir anlama ve açıklama yöntemi olarak tasarladığını” belirtmektedir. Bu bağlamda, insan davranışlarını, beşeri yapıtları ve dokümanları yorumlamayı temele alan yorumlayıcı anlayış, sözlü, yazılı veya görsel dokümanların barındırdığı “doğru ve gerçek anlamın ortaya çıkarılması amacıyla metnin ayrıntılı incelenmesinin önemi üzerinde durur” (Bayhan, 2016, 223).

Eleştirel yaklaşım, pozitivist paradigmaya bir tepki olarak ortaya çıkmıştır. Eleştirel yaklaşımın amacı, mevcut durumu aşmak ve alternatif gerçekliklerin mümkün olduğunu göstermektir (Demirtaş & Özer, 2015). Eleştirel kurama göre araştırmanın amacı, belirli bir olay veya olguyu derinlemesine incelemenin, sorgulamanın, betimlemenin ve anlamının yanında anlayışa göre değişim ve farklılıkları göstermektir (Kartal, 2021). Dolayısıyla eleştirel yaklaşımda araştırmanın birincil amacı sadece sosyal dünyayı incelemek değil, onu değiştirmek olup araştırmacılar sosyal kontrol, güç ilişkileri ve eşitsizliğin altında yatan kaynakları ortaya çıkararak, sosyal ilişkileri eleştirmek ve dönüştürmek için çalışmalar yürütürler (Neuman, 2014). Bu sayede eleştirel anlayış, meydana gelen olayların altında yatan fikirlerin daha derinden anlaşılmasına olanak tanır ve sosyal sistemlerde meydana gelen sorunlar dikkate

<sup>§</sup> “Friedrich Daniel Ernst Schleiermacher 1768-1834 yılları arasında yaşamış Alman ilahiyatçısı, filozofu ve eğitimcisidir” (Taşdelen, 2007, 189).



alınarak çözümler bulunabilir (Akyıldırım, 2020). Eleştirel paradigma, basamaklıp tarihi, ekonomik, sosyal ve kültürel değerlerin normallliğini sorgular (Çıvık & Sezerel, 2018).

Sözü edilen üç araştırma yaklaşımının yukarıda açıklanan özellikleri, eğitim bilimleri araştırmalarında tercih edilen doküman incelemesi yönteminin çeşitliliğinin ve çoğulluğunun nedenini açıklamaktadır. Bununla birlikte, üç araştırma yaklaşımının özel olarak doküman incelemesine nasıl baktığının ortaya koyulması, geçerliğin ve güvenilirliğin sağlanması için kullanılacak kriterlerin açıklanmasını kolaylaştırabilir. Bu doğrultuda, üç araştırma yaklaşımının doküman incelemesine yönelik bakış açıları Jupp ve Norris'in (1993) çalışmasından özetlenerek Tablo 2'de sunulmuştur.

**Tablo 2**

*Pozitivist, Yorumlayıcı ve Eleştirel Paradigmaların Doküman İncelemesine Yaklaşımı*

<b>Araştırma Paradigması</b>	<b>Doküman İncelemesine Yaklaşımı</b>	<b>Araştırma Örnekleri</b>
Pozitivist	Doküman incelemesi, içerik analizi ile örneklendirilir. Sistematiçlik, objektiflik, genellik, nicellik ve belgelerin açık içeriğine odaklanmak doküman incelemesi için en önemli beş özelliktir. Tamamen teknik bir süreç olduğu görüşü vardır.	<ul style="list-style-type: none"><li>▪ Kelimelerin veya temaların bir dokümanda kaç kez geçtiğinin sayılması,</li><li>▪ Belirli türdeki öykülerin sütun sayısının belirlenmesi.</li><li>▪ Öykü türleriyle ilişkili manşetlerin boyutunun ölçülmesi.</li></ul>
Yorumlayıcı	Dokümanlara yönelik yorumlayıcı yaklaşım, sosyal araştırmalardaki etnografik geleneğe çok yakındır. Yorumlayıcı yaklaşım, doküman incelemesinin merkezine sosyal anlamları yerleştirir. Dokümanlardaki anlamların, hem dokümanı üretenler hem de hedef kitle ve araştırmacılar tarafından atfedilme biçimlerine ve sonraki sonuçlarına odaklanılır.	<ul style="list-style-type: none"><li>▪ Bireyler tarafından üretilen dokümanların ve bunların gündelik uygulamalarda oynadıkları rollerin incelenmesi.</li><li>▪ Araştırma hipotezlerini test etmek için bir dizi anket ve kişilik testi kullanan pozitivist araştırma yaklaşımının aksine, bireylerin öznel olarak deneyimlediği yola odaklanarak deneyimleri ve bunlarla ilgili duyguları hakkında yazdıkları denemeler, öyküler ve şiirlerin incelenmesi.</li></ul>
Eleştirel	Eleştirel gelenek içinde önemli bir temsilci söylem analizidir. Eleştirel paradigmanın merkezi kavramları ve vurguları, dokümanlar aracılığıyla incelenebilecek kilit sorular ortaya çıkarır. Bu şekilde, eleştirel paradigma, diğer yaklaşımlardan kendisini ayıran veri kaynağı olarak dokümanların analizine bir farklılık getirir.	<ul style="list-style-type: none"><li>▪ Büyük topluluklar ve toplumsal yapılar düzeyinde analizler yapılması ve dolayısıyla bu düzeyde işleyen resmi, yarı resmi ve diğer kamusal dokümanların incelenmesi.</li><li>▪ Sosyal gruplaşmalar arasındaki çatışmaların ve bunlar arasındaki mücadelelerin dinamiklerinin dokümanlar üzerinden incelenmesi.</li><li>▪ Söylenenleri olduğu gibi kabul etmeme taahhüdünün incelenmesi (ve bu nedenle dokümanların açık içeriğine güvenmeme ve ayrıca ideoloji, güç ve kontrol ile ilgili rollerini inceleyerek dokümanlara eleştirel bir duruş sergileme).</li><li>▪ Mevcut durumu değiştirme taahhüdünün incelenmesi (ve dolayısıyla yalnızca dokümanlara karşı eleştirel bir duruş benimsemekle kalmayıp, aynı zamanda resmi dokümanlarda hakim olanlara alternatif tanımlar, açıklamalar ve çözümler bulma).</li></ul>

*Not:* Araştırma paradigmasının doküman incelemesine yaklaşımları ve araştırma örnekleri Victor Jupp ve Clive Norris tarafından yazılan *Traditions in Documentary Analysis* (1993, 37-51) başlıklı kitap bölümünden uyarlanmıştır.

Tablo 2’de açıkça görüldüğü gibi üç araştırma paradigmasının doküman incelemesi yöntemine yaklaşımları insan doğası, gerçeklik anlayışları veya metodolojik yönelimleri bakımından farklılaşmaktadır. Dolayısıyla, doküman inceleme yöntemi kullanılan eğitim bilimleri araştırmalarında geçerlik ve güvenilirliğin, araştırma paradigmasına göre farklılaşacağı ifade edilebilir. Eğitim bilimlerindeki araştırmalar, geçerlik ve güvenilirlik açısından doküman inceleme yönteminin teorik çerçevesinin dayandığı araştırma paradigmasına özgü araçlar üretme konusunda sorunlar yaşamaktadır. Alan yazındaki çalışmalarda yer verilmeyen ya da oldukça yüzeysel bir şekilde değinilip geçilen geçerlik ve güvenilirlik konusu; hem araştırmanın yöntemiyle ilgisiz, uyumsuz ve kalıplaşmış bilgilerin bir araya getirilmesine yol açmakta hem de işlevsiz kalmaktadır. Bu nedenle bir araştırma yöntemi olarak doküman incelemesinin kullanımı eğitim bilimleri alanıyla her zamankinden daha alakalı ancak bunun yanında kullanımı dikkat gerektiren bir hale gelmektedir.

Dokümanlar, eğitimsel konuları araştırmanın bir aracı olarak kullanılacaksa, araştırma yaklaşımından bağımsız olarak bir takım geçerlik ve güvenilirlik sorunlarının çözülmesi gerekir. En önemli sorunların başında ise dokümanların arşivlerde saklanırken tarafsız davranılmamasıdır. Daha yalın bir şekilde ifade etmek gerekirse bazı dokümanlar arşivlenirken bazıları arşivlenmez (Gidley, 2012). Örneğin okullarda öğretmenlerin öğrencilere uyguladığı fiziksel şiddeti araştıran bir araştırmacı için öğretmenlerin geçirdiği adli ve idari soruşturmaların ya da aldıkları cezaların kayıtları önemli bir veri kaynağıyken, çeşitli nedenlerle bu dokümanlar okul arşivinden çıkarılmış olabilir. Dokümanların geçerliği ve güvenilirliği hakkında büyük özen gösterilmesi gerektiğini vurgulayan Scott (1990, 6), dokümanların “özgünlük”, “inandırıcılık”, “temsil edilebilirlik” ve “anlam” kriterlerine göre değerlendirilmesi gerektiğini önermektedir\*\*. Tablo 3’te Scott (1990) tarafından önerilen kriterler özetlenmiştir.

