

Kavram Haritası Puanlarının Güvenirliğinin İncelenmesi: Genellenebilirlik Kuramında Çaprazlanmış Karışık Desen Örneği

Investigation of Concept Map Scores' Reliability: Example of Crossed Mixed Design in Generalizability Theory

Gülten KAYA UYANIK*, Neşe GÜLER**

ÖZ: Bu çalışmada ölçme ve değerlendirme dersinin “Ölçme ve Değerlendirmeye İlgili Temel Kavramlar” konusunu ölçen iki farklı doldurma kavram haritası tekniğinin genellenebilirlik kuramı ile elde edilen parametreleri araştırılmıştır. Araştırma verileri, 2012-2013 Öğretim Yılı'nda Sakarya Üniversitesi Psikolojik Danışmanlık ve Rehberlik Bölümü 3. sınıfta okuyan 49'u (%53,8) kız 42'si (%46,2) erkek toplam 91 öğrenciden elde edilmiştir. Çalışmada o (öğrenci) x m (madde) x h (harita) çaprazlanmış ve yüzeylerden madde yüzeyi rastgele, harita yüzeyi ise sabit olmak üzere karışık desen kullanılmıştır. Yapılan G çalışması ile oxmxh deseninde var olan varyans kaynaklarının varyans değerleri ve toplam varyansları açıklama yüzdeleri elde edilmiştir. K çalışması ile madde sayılarının artırılması ve azaltılmasıyla G ve Phi katsayılarındaki değişim araştırılmıştır ve madde sayısı arttırıldıkça G ve Phi katsayılarında da artış olduğu gözlenmiştir. Bu çalışma ile genellenebilirlik kuramının gücü, elverişliliği ve esnekliğine dayanarak kavram haritası puanlama güvenilirliğinin elde edilmesinde klasik test kuramına göre tercih edilir olabileceği örneklendirilmeye çalışılmıştır. Diğer yandan bu çalışma, içerdiği *karışık* desen örneği ile literatürde yer alan genellenebilirlik çalışmalarından farklılaşmıştır ve bu nedenle önemli olduğu düşünülmektedir.

Anahtar sözcükler: kavram haritası, genellenebilirlik kuramı, karışık desen

ABSTRACT: This study investigates the parameters obtained through the generalizability theory which measures the topic “basic concepts related with measurement and evaluation” in the measurement and evaluation course. Data for the research were collected from 91 (49 girls 42 boys) students who are studied in the Sakarya University science education department in 2011-2012 semester. The o (student), X m (item) X h (map) were crossed in the study, and due to the fact that the item facet was random but the map facet was fix, the mixed design was used. The variances of the results obtained with filling concept maps and the percentages of explaining the total variances were researched with the G study and with the C and S techniques; and the changes in the G and Phi coefficients were researched by increasing and decreasing the item numbers with the D study. With this study, we aimed to exemplify the preferability compared to the classical test theory in having reliability of the concept map by leaning on the power, suitability and flexibility of the G theory. On the other hand, this current research is different from the generalizability studies available in literature in that it includes the mixed design sample, and therefore, it is believed to be important.

Keywords: concept map, generalizability theory, mixed design

1. GİRİŞ

Dünyanın hızlı değişimi, bilginin ve bilgiye ulaşma hızının artışı, teknolojideki gelişmeler ve nüfustaki değişimler eğitim sisteminde yenilikler yapmayı mecbur kılmıştır. Bu değişimlerle eğitim, öğretmen merkezli olmaktan çıkıp öğrenci merkezli olmuştur. Öğrencilerin, pasif bir bilgi alıcı olmak yerine aktif, yaşam boyu bağımsız öğrenici ve problem çözücü olması hedeflenmiştir. Bu doğrultuda eğitim sürecinde öğrencinin zihinsel becerilerini geliştirmeye ve bilgiyi yapılandırmaya ağırlık verilmektedir ve bu durum öğretim sürecinde anlamlı öğrenmenin gerekliliğini ortaya çıkarmaktadır.

* Yrd. Doç. Dr., Sakarya Üniversitesi, guldenk@sakarya.edu.tr

** Doç. Dr., Sakarya Üniversitesi, nguler@gmail.com

Anlamlı öğrenme ilk olarak 1968 yılında Ausubel tarafından ortaya atılmıştır. Ausubel anlamlı öğrenmeyi “kişinin sahip olduğu kavramlar ve önermeler ile yeni bilgileri ilişkilendirerek bilgiyi oluşturması ve kalıcı öğrenme” şeklinde tanımlamıştır. Anlamlı öğrenmenin gerçekleşmesi için kavramlar arası ilişkilerin kurulması gerekir. Bunun yanında Ausubel (1968) anlamlı öğrenmenin gerçekleşmesi için bir bütün içindeki kavramların belli bir düzen içinde olması ve kavramların sıralandırılarak öğrencilere verilmesi, yeni öğrenilen konunun kendi içinde tutarlı olması veya öğrencinin daha önceki bilgileriyle çelişmemesi, öğrencinin öğrendiği ilkeyi veya bilgiyi farklı durumlara ve problemlere uygulayabilmesi gerektiğini savunur.

Ausubel’in (1968) “Eğer tüm eğitim psikolojisini bir tek ilkeye indirgeyecek olursam: Anlamlı öğrenmeyi etkileyen en önemli yegâne etken öğrenenin hali hazırda ne bildiğidir. Bunun ne anlama geldiğini araştırın ve buna uygun şekilde öğretin.” şeklinde ifade ettiği ilkeye dayanarak Novak ve Gowin (1984) yaptıkları çalışmada öğrencilerin ne bildiklerini ve anlamalarının zaman içinde nasıl değiştiğini araştırırken bir yandan da bir araç geliştirmişlerdir. Geliştirdikleri bu araca önceleri ‘bilişsel haritalar’ daha sonra da ‘kavram haritaları’ adını vermişlerdir. Novak ve Gowin (1984) kavram haritalarını ‘bireylerin önceden edindikleri bilgilerle yeni öğrendikleri arasında köprü oluşturan, bireylerin zihinlerinde kavramları nasıl ilişkilendirdiğini gösteren, bilgiyi organize eden ve sunan araçlar’ olarak tanımlamaktadırlar.

Bilgiyi düzenlemek ve simgeleştirmek için geliştirilmiş olan kavram haritaları; geniş bir başlık altındaki kavramların birbirleriyle ilişkilerini gösteren iki boyutlu bir şemadır. Kutular ve bağlantı çizgileri içeren kavram haritalarında kavramlar kutulara yazılır. Kavramların birbirleri ile olan ilişkileri doğrultusunda bağlantıları çizgilerle belirtilir. İki kavram arasındaki ilişkinin yönüne göre bu çizgilerin ok yönü de belirlenir. Şekil 1’de kavram haritası bileşenleri gösterilmiştir.



Şekil 1. Kavram haritası bileşenleri

Kavram haritaları bilginin zihinde somut ve görsel olarak düzenlenmesini sağlar. Tüm bir öğretim yılında, bir ders için ya da tek bir ünite içinde kullanılan kavram haritaları kavramlar arası önemli ilişkileri şematize etmede etkili bir yoldur. Bunun yanında kavram haritalarının öğrencilerin düşünme, analiz etme, problem çözme gibi yaratıcı yeteneklerini geliştirerek kavramları daha iyi anlamasını sağlaması daha çok fen bilimleri eğitiminde bir öğretim tekniği olarak kullanılmasını sağlamıştır (Novak, Gowin ve Johansen, 1983). Kavram haritalarının teori ve pratiği birleştirmeyi sağlayan etkili bir yöntem olması da fen bilimleri eğitiminde sıklıkla kullanılmasının bir diğer gerekçesi olabilir. Çünkü fen bilimleri özellikle deney çalışmaları ile teori ile pratiği birleştiren bir bilim dalıdır. Ancak, fen bilimleri eğitimi dışında da kavramlar arası ilişkinin var olduğu matematik, sosyal bilgiler gibi farklı alanların öğrenimi ve ölçülmesinde de kavram haritaları uygulanmaktadır (Özsoy, 2004; Altıntaş ve Altıntaş, 2008).

