

Zorlu Kararlar: Durumsal Yargı Testlerinin Değerlendirmede Kullanılması

Yazar: Dan Hughes, a&dc - Ürün Geliştirme Direktörü

Çeviri: Eren Suna, Assessment Systems – Ölçme ve Değerlendirme Yöneticisi

İş hayatında çeşitli pozisyonlarda etkin performans gösterebilmek için yargılama (akıl yürütme) kritik bir beceridir. Alınan bazı kararlar yalnızca formal prosedürlerin izlenmesini gerektirse de çalışanlar, karşılaştıkları zorlu durumlara başa çıkmak için yargılama becerilerini sıklıkla kullanmaktadırlar. Örneğin, çalışanlar kendilerine yeterince yardımcı olmayan bir iş arkadaşlarıyla çalışmak, öncelikler arasında denge kurmak ya da beklenmedik aksilikler karşısında uygun tepkiyi göstermek durumunda olabilirler.

Durumsal Yargı Testleri (DYT, Situational Judgement Tests), adayların ve çalışanların yargılama becerilerini değerlendirmek için etkin bir teknik olarak öne çıkmaktadır. Durumsal Yargı Testlerinin kullanım sıklığı ve popülaritesi giderek artmakta, geleneksel psikometrik araçlar olan Bilişsel Beceri Testleri ve Kişilik Envanterlerinin yanı sıra kullanılmaktadır.

Peki, Durumsal Yargı Testi nedir? Bu ölçme araçlarının kullanmanın sağlayacağı yararlar nelerdir? Nasıl geliştirirler? Bu raporda DYT'lerin öne çıkan özelliklerine ve arka planına, değerlendirilmede kullanılması halinde sağladıkları beş temel faydaya ve bu araçların nasıl geliştirildiklerine odaklanılmaktadır.

Durumsal Yargı Testi Nedir?

Durumsal yargı testi, yanıtlayıcılara iş hayatında karşılaşılabilecekleri gerçekçi ve zorlu durumları yansıtan, varsayımlara dayalı (hipotetik) senaryolar sunan bir araçtır. Her bir senaryo için yanıtlayıcılara, ilgili senaryoda yapılabilecek çeşitli eylemler sunulmaktadır. Yanıtlayıcılardan beklenen ise kendilerine sunulan eylemlerin her birinin senaryodaki duruma ne kadar çözüm getireceğine dair görüşlerini çoktan seçmeli formatta bildirmesidir. Örnek bir DYT senaryosu, söz konusu senaryoda yapılabilecek dört eylemle birlikte aşağıda verilmiştir.

Yöneticiniz tarafından verilen bir kurum içi eğitime katıldınız. Eğitim sonrasında diğer katılımcılarla yaptığınız özel konuşmalarda birçok katılımcının, yöneticinizin eğitimi verme şekliyle ilgili eleştirilerde bulunduğunu gördünüz. Yöneticiniz ise bu geribildirimlerden habersiz ve siz diğer katılımcıların aksine eğitimin verilme şeklinin uygun olduğunu düşünüyorsunuz.

Lütfen yukarıdaki senaryo için aşağıdaki eylemlerin ne kadar etkili olacağını aşağıdaki ölçeğe göre değerlendiriniz.

	Oldukça Etkisiz	Etkisiz	Etkili	Oldukça Etkili
Yöneticinize, eğitimin iyi gittiğini düşündüğünüzü ve diğer katılımcılardan geribildirim almadığınızı söylemek.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Diğer katılımcılara, mutlaka yöneticinize geribildirim vermelerini önermek.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Duyduğunuz geribildirimleri isim bildirmeksizin yöneticinize anlatmak.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Yöneticinize eğitimle ilgili hiçbir geribildirimden bahsetmemek.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

DYT'ler değerlendirme sürecinde yeni bir yaklaşım olarak görülse de, benzer yaklaşımın aslında geçen yüzyılın başlarında kullanıldığı ifade edilebilir (ilk örnekler 1926'da basılan bir testte görülmektedir). 1990'lı yıllarda DYT'lerin değerlendirme sürecindeki etkinliğini gösteren yeni araştırmalarla birlikte DYT'lere yönelik ilgi tekrar artış göstermiştir. McDaniel, Morgeson, Finnegan ve Campion ve Braverman (2001) yaptıkları araştırmada DYT'leri "performansın iyi birer yordayıcısı (göstergesi)" olarak tanımlamışlardır.