**Tablo 3**

*Dokümanların Değerlendirilmesinde Kullanılabilecek Dört Kriter*

<b>Değerlendirme Kriterleri</b>	<b>Değerlendirme Sorusu</b>	<b>Açıklama</b>
Özgünlük	Gerçek mi ve kaynağı belirli mi?	Dokümanların gerçekliği hakkında bir sonuca varılamadığı sürece, nihai olarak oluşturulan verilerin kalitesi hakkında bilinçli bir yargıda bulunma olasılığı yoktur.
İnandırıcılık	Hatasız ve çarpıtmasız mı?	Dokümanların çarpıtılmamış ve katıksız, hata ve manipülasyondan arınmış olma derecesini ifade eder.
Temsil edilebilirlik	Kendi türünde örnek mi ve eğer değilse, örnek olmama derecesi biliniyor mu?	Dokümanın örneksel olup olmadığının değerlendirilmesine ilişkin genel sorunu ifade eder. Ancak araştırmacının her zaman "örneksel" doküman istediği varsayılmamalıdır. Önemli olan, araştırmacının ulaşılan herhangi bir sonucun uygulanmasına sınırlar atayabilmesi için bunun ne kadar örneksel olduğunu bilmesidir.
Anlam	Açık ve anlaşılır mı?	Dokümanın araştırmacı için ne kadar açık ve anlaşılır olduğudur. Doküman nedir ve araştırmacıya ne anlatır?

\*\* Bu kriterler Scott (1990) tarafından yayımlanmadan dokuz yıl önce Platt (1981) tarafından benzer kriterlerin ortaya koyulduğu söylenebilir. Platt’in (1981) çalışması ile Scott’ın (1990) çalışması karşılaştırıldığında “özgünlük” kriterinin her iki yazarda da ortak olduğu, bunun dışında Platt’in inandırıcılık yerine doğruluk, temsil edilebilirlik yerine örnekleme, anlam yerine çıkarım ifadelerini kullandığı görülmektedir. Aynı zamanda, bu ifadelerin açıklamaları da birbirine benzerdir.

*Not:* Dokümanların değerlendirilmesinde kullanılabilecek kriterler ve açıklamaları John Scott tarafından yazılan *A Matter of Record: Documentary Sources in Social Research* başlıklı kitaptan uyarlanmıştır.

Tablo 3'te verilen değerlendirme kriterleri, doküman analizi yönteminin kullanıldığı araştırmalarda geçerlik ve güvenilirlik konularının oldukça ciddiye alınması gerektiğine işaret etmektedir. Bu noktada araştırmacının yapması gereken veri kaynağı dokümanlara ne kadar güvenebileceğini değerlendirmektir. Dokümanların güvenilebilirliği konusunun esas olarak dokümanların gerçeklere dayalı, önyargılı ya da manipülatif bir biçimde üretilmiş olmasıyla ve belgelerin mevcudiyetiyle ilgili olduğunu öne süren McCulloch'a (2004) göre bu potansiyel güven sorununun üstesinden gelmek için çeşitleme yapılmalıdır. McCulloch (2004) çeşitlemeyi, alternatif bakış açılarını ve ilgi alanlarını temsil edecek çok çeşitli türden dokümanların birbirine karşı test edilmesi olarak tanımlamaktadır. Krüger (2010) çeşitlemenin tanımını biraz daha genişleterek farklı verilerin, farklı araştırmacıların veya farklı araştırma ekiplerinin, farklı teorik yaklaşımların veya aynı olgunun farklı analiz yöntemlerinin birleşimini içerdiğini belirtmektedir.

Araştırma paradigmasından bağımsız olarak dokümanların güvenilebilirliğinin değerlendirilmesinin ardından Tablo 2'de açıklanan araştırma paradigmalarının doküman inceleme yöntemine yaklaşımları göz önünde bulundurularak geçerlik ve güvenilirlik kriterleri karşılanabilir. Bengtsson (2016)'a göre pozitivist araştırma paradigmasına uygun bir doküman incelemesi yöntemi kullanan araştırmacılar, güvenilirliği sağlamaya yönelik olarak analiz sürecinde bilişsel değişimi en aza indirmek için kodların açıklamalarını içeren bir kodlama listesi kullanmalıdır. Kodların, araştırma tasarımına bağlı olarak tümevarımlı veya tümdengelimli olarak üretilebileceğini vurgulayan Bengtsson (2016)'ın kodlama listelerinin oluşturulmasına yönelik görüşleri Tablo 4'te özetlenmektedir.

**Tablo 4**

*Pozitivist Araştırma Paradigmasında Kodlama Listelerinin Araştırma Tasarımına Göre Oluşturulması*

<b>Araştırma Tasarımı</b>	<b>Açıklama</b>
Tümdengelimli Tasarım	Araştırmacının analiz sürecine başlamadan önce bir kodlama listesi oluşturması gerekir. Tümdengelim yoluyla oluşturulan kod listeleri ile yüksek güvenilirlik elde etmek çok daha kolaydır. Yardımcı olabilecek bilgisayar programları da vardır. Kullanımları zorunlu olmasa da süreci kolaylaştırabilirler. Bu programlar kodları bularak ve verileri kategoriler halinde gruplandırarak süreci hızlandırır. Bununla birlikte, araştırmacı yine de temaları neyin oluşturduğuna ve bulgulardan hangi sonuçların çıkarılabileceğine karar vermelidir.
Tümevarımlı Tasarım	Araştırmacı, analiz süreci içerisinde kodlama listesi oluşturabilir. Kodlar, çalışma ilerledikçe ve daha fazla veri kullanılabilir hale geldikçe değişebilir. Başlangıçta net görünen anlam birimlerinin yorumları, süreç içinde belirsizleşebilir. Bu nedenle tutarlılığı ve güvenilirliği artırmak için kodlama işlemi metnin farklı sayfalarından başlanarak tekrar tekrar yapılmalıdır.

*Not:* Tümdengelimli ve tümevarımlı tasarıma göre kodların oluşturulması Mariette Bengtsson tarafından yazılan *How to plan and perform a qualitative study using content analysis* başlıklı araştırma makalesinden uyarlanmıştır.

Tablo 4'te açıklanan tümdengelimli veya tümevarımlı tasarıma göre kodların oluşturulması, pozitivist paradigmaya dayalı doküman inceleme yöntemi kullanılan eğitim bilimleri araştırmalarında güvenilirliğin sağlanması için tek başına yeterli değildir. Pozitivist araştırma paradigması ekseninde doküman incelemesi yönteminde güvenilirliği sağlamak için

araştırmacıların atabileceği başka adımlar da vardır. Tablo 5, pozitivist paradigmaya dayalı doküman incelemesinin güvenilirliğini sağlamak için uygulanan stratejileri göstermektedir.

**Tablo 5**

*Pozitivist Paradigma Ekseninde Doküman İncelemesi İçin Güvenirlik Stratejileri*

<b>Güvenirlik Stratejisi</b>	<b>Açıklama</b>
Kararlılık	Kararlılık, bir sürecin zaman içinde değişmeme derecesidir. Bir kodlama prosedürünün tekrarlanan denemelerde ne ölçüde aynı sonuçları verdiği olarak ölçülür. Bu tür değerlendirmeler için veriler, test-tekrar test koşulları altında oluşturulur. Araştırmacı, genellikle bir süre geçtikten sonra aynı dokümanı yeniden inceler, yeniden kodlar veya yeniden analiz eder. Test-tekrar test koşulları altında, güvenilmezlik, araştırmacının performansındaki zaman içindeki değişimlerde kendini gösterir. Araştırmacıların güvensizliği, dikkatsizliği, dikkat dağınıklığına açık olmaları, yazılı talimatları anlamadaki zorluklar veya yorgunken performans standartlarını gevşetme eğiliminde olmaları gibi nedenler güvenilirliği etkiler. Kararlılık, pozitivist paradigma eksenindeki doküman inceleme yöntemindeki en zayıf güvenilirlik ölçüsüdür. Bu nedenle, kararlılık ölçümleri, güvenilirlik için tek kriter olarak yetersiz görünmektedir.
Tekrarlanabilirlik	Tekrarlanabilirlik, bir sürecin, farklı koşullar altında, farklı yerlerde çalışan veya farklı ancak işlevsel olarak eşdeğer ölçüm araçları kullanan farklı araştırmacılar tarafından yeniden üretilebilme derecesinin bir ölçüsüdür. Tekrarlanabilirliğin gösterilmesi, test-test koşulları altında elde edilen güvenilirlik verilerini gerektirir. Örneğin birbirinden bağımsız çalışan iki veya daha fazla kişi, aynı kayıt yönergelerini aynı analiz birimlerine uygular. Kararlılık ile karşılaştırıldığında, kodlayıcılar arası güvenilirlik, bir kavramın ortak tanımı üzerinde uzlaşma ve paralel formlar güvenilirliği olarak da adlandırılan tekrarlanabilirlik, kararlılıktan çok daha güçlü bir güvenilirlik ölçüsüdür.
Doğruluk	Dokümanın sınıflandırılmasının ve/veya kodlanmasının istatistiksel olarak bir standarda veya norma karşılık gelme derecesidir. Doğruluk, en güçlü güvenilirlik testidir. Pozitivist paradigma eksenindeki doküman inceleme yönteminde, doğruluk ölçümleri genellikle kodlayıcıların çalışmalarının deneyimli araştırmacılar tarafından oluşturulan standartlara göre test edilmesini içerir. Bu nedenle, doğruluğun kullanımı, kodlayıcı eğitimi ve nesnel standartların kolayca bulunabildiği pozitivist paradigma eksenindeki doküman inceleme yöntemiyle sınırlıdır.

