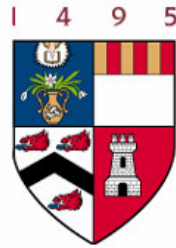
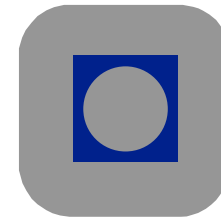


Yaşam kalitesini ölçmek

Peter Fayers



**Department of Public Health
University of Aberdeen, UK**



**NTNU
Trondheim, Norway**

Ölçme teorisi

- Ölçme teorisine pek çok disiplinin katkısı olmuştur
- Farklı ‘paradigmalar’ a ilişkin
 - sıklıkla şiddetli duyguların eşlik ettiği çelişkiler olmuş ve olagelmektedir.
- Bu çelişki, bir dereceye kadar farklı disiplinlerin birbiri ile iletişiminin olmaması nedeniyledir.



Ölçme teorisi

- **Yaklaşımların tarihsel bağlamlarının ve kuramsal temellerinin anlaşılması bunların güçlü ve zayıf yanlarının takdir edilmesini olanaklı kılar.**



Bir ölçme sınıflandırması

| | Amaçlar | Örnekler | Anahtar isimler (e.g.) |
|----------------------|--|---|--|
| Metrik | Ölçmenin temel araştırmaları, özellikle fiziksel bilimlerde | Duyarlılık, güvenilirlik; hata teorisi; Normal dağılım | Carl Gauss, Adolphe Quetelet |
| Biometrik | Değişkenliğin Çoğul kaynaklarının ölçümleri, Tarım ve tıp | Hipotez sorgulama testleri; ANOVA; ölçme hataları kaynakları | Karl Pearson, Ronald Fisher |
| Psikometrik | Kötü-tanımlanmış yapılar; latent değişkenler; zekanın boyutları | Faktör analizi; structural equation models; α Güvenirlik | Charles Spearman, Lee Cronbach Rensis Likert |
| Eğitimsel ölçümler | Duyarlılık ölçümleri skalaları; ikili yanıtlar | Rasch modelleri; item yanıt teorisi (IRT) | Georg Rasch |
| Clinimetrik | tanısal özetler, prognostik göstergeler | Klinik indeksler ve çok boyutlu özetler | Alvan Feinstein |
| Ekonometrik | yararlılık puanları | maliyet-yarar analizleri | George Torrance |
| Sosyoloji | Bireylerin nasıl düşündüğü ve tepki verdiğinin değerlendirilmesi | Soru formları için nitel yöntemler; bilişsel yaklaşımlar | |
| Tarama araştırmaları | Toplum taramaları | Kayıp veriler; imputation | Rubin & Little |



1. METRİK ÖLÇÜMLER



Metrik ölçümlerin geleneksel teorisi

- İlk dönemlerdeki bilim insanları *tek-item* ölçümlerine odaklandılar
 - Örneğin, ağırlıklar, uzunluklar, fizik ve kimya ölçümleri
- Tüm özetleme ölçümlerinde de aynı kural geçerlidir.
- Tüm ölçümler –kendi bağımsız amaçları için, güvenilir ve duyarlı olmalıdır.



Duyarlık

- **İlişkili farklılıkları ayırt etme becerisi:**
 - Bir 30-cm (bir ayak uzunluğunda) cetvel bir çok amaç için uygun olabilir, fakat duman parçacıklarını ölçmek için yeterli değildir.
Hatta Avustralya'nın ebatlarını ölçmek için de.
- **Kesinlik.**
 - “Ağrınız var mı? Evet / Hayır” ağrı düzeyinin kesin puanlamalarını vermez.
 - A 0-100 VAS yüksek düzeyde olası kesinlik sunar, fakat örneğin 56 ile 57 arasındaki farkı kaç hasta gerçekten ayırt edebilir?
 -



Duyarlık

- **Duyarlık:**
 - Eğer hastaların çoğu bir belirti düzeyi olarak (klinik düzeyde anlamlı), aynı yanıt kategorisini seçerse bir görüşme formu duyarlı olmayacaktır.
- **Taban ve tavan etkileri:**
 - Hastalar en az (taban) ve en yüksek (tavan) değerleri kullanarak yanıt vermeye eğilimlidir.
 - Örnek: 4-puanlı ağrı skalası sunulduğunda (ağrı yok, biraz, oldukça, çok fazla), pek çok ileri dönem hastası 'çok fazla' diye yanıtlayacaktır; Skala yoğundur ve farklı yükseklik düzeylerini ayırt edemez.
- **Araçlar, ölçümün ilişkili yelpazesinde yeterli kesinlikte olmalıdırlar.**



Güvenirlilik

- **Skalalar ve ölçümlerin sonuçları yeniden üretilebilir ve tutarlı olmalıdır.**
- **Tekrarlanabilme güvenirliliği:**
 - Test-yeniden test güvenirliliği
 - Görüşmeciler arası güvenirlilik
 - Form iç güvenirlilik
- **“Parazitli” ölçümler güvenilir değildir.**



Duyarlık vs. Güvenirlik

- **Güvenirlik duyarlık için gereklidir:**
 - Güvenilir olmayan bir ölçümde “parazit” vardır; bu durumda duyarsız da olur.
- **Güvenilir bir ölçümün duyarlı olması gerekmez:**
 - Tavan etkisi olan bir görüşme formu duyarlı değildir, buna karşın ölçümler rahatça tekrarlanabilir.



Tanı: Duyarlık ve özgünlük

- **tanı skalalarının yüksek duyarlık ve özgünlüğü olmalıdır.**
- **Birkaç hatalı pozitif, birkaç hatalı negatif.**
- **Örnek:**
 - **Mini-mental durum muayenesi (MMSE) demansı araştırmak için kullanılır.**
 - **MMSE puanları 0 – 30 arasındadır**
 - **24'ün altında puanı olan kişiler 'olası vakalar' olarak sınıflandırılır.**



Duyarlık ve özgünlük

| MMSE | Demans | | |
|------|------------|------------|------------|
| | evet | hayır | |
| <24 | <i>a</i> | <i>b</i> | <i>a+b</i> |
| >=24 | <i>c</i> | <i>d</i> | <i>c+d</i> |
| | <i>a+c</i> | <i>b+d</i> | <i>N</i> |

- Duyarlık = $a/(a+c)$
= doğru tanınmış olgu (+) oranı
- özgünlük = $d/(b+d)$
= doğru negatif oranı
- Kesme noktasını yükseltme duyarlığı arttıracaktır, fakat özgünlüğü azaltacaktır (daha fazla hatalı pozitif *b*)



Normal dağılım

- Ölçüm erken dönem teorilerinin çoğu Normal dağılımlara odaklanmıştır, çünkü fiziksel ve tıbbi alandaki ölçümlerin çoğunda hatalar Normal-dağılımlıdır.
- Hata teorisi:
Gauss (~1800) ve Quetelet (1835), ve diğerleri, dikkati çekenler ...



2. BIYOMETRİK



Biyometrik

- **Biyolojik (ve klinik) ölçümler çoklu kaynaklardan gelen yüksek düzeyde değişkenlik ile karakterizedir.**
- **Tarım araştırmaları yöntemlerin gelişmesinde temel bir rol oynamıştır.**
 - – örneğin, **assessing buğday mahsülü.**
Ronald Fisher (~1930), Frank Yates, etc.
- **Varyans Analizi – ANOVA.**



ANOVA

- **Atipik klinik ölçüm, kan basıncı:**
 - Puanlayıcı (gözlemci) değişkeni.
 - Gece-gündüz ve diğer biyoritim oynamaları.
 - An be an (zaman) değişkeni.
 - Görüşme formu değişkeni.
 - Hasta-hasta değişkeni.

- **... ve, son olarak, sağaltım etkisi değişkeni.**



ölçüm değişkenliği

- Ölçüm değişkenliğine biyometrik yaklaşım tipik olarak:
 - a) ANOVA kullanarak değişkenliğin nedenlerini ortaya çıkarmak.
 - b) Tekrarlayan ölçümler ile kesinliği yükseltmek.

Örn. BP

- Aynı görüşmede tekrarlayan ölçümler.
- Ardışık günlerde tekrarlayan ölçümler.
- Çok sayıda gözlemci, gibi.



ANOVA

- **Yaşam Kalitesi Görüşme formu gelişiminde ANOVAnın basit bir uygulaması, Standart Hata ölçümü (SEM) değerlendirmesidir, test-yeniden test değişkenliği.**
- **SEM, ICC gibi korelasyonlardan daha bilgi vericidir.**



ANOVA

- Ölçümün Standart hatası gözlenen puanların kesin olmamasını belirtir. (CI= gözlenen puan \pm 1.96 x SEM)

$$ICC = \frac{\text{between-subject variance}}{\text{between-subject variance} + \text{within-subject variance}}$$

$$SEM = \sqrt{\text{within-subject variance}} = \sqrt{(\text{total variance})(1 - ICC)}$$

Table 2. Test-retest reproducibility results of the SIP (N = 40)

| Areas of disability ^a | Between-subject variance | Within-subject var | | ICC | SEM (%) |
|----------------------------------|-----------------------------|--------------------|---------|-------|------------|
| | | Week | Residue | | |
| Total SIP (136) | 131.94 | 0.00 | 11.16 | 0.922 | 3.34 |

Beckerman et al. Quality of Life Research 10: 571–578, 2001.

Özet

- **Erken çalışmalar tüm ölçümlere uygulanacak temel özelliklere odaklandı.**
- **Güvenilmez yada duyarsız olan herhangi bir ölçüm, özet puanı yada endeksin bilimsel yada klinik değer taşıması olası değildir.**
- **Biyometristler klinik ve tarım verileri ile ilgili yüksek düzeyde random değişiklikleri gösterdiler.**



3. psikometrik



psikometrik

- İlk ilgi alanları:
 - Zeka
 - Kişilik
- Araştırma tipik olarak:
 - Gözlenemeyen hedef ölçümler – “*latent değişkenler*”
 - Varsayımsal hipotetik yapıların ortaya koyulması.
 - Bir çok durumda, latent değişkenlerin gerçekten var olup olmadığı tartışmalıdır.



zeka - Galton

- **Francis Galton ~ 1800 lerin sonu**
- **Pre-Galton:**
 - Toplumun çoğu eşit düzeyde akıllıdır, fakat çalışarak ve uğraşarak beceriler kazanabilirler.
- **Galton:**
 - Genel zeka, en az zekiden en zeki olana doğru sürekli bir tarzda yayılır.
 - Kalıtılabilir özellikler.



zeka – Spearman

- **Charles Spearman:**
 - “her çeşit zihinsel beceriler bir şeyde iyi isek diğer şeylerde iyi olacağız anlayışı ile pozitif olarak bağlantılıdır.”
- **Korelasyon !!**
- **İki çeşit faktör:**
 - Bir tek genel faktör, G faktörü (yada G) tüm gözlene korelasyonları tanımlayacak,
 - ve numerous specific faktörs, s_1, s_2, \dots that would account for the differences in test scores.