Kavram haritaları yapılarına göre farklılık gösterir. Bu yapısal farklılıklar; kavramların hiyerarşik olup olmaması, kavramların öğrenciye verilir verilmemesi, bağlantı sözcüklerinin verilir verilmemesi ve verilen boşlukları doldurma açısından olabilir. Sıklıkla görülen

yapılardan hiyerarşik yapıdır. Ancak kavram haritasının hiyerarşik olup olmaması içerdiği konu ile ilişkilidir. Eğer konu hiyerarşik bir yapıya sahip değilse kavram haritasının da hiyerarşik bir yapıya sahip olmasına gerek yoktur. Kavram haritalarında görülen yapısal değişikliklerden biri de; kavramların öğrenciye verilmesi ya da öğrencinin kendisi tarafından belirlenmesi şeklindedir. Bu şekilde yapılandırılan kavram haritaları ‘oluşturma kavram haritası’ (construct map) olarak adlandırılır. Kavram haritaları kavramlar arasındaki ilişkilerin verilmesi ya da verilmemesine göre de değişiklik gösterebilir. Bunlar; bağlantı sözcükleri verilmeden yapılması istenen kavram haritaları, bir başka ifade ile ilişkiler yaratılarak oluşturulan kavram haritaları (created map) ve bağlantı sözcükleri verilerek yapılması istenen kavram haritaları, bir başka ifade ile ilişki kümesinden seçilerek oluşturulan kavram haritaları (selected map) olarak adlandırılır. Bir diğer yapısal değişiklik; öğrencilerin verilen kavramları, kavram haritasında bırakılmış olan boşlukları doldurmak için kullanması şeklinde olabilir. Bu şekilde yapılandırılan kavram haritaları ‘doldurma kavram haritası’ (fill-in map) olarak adlandırılır.

Doldurma kavram haritaları uzmanlarca çizilmiş kavram haritalarından kavram ya da bağlantı sözcüklerinin bazılarının ya da tümünün çıkarılmasıyla oluşturulur. Doldurma kavram haritaları öğrencilerin bu boşluklara dolduracakları kavramları seçmeleri ya da üretmelerine göre ikiye ayrılır. Öğrencilerden kavram haritalarındaki bu boşlukları üretmekle doldurmaları isteniyorsa bu tür kavram haritalarına üret ve doldur (create and fill-in) kavram haritaları denir. Öğrencilerden kavram haritalarındaki bu boşlukları verilen kavram ya da bağlantı sözcükleri kümesinde seçerek doldurmaları isteniyorsa bu tür kavram haritalarına da ‘seç ve doldur kavram haritaları’ (select and fill-in) adı verilir (Schau, Mattern, Weber, Minnick ve Witt, 1997). Doldurma kavram haritaları ölçme aracı olarak ilk kez Surber tarafından 1984’te kullanılmıştır. Naveh-Benjamin, Lin ve McKeachie (1995) doldurma yaklaşımını sınıflandırılmamış ilişkileri içeren hiyerarşik kavram haritalarında kullanmışlardır. Bu çalışmada üret ve doldur (create and fill-in) ile seç ve doldur (select and fill-in) türünde oluşturulması istenen iki farklı doldurma kavram haritası kullanılmıştır. Çalışmada, bu haritalar kısaca C ve S tekniği ile hazırlanmış doldurma kavram haritaları olarak isimlendirilmiştir.

Yapılarına göre farklılık gösteren kavram haritaları puanlama yöntemlerine göre de farklılaşır. Novak ve Gowin (1984), sayı ve anlamlılık olarak kavramlar arası bağlantılar, haritada kavramlar arası hiyerarşik yapının gösterilme derecesi, değişik kavram hiyerarşisi arasında bağlantının var oluşu ve uygun örneklerin sağlanması gibi birkaç kriter üzerinde puanlamayı önerir. Bu kriterler yalnızca hiyerarşik kavram haritalarında uygulanır. Ancak sadece bazı kavram haritaları hiyerarşiktir ve hiç hiyerarşik yapı yoksa bu puanlama dikkate alınmaz. Bir diğer puanlama yönteminde bağlanmış kavram çiftleri sayılır. Bu bağlar hiyerarşik, çoklu ya da çapraz olabilir. Puanlar bağların sayısına göre verilir (Shavelson, 1993). Diğer bir yöntemin odağında, kavram haritalarındaki önermeler yer alır. Bu yöntemde önermelerin üç parçası puanlanır. Bu parçalar; kavramlar arasındaki ilişki, etiket, kavramlar arasındaki hiyerarşik ya da nedensel ilişkileri ifade eden okun yönü şeklindedir. Bunun yanında kavram haritaları dereceli puanlama anahtarı (rubrik) kullanılarak da puanlanabilir. Doldurma (fill map) şeklinde hazırlanmış olan kavram haritaları doğru yanıtlara 2, kısmen doğru yanıtlara 1, boş ve yanlış yanıtlara 0 verilerek puanlanabilir. Kısmen doğru yanıtlar doldurma kavram haritası hazırlanacak olan konunun uzmanları tarafından belirlenir.

Literatürde yer alan kavram haritaları üzerine yapılan pekçok araştırmada, kavram haritalarının iç tutarlık güvenirliliği (Herl, O’Neil, Chung, ve Schacter, 1999; Lay-Dopyera ve Beyerbach, 1983; Lomask, Baron, Greig, ve Harrison, 1992, McClure, Sonak ve Suen, 1999; Nakhleh ve Krajcik, 1991); kararlılığı (Lay-Dopyera, ve Beyerbach, 1983); uyum geçerliliği (Anderson, ve Huang, 1989; Baker, Niemi, Novak, ve Herl, 1991; July, Markham, Mintzes, ve Jones, 1994; Novak, Gowin, ve Johansen, 1983; Rice, Ryan, ve Samson, 1998; Schreiber, ve Abegg, 1991); yordama geçerliliği (Acton, Johnson, ve Gollsmith, 1994); farklı puanlama yöntemlerinin denkliği (McClure ve diğerleri, 1999; Rice ve diğerleri, 1998) ve farklı kavram

haritalarının denklięi (Ruiz-Primo, Shavelson, Li, ve Schultz, 2001; Yin, Vanides, Ruiz-Primo, Ayala, ve Shavelson, 2005) konularına ışık tutulmuştur. Bu çalışmalar uygun olmayan seçenekleri elemeye yardımcı olan; farklı kavram haritaları konuları, yanıtlama formatı ve puanlama cetvelinin teknik özellikleri hakkında önemli bilgiler sunmaktadır. Buna rağmen kavram haritalarının değerlendirilmesinde, deęişkenlik kaynakları birden fazla olduęu için klasik test kuramı (KTK) tüm hata kaynaklarını bir arada ve aynı anda ele alan bir model sunmamaktadır.

Başarısı ölçlmek istenen adayın kavram haritasından elde ettięi başarı puanını tüm adayların her türlü deęişen durum altında -konu, cevaplama yöntemi, puanlama sistemi- kavram haritasından elde ettięi puanlar evreninden bir örnek olarak görecektir olursak kavram haritası değerlendirmesi genellenebilirlik kuramı çerçevesinde yapılabilir (Yin ve Shavelson, 2008) .

Genellenebilirlik (G) kuramı ölçme sonuçlarının güvenilirliğinin belirlenmesini, güvenilir gözlemlerin tasarımı, araştırılmasını ve kavramsallaştırılmasını sağlayan istatistiksel bir kuramdır. Bu kuram psikoloji ve eğitim gibi sosyal bilimler için geniş kavramsal çerçeve ve sayısız ölçme durumu için oldukça güçlü bir istatistiksel yol olan varyans analizini kullanarak kapsamlı bir yapı oluşturmaktadır (Brennan,1992). Genellenebilirlik kuramı, hem klasik test kuramının hem de varyans analizinin bir uzantısı olarak, çoklu hata kaynaklarının aynı anda ele alınabildięi bir model olarak karşımıza çıkmaktadır (Gler, 2008).