*Not:* Pozitivist paradigma ekseninde doküman incelemesi için güvenilirlik stratejileri Klaus Krippendorff tarafından yazılan Content analysis: An introduction to its methodology başlıklı kitaptan uyarlanmıştır.

Tablo 5'te verilen açıklamalar incelendiğinde doğruluğun, kararlılıktan ve tekrarlanabilirlikten daha güçlü bir strateji olduğu görülmektedir. Doğruluğu oluşturan veriler, veri oluşturma süreçlerinin karşılaştırıldığı bir standarda atıfta bulunulmasını gerektirir (Krippendorff, 2018). Bu gerekliliğin ardındaki fikir, araştırmanın yalnızca, incelenen analiz birimleri hakkında, bağımsız gözlemciler arasında yeterli uzlaşma olması durumunda yayımlanabileceğidir; böyle bir uzlaşma yoksa, araştırmanın normal olarak yeterince tekrar üretilemeyeceği ve hatta güvenilmez olduğunun düşünülmesidir (Bolognesi, Pilgram ve van den Heerik, 2017). Son yıllarda, eğitim bilimleri alanındaki bilimsel dergilerin, araştırmacılar kodlayıcılar arası güvenilirlik puanlarını bildirmelerini istemesinin altında yatan nedenin de bu olduğu söylenebilir. Kodlayıcılar arası güvenilirliği belirlemek için alan yazında bir dizi güvenilirlik ölçümü önerilmiştir. Bunlardan en basiti, bağımsız gözlemciler arasındaki uzlaşma yüzdesinin belirlenmesidir. Tablo 6'da pozitivist paradigmayı temele alan doküman inceleme yöntemi eğitim bilimleri araştırmalarında kullanılacak güvenilirlik testlerinden bazıları sunulmuştur.

**Tablo 6***Kodlayıcılar Arası Güvenilirliği Belirlemede Kullanılabilecek Bazı Güvenirlik Testleri*

<b>Güvenirlik Testi</b>	<b>Formül</b>	<b>Açıklama</b>
Scott'ın $\pi$ istatistiği (Scott, 1955)	$\pi = \frac{P_o - P_e}{1 - P_e}$ $P_e = \sum_{i=1}^k p_i^2$	P <sub>o</sub> , iki analistin aynı verileri bağımsız olarak kodlarken üzerinde anlaştıkları uyuşmaların yüzdesini temsil eder. Üzerinde anlaşılan kod sayısının toplam kod sayısına bölünmesiyle bulunur. P <sub>e</sub> , şansa dayalı olarak beklenecek anlaşma yüzdesidir. $\pi$ , elde edilen ve tesadüfi anlaşma arasındaki gerçek farkın, elde edilen ve tesadüfi anlaşma arasındaki maksimum farka oranıdır. P <sub>e</sub> formülündeki k, toplam kategori sayısını ve p <sub>i</sub> ise i'inci kategoriye giren tüm örneklemin i=1'e oranıdır.
Cohen'in $\kappa$ istatistiği (J. Cohen, 1960)	$\kappa = \frac{P_o - P_e}{1 - P_e}$ $P_e = \frac{1}{N^2} \sum_k n_{k1} n_{k2}$	$\kappa$ katsayısı, tesadüfi uyuşma değerlendirmeden çıkarıldıktan sonraki uyuşmanın oranıdır. P <sub>o</sub> , analistlerin üzerinde anlaştıkları birimlerin oranıdır ve Scott'ın $\pi$ istatistiğindeki gibi hesaplanır. P <sub>e</sub> , şans eseri uyuşmanın beklendiği birimlerin oranıdır. P <sub>e</sub> formülündeki k, sınıflandırma sayısıdır. N sınıflandırılacak gözlem sayısıdır. n <sub>ki</sub> ise değerlendiricinin k kategorisini tahmin etme sayısıdır.
Fleiss'in $\kappa$ istatistiği (Fleiss, 1971)	$\kappa = \frac{\bar{P} - \bar{P}_e}{1 - \bar{P}_e}$ $\bar{P} = \frac{1}{N} \sum_{i=1}^N P_i$ $P_i = \frac{1}{n(n-1)} \sum_{j=1}^k n_{ij}(n_{ij} - 1)$ $\bar{P}_e = \sum_{j=1}^k p_j^2$ $P_j = \frac{1}{N_n} \sum_{i=1}^N n_{ij}$	Cohen'in $\kappa$ istatistiği her bir denek örneğini nominal bir ölçekte derecelendiren iki değerlendirici arasındaki anlaşma derecesini ölçerken, Fleiss'in $\kappa$ katsayısı ikiden daha çok sayıda değerlendirici için uygulanabilmektedir. $1 - \bar{P}_e$ miktarı, tesadüfen tahmin edilebilecek olanın ötesinde elde edilebilecek uyuşma derecesini ölçer. Gerçekte tesadüfinin üzerinde elde edilen uyuşma derecesi $\bar{P} - \bar{P}_e$ 'dir, öyle ki şans eseri beklenen miktar için düzeltilmiş normalleştirilmiş genel uyuşma ölçüsü $\kappa$ 'dır. Formüllerde kullanılan N değerlendirilecek durumların sayısıdır; n toplam değerlendirici sayısı ve k ise değerlendirmede kullanılacak kategori sayısını ifade etmektedir. i alt indisi (i= 1, . . .,N) durumları, j alt indisi (j=1,...,k) ise kategorileri temsil etmektedir.
Miles ve Huberman'ın kodlayıcılar arası güvenilirlik formülü (Miles ve Huberman, 1994)	$\text{Güvenirlik} = \frac{n}{N}$	Güvenirlik katsayısının %80'den yüksek bir değer olması önerilmektedir. Söz konusu formülde n, kodlayıcıların üzerinde görüş birliğine vardıkları kodlama sayısını, N ise üzerinde görüş birliğine vardıkları ve görüş ayrılığı yaşadıkları toplam kod sayısını göstermektedir.
Krippendorff'un alfa güvenilirlik katsayısı (Krippendorff, 2018)	$\alpha = 1 - \frac{D_o}{D_e}$ $D_o = \frac{1}{n} \sum_{j=1}^N m_j E(\delta_j)$	$\alpha$ , mükemmel güvenilirlikten sapmayı, gözlenen ile beklenen uyuşmazlığın oranına göre değerlendirir. Formülde verilen D <sub>o</sub> gözlenen uyumsuzluk ve D <sub>e</sub> beklenen uyumsuzluktur. Gözlenen uyumsuzluk formülündeki $E(\delta_j)$ , $\delta(v_{ij}, v_{i'j})$ sayısının

$$D_e = \frac{1}{P(n, 2)} \sum_{c \in R} \sum_{k \in R} \delta(c, k) P_{ck} \quad \binom{m_j}{2} \text{ ortalamasıdır. } D_e \text{ denkleminde yer alan } P_{ck} \text{ ifadesi, } (c, k) \text{ çiftlerinin yapılabileceği yolların sayısıdır.}$$

Tablo 6’da verilen Miles ve Huberman’ın güvenilirlik formülü dışında kalan testler kategori sayısı arttıkça kodlayıcıların uzlaşmaya varma şansının azalmasını veya uzlaşmaya varma şansının kategorilerin aşırı temsilinden etkilenmesini çözmek için önerilebilir. Formüller incelendiğinde, söz konusu güvenilirlik testlerinin uzlaşma yüzdesini tesadüfi uzlaşma olasılığına göre düzelttiği söylenebilir. Güvenirliği kontrol etmek için az ya da çok standartlaştırılmış bir ölçüm sağlayan bu tür hesaplamaların hiçbir şekilde doğruluk veya kalitenin doğrudan ölçümleri olmadığını belirtmek gerekir (Bolognesi vd., 2017). Çeşitli nedenlerle kodlayıcılar benzer kodlamaları yapabilirler ve böylece kodlayıcılar arası güvenilirlik yüksek olabilir. Ancak bu durum kodlayıcılar arası güvenilirlik değerlerinin herhangi bir yararının olmadığı anlamına gelmemelidir çünkü yeterince yüksek güvenilirlik değerleri önemli bir gösterge olarak kabul görmektedir. Bununla birlikte, pozitivist paradigmanın hesaplamalara dayalı doğasına rağmen araştırmacıların insani doğası gereği, güvenilirlik her zaman tam olarak sağlanamayabilir, yalnızca güvenilirlik düzeyi yükseltilebilir. Güvenirliğin yanında geçerlik de pozitivist paradigmayı temele alan doküman inceleme yönteminin kullanıldığı eğitim bilimleri araştırmaları için üzerinde önemle durulması gereken bir kavramdır. Doküman inceleme yönteminin kullanıldığı bir eğitim bilimleri araştırmasının temelinde pozitivist bir bakış açısı olduğu durumlarda söz konusu araştırmanın geçerliğini oluşturan kriterler Tablo 7’de belirtilmiştir.