Faktör analizi

- Spearman ~ 1904
- Bir grup bilişsel-beceri testleri sonuçları uyumlu idi.
- İki faktör teorisi:
 - Birinci faktör genel zekadır, g ;
 - Özel beceriler ikinci faktörü oluşturur.
- Faktör analizi!!
- Catell (1956) insan kişiliğini tanımlayan 16 “faktör” önermiştir.



psikometrik modeller

- **Kişilik modelleri, zeka**
 - Karmaşık, çok-etmenli
(= çok-boyutlu)
 - Temel olarak gözlenemez
(= *latent değişkenler*)
 - *Hipotetik yapılar*



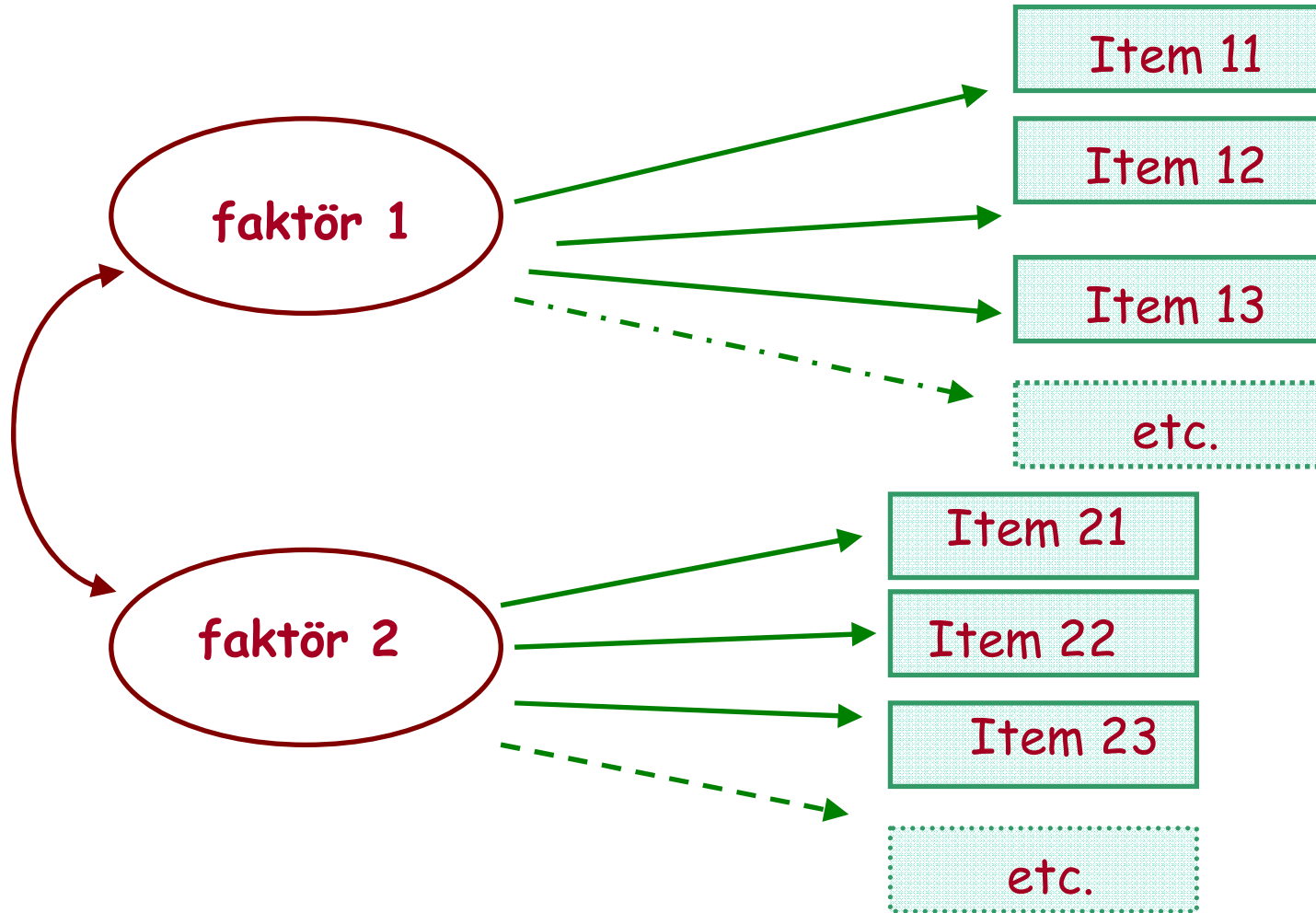
psikometrik soru formları

Bu modelleri keşfetmek için:

- a) Korelasyonları keşfetmeli – Spearman'ın belirttiği gibi;
 - b) Çok-iteimli soru formları geliştirmeliyiz, tek-iteimler test edilebilir modeller sağlamadığı için.
- **Paralel testler**



psikometrik soru formları



faktör analizi

- Patern arayan korelasyonları açıklar (yada kovaryansları).
- Item 11, 12, 13 v.s. aynı latent değişkeni yansıttıkları için güçlü biçimde korele görünebilirler, faktör 1. (eğer faktör 1 sürekli ise, $r = 0$)
- Item 11 ve 21, faktör 1 ve faktör 2 arasındaki korelasyonu yansıttıkları ölçüde korele olmalıdırlar.
- Not: $\text{Corr}(11, 12) < \text{Corr}(\text{faktör 1}, \text{faktör 2})$, ek random hata yüzünden.



biyometrik

- **Fakat açıklayıcı faktör analizi**
 - Tek, biricik matematik çözüme ulaştırmaz – sonsuz sayıda rotasyon vs vardır., ve çözümler üzerinde öznel seçim güçlü bir etkisi vardır.
 - Hipotezi test etmez.

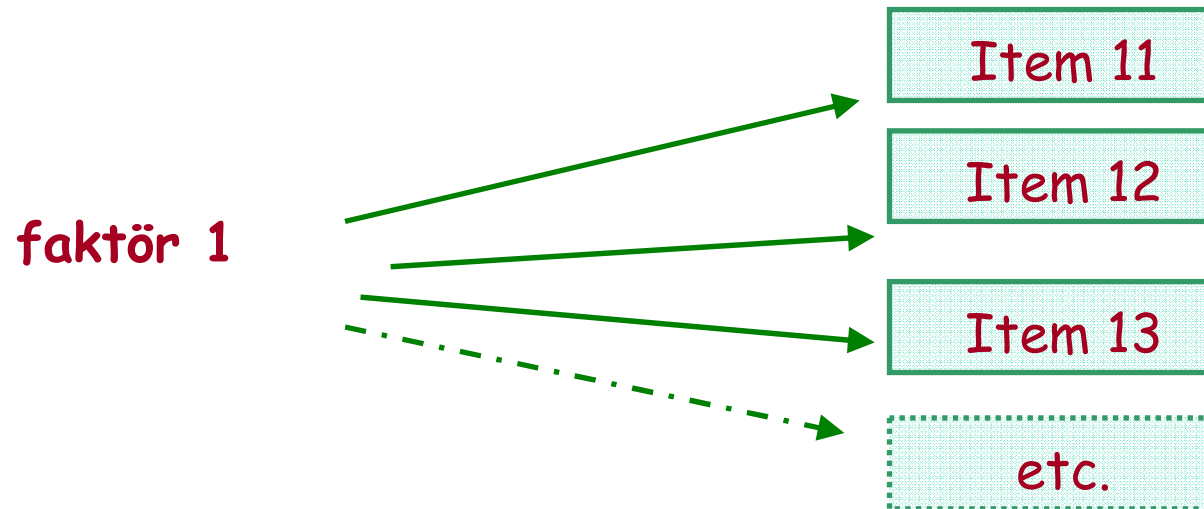


Paradigmaların çarpışması

- **biyometrik ve psikometrik**
 - Farklı hedefler
 - Farklı vurgular
 - Farklı güçlülükler
 - Farklı zayıflıklar **taşır**
- **İki disiplinin taraftarları birbirini ile aynı dilde iletişim kuramaz...**
- **(biyometrik => İstatistik)**



Cronbach α



- Items 11, 12, 13, vs. hepsi aynı faktörü ölçmeli.
- Homojen olmalı.
- “iç tutarlılık” olmalı.

Cronbach α

- 1950’de Cronbach iç tutarlılık yada homojenite ölçümü olarak α -katsayısını önermiştir.
- α aynı zamanda “*Güvenirlik*” ölçer, çünkü beklenen tekrarlanabilirlik derecesini gösterir.
- Biyometrik yaklaşımın tersine psikometristlerin testdeki homojen itemlerin sayısını arttırarak “*Güvenirlik*” i arttırdıklarına dikkat edin



Cronbach α

- Cronbach'ın orjinal 1951 yazısı etkileyicidir ve oft site edilmiştir, fakat site edenlerce gerçekten okunduğu – insan kuşkulaniyor – enderdir.
- Cronbach'ın son yazısı kuvvetle önerilir:

Cronbach LJ (2004). My current thoughts on coefficient α and successor procedures. *Educational ve Psychological Mölçüm*, 64, 391-418.



Cronbach α

- Cronbach α aşırı ve kötüye kullanılmıştır.
- Heterojen skalaların α *sı düşüktür*, yine de tekrarlanabilme bakımından geçerli ve duyarlı ve güvenilir olabilirler.
- Cronbach bir çok durumda ANOVA ve SEM in çok daha uygun yöntemler olduğunu önerir.



Doğrulayıcı faktör analizi

- Çok sonraları, CFA, ve yapısal eşitleme modellerine uzantısı, faktör analizini daha saygın bir istatistik dayanak olmuştur.
- CFA hipotezleri test etmek için olanaklar sağlar.
- CFA yine de kötüye kullanılmıştır –alternatif hipotezlerin gerçekçi bir setini saymak ve ön-belirleme için gereklidir.
- CFA, model 1 in model 2den anlamlı olarak daha iyi uyup uymadığını sorgular, fakat bu ne model 1 in doğru olduğunu ispatlar, ne de model 1 in özgün olmayan modellerden daha iyi olduğunu.
- Genelde, CFA nedenselliği ispat edemez.



Doğrulayıcı faktör analizi

- **CFA, uygunluğun delilleri olup olmadığını kontrol eder.**
- **Buradaki sorun, bir örneklem yeterince büyük ise, karmaşık bir yapı için herhangi basit bir modelin yüksek ‘uygunsuzluk’ puanları olduğu gösterilebilir. İstatistiksel anlamlılık sınırlıdır. ‘Uygunluk iyiliği’ yada uygunluğun yeterliği daha anlamlıdır.**
- **Bu durumda uygunluk makul düzeyde iyi ise gözlenen verilerin belirlenen modelle kabul edilebilir düzeyde uyumlu olduğunu söyleyebiliriz.**



dođrulayıcı faktör analizi

- İstatistikçi bir gözden geçirmeci olarak, CFA yada SEM kullanan bir yazı okuduğumda

1 numaralı sorum :

- Hangi alternatif modeller dikkate alındı?

