

Geçerlilik ve Güvenilirlik Araştırması Özet rapor

26.01.2012

Araştırma Raporuna Genel bakış

Bu raporun amacı discprofil™ davranış envanteri profilleri için geçerlilik ve güvenilirlik araştırması sağlamaktır. Bölüm 1, davranış envanterinin arka planını ve araştırmaları içerir,

İçindekiler

Geçerlilik süreci	4
Güvenilirlik	4
Test-Tekrar Test güvenilirliği	5
İç tutarlılık	6
Geçerlilik	7
Yapı geçerliliği	11
Ölçek iç tutarlılığı	12
Diğer kişilik değerlendirmeleri ile korelasyon	23
Geçerlilik sonuçlarının özeti	25

Geçerlilik Sürecine Bakış

Psikometrik ölçüm araçları bizim dokunamadığımız ya da göremediğimiz soyut nitelikleri ölçmek için kullanılır. Bu zeka, dışa dönüklük, ya da dürüstlük gibi özelliklerdir. Araştırmacılar bu enstrümanların değerlendirmesini nasıl yapar? Bu tarz araçların kişinin davranış özellikleri ile doğru bilgi verdiğini nasıl anlayışılabilir, ya da sadece inandırıcı fakat gelen gelişigüzel bir geribildirim üreten bir araç olup olmadığını bilebilir mi? Bir psikometrik ölçüm aracının gerçekten yararlı ve doğru olarak kabul edilmesi için, bilimsel topluluk tarafından oluşturulmuş bir dizi farklı standartları karşılaması gerekir. Geçerlilik çalışması, söz konusu psikometrik ölçüm aracının bu farklı standartlara göre araştırmacılarca test edilerek kalitesinin değerlendirilmesi sürecidir. Rapor, bu farklı standartları anlamak ve discprofil™ davranış envanterinin inceleme altındaki performansını görmeye yardımcı olmak için tasarlanmıştır. Doğrulama iki temel soru sorar.

- 1. Araç ne kadar güvenilirliir?** Araştırmacılar, değerlendirme aracının ölçümlerini tutarlı ve güvenilir bir şekilde yapıyor olmasını bekler. Sonuçlar rastgele bir çok varyasyon içeriyorsa, değerlendirme aracı daha az güvenilir olarak kabul edilir.
- 2. Araç ne kadar geçerlidir?** Araştırmacılar, değerlendirme aracının tam olarak ölçmek istenileni ölçüp ölçmediğini inceler. Ölçüm aracı ne kadar ölçülmek istenileni ölçmeye yönelik ise o oranda geçerli bir araçtır.

Hiçbir psikometrik aracın mükemmel güvenilir ya da tamamen geçerli olmadığını unutmayın. Tüm psikolojik araçlar çeşitli hata kaynaklarına tabidir. Bir psikometrik ölçüm aracını güvenilir / güvenilmez ya da geçerli / geçerli değil diye kategorize etmek yerine önceden tanımlanmış bir derecelendirme skalası üzerine oturtmak daha doğru olur. Sonuç olarak, bu araç, sağlam ve güvenilir midir? şeklindeki bir soru yerine, bu aracın güvenilirliği için ne kadar kanıt vardır? şeklindeki bir soru daha uygun olur.

Güvenilirlik

Güvenilirlik hakkındaki konuşmalar, DISC gibi profillerle ilişkilendirildiğinde, kısmen değerlendirme aracının kararlılığına ve kısmen de iç tutarlılığına bakılır.

Kararlılık, belli zaman aralıklarınca aynı ölçümleri elde etmek üzere, aracın yeteneğini ifade eder. Bu genellikle aynı kişilerin, İki ölçüm arasındaki uygun bir zaman aralığından sonra ikinci kez, anketi doldurmas suretiyle test edilir. (test-tekrar test olarak adlandırılır) Daha sonra sonuçlar karşılaştırılarak birbirleriyle ilişkisinin ne kadar güçlü olduğu belirlenir.