**Tablo 7**

*Pozitivist Paradigma Ekseninde Doküman İncelemesi İçin Geçerlik Kriterleri*

<b>Geçerlik Kriterleri</b>	<b>Açıklama</b>
Kategorilerin yakınlığı	Bu kriter, özellikle içerik analizlerinde, kategoriler üzerinde anlaşmaya varılan bir tanıma ulaşmak için birden fazla sınıflandırıcı kullanılmasını gerektirir. Birden fazla sınıflandırıcı kullanılarak, açık değişken olabilecek bir kavram kategorisi, eşanlamlıları veya örtük değişkenleri içerecek şekilde genişletilebilir.
Sonuçlar	Dolaylı anlatımlara hangi düzeyde izin verilebilir? Sonuçlar verileri doğru bir şekilde takip ediyor mu? Sonuçlar başka olgularla açıklanabilir mi? Bu kriter, özellikle doküman incelemesinde bilgisayar yazılımları kullanırken ve eşanlamlılar arasında ayırım yaparken sorunlu hale gelir. Örneğin, “yüz” sözcüğü çeşitli anlamları ifade eder. Bilgisayar yazılımları, “yüz” kelimesinin geçtiği yerin ve sıklığının sayısını doğru bir şekilde verebilir, ancak her bir özel kullanımın doğasında var olan anlamın doğru bir açıklamasını üretemez. Bu problem, araştırmacının sonuçlarını geçersiz kılabilir.
Sonuçların bir teoriye genellenebilirliği	Kavram kategorilerinin açık tanımlarına, bunların nasıl belirlendiğine ve kişinin ölçmek istediği fikri ölçmede ne kadar güvenilir olduğuna bağlıdır. Genellenebilirlik, güvenilirlik kriterlerine bağlı olduğundan, güvenilirlikle paralellik gösterir.

*Not:* Pozitivist paradigma ekseninde doküman incelemesi için geçerlik kriterleri, Dr. Gina Lovasi liderliğinde Columbia Üniversitesindeki çeşitli fakülte ve doktora öğrencilerinden oluşan bir ekip tarafından oluşturulan ve <https://www.publichealth.columbia.edu/research/population-health-methods/techniques> adresinde sürekli olarak güncellenen materyallerden uyarlanmıştır.

Tablo 7’de eğitim bilimleri araştırmacılarının doküman inceleme yöntemi kullandıkları pozitivist paradigmaya sahip araştırmalarını gerekçelendirmek için kullanabilecekleri geçerlik kriterlerinin bir tipolojisi sunulmuştur. Ayrıca araştırmacıların geçerliği sağlamaya yönelik çabalarının en azından bir kısmını pozitivist paradigmaya uygun şekilde niceliksel olarak

değerlendirebilecekleri yolları göstermektedir. Bu geçerlik kriterlerinden farklı olarak Krippendorff (2018), bir araştırmamanın mümkün olduğu kadar çok insana hitap etmesi gerektiği fikrini odağa alarak görünüş geçerliği, sosyal geçerlik ve ampirik geçerlik olmak üzere üç kriter öne sürmüştür. Krippendorff (2018) tarafından ortaya koyulan bu üç kriter Tablo 8’de özetlenmiştir.

**Tablo 8**

*Pozitivist Paradigma Ekseninde Doküman İncelemesi İçin Krippendorff’un Geçerlik Kriterleri*

<b>Geçerlik Kriterleri</b>	<b>Açıklama</b>
Görünüş geçerliği	Mantıklı, makul ve bilinen gerçektir. Araştırma bulguları ayrıntılı gerekçeler sunmayı gerektirmediklerinde, mantıklı, makul ve inandırıcı olduklarında görünüş geçerliği başvurulabilir bir kriterdir. Görünüş geçerliğinin kökleri sağduyuya, yaygın olarak paylaşılan fikir birliğine dayansa da temelde herkesin onunla aynı fikirde olacağı varsayımıyla araştırmacının yargısıdır.
Sosyal geçerlik <sup>††</sup>	Sosyal geçerlik, farklı kamusal sorunları inceleyen araştırmaların, bu sorunlar hakkında endişelenen ve araştırma bulgularını eylemlere dönüştürmeye hevesli olan savunucular ve muhalifler tarafından sosyal olarak doğrulanabilmesidir. Araştırma konusunda kamu tarafından kabul edilen otoriteler, bulguların sosyal geçerliği için anahtardır.
Ampirik geçerlik <sup>‡‡</sup>	Alanyazındaki mevcut bulguların ve yerleşik teorinin bir araştırma sürecinin çeşitli aşamalarını ve sonuçlarını destekleme derecesidir. Ampirik geçerlik sezgiyi inkar edemez (görünüş geçerliliği) veya kendisini sosyal, politik ve kültürel faktörlerden tamamen ayıramaz (sosyal geçerlilik). Sonuçta, bilimsel araştırmalar, kendi teorik gündemleri, fikirleri, sosyal ve politik kaygıları olan diğer araştırmacılar tarafından incelenir.

*Not:* Pozitivist paradigma ekseninde doküman incelemesi için Krippendorff’un geçerlik kriterleri, Klaus Krippendorff tarafından yazılan Content analysis: An introduction to its methodology başlıklı kitaptan yararlanmıştır.

Tablo 8’de, geçerliğin üç ana kriter açısından değerlendirildiği görülmektedir. Tablo 8’de verilen açıklamalar incelendiğinde görünüş geçerliği ve sosyal geçerliğin daha çok genel kabul görme ile ilgili olduğu söylenebilir. Ampirik geçerliğin ise büyük ölçüde araştırma sonuçlarına, araştırma sürecine ve verilerin elde edildiği koşullara dayanan ampirik kanıtlar getiren rasyonel argümanlara dayandığı söylenebilir.

Daha çok teknik bir süreç olarak görülen pozitivist paradigmadan farklı olarak yorumlayıcı ve eleştirel paradigmada araştırmamanın merkezine öznel deneyimlerin ve toplumsal dinamiklerin anlamları yerleştirildiğinden, geçerlik ve güvenilirlik kriterlerinin de pozitivist paradigmadan bir miktar farklılık göstermesi normal kabul edilebilir. Bu bakış açısıyla, yorumlayıcı ve eleştirel paradigma eksenindeki doküman inceleme yöntemi kullanılan eğitim bilimleri araştırmalarında geçerliğin temelinde verilerin yerine olayların anlamının olacağı söylenebilir. Tablo 9, yorumlayıcı ve eleştirel paradigmaya sahip bir eğitim bilimleri

<sup>††</sup> Riffe ve diğerleri (2019, 138), sosyal geçerlik kavramını dış geçerlik ile ilişkilendirerek, dokümanlardan keşfedilebilecek içeriğin sosyal önemine ve araştırmacılar tarafından oluşturulan araştırma kategorilerinin “akademik bir izleyici kitlesinin ötesinde alaka ve anlama sahip olma derecesine bağlı olacağını” belirtmektedir.

<sup>‡‡</sup> Campbell (1957, 297), iki ana kriter açısından geçerliği değerlendirdiği çalışmasında “gerçekten de deneysel uyaran bu özel durumda önemli bir fark yarattı mı?” sorusunu ortaya atarak iç geçerliğin değerlendirilebileceğini vurgulamaktadır. Bu sorunun cevabını bulmak için alanyazındaki mevcut bulguların değerlendirilmesinin yapılması gerektiğinden ampirik geçerlik ile iç geçerlik arasında güçlü bir benzerlik olduğu söylenebilir.

araştırmasının doküman incelemesinin geçerliğini sağlamak için uygulanabilecek stratejileri göstermektedir.

**Tablo 9**

*Yorumlayıcı ve Eleştirel Paradigma Ekseninde Doküman İncelemesi İçin Geçerlik Stratejileri*

Geçerlik Stratejileri	Açıklama
Verilerin sistematik hale getirilmesi	Sürekli karşılaştırmalı bir yöntemle analitik bir tümevarım gerçekleştirilerek veriler sistematik hale getirilebilir. Bu süreçte, bilgisayar yazılımlarından yararlanılabilir.
Denetim izi	Yöntemin sistematik sürecinin her adımı belgelenir ve veri analizi sürecinin yazılı kanıtları saklanır.
Dokümanlar için değerlendirme kriterleri	Dokümanların araştırmaya dahil edilmesi ve hariç tutulması için nesnel kriterler önceden belirlenir.
Dokümanların tasdik edilebilirliği	Dokümanların gerçek olduğuna, taklit veya sahte olmadığına dair geçerli kanıtlar sunulur.
Eş kodlama	Araştırma tasarımı gereği dokümanlar üzerinde kodlama işlemi yapılması gerekiyor ise dokümanların tamamı veya bir bölümü araştırma alanındaki bir uzman tarafından kodlanır.
Zaman aralığı	Araştırmaya dahil edilen dokümanların, araştırma konusunun belirlediği zaman dilimine uygun olduğu teyit edilebilir.
Maksimum doküman çeşitliliği	Çeşitli türde dokümanlar bir arada kullanılabilir.