I.e.,

Sonuçları seçilen modelin ‘dođru’ olduğuna ikna edici delil sunuyor mu , yada oldukça sınırlı bir seçilmiş modeller setinin basitce en az kötü olanı mı?



Gözlenebilen sonuçlar özet – psikometrik vs. biyometrik

- **Geleneksel metrikler**
 - Gözlenebilen sonuçlar.
 - Temel ölçüm özellikleri, Duyarlık ve Güvenirlik gibi.
- **biyometrikler**
 - Gözlenebilen sonuçlar, fakat hata kaynağı / değişkenliği konusunda daha fazla dikkate alma.
 - Hata kaynaklarını azaltarak kesinliği geliştirme yöntemleri.
 - Hatalar / belirsizlik olduğunda sonuçların karşılaştırılması (ANOVA, hipotez testleri).
 - (ve tabi ki longitudinal veri analizi, etc., etc. – fakat bu günkü konumuz değil).



özet – psikometrik vs. biyometrik

- **psikometrik**
 - Gözlenemeyen sonuçlar (Latent değişkenler).
 - Bir tek edinilmiş ölçüm kullanmak gerçekçi değil.
 - Vurgu boyutlar / faktörlerin sayısını oluşturmakta.
 - Vurgu hipotetik yapıları incelemekte
 - Yapıları temsil edecek Multi-item skalalar.
 - “Paralel testler.”



Example – psikometrik constructs

- **Fatig ve depresyon.**
- **Fatig (fiziksel, mental, vs.), için boyutların sayısı ile ilgili bir çelişki vardır, çeşitli araçlarca tartışılan 1, 2, 3 yada daha fazla boyutlar.**
- **Depresyonun tek boyut olması ile ilgili de anlaşmazlık vardır, ve tartışma depresyon ile fatigin üst üste bindiği derece düzeyindedir.**
- **Bu yapılar araştırılırken, bir psikometrik model çok uygundur.**



Örnek – psikometrik yapılar

- **Fakat, depresyonun bir item (yada diğer yazılara göre iki) kullanılarak güvenilir olarak değerlendirilebileceği yönünde bir çok yeni yazı vardır.**
- **Bu itemlerin klinik çalışmalarda değerlendirme yada tarama için uygun olduğu söylenmiştir.**
- **(i.e. Terapiye başlamadan ek bir derinlemesine değerlendirme gerekebilir)**
- **Benzer iddialar tek item fatig skalası için yapılmıştır.**



soru

Sonuçlarınız hangisine uyuyor

- a. Psikometrik model?**
- b. Biyometrik model?**

**Ne tip skalalar ve ölçümlerle
ilgileniyorsunuz?**

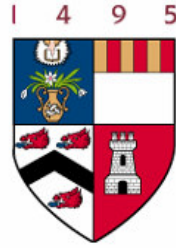
**(.... ignore, for the time being, clinimetrics
ve, e.g., IRT ölçümleri !)**



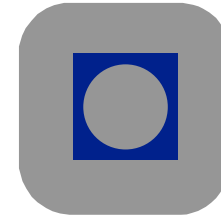
Yaşam kalitesini ölçmek

Eğitimde Ölçüm - IRT

Peter Fayers



**Department of Public Health
University of Aberdeen, UK**



**NTNU
Trondheim, Norway**

4. EĞİTİM SKALALARI, Item Yanıt Teorisi



eđitim – beceri ölçümleri

- **Karakteristik olarak:**
 - sonuçlar direkt olarak gözlenemez (yada güvenilmez kabul edilir, fazla öznel, ve proxy (ikame) puanlama ile yapıldığında bias a açıktır).
 - sonuçlar *tek boyutlu olarak kabul edilir* – örn. Her çalışma alanı için bir tek puan gerekir. (boyutluluk genellikle ya ihmal edilir yada sınavın amacına göre tanımlanır.)
 - Deđerlendirmelerin kapsamlı içeriđi olduđu güvencesi verilmelidir – geçme / kalmada tüm program bilgisinin yalnızca bir parçası etkili olursa zayıf bir sınav olur.
 - Net ve geçerli bir deđerlendirme olmalıdır (öđrencilerin düzeyi).
 - Bias dikkate alınmalıdır.



eđitim metrikleri

- **Bunların hepsi çok-itemli testleri iřaret ediyor, fakat psikimetrik soru formlarından çok farklı nedenlerle.**
- **İlk alıřmaların önemli bir kısmı ikili item arařtırmaları –'dođru' yada 'dođru deđil' yanıtı olan sorular idi.**
- **Çok zor (yada çok kolay) olan bir test itemi bilgi verici olmayacaktır.**



eđitim metrikleri

- Erken evrelerden bu yana, eđitim testleri paralel itemler kullanmamıştır (bir çok psikometrik testde kullanıldığı gibi), fakat *deđişen güçlükteki* sorulara odaklanılmıştır.
- Uzun testler bilgi ve becerinin kapsamlı içerilmesini sağlayabilir.
- Uzun testler, sorular beceri yelpazesinin tüm aralıklarını içeriyorsa becerinin tam deđerlendirilmesini sağlayabilir.



eđitim vs. Psikometri

- Psikometri *yapı geerliđi* ni arařtırmak iin faktör analizi kullanmıřtır;
- Tam tersi, eđitim testleri alanlarını sınav sorularının seđimi ile tanımlar *tanımlar*.
- Psikoojik testler itemlerini benzer güçlüklerle seđerler; eđitim sınavları farklı zorlukta itemler seđer.
- Psikometrik, paralel-item skalaları sıklıkla “*Likert summated skalalar*” olarak puanlanır ; eđitim sınavları – ve sıklıkla hala – “dođru yanıtların sayısı” ile puanlanır.



Item Yanıt Teorisi (IRT) modelleri

- Erken çalışmacılar örn. Thurstone 1925) değişen zorluktaki sorular için “doğru yanıt oranları” için S-şeklinde plot yaratmıştır.
- Biometrisistler lojistik modellerle çalışıyorlardı (Fisher & Yates, late 1930s)
- IRT modelleri “ y ” (bağımlı değişken) in latent ve bilinmeyen bir değişken olduğu durumlarda lojistik modelleri çalışıyorlardı.



Item Yanıt Teorisi (IRT)

- **Geleneksel teori ortalamalara dayanır –**
*Örn. Depresyon için bir hasta 14 puan alır;
yada bir grup hastanın mean depresyon
puanı 25 olabilir.*
- **IRT olasılık üzerine temellenir –**
*örn. Eğer bir hasta depresyon skalasında
14 alırsa onun depresyonda olma
olasılığı nedir?*
*Depresyonu olan bir hastanın
herhangi bir özel itemde endorsing
olasılığı nedir?*



“Hiyerarşik” yada “sıralı” skalalar.

- QoL arařtırmalarında IRTnin ilk örnekleri ADL deęerlendirme ve aęrı arařtırmalarında idi.
- Bu alanlar “*hiyerarşik*” skalalarda özellikle to beceri (ADL) yada řiddet (aęrı) e göre kabul edilebilir.

Örn. Fiziksel işlev

Uzun mesafe kořmakta güçlüęünüz var mı?

Uzun yürüyüşler yapmada güçlüęünüz var mı?

Evden çıkmakta güçlüęünüz var mı?

Yatakta kalmak yada sandalyede oturmak zorunda mısınız?

- beceri (yokluęu), ve řiddet, eęitim testlerinde “güçlük” analogudur.



Kesinlik ve IRT

IRT modeli

- Kesinlik ve güvenilirliđi arttırmak için, üzerinde durulan konularda süreklilikte düzenli dağılmış güçlüklerden ek test itemleri seçilir.

Örn. Fiziksel işlev

Uzun mesafe koşmakta güçlüđünüz var mı?

Uzun yürüyüşler yapmada güçlüđünüz var mı?

Kısa yürüyüşler yapmada güçlüđünüz var mı?

Evden çıkmakta güçlüđünüz var mı?

Yatakta kalmak yada sandalyede oturmak zorunda mısınız?



Test design

- **IRT fazla bilgi vermeyen itemleri ayırt edebilir (düşük diskriminasyon, yada diğer itemlerle benzer zorlukta).**
 - **Optimal etkinlik için, itemler güçlüklerine göre makul ve uniform biçimde yerleştirilmelidir, ve ikincil değişkenin tüm yelpazesini içermelidir.**
 - **Mamafi, eğer yalnızca (örneğin) yüksek düzey becerili olgular ile ilgileniyorsanız, bu güçlük bölgesindeki itemlere konsantre olmak mantıklı olacaktır.**