Bir kişinin DiSC tarzı değişmeden kalırsa, istikrarlı bir aracın iki farklı uygulama arasında oldukça benzer sonuçlar üretmesi gerekir. Ancak gerçekte, söz konusu bireyin ölçülen özelliğin üzerinde değişiklik olmasa da, en sofistike psikolojik test ile bile mükemmel test-tekrar test güvenilirliği elde etmek neredeyse imkansızdır.

Bunun sebebi, test sonuçlarının testin ölçmeye çalışmadığı çeşitli dış faktörlerden de etkileniyor olmasıdır. Mesela ilk seferinde yorgun iken testi uygulayan bir aday, iyice dinlendikten sonra testi ilk seferinden daha farklı şekilde cevaplayabilir. Benzer şekilde başka bir kişi içinde olduğu ruh haline bağlı olarak cevaplayabilir. Genel olarak iki test uygulaması arasındaki uzun zaman aralığının olması, bu gelişigüzel değişkenlerin, ölçüm aracının test tekrar test güvenilirliğini düşürmesine sebep olduğu bilinmektedir. Başka bir şekilde ifade etmek gerekirse, iki uygulama arasındaki sürenin uzunluğu test-tekrar test güvenilirliği ile ters orantılıdır.

Pratikte DISC'in kararlılığının (test-tekrar test) ölçümü için, bir grup katılımcıdan envanteri cevaplamaları istenmiş ve bir süre bu gruptan aynı envanterin tekrar cevaplaması istenmiştir. Bu kararlılık bir enstrüman üzerinde bir grubun ilk puanları ve bunların daha sonraki puanları arasındaki matematiksel ilişkiye bakarak oluşturulan bir istatistik olan bir güvenilirlik katsayısı, şeklinde belirlenebilir.

Güvenilirlik katsayıları 0 ile +1 arasında değişir. korelasyon katsayısı + 1'e daha yakın oldukça, daha stabil bir araç olarak kabul edilir. Araştırmacılar genellikle onları bu test-tekrar test güvenilirlik katsayıları yorumlamaya yardımcı olmak için aşağıdaki yönergeleri kullanır: .70 in üzerindeki katsayılar "kabul edilebilir" olarak kabul edilir, ve .80 üzerindeki katsayılar "çok iyi" kabul edilir. discprofil™ davranış envanterinin kırk farklı DISC Stili için iki üç haftalık süre içinde, test-tekrar test güvenilirliği için ölçülmüş ve aşağıdaki katsayıları bulunmuştur:

Tablo.1 Test-Tekrar Test güvenilirliği

DISC Stili	Güvenilirlik
C	.86
CD	.87
CDI	.85
CDS	.86
CI	.88
CID	.85
CIS	.86
CS	.87
CSD	.85
CSI	.88
D	.85
DC	.86
DCI	.87
DCS	.86
DI	.87
DIC	.85
DIS	.86
DS	.88
DSC	.85
DSI	.86
I	.87
IC	.85
ICD	.88
ICS	.85
ID	.86
IDC	.87
IDS	.86
IS	.87
ISC	.85
ISD	.86
S	.88
SC	.85
SCD	.86
SCI	.87
SD	.85
SDC	.86
SDI	.88
SI	.85
SIC	.85
SID	.86
N= 1274	

Bu sonuçlar discprofil™ davranış envanteri tarafından üretilen sonuçların zamana karşı oldukça istikrarlı olduğunu ortaya koymaktadır. Sonuç olarak, testi yapanlar ve test yöneticileri davranış envanterinin farklı zamanlarda tekrar yapılması durumunda sadece küçük değişiklikler beklemelidirler. Ancak, uygulamalar arasındaki dönem arttıkça, bu uygulamaların sonuçları arasındaki fark daha belirgin olacaktır. Aslında, herhangi bir sofistike psikometrik ölçüm aracının iki farklı zamandaki uygulamalarının aynı sonuçları vermesi olası değildir. Bu tür minör değişiklikler discprofil™ davranış envanteri için de gözlenir.