Tablo 9’da sunulan geçerlik stratejilerinin, yorumlayıcı ve eleştirel paradigmayla uyumlu hale getirilmiş stratejiler olduğu söylenebilir. Bunun nedeni, yorumlayıcı ve eleştirel paradigmalarda temel varsayımlarının ve ilkelerinin doküman incelemesi bağlamında birbirine yakınlık gösterirken<sup>§§</sup> pozitivist paradigmanınkilerle bağdaşmamasıdır. Paradigmadaki farklılıklardan ve bunun yanında yöntemin aynılığından kaynaklanan benzerliklerin ve ayrışmaların olması da normal karşılanabilir.

Doküman incelemesi yönteminin geçerlik stratejilerinde görülen araştırma paradigmalardaki farklılıklarla güvenilirlik stratejilerinde de karşılaşıldığı söylenebilir. Güvenirlik için temel fark, yorumlayıcı ve eleştirel paradigma çerçevesinde, araştırmacının açıklamaları araştırmanın hedef kitlesinin (okuyucular, kullanıcılar), "buna neden inanmalıyım?" sorusunu yanıtlamasına yardımcı oluyorsa çalışmanın güvenilir olduğudur (McGregor ve Murnane, 2010, 422). Bu kapsamda, yorumlayıcı ve eleştirel paradigmaya sahip bir eğitim bilimleri araştırmasının doküman incelemesinin güvenilirliğini<sup>\*\*\*</sup> sağlamak için uygulanabilecek stratejiler Tablo 10’da sunulmuştur.

<sup>§§</sup> Yorumlayıcı ve eleştirel paradigmalarda temel varsayımlarının ve ilkelerinin birbirine olan benzerliği farklı araştırmacılar tarafından da vurgulanmaktadır. Örneğin Morrow (2005), yorumlayıcı paradigmanın eleştirel paradigmayla belirli dünya görüşlerini paylaştığından, bu paradigmalarda güvenirlilik standartlarında önemli ölçüde örtüşme olduğunu belirtmektedir. Taylor ve diğerleri (2012), yorumlayıcı araştırma paradigmasının insanların düşünceleri, inançları, değerleri ve ilişkili sosyal eylemleri hakkında bağlama dayalı bir anlayış oluşturmakla ilgili olduğunu, bunun yanında eleştirel paradigmada ise müdahaleci bir rol üstlenmek, savunuculuk ve diğer aktif katılım biçimleri yoluyla ırk ayrımcılığı ve iklim değişikliği gibi sorunları gidermek için sosyal dünyanın yorumlayıcı anlayışının ötesine geçmeyi arzuladığını vurgulamaktadır. Fossey ve diğerleri (2002), yorumlayıcı ve eleştirel araştırma paradigmalarının ikisinin de insan eylemlerinin ve deneyimlerinin anlamlarını anlamaya ve ilgili kişilerin bakış açılarından anlamlarının açıklamalarını oluşturmaya önem verdiğinin altını çizmektedir.

<sup>\*\*\*</sup>Geçerlik stratejilerinde olduğu gibi yorumlayıcı ve eleştirel paradigmalarda güvenirlilik stratejileri de ortak bir şekilde değerlendirilmiştir. Bunun altında yatan neden yukarıdaki dipnotta belirtildiği gibi iki



**Tablo 10***Yorumlayıcı ve Eleştirel Paradigma Ekseninde Doküman İncelemesi İçin Güvenirlik Stratejileri*

<b>Güvenirlik Stratejileri</b>	<b>Açıklama</b>
Dokümanların derinlemesine anlaşılmasına ve anlam inşası	Dokümanların anlaşılması, çeşitli faktörlere bağlıdır. Bu faktörlerden en önemli iki tanesi bağlam ve kültürdür. Bağlamsal temellendirme, araştırmacıların dokümanlarda yer alan anlamı ortaya çıkarmak için gereklidir. Diğer bir faktör ise kültürel farklılıktır. Dokümanlar araştırmacının yabancı olduğu bir kültürel çevrede üretilmişse araştırmacının bu kültüre olan ilgisi ve yakınlığı açıkça tanımlanmalıdır.
Araştırmacının örtük varsayımlarını ve önyargılarını yönetmesi	Araştırmacılar örtük varsayımlarının, önyargılarının, eğilim ve yatkınlıklarının farkında olarak araştırmayı gereğinden fazla etkilemelerini önlemek için bunları bir kenara bırakmalıdırlar. Bunu gerçekleştirebilmek için araştırmacının varsayımlarının ve önyargılarının ne olduğunu tam olarak bilmesi gerekir ancak pratikte bu çok mümkün değildir. Bununla birlikte, bunları ortaya çıkarmak için bazı stratejiler kullanılabilir. Doküman incelemesinin başlangıcından sonuna kadar araştırmacının deneyimlerini, tepkilerini, varsayım veya önyargılarını not edebileceği, daha sonrasında bu kayıtları bilinçli olarak analizlerine dahil edip etmeme kararı vermesini sağlayabileceği bir seyir defteri kullanılması bu stratejilerden birisi olabilir. Diğer bir strateji ise araştırmacının bu türden öz-anlayışlarını yansıtabilecek alanında uzman ve deneyimli araştırmacılarından bir danışma grubu oluşturmasıdır.
Üretici teyidi	Bu strateji dokümanı üreten kişi veya topluluğa ulaşmanın mümkün olduğunda kullanılabilir. Doküman incelemesi yapan araştırmacılar, dokümanları incelerken, veri toplama sürecinde açıklama istemek ve dokümanlardan çıkardıkları anlamları daha derinlemesine araştırmak üzere dokümanları üreten(ler)den yararlanabilir. Bu sayede araştırmacı, yorumlarının ve/veya eleştirilerinin dokümanı üreten(ler)in kastettiği anlam(lar)ı ne kadar iyi yansıttığını öğrenebilir.
Destekleyici alıntılar	Sayıların pozitivist paradigmadaki bir araştırmanın güvenilirliğine katkıda bulunması gibi dokümanlardan yapılan gerçek alıntılar da araştırmacının yorumlarının ve/veya eleştirilerinin aslında incelenen dokümanlara dayandığına ikna etmek için gereklidir. Bununla birlikte, araştırmacının yorumları ve/veya eleştirileri ile dokümanlardan yapılan destekleyici alıntılar arasında bir denge sağlanmalıdır. Dokümanlardan yapılan alıntılar ile desteklenirse dahi araştırmacının yorumlarına ve/veya eleştirilerine gereğinden fazla yer vermesi yorumların nereden geldiği konusunda çalışmanın hedef kitlesini şüpheye düşürebilir. Bunun yanında, dokümanlardan çok fazla alıntı yapmak, araştırmanın hedef kitlesinin alıntılar içinde kaybolarak çalışmanın özünden uzaklaşmasına neden olabilir.
Derinlik ve zenginlik	Bu strateji, bir doküman incelemesinde derin ve zengin bir anlam elde etmek için her dokümanın birden fazla defa incelenmesini içerir. Bu her zaman mümkün olmasa da veri kaynağı olarak az sayıda dokümandan yararlanıldığı ya da araştırmacının zaman kısıtı olmadığı durumlarda uygulanabilir.
Çeşitleme	Farklı araştırmacılar, farklı teorik yaklaşımlar, farklı analiz yöntemleri kullanılabilir.
Yeniden kodlama	Dokümanların kodlanmasının takip eden belirli sürelerde kodlama süreci tekrarlanır. Araştırmacı, tekrarlanan karşılaştırmalar ve incelemeler yoluyla, dokümanlardan çıkarılan anlamı doğru bir şekilde yansıtmaya kadar temel kategorileri veya temaları gözden geçirebilir.
Meslektaş Teyidi	Bir meslektaş, aynı tümevarım veya tümdengelim sürecini izleyerek doküman incelemesinden elde edilen bulguları doğrular.

paradigmanın temel varsayımlarının ve ilkelerinin birbirine olan benzerliğidir. Bu konuda, Assalahi (2015), eleştirel paradigmadaki güvenilirlik sorunlarının, yorumlayıcı paradigmanınkileri yansıttığını vurgulamaktadır.