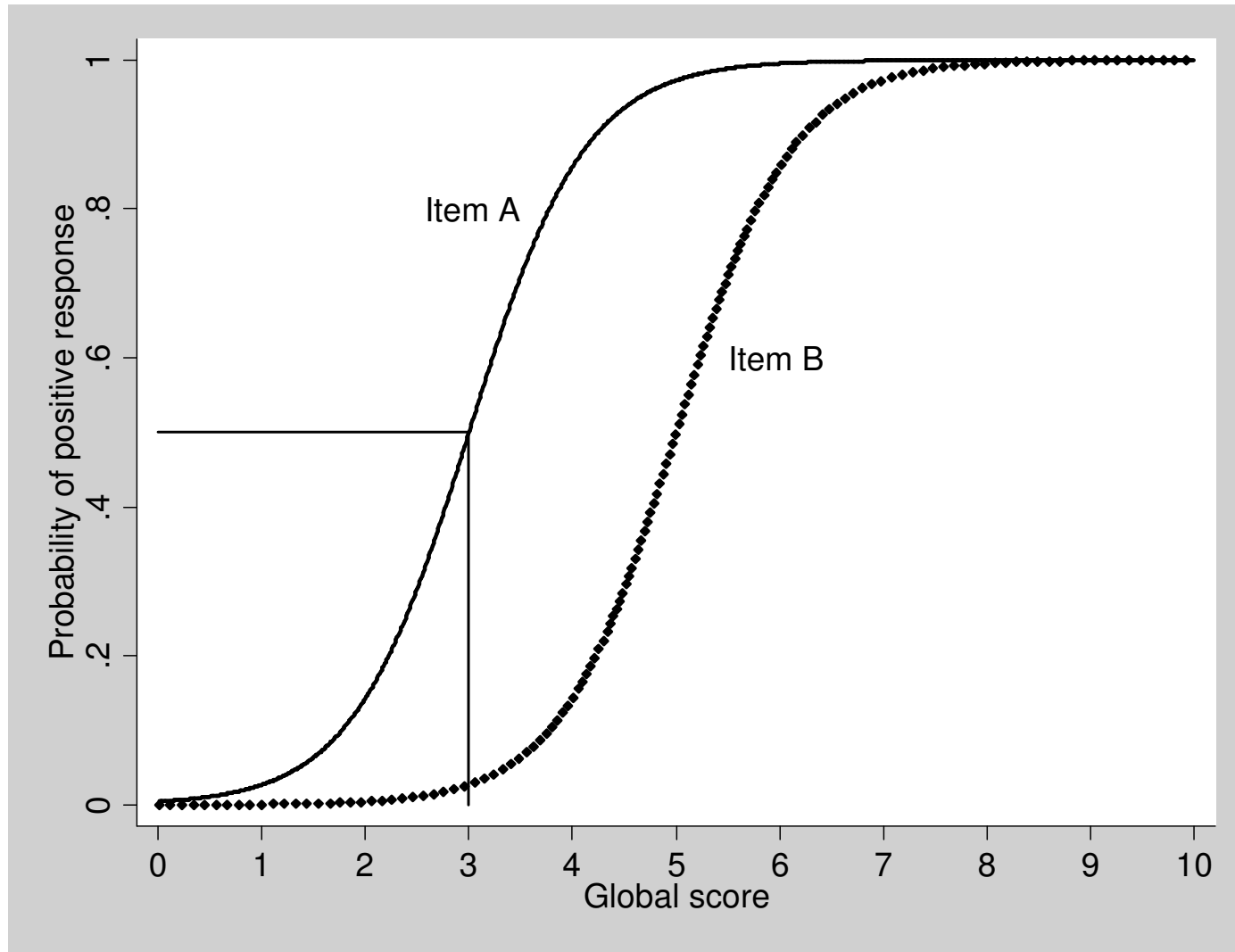


Lojistik item yanıt modelleri

- IRT bir grup farklı model için jenerik isimdir.
- En yaygın formu lojistik item yanıt modelidir. Lojistik eğriler psikolojik, eğitim ve diğer ölçümlere iyi bir uygunluk sağlar.
- $P(\theta_h)$ ise beceri θ_h , ve b olan h hastasının positif yanıt olasılığı item zorluğudur, logit formunda ise:

$$P(\theta_h) = \frac{\exp\{\theta_h - b\}}{1 + \exp\{\theta_h - b\}}$$

Tek-parametre modeli



Fitting Rasch model

- Georg Rasch (1961) popülasyon dağılımını bilmeden model parametrelerini tahmin etmenin mümkün olduğunu göstermiştir.– örn. İtemlerin görece zorluğu farklı popülasyon dağılımlarında değişmezdir.

$$P(\theta_h) = \frac{\exp\{\theta_h - b\}}{1 + \exp\{\theta_h - b\}}$$

- QoL durumu bilinmiyor - θ_h ikincil değişkendir.
- Testde iki yada daha fazla item olduğunda, QoL bilinmediğinde bile görece zorluğu çözümleriz.



IRT hilesi

- Görece item zorluklarını ve item diskriminasyonunu tahmin etmek için θ ları tahmin etmek gereksizdir. (Rasch modeli) (2-parametre IRT modeli) .
- İkincil değişkenin gerçek değerlerini bilmeden ICC eğrilerinin görece pozisyonlarını ayırt etmek mümkündür.
- Parametreleri tahmin ederek, (görünürde dögüsel bir tartışma ile) bir hastanın “becerileri” yada QoL puanları tahmin edilebilir.
- Fakat tahmin için yüksek hasta sayıları gerekir.
- Tek parametre modeli için sıklıkla en az 400 örneklem büyüklüğü önerilir, ve daha karmaşık modeller için 1000 düşünülür.



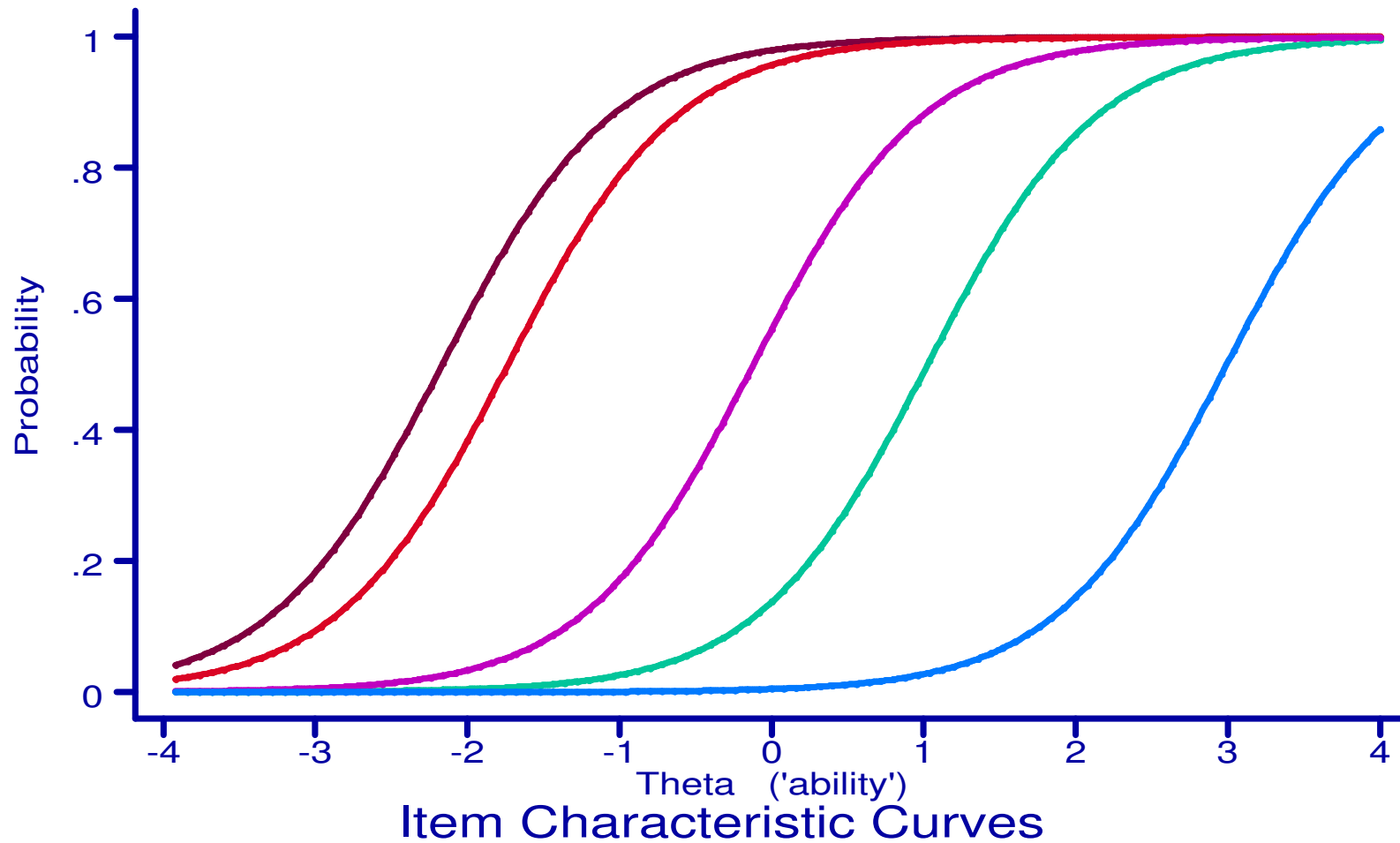
EORTC QLQ-C30 Fiziksel İşlevler

1. versiyon

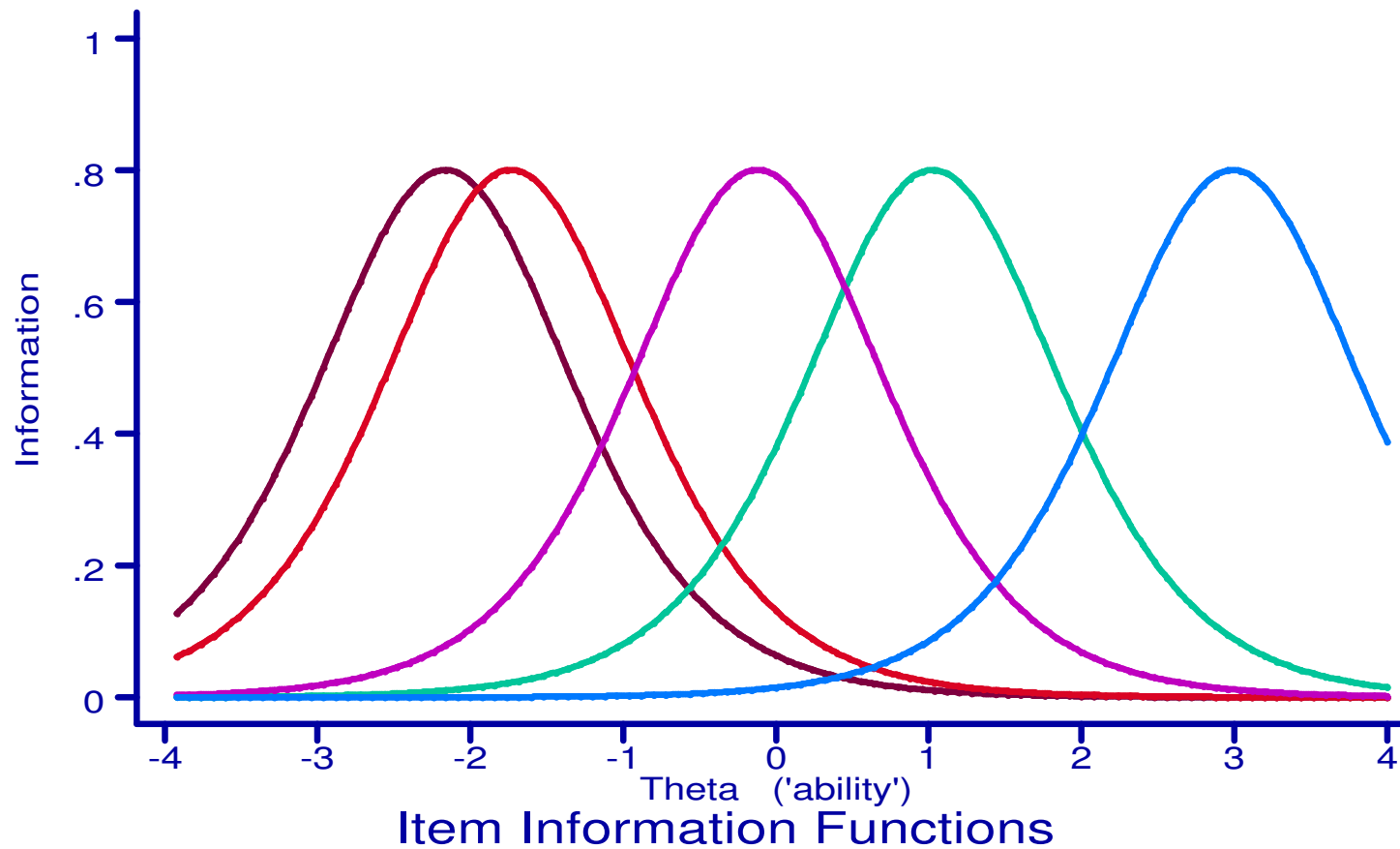
| | hayır | evet |
|--|-------|------|
| 1. Ağır alışveriş çantası yada bavul taşımak gibi yorucu işleri yapmaktan zorluğunuz var mı? | 1 | 2 |
| 2. Uzun bir yürüyüş yapmada zorluğunuz var mı? | 1 | 2 |
| 3. Evin dışında kısa bir yürüyüş yapmada zorluğunuz var mı? | 1 | 2 |
| 4. Günün önemli kısmında yatakta yada sandalyede kalmak zorunda mısınız? | 1 | 2 |
| 1. Yemek yemek, giyinmek, yıkanmak tuvalete gitmek için yardıma ihtiyacınız var mı? | 1 | 2 |