İç tutarlılık

İç tutarlılık aynı şeyi ölçmek için sorulan sorular arasındaki korelasyon derecesini değerlendirir. Bu DiSC modelindeki ölçeklerin her biri bir dizi farklı öge ve öge grubu kullanılarak ölçülür. (öneğin S için nazik, I için inandırıcı, C için alçak gönüllü, D için orijinal, kendine has gibi)

Tüm bu öge ve öge gruplarının hepsi bir dereceye kadar birbiri ile ilişkili olmalıdır. Diğer bir deyişle, bir ölçekte tüm öğeler birbiriyle uyumlu olmalıdır. Cronbach Alfa adlı bir istatistik genellikle iç tutarlılığı değerlendirmek en iyi yöntem olarak kabul edilir. Cronbach Alfa genellikle 0.0 ile 1.0 arasında değişen belirli bir sayıda olarak korelasyon derecesini ifade eder. Alpha değeri 0.0 ise, belirli bir ölçekte ürün / ifadeleri arasında bir ilişki yoktur.

Öte yandan, eğer bir anketteki tüm ifadeler özdeş bir tarzda ölçüyorsa, Alpha değeri 1.0 olacaktır, bu da mutlak iç tutarlılığı gösterir.

Cronbach Alfa değerlendirmenin ölçeklerinin her biri için ayrı ayrı hesaplanır.

Aşağıdaki yönergeler bir ölçeğin iç güvenilirlik kalitesini değerlendirmek için kullanılır:

.70 üzerinde alfa değerleri genellikle kabul edilebilir ve tatmin edici olarak kabul edilir, .80 üzerinde alfa değerleri genellikle oldukça iyi kabul edilir ve .90 üzerindeki değerler olağanüstü iç tutarlılığı yansıttığı kabul edilir.

Aslında, bir ölçekte çok yüksek alfa değerleri onun diğerleri ile çok benzer olduğu için gereksiz olduğunu gösterebilir. Bu gibi durumlarda, değerlendime aracının maddelerin birçoğu bir ölçülen konu hakkında çok az yeni bilgiler sağlayabilir.

Alfa katsayıları 1752 katılımcıdan oluşan bir örnek grup için hesaplandı. Bu örnek grubun demografik özellikleri Ek 1'de yer almaktadır. discprofil™ davranış envanterinin tüm ölçeklerinin güvenilirlikleri Tablo 2'de listelenen Alpha değerleri ile iyi ve mükemmel aralığında iç tutarlılık gösterirler. Tüm güvenilirlikler 0,87 medyan ile .70 üzerindeki değerlerle iyidir.

Tablo.2 discprofil™ ölçekleri iç tutarlılığı

DISC Stili	Öge sayısı	Cronbach's Alpha
C	8	.90
CD	5	.87
CDI	4	.91
CDS	4	.88
CI	5	.86
CIID	4	.79
CIS	4	.77
CS	5	.81
CSD	5	.89
CSI	5	.90
D	8	.88
DC	5	.86
DCI	4	.79
DCS	4	.77
DI	5	.81
DIC	4	.90
DIS	4	.87
DS	5	.91
DSC	5	.88
DSI	4	.86
I	8	.91
IC	5	.88
ICD	4	.86
ICS	4	.79
ID	5	.77
IDC	5	.81
IDS	4	.89
IS	5	.90
ISC	4	.83
ISD	4	.87
S	8	.88
SC	5	.86
SCD	4	.91
SCI	4	.88
SD	5	.86
SDC	4	.79
SDI	4	.77
SI	5	.81
SIC	4	.89
SID	4	.90

N= 1274

Geçerlilik

Belirtildiği gibi, geçerlilik değerlendirme aracının ölçmesi için tasarlandığı konuyu ölçebilme derecesini gösterir. (Zeka, dışa dönüklük, ya da dürüstlük gibi) soyut nitelikleri ölçen psikolojik bir aracı geçerliliğini değerlendirmek zor olabilir. Bununla birlikte, araştırmacıların soruyu cevaplamak için kullandığı birkaç temel strateji vardır. Burada doğrulama stratejileri yapı geçerliliği başlıkları altında tartışıldı.