Tablo 10'da verilen yorumlayıcı ve eleştirel paradigmanın doküman inceleme yöntemine ilişkin güvenilirlik stratejilerinin, özellikle bu niteliklerin nasıl sağlanacağı söz konusu olduğunda, pozitivizminki kadar net ve yerleşik olmadığı görülmektedir. Bu makalede dokümanların derinlemesine anlaşılması ve anlam inşası olarak adlandırılan güvenilirlik stratejisinin, Guba ve Lincoln'ün (1989, 248) “ontolojik özgünlük” ve “eğitsel özgünlük” kavramları<sup>†††</sup> ve Patton'ın (2015) “verstehen (idrak etmek)” ve “diyalog” kavramları<sup>†††</sup> ile ilişkili olabileceği düşünülse de bu strateji daha derine inmektedir. Bir başka strateji olan üretici teyidinde ise dokümanların incelenmesinin ardından yapılan transkripsiyonun ya da çıkarımların yalnızca doğruluğunu kontrol etmek için üreten(ler)e gönderilmesinin vurgulanmadığı açıkça görülmektedir. Üretici teyidi stratejisinde araştırmacıların doküman incelemesi neticesinde ortaya çıkardıkları ürünle ilgili olarak üreten(ler)den düzeltme, yönlendirme ve geri bildirim alma sorumluluklarını yerine getirmelerine odaklanılmaktadır. Tablo 10'da verilen güvenilirlik stratejilerinden hangilerinin kullanılacağına karar vermek ise çalışma boyunca geçerlik ve güvenilirliği sağlayarak sürecin kalitesini korumak zorunda olan araştırmacının sorumluluğundadır.

Sonuç olarak, Guba'nın (1981) nitel araştırmanın geçerlik ve güvenilirliğini değerlendirmek için özetlediği kriterlere doküman incelemesi yöntemini kullanan eğitim bilimleri çalışmalarında hala atıfta bulunulmasına rağmen son zamanlarda doküman inceleme yönteminin kendine has felsefi duruşu ve amaçları göz önüne alınarak daha tutarlı kriterlere göre değerlendirilmesi gerektiği söylenebilir. Doküman inceleme yöntemini kullanan eğitim bilimleri alanındaki araştırmacılar, araştırmalarının geçerliğini ve güvenilirliğini sağlamak için paradigmaya dayalı stratejiler belirlemenin yanı sıra doküman inceleme yönteminin kendisinden ortaya çıkan bir dizi özel kaygıyı da dikkate almalıdırlar. Doküman inceleme yöntemi kullanılan eğitim bilimleri araştırmalarında geçerlik ve güvenilirlik stratejilerini açıklamadan çalışmanın geçerli ve güvenilir olduğunu basitçe kabul etmek, araştırmanın bilimsel değerini önemli ölçüde baltalayabilir. Bu nedenle, doküman inceleme yönteminin geçerlik ve güvenilirliğini sağlamak için, Tablo 3-10'da ana hatları çizilen kriterler, kural koyucu olarak alınmak yerine, bağlamı ve amacı göz önüne alındığında, bir eğitim bilimleri çalışmasının öğelerinin gözden geçirilmesine ve değerlendirilmesine rehberlik etmek için kullanılabilir. Doküman inceleme yönteminde geçerlik ve güvenirlığe ilişkin bu çalışmada, eğitim bilimleri alanında araştırma yapmaya yeni başlayanlar için yararlı yönergeler sağlanmaya ve aynı zamanda deneyimli araştırmacıların üzerinde hemfikir olabilecekleri kriterlerle ilgili alanda bir tartışma başlatılmaya çalışılmıştır. Dergi editörleri, makale değerlendiricileri ve eğitim bilimlerinin farklı kurullarındaki jüri üyelerinin bu çalışmayı faydalı bir araç olarak görmeleri ve doküman inceleme yönteminin kalitesinin korunmasında yardımcı olması umulmaktadır.

## KAYNAKÇA

Akyıldırım, E. (2020). “Tüketim Toplumu” kavramının pozitivist, eleştirel, yorumsamacı yaklaşımlarla irdelenmesi. *Asya Studies*, 4(12), 118-127. <https://doi.org/10.31455/asya.617128>

Asdal, K. & Reinertsen, H. (2022). *Doing document analysis: A practice-oriented method*. Sage.

<sup>†††</sup> Guba ve Lincoln (1989, 248) iki kavramın açıklamasına kitaplarında ayrıntılarıyla yer vermektedirler. Ontolojik özgünlük, bireyin (veya grubun) dünyayı bilinçli olarak deneyimlemesindeki gelişmedir. Eğitsel özgünlük ise bireysel katılımcıların kendi paydaş gruplarının dışındaki diğer kişilerin yapılarını anlama ve takdir etme derecesini temsil eder.

<sup>†††</sup> Bu iki kavramı Patton (2015, 56 ve 485) kitabında açıklamaktadır. Verstehen, olayları bir başkasının bakış açısıyla görme kapasitesini içerir. Diyalog ise olayı neyin kritik hale getirdiğini anlamak için başlayan süreçtir.

- Assalahi, H. (2015). The philosophical foundations of educational research: A beginner's guide. *American Journal of Educational Research*, 3(3), 312-317.
- Balaban Salı, J. (2018). Verilerin toplanması. In A. Şimşek (Ed.), *Sosyal bilimlerde araştırma yöntemleri* (pp. 134-161). Anadolu Üniversitesi.
- Baltacı, A. (2017). Avrupa okullarında dini simgelerin yasallığı. *Cumhuriyet İlahiyat Dergisi*, 21(1), 45-80. <https://doi.org/10.18505/cuid.286581>
- Baltacı, A. (2019). Nitel araştırma süreci: Nitel bir araştırma nasıl yapılır? *Ahi Evran Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 5(2), 368-388. <https://doi.org/10.31592/aeusbed.598299>
- Baş, T. & Akturan, U. (2017). *Nitel araştırma yöntemleri, Nvivo ile nitel veri analizi, örnekleme, analiz, yorum* (3. Baskı). Seçkin Yayıncılık.
- Bayhan, V. (2016). Sosyal bilimlerde "objektiflik" efsanesi. *Mukaddime*, 7(2), 217-242.
- Bengtsson, M. (2016). How to plan and perform a qualitative study using content analysis. *NursingPlus Open*, 2, 8-14. <https://doi.org/10.1016/j.npls.2016.01.001>
- Bolognesi, M., Pilgram, R. & van den Heerik, R. (2017). Reliability in content analysis: The case of semantic feature norms classification. *Behavior Research Methods*, 49, 1984-2001. <https://doi.org/10.3758/s13428-016-0838-6>
- Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin*, 54(4), 297-312. doi:<https://doi.org/10.1037/h0040950>
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37-46. <https://doi.org/10.1177/001316446002000104>
- Cohen, L., Manion, L. & Morrison, K. (2005). *Research methods in education*. RoutledgeFalmer.
- Çıvık, B. & Sezerel, H. (2018). Araştırma paradigmaları ve turizm yazını. *Turizm Akademik Dergisi*, 5(1), 1-14.
- Demirtaş, H. & Özer, N. (2015). Eleştirel kuram açısından eğitim ve eğitim yönetimi. *Journal of Teacher Education and Educators*, 4(2), 206-227.
- Doğanay, A. (2018). Bilimsel yönetime giriş. In A. Şimşek (Ed.), *Sosyal bilimlerde araştırma yöntemleri* (pp. 2-29). Anadolu Üniversitesi.
- Endacott, R. (2005). Clinical research 4: Qualitative data collection and analysis. *Intensive and Critical Care Nursing*, 21(2), 123-127.
- Erlanson, D. A., Harris, E. L., Skipper, B. L. & Allen, S. D. (1993). *Doing naturalistic inquiry: A guide to methods*. Sage.
- Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5), 378-382. <https://doi.org/10.1037/h0031619>
- Fossey, E., Harvey, C., McDermott, F. & Davidson, L. (2002). Understanding and evaluating qualitative research. *Australian & New Zealand Journal of Psychiatry*, 36(6), 717-732. <https://doi.org/10.1046/j.1440-1614.2002.01100.x>
- Gidley, B. (2012). Doing historical and documentary research. In C. Seale (Ed.), *Researching society and culture 3rd Edition* (pp. 263-282). SAGE Publications Ltd.
- Gomm, R., Hammersley, M. & Foster, P. (2000). Case study and generalization. In R. Gomm, M. Hammersley, & P. Foster (Eds.), *Case study method: Key issues, key texts* (pp. 98-115).

- Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Communication and Technology*, 29(2), 75-91.
- Guba, E. G. & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Sage Publications, Inc.
- Hammersley, M. (2006). *What's wrong with ethnography?: Methodological explorations*. Routledge.
- Hayashi Jr, P., Abib, G. & Hoppen, N. (2019). Validity in qualitative research: A processual approach. *The Qualitative Report*, 24(1), 98-112.
- Heale, R. & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-based Nursing*, 18(3), 66-67.
- Jupp, V. & Norris, C. (1993). Traditions in documentary analysis. In M. Hammersley (Ed.), *Social research: philosophy, politics and practice* (pp. 37-51). SAGE Publications.
- Kabakçı Yurdakul, I. (2013). Veri toplama araçlarında bulunması gereken nitelikler. In A. Şimşek (Ed.), *Bilimsel araştırma yöntemleri* (pp. 117-138). Anadolu Üniversitesi.
- Kartal, R. N. (2021). Gerçekler ve illüzyonlar ekseninde eleştirel (Critical) teori/yaklaşım. *Asya Studies*, 5(16), 195-213. <https://doi.org/10.31455/asya.895617>
- Kaya, F. & Büyükekmekci, B. (2022). Sosyal bilimlerde iki farklı yöntem: Pozitivizmin “açıklama”sından hermeneutiğin “anlama”sına. *Akademik Hassasiyetler*, 9(19), 415-442.
- Kıral, B. (2020). Nitel bir veri analizi yöntemi olarak doküman analizi. *Siirt Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 8(15), 170-189.
- Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. Sage Publications.
- Krüger, H. H. (2010). The importance of qualitative methods in the German educational science. In R. Bohnsack, N. Pfaff, & W. Weller (Eds.), *Qualitative analysis and documentary method in international education research* (pp. 53-74). Barbara Budrich Publishers.
- Leung, L. (2015). Validity, reliability, and generalizability in qualitative research. *Journal of Family Medicine and Primary Care*, 4(3), 324-327.
- Long, T. & Johnson, M. (2000). Rigour, reliability and validity in qualitative research. *Clinical Effectiveness in Nursing*, 4(1), 30-37. <https://doi.org/10.1054/cein.2000.0106>
- McCulloch, G. (2004). *Documentary research: In education, history and the social sciences*. RoutledgeFalmer.
- McGregor, S. L. T. & Murnane, J. A. (2010). Paradigm, methodology and method: Intellectual integrity in consumer scholarship. *International Journal of Consumer Studies*, 34(4), 419-427. <https://doi.org/10.1111/j.1470-6431.2010.00883.x>
- Miles, M. B. & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2 ed.). Sage.
- Miller, P. (2008a). Validity. In L. M. Given (Ed.), *The Sage encyclopedia of qualitative research methods* (pp. 909-910). SAGE Publications, Inc.
- Miller, P. (2008b). Reliability. In L. M. Given (Ed.), *The Sage encyclopedia of qualitative research methods* (pp. 753-754). SAGE Publications, Inc.
- Morrow, S. L. (2005). Quality and trustworthiness in qualitative research in counseling psychology. *Journal of Counseling Psychology*, 52(2), 250-260. <https://doi.org/10.1037/0022-0167.52.2.250>

- Namlı Altıntaş, İ. & Erdem, C. C. (2021). Hermeneutik ve nitel araştırma ilişkisi üzerine bir değerlendirme. *Uluslararası Türkoloji Araştırmaları ve İncelemeleri Dergisi*, 6(1), 37-45. <http://dx.doi.org/10.29228/uluturkad.49394>
- Neuman, W. L. (2014). *Social research methods: Qualitative and quantitative approaches* (7 ed.). Pearson Education Limited.
- Noble, H. & Smith, J. (2015). Issues of validity and reliability in qualitative research. *Evidence-based Nursing*, 18(2), 34-35.
- Özkan, U. B. (2021). *Eğitim bilimleri araştırmaları için doküman inceleme yöntemi*. Pegem Akademi Yayıncılık.
- Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4 ed.). Sage Publications, Inc.
- Platt, J. (1981). Evidence and proof in documentary research: 1 Some specific problems of documentary research. *The Sociological Review*, 29(1), 31-52. <https://doi.org/10.1111/j.1467-954X.1981.tb03021.x>
- Rapley, T. (2008). *Doing conversation, discourse and document analysis*. Sage.
- Riffe, D., Lacy, S., Watson, B. & Fico, F. (2019). *Analyzing media messages: Using quantitative content analysis in research* (4 ed.). Routledge.
- Sak, R., Şahin Sak, İ. T., Öneren Şendil, Ç. & Nas, E. (2021). Bir araştırma yöntemi olarak doküman analizi. *Kocaeli Üniversitesi Eğitim Dergisi*, 4(1), 227-256. <http://doi.org/10.33400/kuje.843306>
- Scott, J. (1990). *A matter of record: Documentary sources in social research*. Polity Press.
- Scott, W. A. (1955). Reliability of content analysis: The case of nominal scale coding. *The Public Opinion Quarterly*, 19(3), 321-325.
- Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. 22(2), 63-75.
- Silverman, D. (2014). *Interpreting qualitative data*. SAGE.
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of business research*, 333-339.
- Şişman, M. (1998). Eğitim yönetiminde kuram ve araştırmada alternatif paradigma ve yaklaşımlar. *Kuram ve Uygulamada Eğitim Yönetimi*, 16(16), 395-422.
- Taş, K. (2018). Sosyolojik metodolojinin ontolojik ve epistemolojik temelleri. *Turkish Studies*, 13(2), 49-70.
- Taşdelen, V. (2007). Çağdaş hermeneutige doğru: Schleiermacher ve genel hermeneutik. *Felsefe Dünyası*(46), 189-203.
- Taylor, P. C., Taylor, E. & Luitel, B. C. (2012). Multi-paradigmatic transformative research as/for teacher education: An integral perspective. In B. J. Fraser, K. Tobin, & C. J. McRobbie (Eds.), *Second international handbook of science education* (Vol. 1, pp. 373-387). Springer.
- Thanh, N. C. & Thanh, T. T. L. (2015). The interconnection between interpretivist paradigm and qualitative methods in education. *American Journal of Educational Science*, 1(2), 24-27.

Tosh, J. (2002). *The Pursuit of history: Aims, methods and new directions in the study of modern history (Revised Third Edition)*. Longman Pearson Education Limited.

Whittemore, R., Chase, S. K. & Mandle, C. L. (2001). Validity in qualitative research. *Qualitative Health Research*, 11(4), 522-537. <https://doi.org/10.1177/104973201129119299>

Yıldırım, A. & Şimşek, H. (2018). *Sosyal bilimlerde nitel araştırma yöntemleri (11. Baskı)*. Seçkin Yayıncılık.

## **EXTENDED ABSTRACT**

### **Introduction**

The use of documents is a common practice in educational research and can be utilized to construct the research process. Documents are used in educational sciences as a literature review tool and as a scientific research method falling under the scope of theoretical examination. The use of document examination as a scientific research method is becoming an increasingly complex process, as educational documents often contain fragmented data rather than a whole. It is recommended that researchers ensure the reliability and validity of the documents they use, try to access primary sources, and be cautious about the issues to be considered during document analysis.

Validity in qualitative research refers to the accuracy of the findings and means the suitability of the instruments, processes, and data. The concept of reliability is closely related to the validity concept. The basis of this connection is that reliability provides evidence about validity and informs about the usefulness of research strategies. The reliability in qualitative research evaluates the consistency of results over time. The discussions about reliability and validity in qualitative research show a confusing variety of ideas rather than a clear series of definitions. There are some differences between the definitions of various authors, but they are similar to a great extent.

The aim of this study is to evaluate the discussions on validity and reliability criteria in document analysis method in the context of educational sciences, by examining the efforts made to ensure them and synthesizing the views, which differs from a limited number of studies in the literature. In this context, the study will provide self-knowledge about the strategies for validity and reliability in educational sciences, errors made regarding validity and reliability, and situations that could be traps for researchers in a document-based research.

### **Method**

This study is qualitative research designed in a conceptual survey model. A conceptual survey examines different aspects of a concept within different research areas and disciplines. In this context, this study designed in the conceptual survey model aims to guide educational science researchers on validity and reliability in document analysis methods and to address the gaps found in the literature. Additionally, to demonstrate the necessity of the theoretical review taken as the basis in the study, descriptive analysis was conducted on the articles published in the Journal of Buca Faculty of Education in 2020, 2021, and 2022. In this context, the open access of 161 articles was accessed from the internet site <https://dergipark.org.tr/tr/pub/deubefd/archive> of the journal, and descriptive analysis was conducted. Among these articles, those that used the document examination method were identified. The results of the descriptive analysis were tabulated in the findings section to demonstrate the current situation. The theoretical review as the focus of the study was extensively discussed in the Discussion and Conclusion section in light of the findings, and recommendations were made. The research method used in this study is document examination due to the nature of the research.

## Results

According to the findings obtained from the descriptive analyses, it is observed that in 70% of educational science research studies that used the document examination method in the Buca Faculty of Education Journal, no explanation was made regarding the validity of the study. However, reliability was not mentioned in approximately half of these studies (46%). Based on the results of the descriptive analysis, it can be said that in a large part of educational science research studies that used the document examination method, validity and reliability issues were not addressed. Therefore, the current situation reveals that these issues need to be discussed. In the following section of the study, validity and reliability issues in educational science research studies that used document examination will be discussed in the methodological context, and recommendations will be made.

## Discussion and Conclusion

To understand, interpret, and establish a theoretical basis for validity and reliability debates in document review methodology, it is necessary to look at the research paradigms on which the theoretical framework of document review methodology is based. When examining different research paradigms, the differences and similarities in validity and reliability approaches can be seen more clearly. The steps that researchers can take to ensure reliability in document analysis method used in a positivist research paradigm are presented in Table 1.