DYT'lerin formatı çeşitlilik gösterebilmektedir. İlk olarak, senaryoların türleri DYT'lerde farklılık arz edebilir. Birçok DYT'de kısa senaryolar (iki ile dört cümle aralığında) kullanılırken bazılarında daha detaylı ve derinlemesine bilgi veren uzun senaryolar kullanılabilir. Ek olarak, DYT'lerdeki senaryolar karmaşıklık düzeyi olarak da farklılık gösterebilmektedir. Bazı senaryolar oldukça temel ve daha basit durumlar içerirken bazıları yanıtlayıcının birden çok ve karmaşık taleple ilgilenmesini gerektiren durumlar içerebilmektedir.

DYT'ler ayrıca, sunum formatının gerçekçi olma düzeyi (aslına benzerlik) olarak da birbirinden ayrılmaktadırlar. Yazılı DYT'ler en sık kullanılan türdür; senaryoların ve eylemlerin yazılı açıklamalarından oluşurlar. DYT'lerin daha gerçekçi türü olarak video formatındaki DYT'ler gösterilebilir: yanıtlayıcılar senaryoyu ve diğer bireylerin davranışlarını gündelik hayatta olduğu gibi görsel olarak algılayabilirler. Gerçekçilik algısını artıran diğer bir durum da senaryoları, Bekleyen İşler (Inbox) egzersizlerinde olduğu gibi, yanıtlayıcıdan yanıt bekleyen e-postalar halinde sunmaktır. Daha gerçekçi bir sunumla yapılan DYT'lerin gerçek iş tecrübelerini daha iyi yansıtmaları, dolayısıyla işle ilgili davranışları daha iyi tahmin etmesi beklenmektedir. Örneğin video formatında kullanılan bir DYT'nin, aynı içerikle yazılı uygulanan DYT'ye kıyasla daha yüksek bir yordama geçerliliğine (tahmin geçerliği) sahip olduğu belirlenmiştir (Lievens ve Sackett, 2006).

DYT'lerin birbirinden farklılık gösterdiği diğer bir özellik de yanıtlanma şekilleridir. Yanıtlayıcılardan bir senaryodaki eylemlerin ne kadar etkili olacağına ilişkin karar vermesi beklenebilir (bilgi temelli yanıtlanma). Bu durumda yanıtlayıcı, farklı eylemlerin duruma etkisini hakkında çıkarım yapmasını gerekir. Alternatif olarak yanıtlayıcılardan, her bir eylemi yapma eğilimlerini derecelendirmeleri istenebilir (eğilim temelli yanıtlanma). Bu durumda ise yanıtlayıcıdan, verilen durumda tipik olarak nasıl davranacağını bildirmesi beklenmektedir. Yanıtlanma şeklinin, DYT ile ölçülen kişisel özelliklerin türü üstünde etkili olduğunu araştırmalar göstermiştir (McDaniel, Hartman, Whetzel ve Grubb, 2007). Eğilimlere yönelik yanıtlanma, kişilik eğilimleriyle ilişkiliyken bilgi temelli yanıtlanma ise kişinin bilişsel yetenekleriyle ilişkilidir. Yanıtlanma davranışı (faking) ile ilgili endişeler doğrultusunda, bilgi temelli yanıtlanma türünde DYT'lerin yüksek riskli test (high stakes testing) durumlarında (işe alım, terfi vb.); eğilim temelli yanıtlanan DYT'lerin ise daha az riskli test durumlarında (kişisel gelişim vb.) kullanılması önerilmektedir.