Yapı geçerliliği

Yapı geçerliliği çok teorik düzeyde bir araç geçerliliğini inceler. Yapı deneyimimizi anlamlı kılmak için kullanılan soyut bir fikir ya (örneğin zeka, egemenlik, veya dürüstlük gibi) bir kavramdır. discprofil™ davranış envanterinin DI ölçeği, örneğin, belirli bir yapı (örneğin gözü kara olmaya eğilim, maceracı ve hızlı olması) ölçer. Bu, "cesaret" bir yapısı, teorik olarak, diğer birçok yapı ile ilgilidir. Örneğin, çok cesur birisi doğada özellikle dikkatli olmayacağını varsaymak mantıklıdır. Böylece, cesur eğilimler ve tedbirli eğilimler teorik olarak olumsuz bir şekilde bağlantılıdır. Sonuç olarak, bir cesaret eğilim ölçümünün yüksek geçerliliği varsa, örneğin DI ölçeğinde yüksek puanlamaya sahip insanlar, "ihtiyatlı" yapısı ölçeğinde nispeten düşük puanlı olmalıdır, SC ölçeği gibi.

Araştırmacılar yapı geçerliliği incelediği zaman esasen, ilk olarak, bir dizi teorik ilişki belirtirler. (örneğin: cesaret yapısı teorik olarak x,y ve z yapılarıyla ilişkilidir.) sonra bu teorik ilişkilerin gerçekte var olup olmadığını ampirik olarak test ederler. Eğer öne sürülen ilişki var ise değerlendirme aracı yüksek geçerliliğe sahip demektir.

Ölçek iç tutarlılığı

Tahmin edebileceğiniz gibi, yapı geçerliliğini test etmek için kullanılan farklı yöntemler vardır. İlk olarak, bir bütün olarak bir değerlendirme aracı geçerliliğini inceleyebilirsiniz. DISC gibi enstrümanlar arka planda ölçeklerin birbirine belirli bir ilişkisi olduğu bir model önerir. Araştırmacılar model tarafından önerilen teorik ilişkinin var olup olmadığını görmek için ölçekler arasındaki gerçek ilişkiyi incelerler. DiSC modeli birbirine komşu (örneğin, Di ve i) orta korelasyona sahip olacağını önermektedir. Yani, bu korelasyonlar tekli ölçeklerin alfa güvenirliğinden önemli ölçüde daha küçük olmalıdır. Örneğin, DI ve I ölçekleri arasındaki korelasyon (0,50), DI (0,81) ve I (0,91) Alfa güvenirliğinden önemli ölçüde düşük olmalıdır. Diğer taraftan teorik olarak ters tarafta olan ölçekler (örneğin I ve C) güçlü negatif korelasyonlara sahip olmaları gerekir. Tablo 3 discprofil™ davranış envanterini tamamlamış 1274 katılımcıdan oluşturulan örnek gruptan alınan veriyi göstermektedir.