**Table 1**

*The Positivist Paradigm's Reliability Strategies for Document Analysis*

<b>Reliability Strategy</b>	<b>Description</b>
Stability	Stability is the degree to which a process remains unchanged over time. It measures how consistently a coding procedure produces the same results in repeated trials.
Reproducibility	Reproducibility is a measure of the extent to which a process can be reproduced by different researchers using different but functionally equivalent measurement tools under different conditions or at different locations. Demonstrating reproducibility requires reliability data obtained under test-retest conditions.
Accuracy	Accuracy is the degree to which the classification and/or coding of a document corresponds statistically to a standard or norm. Accuracy is the strongest reliability test. In the document analysis method within the positivist paradigm, accuracy measurements usually involve testing the work of coders against standards created by experienced researchers.

*Note:* The reliability strategies for document analysis in the positivist paradigm are adapted from the book "Content Analysis: An Introduction to Its Methodology" by Klaus Krippendorff.

The validity, in addition to reliability, is a concept that needs to be emphasized for educational science research that utilizes the document analysis method based on the positivist paradigm. In cases where there is a positivist perspective at the foundation of educational sciences research that utilizes the document analysis method, the criteria that constitute the validity of said research are listed in Table 2.

**Table 2**

*Validity Criteria for Document Analysis in the Positivist Paradigm*

<b>Validity Criteria</b>	<b>Description</b>
The proximity of categories	This criterion, especially in content analysis, can be achieved by using multiple classifiers to reach a consensus definition for each specific category. By using

Results	multiple classifiers, a concept category that can be an open variable can be expanded to include synonyms or implicit variables. To what extent can indirect expressions be allowed? Are the results accurately tracking the data? Can the results be explained by other factors? This criterion becomes problematic, especially when using computer software in document analysis and distinguishing between synonyms. For example, the word "face" expresses various meanings. Computer software can accurately provide the number and frequency of occurrences of the word "face," but it cannot produce a correct explanation of the nature of each specific usage. This problem can invalidate the researcher's results.
Generalizability of results to a theory	It depends on the explicit definitions of concept categories, how they were determined, and how reliable they are in measuring the intended idea. <u>Generalizability is parallel to reliability criteria since it is dependent on reliability.</u>

*Note:* The validity criteria for document analysis in the positivist paradigm were adapted from materials created by a team of faculty and doctoral students at Columbia University, led by Dr. Gina Lovasi and continually updated at <https://www.publichealth.columbia.edu/research/population-health-methods/techniques>.

Klaus Krippendorff proposed three criteria for research, namely face validity, social validity, and empirical validity, focusing on the idea that a study should appeal to as many people as possible, unlike validity criteria in Table 2. These criteria are summarized in Table 3.

**Table 3**

*Validity Criteria for Document Analysis within the Positivist Paradigm: Krippendorff's Criteria*

<b>Validity Criteria</b>	<b>Description</b>
Face validity	Refers to whether the research findings are logical, reasonable, and known to be true. When research findings do not require detailed justifications, face validity can be applied as a criterion if they are logical, reasonable, and convincing. While the roots of face validity lie in common sense and widespread agreement, it is ultimately the researcher's judgment based on the assumption that everyone will agree with it.
Social validity	Social validity is the extent to which research examining different public issues can be socially confirmed by advocates and opponents who are concerned about these issues and eager to translate research findings into action. Authorities accepted by the public regarding the research topic are key to the social validity of the findings.
Empirical validity	The degree to which the existing findings in the literature and established theory support the various stages and outcomes of a research process. Empirical validity cannot deny intuition (face validity) or separate itself completely from social, political, and cultural factors (social validity). In the end, scientific research is examined by other researchers who have their own theoretical agendas, ideas, and social and political concerns.

*Note:* Krippendorff's validity criteria for document analysis within the positivist paradigm were adapted from the book *Content analysis: An introduction to its methodology*, written by Klaus Krippendorff.

In contrast to the positivist paradigm, which is seen as a more technical process, the interpretive and critical paradigm places the meanings of subjective experiences and social dynamics at the center of research. Therefore, it is acceptable to assume that the validity and reliability criteria may differ somewhat from the positivist paradigm. Table 4 shows strategies that can be applied to ensure the validity of document analysis in educational research with an interpretive and critical paradigm. .



**Table 4***Validity Strategies for Document Analysis in the Interpretive and Critical Paradigm*

<b>Validity Strategies</b>	<b>Description</b>
Systematization of data	Data can be systematized through an analytical deduction process using a continuous comparative method. Computer software can be used in this process.
Audit trail	Every step of the systematic process is documented, and written evidence of the data analysis process is kept.
Assessment criteria for documents	Objective criteria are predetermined for the inclusion and exclusion of documents in the research.
Verifiability of documents	Valid evidence is presented to confirm the authenticity and non-counterfeit nature of the documents.
Coding by multiple coders	If coding of documents is necessary for the research design, the entire document or a portion of it is coded by an expert in the research field.
Time interval	The documents included in the research are confirmed to be appropriate for the time period determined by the research topic.
Maximum document variety	Documents of various types can be used together.

The validity strategies presented in Table 4 demonstrate that, in the context of document analysis, they are more compatible with the interpretive and critical paradigms than with the positivist paradigm due to their proximity to the basic assumptions and principles of these paradigms. Differences and similarities between research paradigms are also encountered in reliability measures in document analysis. Therefore, in order to ensure the reliability of document analysis in educational research with an interpretive and critical paradigm, the appropriate strategies presented in Table 5 can be applied.

**Table 5***Reliability Strategies for Document Analysis on the Interpreter and Critical Paradigm*

<b>Reliability Strategies</b>	<b>Description</b>
In-depth understanding and meaning construction of documents	Understanding documents depends on various factors. The two most important ones are context and culture. Contextual grounding is necessary for researchers to uncover the meaning in documents. Another factor is cultural difference. If the documents were produced in a cultural context foreign to the researcher, their interest and familiarity with that culture should be clearly stated.
Managing the researcher's implicit assumptions and biases	Researchers should set aside their implicit assumptions, biases, tendencies, and inclinations, being aware of how they could unduly influence the research. To do so, they need to be fully aware of their assumptions and biases, but in practice, this is not always possible. Nevertheless, some strategies can be used to bring them to the fore. One strategy could be to use a logbook that enables researchers to note down their experiences, reactions, assumptions, or biases from the beginning to the end of the document review and to decide whether to include them in their analysis consciously. Another strategy is to form a consultation group of expert and experienced researchers who can reflect on these self-awarenesses.
Producer confirmation	This strategy can be used when it is possible to reach the person or community who produced the document. While reviewing documents, researchers can benefit from seeking explanations and exploring the meanings intended by the producers for a more in-depth analysis. In this way, the researcher can learn how well their comments and/or criticisms reflect the meaning(s) intended by the producers.
Supportive quotations	Just as numbers contribute to the reliability of a positivist study, actual quotations from documents are necessary to convince the researcher's comments and/or criticisms are based on the documents under investigation. However, a balance

	should be struck between the researcher's comments and/or criticisms and the supportive quotations from the documents. Giving too much weight to the researcher's comments and/or criticisms supported by quotations from the documents can raise doubts about where the comments come from among the target audience. Additionally, quoting from the documents excessively can lead to the target audience losing sight of the essence of the study.
Depth and richness	This strategy involves examining each document multiple times to obtain a deep and rich understanding of it. While this may not always be possible, it is necessary when the research relies on a small number of documents as a data source or when the researcher's goal is to explore complex meanings in the documents.
Variation	Different researchers can use different theoretical approaches and different analysis methods.
Recoding	The coding process of documents is repeated at certain intervals. The researcher can review the basic categories or themes through repeated comparisons and examinations until the meaning extracted from the documents is accurately reflected.
Peer Confirmation	A colleague verifies the findings obtained from document review by following the same inductive or deductive process.

It can be observed that the reliability strategies related to the document analysis method in the interpretive and critical paradigms given in Table 5 are not as clear and established as those in positivism, especially when it comes to how these qualities will be ensured. It is the responsibility of the researcher, who must maintain the quality of the process by ensuring validity and reliability throughout the study, to decide which of the reliability strategies given in Table 5 to use.

Considering the unique philosophical stance and aims of the document analysis method, it can be said that it needs to be evaluated according to more consistent criteria. Researchers in the field of educational sciences who use the document analysis method should take into account a number of specific concerns that arise from the method itself, in addition to developing paradigm-based strategies to ensure the validity and reliability of their research. Simply accepting a study as valid and reliable without explaining validity and reliability strategies in educational research using the document analysis method can significantly undermine the scientific value of the research. Therefore, the criteria outlined in Table 1-5, which can guide the review and evaluation of the elements of an educational science study when considering the context and purpose, can be used to guide the process of ensuring the validity and reliability of the document analysis method, rather than being taken as prescriptive rules.