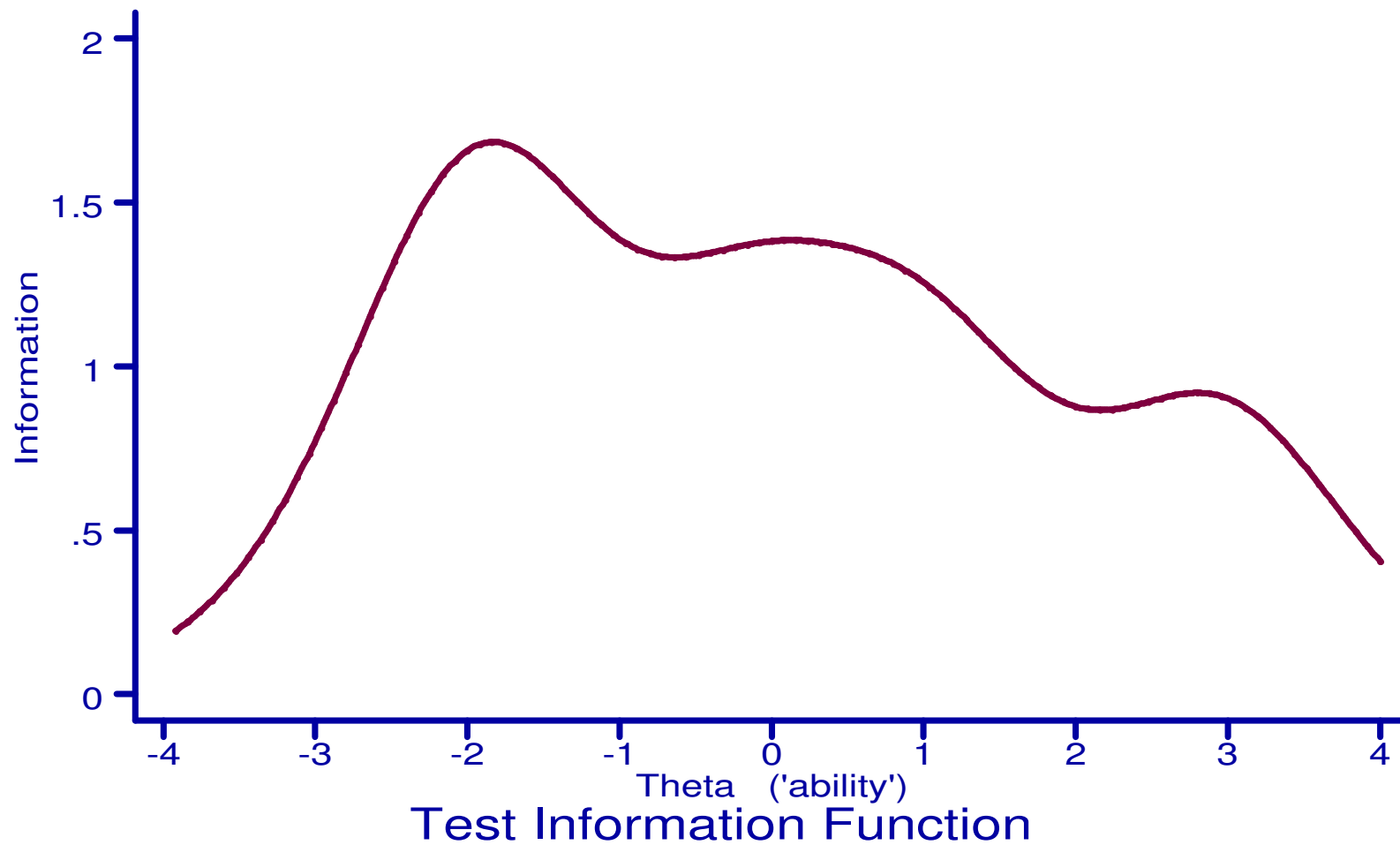
QLQ-C30 Fiziksel İşlevler – item yanıtları



QLQ-C30 PF – item bilgisi



QLQ-C30 PF – test bilgisi



Polytomous yanıtlar

| | Not at all | A little | Quite a bit | Very much |
|---|---------------|-------------|----------------|--------------|
| 1. Ağır alışveriş çantası yada bavul taşımak gibi yorucu işleri yapmakta zorluğunuz var mı? | 1 | 2 | 3 | 4 |
| Uzun bir yürüyüş yapmada zorluğunuz var mı? | 1 | 2 | 3 | 4 |

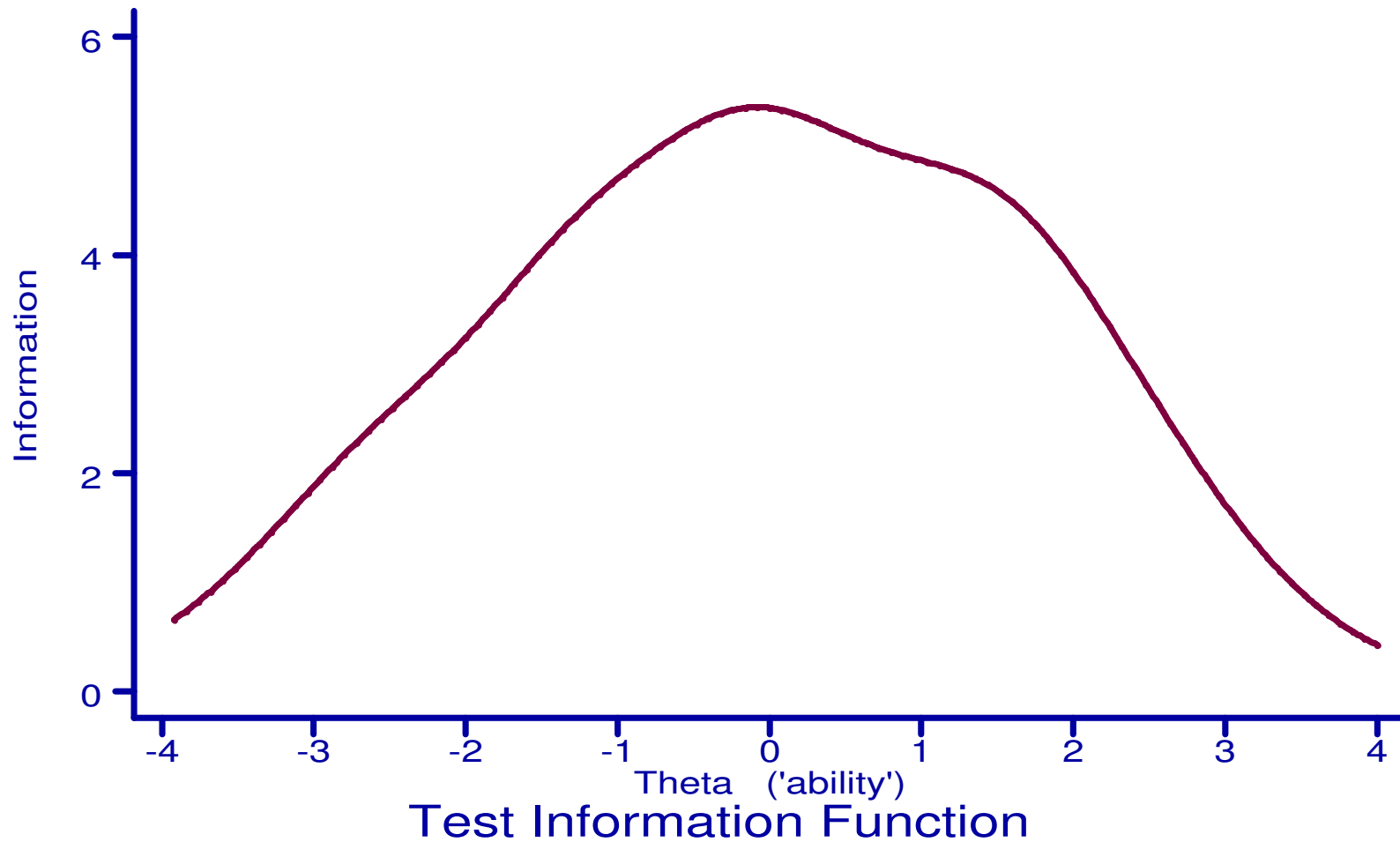
- **Genelleştirilmiş parsial kredi modeli**
 - item i , için kategori $h-1$ yerine kategori h nin seçilme olasılığı:

$$\log\left(\frac{P_{ik|k-1,k}(\theta)}{1 - P_{ik|k-1,k}(\theta)}\right) = \log\left(\frac{P_{ik}(\theta)}{P_{ik-1}(\theta)}\right) = a_i(\theta - b_{ik})$$

- **Bir çok farklı model vardır.**
Örn. Samejima'nın Graded yanıt modeli.



QLQ-C30 Fiziksle işlev



Bilgisayar uyumlu test (CAT)

- **Özelleştirilmiş” yada uyumlandırılmış testler.**
Ör. Fiziksel işlevler.
Bir hasta kısa yürüyüş yapamıyorsa neden uzun yürüyüşler yada koşma ile ilgili sorular sorulsun?
- **Soru formunun özel çeşitleri.**
 - Farklı bir hasta alt grubunda kullanmak üzere.
Örn. çocuklar.
- **Bilgisayar destekli soru formları.**
 - Önceki yanıtlar temelinde uygun zorlukta seçilmiş sorular.
 - Bu
 - (a) becerinin daha kesin sınıflanmasını
 - (b) her kişiye daha az soru yöneltilmesini
 - *sağlar*



IRT modellerinin sınanması

- **IRT modelleri**
 - Lokal bağımsızlık
 - Lojistik dağılımlar

A ilişkin güçlü varsayımlar sağlar

... bu varsayımlar sınanmalıdır!
- **AYNI ZAMANDA, IRT çoğul-itemli skalalar hakkında diğer varsayımları da test etmemize yarar.**



IRT versus Psikometri

- IRT psikometristlerce kucaklanmıştır!
- IRT bazen Klasik Test Teorisine (CTT) karşı olarak Modern Psikometri olarak adlandırılır.
- CTT tartışmalı olarak faktör yapısının boyutluluk ve yapısal geçerlik araştırılması için daha iyidir.
- IRT (bugünkü biçimlerinde)tek boyutlar için daha iyidir.
- IRT alet çantasındaki yeni alettir, fakat çoğumuz hala CTTyide kullanıyoruz.
- İki tamamlayıcı yaklaşım.