DYT'lerin Beş Önemli Yararı

1 İş Benzeri Yapısı

DYT'ler, çalışanların işte karşılaşmaları muhtemel olan gerçekçi senaryolar sunmaktadır. Senaryoların iş gerekliliklerini yansıtmayı yansıtmadığını belirlemek için DYT senaryoları geliştirilirken genellikle, ilgili işte uzman kişiler ya da bölüm yöneticileriyle mülakat yapılmaktadır. Bu sayede katılımcılar, DYT'lerin iş benzeri yapısını hızla kavrar ve neden bu şekilde bir değerlendirmeye tabi olduklarını anlarlar. Seçme sürecinde bu durum, kuruma karşı olumlu algı oluşturulmasını sağlamaktadır.

DYT'lerin iş benzeri yapısının ikinci bir avantajı, adaylar için gerçekçi bir öngörü sağlamasıdır. Böylece değerlendirme süreci, iki yönlü bir öğrenme süreci haline gelmektedir. DYT sonuçları doğrultusunda kurumlar, adayların yargılama ve karar verme becerileri hakkında önemli çıkarımlara sahip olabilirler. Diğer taraftan yanıtlayıcılar ise, aday oldukları pozisyonda karşılaşabilecekleri durumlar ve zorluklar hakkında bilgi edinebilmektedirler.

DYT'ler bazı durumlarda seçim sürecinin zorunlu bir parçası olarak değil, tercihe bağlı olarak uygulanabilmektedir. Bu tür uygulamalarda DYT'ler, adayların tamamlaması için kurumun web sitesinde yer almakta ve test tamamlandığında adaylara geribildirim sağlanmaktadır. Bu sayede yanıtlayıcı, aday olduğu iş ile kişisel beklentilerinin uyuşup uyuşmadığını değerlendirebilmekte ve sürece devam edip etmeyeceğine karar verebilmektedir. DYT'lerin kullanılması, pozisyon için uygun olmayan adayların sürece dahil olma sıklığını azaltmada yardımcı olmakta, kurumlara ise pozisyonun zorluklarından motive olan adaylara odaklanma şansı vermektedir.



2 Uygulama Kolaylığı

DYT'lerde çoktan seçmeli yanıt formatı kullanılmaktadır. Katılımcılardan, her bir senaryoda yapılabilecek eylemlerin ne kadar etkili olacağına yönelik ortak bir skalada (sıklıkla dört, beş ya da altılı dereceleme skalası kullanılmaktadır) puanlama yapması istenmektedir. Alternatif olarak, katılımcıdan her bir eylemi, etkililik düzeyine göre sıralaması da istenebilir.

Katılımcının yanıtları, önceden tanımlanmış yanıt anahtarları ile karşılaştırılır ve test puanları hesaplanır. Ardından bir norm grubu (katılımcıların temsili bir örnekleme) kullanılarak, katılımcının diğerlerine kıyasla nasıl performans gösterdiğini belirlemek için puanlar standartlaştırılır. Çoktan seçmeli yanıt formatına sahip olduğundan raporlama ve hesaplama bilgisayar ile yapılabilir; bu şekilde uygulayıcıların zaman kazanması mümkün olmaktadır.

DYT'lerin formatı, internet üzerinden online olarak uygulanması için oldukça elverişlidir. DYT'ler sıklıkla, pozisyon için uygun olmayan adayların erken aşamalarda, özellikle yüz yüze mülakat aşamasından önce, elenmesi için kullanılmaktadır. DYT'lerin gözetimsiz bir ortamda online uygulanması kopya olasılığını açık bir hale getirmekle birlikte, adayların yanıt anahtarına sahip olmaması bu olasılığı düşürmektedir. Dolayısıyla, bir adayın DYT için "mükemmel" bir yanıt anahtarı oluşturarak diğerleriyle paylaşması mümkün olmamaktadır. DYT'den başarılı olan adayların yargılama becerileri, benzer yetkinliklerin davranışsal göstergelerine odaklanan gözetimli ortamlarda tekrar değerlendirilebilir.