Tablo 3 Ölçeklerin iç korelasyonları

	D	DI	I	IS	S	SC	C	CD
D	.88							
DI	.46	.81						
I	.14	.50	.91					
IS	-.37	.04	.47	.90				
S	-.69	-.31	.03	.57	.88			
SC	-.62	-.73	-.56	-.13	.34	.86		
C	-.19	-.43	-.70	-.49	-.18	.45	.90	
CD	.42	-.14	-.37	-.68	-.66	-.08	.26	.87

Cronbach Alfa güvenirliği diyagonal boyunca kalın rakamlarla gösterilirken, ölçekler arasındaki korelasyon katsayıları tablonun gövdesinde gösterilmiştir. Korelasyon katsayıları -1 den +1 arasında değişir. +1 Bir ilişki iki değişkenin mükemmel pozitif korelasyon olduğunu gösterir, öyle ki bir değişken artarken, karşılık gelen miktarda diğer değişken de artar. -1 Bir ilişki iki değişkenin arasında mükemmel bir negatif ilişki olduğunu gösterir, öyle ki bir değişken artarken, karşılık gelen miktarda diğer değişken de azalır. Bir korelasyon 0 ise, iki değişkenin tamamen ilgisiz olduğunu gösterir; N = 1274 .

discprofil™ davranış envanteri modeli 40 ölçeğin, iki boyutlu bir dairesel alana dağıldığını önerdiği için, katı istatistiksel varsayım setleri, veri gerektirmektedir. Bir ölçeğin korelasyonlarının modellenmesinde, belli bir sırada düzenlenmiş olması beklenmektedir.

Tablo 4 de görülebileceği gibi belirli bir ölçek için güçlü teorik korelasyon r1 ile etiketlenmiştir. ikinci en güçlü r2 ile etiketlenmiştir ve bu şekilde devam eder. Bu durumda, R4, bir teorik ters ölçekli korelasyon gösterir. Sonuç olarak, r4 oldukça güçlü bir negatif korelasyon olmalıdır. Eğer ölçekler iki boyutlu dairesel alan yapısını destekliyorsa, her ölçek için, aşağıdaki ilişki gözlemlenebilir $r1 > r2 > r3 > r4$

Tablo 4 Beklenen ölçek iç korelasyonu

	D	DI	I	IS	S	SC	C	CD
D	1.00							
DI	r_1	1.00						
I	r_2	r_1	1.00					
IS	r_3	r_2	r_1	1.00				
S	r_4	r_3	r_2	r_1	1.00			
SC	r_3	r_4	r_3	r_2	r_1	1.00		
C	r_2	r_3	r_4	r_3	r_2	r_1	1.00	
CD	r_1	r_2	r_3	r_4	r_3	r_2	r_1	1.00

Diğer Kişilik Değerlendirmeleri ile korelasyon

Yapı geçerliği kanıt sağlamak için kullanılan diğer bir yöntem benzer özelliklerin diğer saygın davranış envanterleri ile bir değerlendirme ilişkilendirilmesi içerir. Bu amaçla bir grup katılımcı discprofil™ davranış envanterini ve diğer iki saygın davranış envanterini yapmışlardır, (NEO PI-R™) ve (16PF®). NEO PI-R, 240-maddelik bir değerlendirmedir. Kişiliği 5 Faktör modeline göre (Dışa dönüklük, Uyumluluk, Vicdanlı, Nevrotiklik ve Açıklık Deneyimi) ölçmek amacıyla tasarlanmıştır. 16PF 185 maddelik bir değerlendirmedir. 5 Faktör Kişilik modelinin yanı sıra 16 öncelikli kişilik özelliğini ölçmek amacıyla tasarlanmıştır. Değerlendirme ayrıca ilave olarak benlik saygısı ve ayarlama, mesleki ilgi, sosyal beceriler, liderlik ve yaratıcılık alanlarında ondokuz ilave ölçek hakkında istatistiki puantaj sağlamaktadır. discprofil™ davranış envanterinin, NEO PI-R ve 16PF ile korelasyonu kullanım amacına bağlı olarak talep üzerine gönderilebilir, burada yorumlanması için bir özeti verilmiştir. Bazı discprofil™ ölçekleri için NEO PI-R ve/veya 16PF ile ilgili en güçlü 10 korelasyon listelenmiştir.

I ölçeği

I ölçeği ile ilgili en yüksek 10 korelasyona aşağıda listelenmektedir.