IRT nin QoL arařtırmasında kullanılması

- **IRT QoL verilerine ancak son dönemde uygulanmıřtır:**
 - Bir ok modelin matematik karmařıklıęı,
 - Geniř rneklem byklę ihtiyaı,
 - oklu-kategori itemlerini ierme sorunları
 - zelleřmiř yazılım gerekmesi,
 - İtemlere yanıt kalıplarına iliřkin gl varsayımlar, modelden ayrılmalara duyarlık,
 - Bir ok QoL verisi eęitim sınavlarında olduęu kadar derli toplu ve homojen deęil.



IRT nin avantajları

- **Modellerin matematik doğası doğal skala-yapısı varsayımlarını test etmeyi mümkün kılar.**
 - *örn. Uygunluk testlerinin iyiliği*
- **En yararlı skala itemlerinin seçimi.**
 - *örn. İtem güçlüklerinin yayılımının, diskriminasyon katsayısının , bilgi verme işlevlerinin sınılanması Sorun yaratan itemlerin saptanması – İtem biası yada DIF.*
 - *(Fakat diğer sağlıklı, etkili yöntemler de mevcuttur)*
- **İtemlerin görece zorluğunun öngörülmesi.**
 - (bu makul düzeyde sağlıklı görünüyor ...)
- **Araçların birbirine karşı standardizasyonu –örneklem-**
 - sample-invariant cross-calibration.
- **“Özelleştirilmiş” bilgisayar-destekli uyumsal testler (CAT).**



IRT nin yanlış kullanımı

- IRT her derde çare DEĞİLDİR.
- IRT geleneksel yöntemlerden daha karmaşık ve kötüye kullanıma daha açıktır.
- Too many publications of IRT analizleri ile ilgili bir çok yayın tanımlamayı başaramamış hatta yaklaşılamamıştır:
 - Lokal bağımsızlık,
 - Item-uygunsuzluk testleri,
 - Kişi uygunsuzluk testleri,
 - Örneklem büyüklüğü konuları
 - ve (DIF için) etkinin klinik anlamlılığı.

HRQL arařtırmacılarınca bildirilen artan sıklıkta IRT model ihlalleri raporları vardır.



Klinik skalalar vs. eğitim skalaları

- **eğitim ölçümleri:**
 - Eğitim testlerinde itemler sıklıkla ikili (Evet/Hayır yanıtları)
 - Eğitim sınavları genellikle uzundur (>20 itemden fazla)
 - yanıtlayanlar uzun testleri tolere etmeye gönüllüdür!
 - eğitim itemleri daha fazla bağımsızdır (*lokal bağımsızlık*).
- **Klinik skalalar:**
 - Sıralı kategorili polytomous itemler
 - *Kısa testler* – yanıtlayanın sıkılmasından, yetersiz işbirliğinden kaçınmak için.
 - Lokal bağımsızlıkla ilgili tereddütler.



Klinik skalalar vs. eğitim skalaları

- **eğitim ölçümü:**
 - IRT doğal bir yaklaşımdır.
 - IRT aynı zamanda skala eşitlemeye izin verir, DIF analizi, vs.
- **Klinik skalalar:**
 - IRT kısa skalalar için daha az olanak sunar.
 - Varsayımlar daha fazla çığnenir.
 - DIF tartışmalı olarak daha az ilgi konusu.
 - IRT mamafi, item-seçimi için değerlidir.
 - CAT için, item havuzları büyük olacaktır ve IRT ana yarar sağlayıcıdır.



Klinik skalalar

- **Tıbbi istatistiklerde güçlü bir gelenek vardır – modeller verileri tanımlamadaki becerilerine göre seçilmelidir.**
- **Rasch-taraftarları modele uyan verileri seçer.**
- **2-parametre lojistik model (2PL), farklılaşan tepe eğrileri ile, pek çok tıbbi veri için çok daha iyi uyan model sağlar.**



sonular – eđitim metrikleri

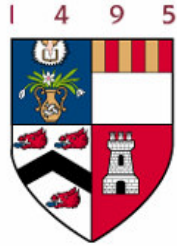
- IRT ok yaralı bir ara olabilir, ve dengeler, fakat ne geleneksel yntemlerin ne de klinimetrik yntemlerin yerini almaz.
- IRT dikkatli kullanılmalıdır ve yaygın Őekilde ktye kullanılmıŐtır.
- IRT DIFi sınamak iin zellikle avantajlıdır, kltrel farklılıkları araŐtırmak, araları birbirine gre dzenlemek, CAT, vs.
- IRT en fazla geniŐ item havuzu olduđunda yararlıdır (rn. CAT, mmkn test dizaynı / item seimi), yada uzun testlerde.



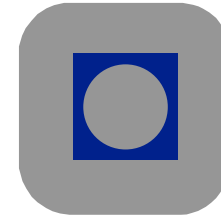
Yaşam Kalitesini Ölçmek

Kalitatif yöntemler, ve Klinimetrik

Peter Fayers



**Department of Public Health
University of Aberdeen, UK**



**NTNU
Trondheim, Norway**

5. KALİTATİF YÖNTEMLER



kalitatif yöntemler

- **Sosyal bilimciler ve araştırma yöntemcileri soru formu düzenlenmesi ve sinanması için kapsayıcı kalitatif yöntemler geliştirmişlerdir.**
- **EORTC QLQ gibi gruplar bu yöntemlerin kullanımını vurgulamışlardır, ve şimdi bunlar HRQoL soru formu geliştirilmesinin standart özellikleridir.**



kalitatif yöntemler

- İlişkili konuların (boyutlar) tanımlanmasında hastaların yada sağlıkçıların kullanılması.
- İlişkili itemlerin ayırt edilmesinde bu grupların kullanılması.
- kalitatif yöntemler şunlardır, örn, odak grupları, ve kognitif görüşme teknikleri ile görüşmeler.
- Psikometrik ve eğitim literatürü genelde düzenlemenin bu yönünü görmezden gelir.
(kişi bir sınavda hangi itemlerin olması gerektiğini öğrenciye sormaz)



kalitatif yöntemler

- **Hedefler:**
 - **Yüz ve içerik geçerliği.**
 - **Klinik olarak duyarlı skalalar.**
- **Araç düzenlemesinin bu ilk evreleri en hayati, ve başarılı bir aracı diğerlerinden ayıran şeydir.**
- **kötü oluşturulmuş soru formunu hiç bir psikometrik analiz tamir edemez – ve, gerçekten, düzenleyiciler hedeflerini ıskaladıklarının şaşkırtıcı şekilde farkında olmayabilirler.**



6. KLINIMETRİK İNDEKSLER



Klinimetrik

- Alvan Feinstein 1987.
- “Klinimetrik”
Bir çok klinik skala ve puanlar korkunç psikometrik özelliklerine karşın klinisyenler tarafından ‘mantıklı’ ve hedeflerine uygun performans gösterdikleri şeklinde değerlendirilmişlerdir.
- Tipik örnek neonatal durumu değerlendirmede kullanılan Apgar puanıdır (1953).
- Apgar puanı şunlardan oluşur
 - Kalp atım hızı
 - Solunum oranı
 - Refleks yanıtlar
 - Deri rengi
 - Kas tonusu
- Cronbach alpha değeri çok düşüktür, yinede puan sonucu iyi öngörmede çok etkilidir.



Klinik endekslerin karakteristikleri

1. Öngörücü yada Prognostik endeksler:

- **Dış bir kriter vardır-sonuç öngörülmüştür.**
- **Regresyon (yada lojistik regresyon) analiz ve skala geliştirmenin doğal yöntemidir.**
- **Psikometrik özellikler ilgisizdir.**
- **İtemler duyarlık, özgünlük, ROC eğrileri vs. ile değerlendirildikleri gibi öngörme oranlarını geliştirdiklerinde dahil edilirler**



Klinik endekslerin karakteristikleri

2. Özet puanları ve endeksler:

- **Uygun bir klinik özet**
- **Örn, “semptom yükü”, “yeti yitimi endeksi”.**
- **Endeksin parçaları endeksin kapasitesini ve amacını tanımlamaya hizmet eder.**
- **Genellikle heterojen itemler içerir.**
- **Psikometrik özellikler ilişkili olabilir, fakat klinik içerik genelde belirleyicidir.**



Skalalar ve endeksler

- **Klinimetrik puanlar için “endeks” terimini kullanmayı tercih ediyoruz.**
- **Psikometristler: çok itemle bir tek atfı ölçerler;**
- **Klinisyenler: çok atfı tek endeksle özetlerler.**
- **farklı hedefler, farklı yöntemler.**



Örnek
– Astım QoL soru formu yapısı

- **Juniper ve ark. astım yaşam kalitesi soru formu geliştirmiştir.**
- **Kapsayıcı bir potensiyel QoL item havuzu yaratıldı.**
- **Soru formunun son hali için havuzdan itemler seçildi.**
- **Juniper ve ark. psikometrik yöntemler (faktör analizi) ile hastanların önemli fikirlerini karşılaştırdı (“etki”).**



Örnek

–bir astım soru formunun yapısı

- **Psikometrik analiz 36-item araç şeklinde sonuçlandı, impakt yöntem ile oluşan 32lik formla karşılaştırıldı.**
- **İkisinde de ortak olan yalnızca 20 item vardı.**

Juniper *et al.* Clinical impact versus faktör analizi for QoL soru formu construction.
J. Clin. Epidemiol. 50: 233-238 (1997)



***Neden ve Ne zaman
psychometric yöntemler uygulanamıyor?***

- **Bir çok QoL skalaların ne sadece klinimetrik endeksler ne de psikometrik skalalardır– değişen düzeylerde doğaları ikisinin arasındadır.**
- **Ne yapalım, tanımlanmış olan hedeflerden ayrı olarak, yaklaşımları ayırt eder?**



İndikatör deęişkenler

- **Kişilik testlerindeki bir çok item, zeka testleri, eğitim kazanımları testleri ve diğer psikometrik değerlendirmeler bir beceri düzeyini yada bir zihin durumunu yansıtır.**
- **Böyle itemler ölçtüklerinin ikincil yapısını deęiştirmez yada etkilemez.**



İndikatör deęişkenler

- Genelde bu itemlere bir çok farklı tanımlayıcı isim verilir:
 - ‘*etki indikatörleri*’, ikincil deęişkenin düzeyini gösterdikleri içinbe;
 - ‘*açık yanıt deęişkenleri*’ çünkü testdeki itemlere gözlemlenmiş yanıtlardır;
 - ‘*indikatör deęişkenleri.*’