3 Esneklik

DYT'lerin ölçtüğü özelliğin ne olduğu hakkında bazı tartışmalar olsa da DYT'lerin bir değerlendirme tekniği olduğu hakkında genel bir görüş birliği bulunmaktadır. Diğer bir ifadeyle DYT de Grup Tartışması, Bekleyen İşler Egzersizi, Analiz Egzersizi türündeki simülasyon egzersizleri gibi bir değerlendirme tekniğidir. Bazı çalışmalar (Stemler ve Sternberg, 2006) DYT'lerin genel olarak "Pratik Zekâyı" ölçtüğünü savunsa da birçok araştırma bu görüşü desteklememektedir. DYT ile ölçülen psikolojik yapı, testin içeriğine dayanmaktadır.

Bu sonuç dolayısıyla DYT'ler aracılığıyla, test içeriğine bağlı olarak geniş bir skalada farklı psikolojik özellikler değerlendirilebilmektedir. Örneğin bir DYT, uygun senaryolar geliştirilerek takım çalışması, kişilerarası duyarlılık ya da çatışma yönetimi gibi kişilerarası yetkinlikleri ölçecek şekilde yapılandırılabilir. Alternatif olarak azim, kararlılık ya da inisiyatif alma gibi başarı odaklı yetkinlikleri ölçmek için de DYT geliştirilebilir. DYT'ler, odaklanılan işte etkin performans göstermek için kritik olan yetkinlikleri yansıtacak şekilde yapılandırılırlar.

Sözel ve sayısal muhakeme testleri eleme amaçlı kullanılan, gelecekteki iş performansının güçlü yordayıcıları olsalar da bu ölçme araçları yalnızca adayın bilişsel becerilerine ve problem çözme becerilerine odaklanmaktadır. DYT'lerin de sürece dâhil edilmesiyle seçme sürecinin erken aşamalarında daha geniş perspektifte yetkinlikler ölçülebilir. Bu durumda da Değerlendirme Merkezine girecek adayların seçiminde daha isabetli tercihlerde bulunulmasına ve Değerlendirme Merkezine giren adayların başarı oranının artmasına yardımcı olacaktır.



4 Geçerlik

Seçme sürecinde kullanılan araçların temel amacı, bireyin işte nasıl performans göstereceğini tahmin edebilmektir. Etkin bir araç olarak DYT'lerin bu nedenle iş performansının göstergeleriyle istatistiksel bir korelasyon (ilişki) göstermesi gerekmektedir (ölçüt-dayanaklı geçerlik).

Yapılan iki araştırma (McDaniel ve ark., 2007; McDaniel ve ark., 2001) DYT'lerin ölçüt dayanaklı geçerliğine odaklanan çalışmaları incelemiştir. Bu araştırmaların bulguları DYT'lerin, iş performansını tahmin etmede güçlü yordayıcılar olduğunu göstermiştir. Ayrıca DYT'lerin geçerlik düzeylerinin, yanıtlama tarzlarına ve geliştirme sürecine odaklanan araştırmaların (örneğin testin, işin gereklerine uygunluğunun artırılması) artması ile daha da yükselmesi olasıdır.

DYT'lerin ölçüt-dayanaklı geçerliğini değerlendirmede önemli bir nokta da DYT'lerin, bilişsel beceri testleri ve kişilik envanterleri gibi geleneksel psikometrik yöntemlerle birlikte kullanıldığında geçerliği artırıcı bir etkisi olmasıdır (artımsal geçerlik). McDaniel ve arkadaşları (2007), DYT'lerin, kişilik faktörleri ve bilişsel beceriler üstünde istatistiksel olarak anlamlı geçerlik artışları sağladığını göstermiştir.