G kuramı KTK'nı geçersiz kılmamakla birlikte karşılaştırıldıklarında aralarında bazı farklılıklar bulunmaktadır (Shavelson ve Webb, 1991). G kuramı, KTK'dan öncelikle hata kaynaklarını ele alışı bakımından ayrılır (Cronbach, Gleser, Nanda, ve Rajaratnam,1972). Hata terimini, birçok deęişkenlik kaynağına ayırtırmakla birlikte bu hata kaynaklarının (puanlayıcı, madde, zaman vb.) ayrı ayrı ve etkileşimlerinin birlikte değerlendirilmesine olanak tanır (Shavelson ve Webb 1991). Böylece G kuramıyla KTK'da olduęu gibi farklı hata kaynaklarına göre farklı güvenilirlik katsayıları hesaplamak yerine tek bir güvenilirlik katsayısı hesaplanabilir. Bir başka deyişle, klasik test kuramına göre güvenilirlik hangi amaç için kullanılacağı göz önüne alınarak farklı hesaplanma yöntemleri seçilirken (Baykul, 2000); genellenebilirlik kuramı, güvenilirlik yöntemlerinin tümünü içine almaktadır (Eason, 1989). Bunun yanı sıra genellenebilirlik kuramında klasik test kuramından farklı olarak, görel ve mutlak değerlendirmeler arasında bir ayırım söz konusu olup, güvenilirlik hesaplanırken hata varyansları arasında görel ve mutlak hatalara göre farklılık bulunmaktadır (Brennan, 2001). Görel değerlendirilmede G-katsayısı kullanılır ve gerçek varyansın, gerçek varyans ile görel varyansın toplamına bölünmesiyle hesaplanır. Güvenirlik indeksi (Phi-katsayısı) ise mutlak değerlendirmeler için kullanılmaktadır. Güvenirlik indeksi gerçek varyansın, gerçek varyans ile mutlak varyansın toplamına bölünmesiyle elde edilir. Dięer bir deyişle, bu iki katsayı hatanın nasıl ele alındığına baęlı olarak deęişmektedir (Brennan,1992; Gler, Uyanık ve Teker, 2012).

G Kuramında, güvenilirliğin araştırılmasında iki aşamadan söz etmek mümkündür. Bunlardan ilki genellenebilirlik çalışması (G çalışması) ve ikincisi karar çalışmasıdır (K çalışması). G kuramının istenilen düzeyde güvenilirlik elde etmek için karar çalışması (K çalışması) sürecine sahip olması, G kuramını klasik test kuramından ayıran başka bir özelliktir. G kuramı ile K çalışması yapılarak sonraki çalışmalar için oluşturulan senaryolarla gelecek çalışmalar için bilgiler elde edilebilmektedir. (Goodwin, 2001). G çalışmasında araştırmacı, üzerinde ölçme yaptığı örnekleme ölçmenin evrenine genellemekle ilgilenir. K çalışması ise, G çalışmasından elde edilen bilgiyi belli bir amaç doğrultusunda sosyal bilimlerdeki ölçmelerde mümkün olan en iyi deseni oluşturmak için kullanır (Shavelson ve Webb, 1991).

G kuramında öğrencinin bir görevde gösterdiği performansı; puanlayıcı, görev, zaman gibi olası tüm deęişkenlik kaynaklarının bir arada bulunduęu karmaşık bir evrendeki performansından çekilen bir örnekleme olarak görülür. Bu nedenle G kuramı, performans ölçmlerinin güvenilirliğinin kestirilmesine esnek bir yaklaşım sunmaktadır. Aynı zamanda G

kuramı, güvenilirlik ile geçerlik arasındaki farkın, güvenilir ölçmeler düzenleyerek nasıl ortadan kaldırılabileceğini ortaya koymaktadır. Genel anlamıyla yapı geçerliği, ölçülmek istenilen bir yapının; bir başka deyişle bireylerde var olduğu kabul edilen bir özelliğin, ölçme sonucunda ortaya konulma derecesi olarak yorumlanabilir (Baykul, 2000). G kuramında yer alan “evren” kavramı; tüm gözlem koşulları ve değişkenlik kaynaklarını kapsamaktadır ki bu da klasik kuramdaki geçerlikte yer alan “yapı” kavramını tanımlamaktadır. G kuramı, bu örtük yapıya (kabul edilebilir gözlemlerin evreni) ilişkin kestirimlerin doğru olarak elde edilmesini sağlarsa, güvenilir sonuçlara ulaşılabileceğini ifade eder. Böylelikle G kuramı, güvenilirlik ile geçerlik arasındaki geleneksel ayrımı ortadan kaldırır (Brennan, 2001).

G kuramına göre evrende bir ya da birden çok varyans kaynağı bulunmaktadır. Bu varyans kaynakları genellenebilirlik kuramında yüzey (facet) olarak adlandırılır (Brennan 2001). Örneğin; beş maddeden oluşan açık uçlu bir matematik sınavında her bir öğrencinin cevapları farklı iki puanlayıcı tarafından puanlanıyor olsun. Bu ölçme sürecinde öğrenciler *ölçme nesnesi*, maddeler ve puanlayıcılar ise birer değişkenlik kaynağı olup yüzeyleri oluşturmaktadır. Ölçme nesnesine (öğrencilere) bağlı varyans istenilen bir durum olduğu için bireylerin doğası gereği var olan farklılıklar klasik test kuramında “hata kaynağı” olarak tanımlanmadığı gibi G kuramında da yüzey olarak düşünülmez. Her bir yüzeyin farklı düzeylerine koşul (condition) adı verilir. Yukarıdaki örnekte, madde yüzeyi beş, puanlayıcı yüzeyi iki koşuldan oluşmaktadır.

G kuramına göre yüzeyler çapraz (crossed) ya da yuvalanmış (nested) şekilde olabilir (Rentz 1987). Bir yüzeyin tüm koşullarının başka bir yüzeyin her bir koşuluyla örtüştüğü duruma çapraz (crossed) desen adı verilmektedir. Çapraz olarak tasarlanmış bir desende yüzeyler arasında “x” işareti kullanılır. Örneğin; matematik sınavı örneğinde 10 öğrencinin yer aldığı bu ölçme sürecinde, her öğrenci (o) aynı beş maddenin (m) tümünü cevaplamış ve her öğrencinin her beş maddesi aynı iki puanlayıcı (p) tarafından puanlanmış ise buradaki desen *çaprazlanmış* desen olacaktır ve bu desen $oxm \times p$ ile gösterilir. Kimi zaman yüzeylerin diğer yüzeyler ile örtüşmediği (çaprazlanmadığı) durumlar da söz konusudur. Bir yüzeyin koşulları diğer yüzeyin her bir koşulu ile örtüşmüyorsa, yüzeyler yuvalanmış (nested) desen oluşturmaktadır. Yuvalanmış olarak tasarlanmış bir desende yüzeyler arasında “:” işareti kullanılır. Örneğin; matematik sınavı örneğinde; 10 öğrencinin yer aldığı bir ölçme sürecinde her bir öğrenci (ö) farklı beş madde (m) cevaplıyor ve her bir öğrencinin her maddesi farklı iki puanlayıcı (p) tarafından puanlanıyor olsun (ki bu durumda $10 \times 5 = 50$ farklı madde ve $10 \times 5 \times 2 = 100$ farklı puanlayıcı gerekmektedir). Bu ölçme sürecindeki desen *tümüyle yuvalanmış* desen olacaktır ve bu desen $ö:m:p$ ile gösterilir. Aynı örnek üzerinden; eğer öğrencilerin her biri aynı 5 maddeyi cevaplıyor ancak tüm öğrencilerin her bir maddesi farklı iki puanlayıcı tarafından puanlanıyor ise (ki bu durumda 5 farklı madde ve $10 \times 2 = 20$ farklı puanlayıcı gerekmektedir) bu ölçme sürecindeki desen *yuvalanmış* desen olup, $öx(m:p)$ ile ifade edilir. Örneklerden de anlaşılacağı üzere *tümüyle yuvalanmış* desen pratikte çok da uygulanabilir bir durum olmayıp, ölçme süreçleri daha çok *çapraz* ve *yuvalanmış* desenler içermektedir. (Güler, Uyanık, ve Teker, 2012; Shavelson ve Webb 1991).