Nedensel Değişkenler

- **QoL skalalarında değerlendirilen belirtiler QoL da *değişikliğe neden olabilirler.***
- **Ciddi belirtiler olasılıkla QoL etkiler. .**
 - ***(Belirtilerin QoL görüşme formunda yer almasının ana nedeni onların QoL etkiliyor almalarıdır.)***
- **MAMAFİ, Yaşam kalitesinin kötü olması özgün belirtilerin olduğu anlamına gelmez.**
 - ***En yüksek becerisi olan bir kişinin tüm soruları yanıtlama olasılığının en yüksek olduğu eğitim testlerinin aksine, kötü yaşam kalitesi olan bir hastanın tüm belirtilerinin olması gerekmez.***



Nedensel Değişkenler

- **Belirtiler ve benzer itemler nedensel değişkenler olarak kabul edilebilirler; “belirtiler” tüm yaşam kalitesi ile ilişkili olarak nedensel değişkenlerin bir örneğidir.**
- **Tedavi yan etkileri eşit iyilikte bir örnek sağlar.**
- **Belirtiler hastalığın indikatör (gösterge) değişkenidir, yan etkiler tedavi konusunda bir şey gösterir. Fakat ne tedavi ne de hastalık QoLı değerlendirirken ana yapımız değildir; amacımız açısından bunlar nedensel değişkenlerdir.**



Nedensel Değişkenler

- ***Nedensel itemlerin atipik özellikleri:***
 - Bir tanesi kendi başına ikincil değişkeni değiştirmekte yeterli olabilir;
 - Hastaların tüm itemleri yaşamış olması gereksiz – sıklıkla da oldukça nadir-dir.
- **Örneğin, az sayıda hasta tüm olası belirtiler ve yan etkilere maruz kalacaktır**
Ciddi bir belirti – örn. ağrı – tüm yaşam kalitesini düşürmekte yeterli olabilir.



nedensel itemler ve psikometri

- **nedensel itemler geleneksel psikometri yöntemlerin kabullenimlerini ihlal eder.**
 - **Ne belirtilerin ortalama düzeyi ne de onların değişkenliği ikincil değişkence belirlenir.**
 - **Belirtiler tedavi ve hastalığın olumsuz sonuçlarıdır, mutlaka değilse de olasılıkla θ değiştirmeye yeterlidir.**



nedensel itemler – korelasyonlar

- **Bir kümedeki belirtiler arası korelasyonlar da aynı zamanda hastalık tipinin yada durumunun, yada verilen tedavinin olumsuz sonucudur.**
- **Korelasyonlar ikincil deęişkenle açıklanamaz.**
- ***θ nun herhangi bir verili deęeri ile korelasyon kağıpları tedavi ve hastalık durumlarına göre olasılıkla deęişirler,***



Korelasyon yapıları

- **En önemli ihlal şudur, nedensel değişkenler yaşam kalitesinin farklı düzeyleri üzerinden birbirleri ile korelasyon göstermezler. Paralel doğalarından gelen korelasyonları yoktur.**
- **Bunun yerine, korelasyonları temel olarak altta yatan değişkenin nedensel itemi yükseltme nedeni ne ise –tedavi yada hastalığın evre yada derecesi gibi- odur.**
- **Yani, nedensel değişkenler yaşam kalitesindeki değişikliklerle hiç ilgisi olmayan tuhaf görünen korelasyonlar sergileyebilirler.**

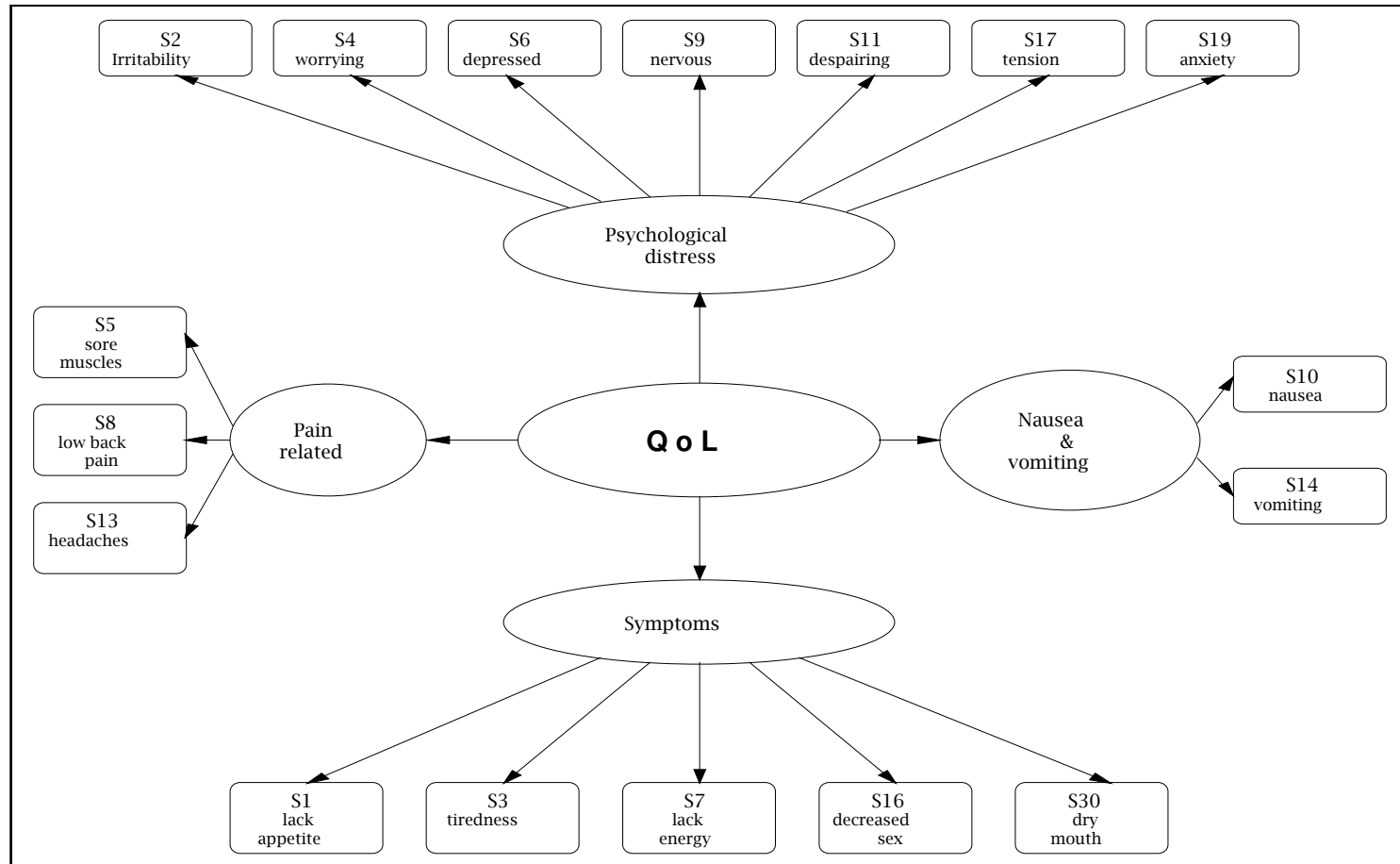


nedensel itemler – korelasyonlar

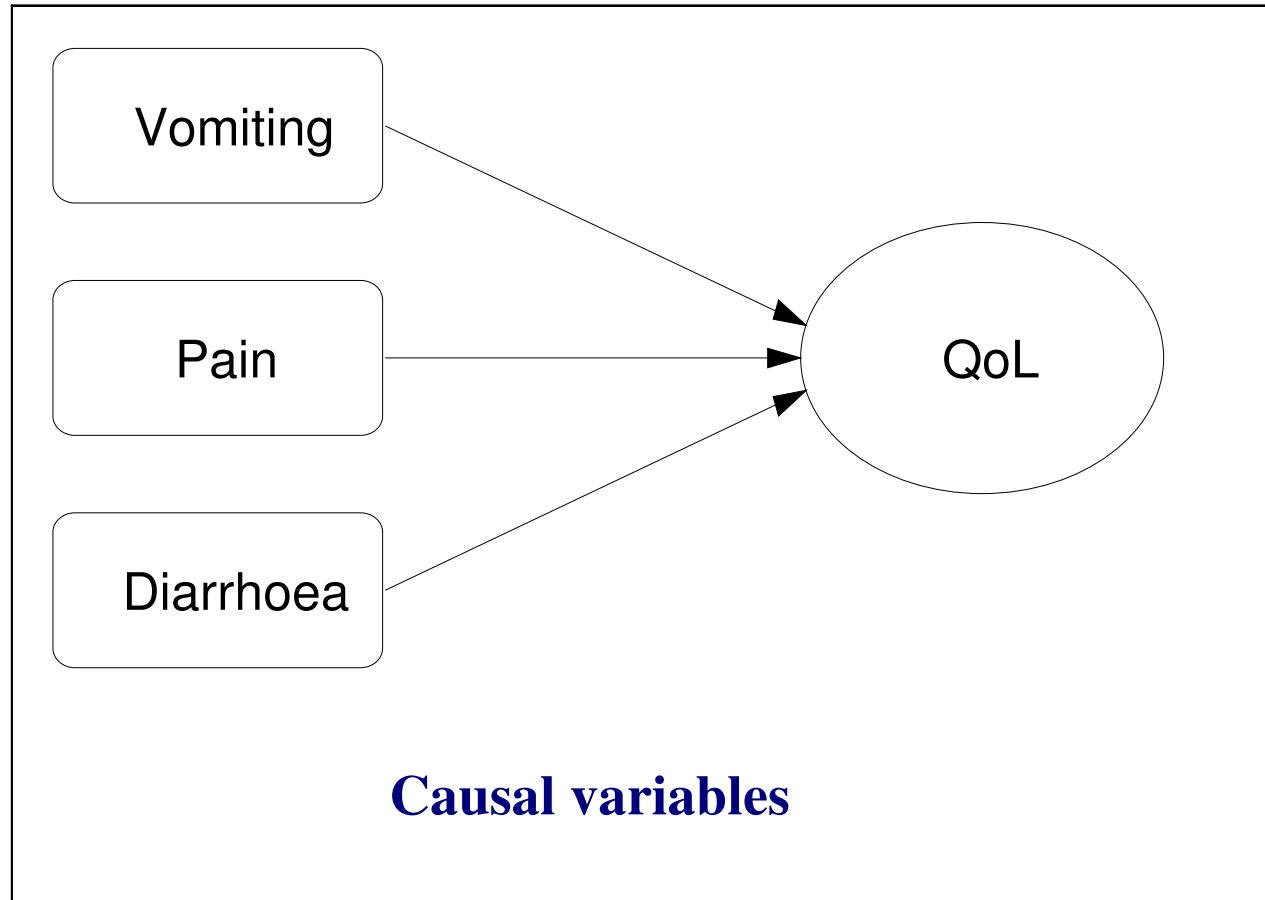
- **nedensel itemler yaşam kalitesini yansıtmazlar onu etkilerler.**
- **Nedensel değişkenlerin item-item ve item-skala *korelasyonları* *altta yatan yaşam kalitesi yapılarını* yansıtmaz.**



Postulated factor structure of RSCL (Simplified, 17 items)



Nedensel deęişkenler



faktör analizi

- “araştırmacı” faktör analizi (EFA) tümüyle korelasyon matris analizine (yada kovaryans matris) dayanır.
- Yani, faktör analizi nedensel değişkenlerin varlığında uygun değildir.
- Örnek 1 de yanıtıcı sonuçları görebilirsiniz
- Yapısal eşitleme modelleri daha uygun olabilir, fakat EFAdan daha büyük örneklem gerekir ve özelleştirmek ve uyumlamak çok daha karmaşıktır.