5 Yansızlık

DYT'lere artan ilginin temel gerekçelerinden biri de, bu araçların azınlık gruplara sağladığı yansızlıktır. Bilişsel beceri testleri, iş performansının güçlü birer göstergesi olsalar da genellikle azınlık/dezavantajlı grupların bu araçlardan elde ettiği puanlar, çoğunluğu oluşturan gruplara kıyasla daha düşük kalmaktadır. Uygulamada bilişsel beceri testlerinin eleme aracı olarak kullanılması bu nedenle etnik açıdan azınlığı oluşturan gruplar için olumsuz sonuçlar doğurabilmektedir (özellikle azınlık gruplarından testi geçen aday oranının diğerlerine kıyasla anlamlı ölçüde düşük kaldığı durumlarda).

Araştırmalar DYT'lerin, bilişsel beceri testlerine kıyasla bu istenmeyen etkiye daha düşük düzeyde sahip olduğunu göstermektedir (Nguyen, McDaniel ve Whetzel, 2005). Ek olarak, DYT'lerde kadın ve erkek katılımcıların performansları arasındaki fark minimal düzeydedir.

Daha yüksek yansızlık düzeyleri, DYT'lerin en temel yararlarından birisidir. Bu nedenle DYT'ler yansızlığı artırmak amacıyla sıklıkla bilişsel beceri testleri ile birlikte kullanılmakta; geçerliğin artmasını ve değerlendirilen yetkinlik skalasının genişlemesini sağlamaktadır.

Durumsal Yargı Testleri Nasıl Geliştirilir?

Test geliştiriciler, DYT'lerin geliştirilmesi sürecinde farklı teknikler kullansa da, tasarım ve geliştirme sürecinin dört temel basamağı olduğu ifade edilebilir.

1 Senaryoların Geliştirilmesi

DYT'ler için senaryo geliştirme sürecinde, odaklanılan pozisyonda uzman kişiler ve bölüm yöneticileriyle mülakat yapılır. Bu mülakatta özellikle, ilgili pozisyonda kişinin davranışının kritik olduğu spesifik durumlara odaklanılmaktadır. Bu şekilde test içeriğinin, odaklanılan pozisyonda başarılı olmak için kritik durumları yansıttığı garanti altına alınır. McDaniel ve arkadaşları (2001), iş analizi yapılarak geliştirilen DYT'lerin, iş analizi yapılmaksızın geliştirilenlere kıyasla daha yüksek geçerliğe sahip olduklarını göstermiştir.

Mülakatlarda toplanan spesifik veriler doğrultusunda test içeriğini oluşturacak senaryolar test geliştiriciler tarafından oluşturulur. Senaryolar oluşturulurken içeriğin net, yanlışlık oluşturmayacak şekilde tüm katılımcılar tarafından kolay anlaşılabilir olması, muğlak olamaması beklenir. Genellikle testte kullanılması planlanan senaryo sayısından daha fazla senaryo geliştirilir, ardından ölçütler doğrultusunda en iyi olduğuna karar verilen senaryolar nihai test formunda kullanılır.

2 Eylemlerin Geliştirilmesi

DYT'lerde yer alan senaryolarda yapılabilecek eylemleri belirlemek amacıyla, mülakat yapılan iş uzmanları ve bölüm yöneticilerinden, senaryodaki durumda verilebilecek tepkiler açısından çeşitli alternatifler üretmeleri beklenmektedir. Bu olası eylemlerin oluşturulması sürecinde uzmanlardan bireysel olarak ya da bir odak grup çalışması yoluyla veri toplanabilir.

Toplanan veriler test geliştiriciler tarafından incelenerek hem senaryonun hem de eylemlerin anlaşılabilirliği ve yanıtlayıcıya yeteri sayıda alternatif eylem sunulduğundan emin olunur. Senaryolara benzer şekilde, planlanan sayıdan daha fazla olası eylem oluşturulmakta ve gerekirse nihai formda eylemlerin sayısı azaltılmaktadır.