Benzer şekilde G kuramında yüzeyler rastgele ya da sabit olabilir. Bir yüzeyin rastgele olarak mı yoksa sabit olarak mı ele alınacağı tamamıyla araştırmacının kararına bağlıdır. Eğer araştırmacı, örneklemin ötesinde genelleme evrenine genelleme yapmak istiyorsa, bu durumda yüzey rastgele olarak ele alınacaktır. Diğer taraftan, eğer araştırmacı örneklemin ötesinde bir genellemeye gitmek istemiyorsa, o zamanda yüzey sabit olarak düşünülecektir. Yüzeyin rastgele ya da sabit olarak ele alınması güvenilirliğin kestirimini etkilemektedir. G kuramı temel olarak rastgele yüzeyli bir ölçme kuramıdır ve en azından bir yüzeyin rastgele olması gerekmektedir. G kuramında ya tüm yüzeyler rastgeledir ya da rastgele ve sabit yüzeyler birlikte bulunmaktadır. Bir başka deyişle, tüm yüzeylerin sabit olması mümkün değildir. Bu sebeple en azından bir sabit yüzeyi bulunan desenler *karışık (mixed) desenler* olarak tanımlanmaktadır (Shavelson ve Webb,1991; Brennan,1992). Bu çalışmada vurgulanmak istenilen bir başka nokta, G kuramında

yer alan *karıık* desenin apraz ve yuvalanmı yzeylerin birlikte yer aldıđı “*yuvalanmı desenler*” iin kullanılıyor olmasıdır. Ancak bu kullanım hatalı olup, “*karıık desen*” kavramının dođru kullanımı sabit ve rastgele yzeylerin birlikte yer aldıđı desenler iin olmalıdır (Gler, Uyanık, ve Teker, 2012; Brennan,1992).

Bu alıma ile lme ve Deđerlendirme dersinin “lme ve deđerlendirmeyle ilgili temel kavramlar” konusunu len iki farklı doldurma kavram haritası tekniđinin genellenebilirlik kuramı ile elde edilen parametreleri aratırılmıtır. alımada o (đrenci) x m (madde) x h (harita) aprazlanmı ve yzeylerden madde yzeyi rastgele, harita yzeyi ise sabit olduđu iin karıık desen kullanılmıtır. G alıması yapılarak C ve S tekniđi ile yapılandırılmı doldurma kavram haritaları uygulamalarından elde edilen sonuların varyansları ve toplam varyansları aıklama yzdeleri, K alıması ile madde sayılarının arttırılması ve azaltılmasıyla G ve Phi katsayılarındaki deđiim aratırılmıtır.

Bu aratırmayla, klasik test yntemiyle mmkn olmayan; kavram haritası puanlarının ve puanlama gvenirliđinin genellenebilirlik kuramı ile tek bir analiz yaparak mmkn olabileceđi gsterilmeye alıılmıtır. alımanın temel amacı kavram haritalarından elde edilen lme sonularının gvenirliđinde, genellenebilirlik kuramının klasik test kuramına gre daha tercih edilir olabileceđinin rneklendirilmesidir. Ayrıca doldurma kavram haritası trlerinin karılatırılması ve bu dođrultuda uygulamada hangisinin tercih edilebilir olduđunun belirlenmesi de alımanın diđer bir amacıdır. Diđer yandan bu alıma, ierdiđi *karıık* desen rneđi ile literatrde yer alan genellenebilirlik alımalarından farklılamıtır. Karıık desen kavramının daha anlaılır ve ilerideki alımalara rnek olması aısından nemli olduđu dnlmektedir.

2. YNTEM

Bu alımada “lme ve deđerlendirmeye ilikin temel kavramlar” konulu iki farklı kavram haritası tekniđinin genellenebilirlik kuramı kullanarak parametrelerinin elde edilmesi incelendiđinden bu ynyle alıma betimsel bir aratırma niteliđindedir.

2.1. alıma Grubu

Bu alımada, 2012-2013 đretim yıllarında Sakarya niversitesi Psikolojik Danımanlık ve Rehberlik Blm 3. sınıfta đrenim gren 49’u (%53,8) kız, 42’si (%46,2) erkek toplam 91 đrenci yer almıtır. đrencilerin tm đretim yntem ve teknikleri dersi kapsamında kavram haritası hakkında bilgilendirilmı ve kavram haritası hazırlamada deneyim edindirilmılerdir.

2.2. Veri Toplama Aracı

alımada veri toplama aracı olarak kullanılan C ve S tekniđi ile hazırlanmı doldurma kavram haritaları lme ve deđerlendirme dersinin ilk konusu olan “lme ve deđerlendirmeyle ilgili temel kavramlar” konusunu kapsamaktadır. Haritalar iki lme ve deđerlendirme uzmanı ve đretim yntem ve teknikleri dersinin sorumlusu olan đretim yesinden oluan bir grup tarafından oluturulmutur. Hazırlanan kavram haritaları daha sonra iki farklı lme ve deđerlendirme uzmanı tarafından incelenmitir. Uzmanlar kavram haritalarını; verilen kavramlar ve ilikilerin birbirlerine gre uygunluđu, haritanın genelinin đrencilerin sınıf dzeyine uygunluđu ve puanlama sisteminin objektifliđi aısından incelemi ve uygun bulmutur. Kullanılan kavram haritaları ek-1’de verilmitir.

Bu alımada kullanılan doldurma kavram haritaları dođru yanıtla 2, kısmen dođru yanıtla 1 ve bo ve/veya yanlış yanıtla 0 verilerek puanlanmıtır. İki kavram arasındaki ilikinin var olduđunu bilen ancak aradaki bađlantı szcđn yazamamı olan đrencinin bu cevabı kısmen dođru kabul edilmitir. Puanlamanın oktan semeli sınavlarda olduđu gibi objektif olmasından dolayı sadece bir puanlayıcı kullanılmıtır.

Çalışmanın problemi olan $oxmxh$ çaprazlanmış karışık deseninden elde edilen parametrelerin ne olacağını yanıtlamak için gerekli olan veriler biri S tekniği diğeri C tekniği ile hazırlanmış olmak üzere iki farklı doldurma kavram haritası ile toplanmıştır. Uygulamalar için 91 kişiden oluşan gruba önce C tekniği ile hazırlanmış doldurma kavram haritası daha sonra S tekniği ile hazırlanmış doldurma kavram haritası uygulanmıştır. Uygulamalar arasında iki hafta ara verilmiştir. Önce C tekniği ile hazırlanmış doldurma kavram haritası uygulayarak ve uygulamalar arasında iki hafta ara vererek hatırlama etkisinin ortadan kalkması sağlanmaya çalışılmıştır.

Bu çalışmada yer alan harita yüzeyi (C ve S olmak üzere iki koşula sahip) sabittir. İstatistiksel olarak G kuramı, sabit yüzeyin koşullarının ortalamasını alarak sabit yüzey ile ilgili işlem yapmaktadır. Yüzeyin koşulları üzerinden ortalama almak kuramsal bir anlam taşımıyorsa veya böyle bir ortalama ile ilgilenilmiyorsa, sabit yüzeyin her bir koşulu için ayrı ayrı G çalışması yürütülebilir. Hangi yaklaşımın kullanılacağı üzerine seçim yaparken bunun kuramsal temellere dayandırılması gerekmektedir. Araştırmacı ne tür bilgiye ihtiyaç duyduğunu belirlemeli ve analizi bu doğrultuda yürütmelidir. Eğer araştırmacı, sabit yüzeyin koşulları üzerinden ortalama almayı seçmişse, sabit yüzey ile ilişkili varyans bileşenlerini de incelemelidir. Elde edilen büyük varyans değerleri, sabit yüzeyin koşullarının ihmal edilemeyecek ölçüde farklılaştığını belirtir. Bu durumda da her bir koşulun ayrı ayrı analiz edilmesi daha anlamlı olacaktır. Ancak bu çalışmada yer alan sabit yüzeyin ortalamasını almak anlamlı olduğu gibi sabit yüzey ile ilişkili varyans bileşenlerini de incelendiğinde bu varyans değerlerinin de küçük olduğu görülmüştür. Bu nedenle bu çalışmada sabit yüzeyin ortalaması alınarak işlemler gerçekleştirilmiştir.