Tablo 5. I ölçeği ile NEO PI-R ve 16PF arasındaki en Güçlü korelasyonlar

Ölçek	Araç	r
Sosyal Etkileyicilik	16PF	.74
Dışa dönüklük	16PF	.70
Sosyal Cesaret	16PF	.70
Dışadönüklük	NEO PI-R	.69
Sosyal Uyum	16PF	.68
Gruba Katılım	NEO PI-R	.65
Sosyal Kontrol	16PF	.62
Canlılık	16PF	.62
Sıcaklık	NEO PI-R	.60
liderlik potansiyeli	16PF	.60

Tablo 5 te listelenen ölçekler, I ölçeği ile ölçülebilen dönüklük ve canlılık yanı sıra sosyal denge ve kendi kendine yeterlilik durumunu yansıtır. Ayrıca yukarıdaki listede yer almamakla beraber yüksek korelasyona sahip NEO PI-R ölçekleri Olumlu duygular ($r = .50$) ve Bireysel farkındalık ($r = -.48$). I ölçeği mesleki eğilim ölçekleri Sosyallik ($r = .56$) ve Girişkenlik ($r = .53$) ile yüksek korelasyona sahiptir.

IS ölçeği

IS ölçeği ile ilgili en yüksek 10 korelasyon aşağıda listelenmektedir.

Tablo 6. IS ölçeği ile NEO PI-R ve 16PF arasındaki en güçlü korelasyonlar.

Ölçek	Araç	r
Sıcaklık	NEO PI-R	.61
Pozitif Duygular	NEO PI-R	.57
Empati	16PF	.56
Güven	NEO PI-R	.55
Fedakarlık	NEO PI-R	.53
Uyumluluk	NEO PI-R	.52
Dışadönüklük	NEO PI-R	.52
Dışa dönüklük	16PF	.51
Sıcaklık	16PF	.49
Uyum	NEO PI-R	.47

Tablo 6 da listelenen ölçekler, IS ölçeği ile ölçülebilen Sıcaklık, kabul eden ve empatik olamaya eğilim durumunu yansıtır.

SC ölçeği

SC ölçeği ile ilgili en yüksek 10 korelasyon aşağıda listelenmektedir.

Tablo 7. SC ölçeği ile NEO PI-R ve 16PF arasındaki en güçlü korelasyonlar.

Ölçek	Araç	r
Baskınlık	16PF	-.63
sosyal Uyum	16PF	-.64
Girişimcilik	16PF	-.65
Sosyal Cesaret	16PF	-.66
Sosyal Etkileycilik	16PF	-.67
Sosyal Kontrol	16PF	-.67
Duygusal Etkileycilik	16PF	-.69
Bağımsızlık	16PF	-.71
Yaratıcılık Potansiyeli	16PF	-.72
Kendine güven	NEO PI-R	-.75

Tablo 7 da listelenen ölçekler, SC ölçeği ile kendinden kontrollü, dikkatli ve pasif eğilim durumunu yansıtır.

Geçerlilik Sonuçlarının Özeti

discprofil™ davranış envanterinin değerlendirilmesi bu aracın güvenilirliği ve geçerliliği için güçlü bir destek olduğunu göstermektedir.. Analizler 87 medyan katsayısı alfa ve .86 medyan test-tekrar test güvenilirliği ile ölçek güvenilirliğinin iyi-mükemmel aralığında olduğunu göstermektedir. Aracın geçerliliğini inceleyen analizler de çok olumludur. discprofil™ davranış envanterinin iki boyutlu yapısı, çok boyutlu ölçeklendirme, ölçek iç korelasyonları ve faktör analizi ile değerlendirilendiğinde, beklentileri oldukça iyi bir şekilde karşıladığı gözlenmiştir. Ayrıca discprofil™ davranış envanterinin ölçekleri ile NEO PI-R ve 16PF® ölçekleri arasındaki korelasyonlar değerlendirmenin geçerliliği için ek destek sağlar.