3 Yanıt Anahtarının Oluşturulması

Senaryolar ve ardından eylemlerin geliştirilmesiyle birlikte test geliştirme sürecinde sıra yanıt anahtarlarının oluşturulmasına gelmektedir. İş uzmanları ve bölüm yöneticileri daha önceki adımlarda olduğu gibi yanıt anahtarının oluşturulmasında da görev almaktadır. Her bir uzman/yönetici testi bireysel olarak tamamlar ve senaryolarda verilen durumlarda yapılabilecek eylemlerin etkililiğini değerlendirir. Bu değerlendirmeden elde edilen veriler, her bir eylemin ideal puanlamasını yapma amacıyla kullanılır. Uzman/yöneticilerin görüşleri arasında ciddi farklılıklar olan eylemler testten çıkarılarak yalnız görüş birliği sağlanan eylemler nihai formda kullanılır.

4 Testin Deneme Uygulaması

DYT'lerin, birçok diğer psikometrik araçta olduğu gibi, aracı gerçekte kullanacak aday profile mümkün olduğunca yakın, temsili bir örnekleme deneme uygulamasının yapılması gerekmektedir. Deneme uygulamasından elde edilen veriler doğrultusunda senaryoların ve eylemlerin güçlük düzeyleri belirlenebilir ve gruba çok kolay ya da çok zor gelen senaryolar ve eylemler testten çıkarılabilir. Ek olarak, deneme uygulamasına katılan gruba ait performans verilerine ulaşılabilirse bu veriler testin geçerliği için kanıt oluşturma amacıyla kullanılabilir.

Deneme uygulaması basamağında toplanan veriler, norm grubu oluşturmada bir başlangıç olarak da kullanılabilir. Böylece bir yanıtlayıcının sonuçları değerlendirilirken kişisel performansının, norm grubundan elde edilen puanlarla karşılaştırılarak belirlenmesi mümkün olmaktadır.

Özet

DYT'ler bireyin yargılama becerisini değerlendirmek için hızlı ve etkili bir teknik olarak iş performansını tahmin etmede önemli göstergeler sunmaktadır. Bu araçlar işle ilgili birçok yetkinliği ve işte başarılı olmak için kritik öneme sahip 'teknik olmayan' becerileri ölçmek için kullanılabilir. DYT'lerin geliştirilme sürecinde iş uzmanları ve bölüm yöneticilerinin kullanılması test içeriğinin, ilgilenilen pozisyonun gerekliliklerini ve zorluklarını yansıtmayı sağlamakta, pozisyonla ilgili gerçekçi bir öngörü edinilmesine yardımcı olmaktadır. Ek olarak, DYT'ler bilişsel beceri testlerine kıyasla azınlık gruplar açısından daha yansız ve adil sonuçlar sağlamaktadır.

Bu yararların birçok kurum tarafından fark edilmesi dolayısıyla DYT'lerin değerlendirme sürecinde kullanım sıklığı artış göstermektedir. DYT'lerin seçme sürecine eleme açısından değer katması, bilişsel beceri testleriyle birlikte kullanımları yanlılığı azaltmakta ve geçerliliği artırmaktadır.

Kaynaklar

Lievens, F., & Sackett, P.R. (2006). Video-based versus written situational judgment tests: A comparison in terms of predictive validity. *Journal of Applied Psychology*, 91, 1181-1188.

McDaniel, M.A., Hartman, N.S., Whetzel, D.L. & Grubb, W.L. (2007). Situational judgement tests, response instructions, and validity: A meta-analysis. *Personnel Psychology*, 60, 63-91.

McDaniel, M.A., Morgeson, F.P., Finnegan, E.B., Campion, M.A. & Braverman, E.P. (2001). Predicting job performance using situational judgement tests: A clarification of the literature. *Journal of Applied Psychology*, 86, 730-740.

Nguyen, N.T., McDaniel, M.A. & Whetzel, D.L. (2005, April). Subgroup differences in situational judgement test performance: A meta-analysis. Paper presented at the 20th Annual Conference of the Society for Industrial and Organizational Psychology, Los Angeles, CA.

Schmidt, F.L. & Hunter, J.E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-274.

Stemler, S.E. & Sternberg, R. (2006). Using situational judgement tests to measure practical intelligence. In Weekley, J.A. & Ployhart, R.E. (Eds.) *Situational Judgement Tests: Theory, Measurement and Application*. London; Lawrence Erlbaum.