2.3. Verilerin Analizi

Uygulamalar sonucunda genellenebilirlik kuramına göre ana ve ortak etkiler için varyans değerlerinin kestirilmesinde, puanların güvenirliliğinin G ve Φ (Phi) katsayılarıyla hesaplanmasında ve K çalışmalarında EduG programı kullanılmıştır.

3. BULGULAR

C ve S tekniği ile hazırlanmış 10 maddeli kavram haritalarına 91 öğrencinin vermiş olduğu cevaplara göre elde edilen toplam puanların betimsel istatistiği Tablo 1’de verilmiştir.

Tablo 1: Farklı teknikle hazırlanmış kavram haritalarında yer alan 10 maddeye öğrencilerin verdikleri cevaplardan elde edilen toplam puanların betimsel istatistikleri

İstatistikler	S Tekniği	C Tekniği
N	91	91
Ortalama	15,53	12,78
Ortanca	16,00	14,00
Tepe değer	20	14
Std. Sapma	3,987	4,944
Çarpıklık	-1,121	-,197
Basıklık	-1,273	-1,091

Tablo 1’de görüldüğü gibi 91 öğrencinin 10 maddeden 0-2 puan ölçeği üzerinden aldıkları toplam puanlara ilişkin ortalamalar S uygulaması için 15,53; C uygulaması için ise 12,78’dir. Her iki harita tekniğine ilişkin ortanca aritmetik ortalamadan büyüktür ve harita tekniğine ait puanlar sola çarpık dağılım göstermektedir. Bu durum, her iki harita tekniği için çarpıklık katsayısının negatif çıkmasıyla da görülebilmektedir. Basıklık katsayılarına bakıldığında, her iki

uygulamadan elde edilen basıklık katsayısı 0'dan küçük bir değer olup puanlar normalden daha basık dağılım göstermektedirler. Genellenebilirlik kuramında normallik ve puan etkilerinin bağımsızlığı varsayımlarının belirlenmesi zorunluluğu yoktur (Brennan, 1992). Bu nedenle elde edilen betimsel istatistiklerin normalden farklı dağılması istatistiksel olarak bir engel oluşturmamaktadır.

Farklı tekniklerle elde edilmiş 10 maddelik kavram haritası uygulamasından elde edilen oxmxh çaprazlanmış karışık desene ilişkin varyans değerlerini hesaplamak için aşağıdaki adımlar yürütülmüştür (Shavelson ve Webb, 1991; Güler, Uyanık ve Teker, 2012).

Adım 1: Tüm değişkenlik kaynakları rastgele olarak ele alınarak varyans analizi yürütülüp varyans bileşenleri kestirilir. oxmxh deseninde öğrenciler (*o*), maddeler (*m*) ve harita teknikleri (*h*) ile çaprazlanmışlardır. Öğrenciler ve maddeler rastgele, harita teknikleri ise sabittir. Ancak, bu adım için harita teknikleri de dâhil bütün değişkenlik kaynakları rastgele olarak ele alınır. Tamamıyla rastgele oxmxh deseni için kestirilen varyans bileşenleri $\sigma_o^2, \sigma_m^2, \sigma_h^2, \sigma_{om}^2, \sigma_{oh}^2, \sigma_{mh}^2, \sigma_{omh,e}^2$ 'dir. Kestirilen bu varyans bileşenlerine ilişkin değerler Tablo 2'de verilmiştir.

Adım 2: Karışık (mixed) desenin rastgele kısmı ile hesaplanacak ortak varyans bileşenleri belirlenir. oxmxh deseninde, öğrenciler ve maddeler rastgele olup desenin rastgele kısmını öğrencilerin maddeler ile çaprazlanması oluşturur. Bu durum için öğrenciler (*o*), maddeler (*m*) ve bunların etkileşimleri ile hata (*om,e*) hesaplamaları yapılır. Bu varyans bileşenlerini tamamıyla rastgele desendeki bileşenlerden ayırabilmek adına $\sigma_{o*}^2, \sigma_{m*}^2, \sigma_{om,e*}^2$ şeklinde gösterilir.

Tablo 2: Oxmxh desenine ilişkin g çalışması ile kestirilen varyans değerleri ve toplam varyansı açıklama oranları (tüm yüzeyler rastgele)

Varyans Kaynağı	Sd	Kareler Toplamı	Kareler Ortalaması	Varyans	%
o	90	186,869	2,076	0,073	11,1
m	9	158,903	17,656	0,089	14
h	1	9,573	9,574	0,009	0,7
om	810	471,196	0,582	0,195	43,3
oh	90	21,326	0,237	0,004	0,7
mh	9	10,152	1,128	0,010	1,5
omh,e	810	155,948	0,193	0,193	28,7
Toplam	1819	1013.969			100%

Adım 3: Belirlenen karışık desenin rastgele kısmına ait varyans bileşenleri aşağıda verilen eşitlikler yardımıyla hesaplanır.

$$\sigma_{o*}^2 = \sigma_o^2 + \frac{1}{n_h} \sigma_{oh}^2 \quad (1) \quad \sigma_{p*}^2 = \sigma_p^2 + \frac{1}{n_h} \sigma_{ph}^2 \quad (2)$$

$$\sigma_{op,e*}^2 = \sigma_{op}^2 + \frac{1}{n_h} \sigma_{oph,e}^2 \quad (3)$$

Karışık deseninin varyans bileşenlerinin hesaplanması için gerekli olan değerler yukarıda verilen denklemlerde yerine konularak elle hesaplanmıştır. Dikkat edilirse, eşitliklerin sağ tarafında sadece sabit yüzey ile etkileşimi içeren varyans bileşeni yer almakta, rastgele yüzey

etkileşimi bulunmamaktadır. Örneğin; öğrenciler için kestirilen varyans bileşeninde öğrencilerin maddeler ile etkileşimini içeren σ_{om}^2 varyans bileşeni yer almaz. Elde edilen sonuçlar Tablo 3'te yer almaktadır.

Tablo 3: Omxh desenine ilişkin g çalışması ile kestirilen varyans değerleri ve toplam varyansı açıklama oranları (harita tekniği yüzeyi sabit)

Varyans Kaynağı	Sd	Kareler Toplamı	Kareler Ortalaması	Varyans	%
o*	90	186,869	2,076	0,075	11,1
m*	9	158,903	17,656	0,094	14
h	1	9,573	9,574	0,009	1
om*	810	471,196	0,582	0,291	43
oh	90	21,326	0,237	0,004	0,6
mh	9	10,152	1,128	0,010	0,2
omh,e	810	155,948	0,193	0,193	29
Toplam	1819	1013,969			

*Sabit yüzeyin varlığı sebebiyle varyans değerleri hesaplanan varyans kaynakları

Tablo 3'te verilen omxh çaprazlanmış karışık deseni G çalışması sonucunda kestirilen varyans ve toplam varyansı açıklama oranları incelendiğinde, öğrenci (o) ana etkisi için kestirilen varyans bileşeninin (0,075) toplam varyansın %11'ini açıkladığı görülmektedir. Genellenebilirlik çalışmalarında, öğrenci ana etkisi evren puanı varyansı olarak değerlendirilir ve ölçülen özellik açısından öğrenciler arası farklılaşmayı ifade eder (Shavelson ve Webb, 1991; Brennan 2001). Öğrenciler için kestirilen varyansın toplam varyans içindeki oranının büyük olması istenilen bir durumdur. Bu durum, ölçme ile elde edilen boyutta öğrenciler arası farklılıkların ortaya çıkarılabildiğinin bir göstergesidir (Güler, 2008). Bu çalışmada elde edilen sonuca göre, kavram haritaları ile yapılan ölçme sürecinin öğrenciler arasındaki farklılığı çok fazla açığa çıkaramadığı söylenebilir. Madde (m) ana etkisi için yapılan G çalışmasından kestirilen varyans bileşeni (0,094) toplam varyansın %14'ünü açıklamaktadır. Madde ana etkisi kavram haritaları içerisinde yer alan her bir ölçme biriminin (madde) güçlük düzeyinin farklılaşma derecesini gösterir. Elde edilen sonuca göre, öğrencilere uygulanan kavram haritalarında farklı güçlük düzeylerinde maddelerin yer aldığı yorumu yapılabilir. Harita tekniği (h) ana etkisi için kestirilen varyans bileşeni (0,009) toplam varyansın %1'ini açıklayarak, ana etkiler içinde en düşük varyans bileşeni olmaktadır. Harita tekniği ana etkisinin G çalışması ile kestirilen varyans oranının düşük çıkması tekniklerin arasında bir tutarlılığın söz konusu olabileceği diğer bir ifade ile harita teknikleri arasında önemli bir farklılaşmanın olmadığı şeklinde yorumlanabilir.

Öğrenci x madde (om) ortak etkisi (0,291) toplam varyansın % 43'ünü açıklamaktadır ve elde edilen en büyük varyanstır. Bu durum belli öğrencilerin bağıl durumlarının bir maddeden diğerine çok farklılaştığını göstermektedir. Benzer şekilde bu durum için kavramların anlaşılma ve doğru olarak cevaplandırma durumlarının öğrenciden öğrenciye farklılık gösterdiği söylenebilir. Öğrenci x harita tekniği (oh) ortak etkisi (0,004) toplam varyansın % 0,6'sını açıklamaktadır. Bu değer elde edilen en küçük ikinci varyanstır. Bu sonuç için harita tekniğine göre öğrencilerin ölçme sonuçları arasında bir farklılık oluşmadığı yorumu yapılabilir. Madde x harita tekniği (mh) ortak etkisi (0,010) toplam varyansın %0,2'sini açıklamaktadır. Bu varyans değeri, toplam varyans değeri açıklamada en küçük orana sahiptir. Bu sonuca göre, madde x harita tekniği ortak etkisinden kaynaklanan farklılığın neredeyse hiç olmadığı yorumu yapılabilir. C ve S tekniği ile yapılandırılmış kavram haritalarında madde olarak tanımlanan kavramların iki harita tekniğinde de aynı olmasından dolayı madde x harita ortak etkisinin varyans değerinin düşük olması beklenen bir durumdur. Öğrenci x madde x harita tekniği (artık) ortak etkisi varyans bileşeninin (0,193) toplam varyansı açıklama oranı %23 olarak görülmektedir. Bu oran tablodaki ikinci en büyük değerdir. Öğrenci x madde x harita tekniği

(artık) varyansın büyük olması; öğrenci, madde, harita tekniği etkileşiminin, ölçülemeyen değişkenlik kaynaklarının ve/veya tesadüfi hataların büyük olduğunun bir göstergesidir.

G çalışmasında kullanılan veriler üzerinden hesaplanan varyans değerleri K çalışmasında alınacak kararlar için kullanılır. K çalışması, G evreninde yer alan yüzeylerin koşullarının azaltılıp artırılmasıyla göreceli ve mutlak kararlar için sırasıyla G ve Phi katsayılarının kestirilmesine izin verir. Aynı zamanda G çalışmasında yer alan yüzey koşullarıyla aynı sayıda koşul içeren yüzeyler üzerinden G ve Phi katsayıları hesaplanırsa G çalışması verilerinin de güvenilirlik değerleri elde edilmiş olur. G ve Phi katsayılarının nasıl hesaplandığı sırasıyla, aşağıda verilen 1 ve 2 eşitlikleri ile gösterilmiştir. Tablo 4'te, yapılan K çalışmasında sabit yüzey olan harita tekniği yüzeyinin sabit tutulup (Tablo 3'te yer alan değerler üzerinden) madde sayısının beş azaltılıp beş artırılmasıyla hesaplanan G ve Phi katsayıları değerleri verilmiştir.

$$G \text{ katsayısı} = \frac{\sigma_{o^*}^2}{\sigma_{o^*}^2 + \frac{\sigma_{m^*}^2}{n_m} + \frac{\sigma_{h^*}^2}{n_h} + \frac{\sigma_{mh}^2}{n_m n_h}} \quad (1)$$

$$\Phi \text{ katsayısı} = \frac{\sigma_{o^*}^2}{\sigma_{o^*}^2 + \frac{\sigma_{m^*}^2}{n_m} + \frac{\sigma_{h^*}^2}{n_h} + \frac{\sigma_{m^*}^2}{n_m} + \frac{\sigma_{h^*}^2}{n_h} + \frac{\sigma_{mh}^2}{n_m n_h} + \frac{\sigma_{mh}^2}{n_m n_h}} \quad (2)$$

Tablo 4: Omxh çaprazlanmış karışık desen k çalışması sonuçları

Harita Tekniği	2	2	2	2	2
Madde Sayısı	5	10	15	20	25
G Katsayısı	0,49	0,65	0,73	0,78	0,81
Φ Katsayısı	0,47	0,62	0,70	0,74	0,77

Tablo 4'te görüldüğü gibi, oxmh deseninde yer alan harita yüzeyi sabit tutularak çalışmada yer alan 10 madde sayısına göre elde edilen G ve Phi katsayısı sırasıyla, 0,65 ve 0,62 olarak hesaplanmıştır. Göreceli ya da mutlak değerlendirmeler için çalışmada elde edilen bu değerler yeterli sınırı olarak kabul edilen 0,70 değerinin altında kaldığı için yeterince yüksek değildir. Literatürde yer alan kavram haritaları çalışmalarına bakıldığında genel olarak kavram haritaları için elde edilen G ve Phi katsayıları düşük değerdedir (Yin ve Shavelson, 2008). Tablo 3'te gözlenen öğrencilere ilişkin varyans değerinin oranının düşük olması kestirilen G ve Phi katsayılarının neden beklenen düzeyde çıkmadığının bir göstergesi olup, bu değerlerin düşüklüğü şaşırtıcı bir sonuç olmamaktadır. Ayrıca K çalışması sonuçlarına göre de madde sayısının yeterli düzeyde olmadığı, madde sayısı arttırıldıkça güvenilirlik değerlerinin de yükseldiği açıkça gözlenmektedir.

4. TARTIŞMA, SONUÇ VE ÖNERİLER

Bu çalışma; 91 öğrenci, 10 kavramdan oluşan üret ve doldur (create and fill-in) ve seç doldur (select and fill-in) türünde iki farklı doldurma kavram haritasını kapsayan ölçme sürecinden elde edilen veriler ile gerçekleştirilmiştir. Bu veriler üzerinden, öncelikle yüzeylerin tümü rastgele kabul edilerek oxmh çaprazlanmış desenin varyans değerleri belirlenmiş, daha sonra bu değerler üzerinden harita yüzeyi sabit kabul edilerek oluşturulacak varyans değerleri ayrıca elle hesaplanmış, hesaplanan bu yeni değerler üzerinden K çalışması yapılmış ve G ve Φ katsayıları elde edilmiştir.

Genellenebilirlik çalışmalarında, ölçülen özellik açısından bireyler arası farklılaşmayı ifade eden öğrenci ana etkisi, oxmh çaprazlanmış *karışık* deseni G çalışması sonucunda kestirilen varyans ve toplam varyansı açıklama oranları incelendiğinde toplam varyansın

%11'ini açıkladığı görülmektedir. Öğrenciler için kestirilen varyansın toplam varyans içindeki oranının yeterince büyük olmaması C ve S tekniği ile hazırlanmış doldurma kavram haritalarının öğrenciler arası farklılıkları çok fazla ortaya çıkaramadığını göstermektedir. Madde (m) ana etkisi için yapılan G çalışmasından kestirilen varyans bileşeni toplam varyansın %14'ünü açıklamaktadır. Madde ana etkisinin aldığı bu değer kavram haritaları içerisinde yer alan her bir kavramın farklı güçlük düzeylerinde olduğunu gösterir. Bu durumun ölçme ve değerlendirme ile ilgili temel kavramlar konusunun 10 temel kavramını içeren kavram haritalarında özellikle bazı kavramların (ölçüm, ölçüt, ölçek ve türleri) öğrencilere karmaşık gelmesinden kaynaklandığı düşünülmektedir. Harita tekniği ana etkisi için kestirilen varyans bileşeni ise toplam varyansın %1'ini açıklayarak, ana etkiler içinde en düşük varyans bileşeni olmaktadır. Harita tekniği ana etkisinin G çalışması ile kestirilen varyans oranının düşük çıkması iki farklı harita tekniği arasında bir tutarlılığın söz konusu olduğunu diğer bir ifade ile harita teknikleri arasında önemli bir farklılaşmanın olmadığını göstermektedir.

Çalışmanın sonucunda oxmxh çaprazlanmış karışık deseninin en büyük varyans kaynağı öğrenci x madde (om) ortak etkisi (%43) olarak bulunmuştur. Bu durum, öğrencilerin verdikleri doğru cevapların bir maddeden diğerine çok farklılaştığını göstermektedir. Diğer bir deyişle, bu durum için kavramların anlaşılma ve doğru olarak cevaplandırma durumlarının öğrenciden öğrenciye farklılık gösterdiği yorumu yapılabilir. Öğrencilerin kavram haritası konusu olan 'ölçme ve değerlendirmeye ilgili temel kavramlar' konusu ile ilgili ön bilgilerinin eşit olmaması bu varyans kaynağının büyük olmasını açıklayabilecek nedenler arasında gösterilebilir. Öğrenci x harita tekniği (oh) ortak etkisi toplam varyansın % 0,6'sını açıklamaktadır. Bu değer elde edilen en küçük ikinci varyanstır. Bu sonuca göre, belli öğrencilerin doğru cevap verme düzeylerinin bir teknikten diğerine çok fazla farklılaşmadığı yorumu yapılabilir. Bu durumda iki farklı kavram haritası türünün kavramlar arasındaki ilişkiyi bilen ve bilmeyen öğrenciyi ayırt etmede farklılık oluşturmadığı söylenebilir. Madde x harita tekniği (mh) ortak etkisi toplam varyansın %0,2'sini açıklamaktadır. En düşük orana sahip olan bu varyans değeri C ve S tekniği ile hazırlanmış haritalarda kullanılan kavramların birebir aynı olmasından kaynaklanıyor olabilir. Ancak, aynı kavramların kullanıldığı haritalarda düşükte olsa bir varyansın olması ise S tekniğinde kavramların öğrenciye verilirken C tekniğinde verilmemesinin bir sonucudur. Öğrenci x madde x harita tekniği (artık) ortak etkisi varyans bileşeninin toplam varyansı açıklama oranı %29'dur. Bu oran tablodaki ikinci en büyük değerdir. Öğrenci x madde x harita tekniği (artık) varyansın büyük olması; öğrenci, madde, harita tekniği etkileşiminin, ölçülemeyen değişkenlik kaynaklarının varlığının ve/veya tesadüfi hataların büyük olduğunun bir göstergesidir.

Çalışmada yer alan oxmxh çaprazlanmış karışık desenine ait G ve Phi katsayıları sırasıyla, 0,65 ve 0,62 olarak kestirilmiştir. Harita tekniği yüzeyinin sabit olduğu K çalışmasıyla, madde sayısının beş azaltılıp arttırılması senaryolarına göre elde edilen G ve Φ katsayılarında artış görülmüştür. Bu sebeple emek ve zaman açısından ekonomik olma durumu da göz önünde bulundurularak, istenilen güvenilirlik düzeyini elde edebilmek için daha fazla sayıda madde kullanılması önerilebilir.

Çalışmanın sonuçları göz önüne alındığında, kavram haritası gibi farklı değişkenlik kaynaklarının bir arada bulunduğu ölçme sonuçlarının güvenilirliğinin belirlenmesinde, tek bir analizle aynı anda ayrıntılı ve açıklayıcı sonuçlar sunduğundan genellenebilirlik kuramına dayalı G ve K çalışmalarının yapılması önerilebilir. Bu çalışma sadece ölçme ve değerlendirme dersi 'ölçme ve değerlendirmeye ilişkin temel kavramlar' konusu düşünülerek gerçekleştirilmiştir. Daha genelleştirici sonuçlara ulaşabilmek amacıyla, farklı derslerde ve farklı konularda doldurma kavram haritalarının güvenilir ve geçerli sonuçlar verip vermediği araştırılabilir.

Bu çalışmada, doldurma kavram haritasının iki farklı tekniği üzerinden elde edilen sonuçların güvenilirliği genellenebilirlik kuramının oxmxh çaprazlanmış karışık deseni ile analiz

edilmiştir. Benzer çalışmalar genellenabilirlik kuramının farklı desenleri ve farklı değişkenlik kaynakları ele alarak yapılabilir.

5. KAYNAKLAR

- Acton, W., H., Johnson, P. J., & Goldsmith, T. E. (1994). Structural knowledge assessment: Comparison of referent structures. *Journal of Educational Psychology*, 86, 303–311.
- Altıntaş, G. ve Altıntaş, S. U. (2008). İlköğretim 5. sınıf sosyal bilgiler dersinde “kavram haritası” kullanımının öğrenci akademik başarısı üzerindeki etkisi. *Kastamonu Eğitim Dergisi*, 16(1), 61-66.
- Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. New York, Holt, Rinehart and Winston.
- Anderson, T. H., & Huang, S. C. C. (1989). *On using concept maps to assess the comprehension effects of reading expository text* (No. 483). Urbana-Champaign: Center for the Studying of Reading, University of Illinois at Urbana-Champaign.
- Baker, E. L., Niemi, D., Novak, J. D., & Herl, H. (1991, July). *Hypertext as a strategy for teaching and assessing knowledge representation*. Paper presented at the NATO Advanced Research Workshop on Instructional Design Models for Computer-Based Learning Environments, The Netherlands.
- Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. Ankara: ÖSYM Yayınları.
- Brennan, R. L. (2001). *Generalizability theory*. New York: Springer-Verlag.
- Brennan, R. L. (1992). *Elements of generalizability theory*. New York: Springer-Verlag.
- Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: theory of generalizability of scores and profiles*. New York: John Wiley.
- Eason, S. H. (1989). Why generalizability theory yields better results than classical test theory. *Mid- South Educational Research Association Annual Meeting*: 8-10 November 1989- Little Rock, AR.
- Goodwin, L. D. (2001). Interrater agreement and reliability. *Measurement in Psychological Education and Exercises Science*, 5 (1), 13-14.
- Güler, N. (2008). *Klasik test kuramı, genellenebilirlik kuramı ve rasch modeli üzerine bir araştırma*. Yayınlanmış doktora tezi, Hacettepe Üniversitesi, Ankara.
- Güler, N., Uyanık, K. G. ve Teker, T. G. (2012). *Genellenebilirlik kuramı*. PegemA Yayıncılık. Ankara.
- Herl, H. E., O’Neil, H. F., Chung, G. K. W. K., & Schacter, J. (1999). Reliability and validity of a computer-based knowledge mapping system to measure content understanding. *Computer in Human Behavior*, 15, 315–333.
- Lay-Dopyera, M., & Beyerbach, B. (1983). *Concept mapping for individual assessment*. Syracuse, NY: School of Education, Syracuse University.
- Lomask, M., Baron, J. B., Greig, J., & Harrison, C. (1992). *ConnMap: Connecticut’s use of concept mapping to assess the structure of students’ knowledge of science*. Paper presented at the National Association of Research in Science Teaching, Cambridge, MA.
- Markham, K. M., Mintzes, J. J., & Jones, M. G. (1994). The concept map as a research and evaluation tool: Further evidence of validity. *Journal of Research in Science Teaching*, 31, 91–101.
- McClure, J. R., Sonak, B., & Suen, H. K. (1999). Concept map assessment of classroom learning: Reliability, validity, and logistical practicality. *Journal of Research in Science Teaching*, 36(4), 475–492.
- Nakhleh, M. B., & Krajcik, J. S. (1991). *The effect of level of information as presented by different technology on students’ understanding of acid, base, and pH concepts*. Paper presented at the annual meeting of the National Association for the Research in Science Teaching, Lake Geneva, WI.
- Naveh-Benjamin, M., Lin, Y., & McKeachie, W. J. (1995). Inferring students’ cognitive structures and their development using the “fill-in-the-structure” (fits) technique. In P. D. Nichols, s. F. Chipman,&r.l. Brennan (eds.), *cognitively diagnostic assessment* (pp. 279- 304). Hillsdale, NJ: Lawrence Erlbaum.

- Novak, J. D., Gowin, D. B., & Johansen, G. T. (1983). The use of concept mapping and knowledge vee mapping with junior high school science students. *Science Education*, 67, 625–645.
- Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. New York: Cambridge University Press.
- Özsoy, N. (2004). Kavram haritalarının ve vee diyagramlarının fonksiyonlar ünitesinin öğretilmesinde ve öğrenilmesinde kullanılması. *GÜ Gazi Eğitim Fakültesi Dergisi*, 24(2), 15-24.
- Rentz, J. O. (1987). Generalizability theory: A comprehensive method for assessing and improving the dependability of marketing measures. *Journal of Marketing Research*, 24(1), 19-28.
- Rice, D. C., Ryan, J. M., & Samson, S. M. (1998). Using concept maps to assess student learning in the science classroom: Must different methods compete? *Journal of Research in Science Teaching*, 35(10), 1103–1127.
- Ruiz-Primo, M. A., Schultz, S. E., Li, M., & Shavelson, R. J. (2001). Comparison of the reliability and validity of scores from two concept-mapping techniques. *Journal of Research in Science Teaching*, 38(2), 260–278.
- Schau, C., Mattern, N., Weber, R.W., Minnick, K., & Witt, C. (1997). *Use of fill-in concept maps to assess middle school students. connected understanding of science*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.
- Schreiber, D. A., & Abegg, G. L. (1991). *Scoring student-generated concept maps in introductory college chemistry*. Paper presented at the annual meeting of the National Association for the Research in Science Teaching, Lake Geneva, WI.
- Shavelson, R. J., & Webb, N. M. (1991). *Generalizability theory: A primer*. Newbury Park, CA: Sage.
- Shavelson, R., J. (1993). On concept maps as potential “authentic” assessments in science. indirect approaches to knowledge representation of high school science. *ERIC Document Reproduction Service No. ED 367 691*, 1993.
- Yin, Y., Vanides, J., Ruiz-Primo, M. A., Ayala, C. C., & Shavelson, J. R. (2005). A comparison of two concept-mapping techniques: Implications for scoring, interpretation, and use. *Journal of Research in Science Teaching*, 42(2), 166–184.
- Yin Y., & Shavelson R. J. (2008). Application of Generalizability Theory to Concept Map Assessment Research, *Applied Measurement in Education*. 21, 273-291

Extended Abstract

Meaningful learning was first defined as ‘an individual’s forming the knowledge by associating the owned concepts and propositions with new knowledge and as permanent learning’ by Ausubel in 1968. In addition to that, Ausubel (1968) suggests that in order for a meaningful learning to occur, the concepts in a whole should be in a form of regularity and the concepts should be presented to students in an order, the newly learnt topic should have internal consistency or it should not be in contradiction with the previous knowledge, and that students should be able to apply the knowledge and principles they have learnt to different situations. Novak and Gowin (1984), while researching what students already know and how their comprehension differs through time, developed a tool at the same time; and they first called it as ‘cognitive maps’, and later as ‘concept maps’ method.

The facets can be at random or fix in the Generalizability Theory (G), according to which the reliability of concept maps is analyzed in this research. Whether to handle a facet at random or fix depends on the researcher’s choice. If the researcher wishes to make generalizations beyond the sample into the population of generalization, then the facet will be handled at random. Yet, if the researcher does not want to reach a generalization beyond the sample, then the facet will be considered as fix. Handling the facet randomly or fix affects the prediction of reliability. The G theory is basically a measurement scale with random facets, and at least one facet needs to be random. In the G theory either all the facets are random, or random and fix facets co-occur. Therefore, the designs with at least one fix facet are defined as mixed designs (Shavelson and Webb, 1991; Brennan, 1992). Another point on which we would like to lay emphasis in this research is that the concept of “mixed design” should be used for the designs in which fix and random designs co-occur (Güler, Uyanık and Teker, 2012; Brennan, 1992).

This study investigates the parameters obtained through the G theory which measures the topic “basic concepts related with measurement and evaluation” in the measurement and evaluation course. The o (student), X m (item)X h (map) were crossed in the study, and due to the fact that the item facet was random but the map facet was fix, the mixed design was used. The variances of the results obtained with filling concept maps and the percentages of explaining the total variances were researched with the G study and with the C and S techniques; and the changes in the G and Phi coefficients were researched by increasing and decreasing the item numbers with the D study.

The variance values for the crossed mixed design *oxmxh*, which was obtained with the application of a 10-item concept map which was formed through different techniques was obtained in three steps (Shavelson and Webb, 1991; Güler, Uyanık and Teker, 2012). Firstly all the sources of variability are handled randomly, variance analysis is conducted, and the variance components are predicted. Secondly, the shared variance components to be calculated with the random part of the mixed design are determined. In the *oxmxh* design, students and items are at random, and the random part of the design is composed by crossing the students with the items. For this state, students (o), items (m) and the error calculation (om,e) with the interaction of these is made. These variance components are shown as $\sigma_{o^2}, \sigma_{m^2}, \sigma_{om,e^2}$ in order to distinguish them from the components in the random design. Lastly, the variance components of the random part of the mixed design are calculated through equations.

According to the values, the main student effect, which represents differentiation between students, accounts for the 11% of the total variance. The fact that the proportion of the variance calculated for students in the total variance is not big enough shows that the filling concept maps that are prepared through the C and S techniques cannot reveal the differences between students adequately. The item (m) main effect accounts for the 14% of the total variance. It is thought that this situation stems from the fact that some of the concepts constituting the 10 basic concepts of the ‘item’ topic seem complex to students. The variance components predicted for the main effects of map technique explains the 1% of the total variance, which is the lowest variance component in the main effects. That the variance proportion predicted for the effects of map technique is low demonstrates that there is consistency between the two different map techniques. The biggest source of variance was found as the interaction effect of student x item (om) (43%). This case displays that certain students’ correct answers change from item to item. The interaction effect for student x map technique (oh) accounts for the 0.6% of the total variance. Accordingly, it may be interpreted that the level of correct answers given by certain students changes greatly from one technique to another. The interaction effect of item x map technique accounts for the 0.2% of the total variance. This variance value, with the lowest proportion, may stem from the fact that the concepts used in maps which are prepared with the C and S techniques are the same. The interaction effect for the “student x item x map technique” accounts for the 29% of the total variance. The bigness of this value is the indicator of the student, item, map technique interaction, of the availability of immeasurable sources of variability, and/or of big random errors.

The G and Phi coefficients for the crossed *oxmxh* mixed design, which is available in the study, were predicted as 0.65 and 0.62, respectively. Increases were seen in the G and Φ coefficients which were obtained by increasing and decreasing the number of items 5 by 5 with the D study. Thus, by considering its economy in terms of labor and time, it may be recommended that more items should be used so as to be able to achieve the desired level of reliability.

Taking the results of the research into consideration, it may be recommended that the G and D studies based on the G theory should be done in determining the reliability of measurement results in which different sources of variability such as concept maps are available; because it presents detailed and explanatory results with one single analysis. In this study, the reliability of the results obtained through two different techniques of filling concept map was analyzed with the crossed *oxmxh* mixed design of the G theory. Similar studies may be performed by using the different designs of the G theory and the different sources of variability.

Kaynakça Bilgisi

Kaya Uyanık, G., & Güler, N. (2016). Kavram haritası puanlarının güvenilirliğinin incelenmesi: Genellebilirlik kuramında çaprazlanmış karışık desen örneği. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi [Hacettepe University Journal of Education]*, 31(1), 97-111.

Citation Information

Kaya Uyanık, G., & Güler, N. (2016). Investigation of concept map scores' reliability: Example of crossed mixed design in generalizability theory [in Turkish]. *Hacettepe University Journal of Education [Hacettepe Üniversitesi Eğitim Fakültesi Dergisi]*, 31(1), 97-111.