Örnek **–bir astım soru formu yapısı**

astım yaşam kalitesi soru formu.

- **Fayers ve arkadaşlarının işaret ettiği gibi “psikometrik yöntemler kullanılmış olsaydı dahil edilmeyecek” itemler, istisnasız, doğası kuvvetle nedensel olanlardır.**
 - **Örn. “sigara içmeye maruz kalma ile ilgili sorunlar”**
- **Şaşırtıcı olmayan biçimde, bu itemler faktör analiz modeline uymazlar.**

“Ndensel değişkenler var olduğunda klinimetrik yaklaşım geleneksel psikometrik yaklaşımdan daha uygundur.”

Fayers et al. Letter to the Editors. J. Clin. Epidemiol. 51: 285-286 (1998)



Item Yanıt Teori modelleri

- **IRT modellerini uygunlaştırmak matematik olarak karmaşıktır, bilgisayar algoritminden biricik bir çözüm elde etmek için genellikle bir dizi istatistik varsayım yapılır.**
- **Bunların arasındaki önemli varsayım “*lokal bağımsızlık*”tır.**



IRT - Lokal bağımsızlık

- **Local bağımsızlık algoritmelerin çözümü için belki de en kritik varsayımdır.**
“İkincil değişken gözlenen itemlerdeki tüm sistemik varyasyonları açıklamak için yeterli olmalıdır.”

Örn. Belirli bir ikincil değişken değeri için gözlenen itemler birbirinden bağımsız olmalıdırlar.



Lokal bağımsızlık

- **Nedensel itemler tedavi gibi dış etmenlerle etkilenirler. Dış değişken değişirse onlar da korele olacaklardır.**
 - **Örn. Kanser hastalarında, kusma saç dökülmesi ile korele olabilir, çünkü her ikisinde kemoterapi yan etkisi olarak artarlar. Lokal bağımsızlık ihlal edilmiştir.**
- **Lokal bağımsızlık IRT modelleri uygunluğu için yaşamsal bir varsayımdır, fakat sıklıkla nedensel değişkenlerle ihlal edilirler.**

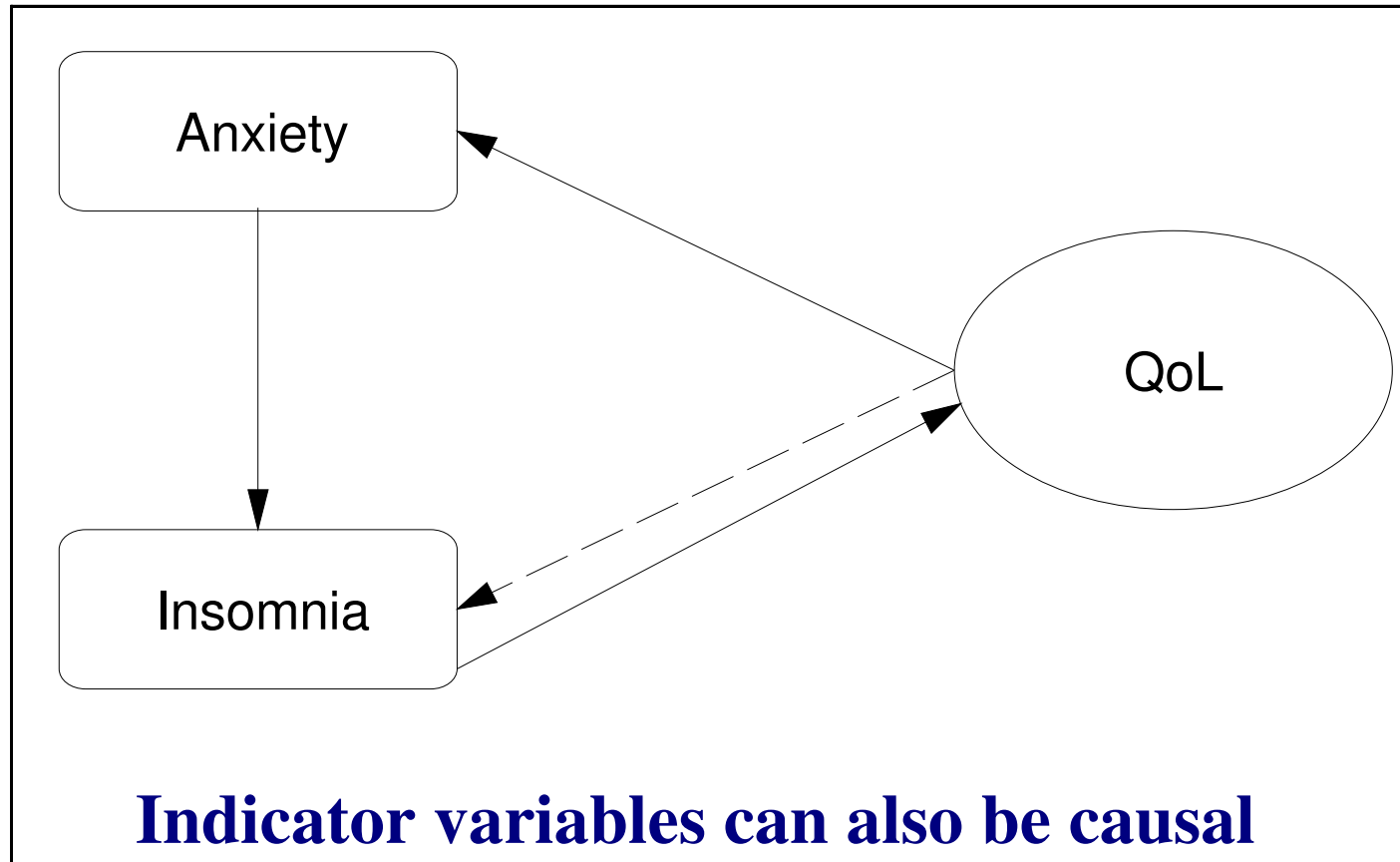


Karışık nedensel ve indikatör değişkenler

- Şimdiye kadar itemleri “nedensel” yada “indikatör” olarak tanımladık.
- Uygulamada, birçok item karışık doğadadır.
 - Örn. kusma: bir çok araştırmacı ‘kusmanın’ nedensel olabileceğinde anlaşır.
 - Şiddetli kusma, bazı kanser tedavilerinde sıktır, büyük oranda hastanın yaşam kalitesini etkileyecektir.
 - Mamafi, hastalar tedavi korkuları nedeni ile bulantı ve kusma yaşayabilirler.
 - Bazı kanser hastaları tedavi verilmeden önce ‘beklentisel bulantı ve kusma’ yaşayabilir.
 - Kusma yaşam kalitesi için temel olarak nedensel görülsede kısmen indikatör değişken olabilir.



Karışık Değişkenler



Karışık nedensel ve indikatör değişkenler

- **Eğer bir item büyük oranda nedensel ise, hatta kısmen indikatör ise de, ona ‘nedensel’ olarak davranmalıyız, çünkü item-item korelasyon yapısı etkilenebilir.**



nedensel itemleri ayırt etme

“Düşünce testi”

- **Hedef popülasyondan tipik bir hasta düşün. Bir item adlandırıp, itemX diyelim:**
- **a) İtemX düzeyi değişirse hastanın yaşam kalitesi değişecektir mi?**
- **b) Eğer hastanın yaşam kalitesi düzelirse (yada kötüleşirse) bunun item X de bir değişiklik olarak yansıtacağını düşünüyor muyuz?**
- **Eğer yanıt (a) ya “evet” ve (b) ye “hayır” ise, itemin nedensel komponenti olması beklenir.**



nedensel itemleri ayırt etme - “Düşünce testi”

Örn. kansere özgü soru formunda kusma.

- **Bir hastanın şiddetli kusması varsa, yaşam kalitesinin azalacağı beklenir
(a) =EVET**
- **Hastanın yaşam kalitesi kötülerse, onların daha şiddetli kusmasını bekliyoruz? Şart değil, belki daha fazla ağrıları var yada başka bir şey.
(b) =HAYIR, şart değil.**

⇒ kusma olasılıkla nedensel item.

***Düşünce testleri basittir,
fakat sıklıkla çok etkindir!***



psikolojik boyutlar

- **Anksiyete, depresyon ve kronik yorgunluk için araçlar olasılıkla yüksek oranda indikatör değişkenler içeriyorlar.**
- **Faktör analizi, iç güvenirlik ve IRT gibi psikometrik yöntemler uygundur.**
- **Mamafi, hala tek tek soruları gözden geçirmek, tüm itemlerin indikatör değişkenler olarak ele alınması mantıklıdır diye kontrol etmek önemlidir.**
- *Yaşam kalitesinin ruhsal yönleri ile ilgili skalalar geliştirirken psikometrik yöntemler sıklıkla uygundur, örn., duygusal işlevler, kronik yorgunluk, depresyon.*



Hastalığa özel araçlar

- **Hastalığa özel yaşam kalitesi araçları bir kaç tane hastalıkla ilişkili belirtiler yada tedavi ile ilişkili yan etkiler içermeye eğilimindedir.**
- **Bunlar olasılıkla yaşam kalitesini etkileyen nedensel itemlerdir.**
- **Bu itemler daha az sıklıkla yaşam kalitesi düzeyini yansıtan indikatörlerdir.**
- **Hastalığa-özel yaşam kalitesi araçları için, hastanın durumunu özetleyen endeksleri tanımlayan klinik duyarlılığa dayanan *klinimetrik yöntemler*, psikometrik yöntemlerin *eleştirel olmayan* kullanımına üstündür.**



-
- **Hastalığa-özel soru formu –**
 - **Klasik ölçüm teorisi**
 - + (Geleneksel psikometrik yöntemler)
 - + ***klinimetrik***

 - **Boyut-özel soru formu –**
 - ***Örn. fatig (2 – 5 boyutlar?),***
 - ***Fiziksel işlevler (hiyerarşik)***
 - **Klasik ölçüm teorisi**
 - **+ IRT**
 - + **Geleneksel psikometrik yöntemler**
(örn. SEM boyutluluğu araştırmak için)
 - + ***klinimetrik***



Sonuç

- “nedensel deęişkenler” kavramı:
- **klinimetrik psikometrik yöntemler karşılaştırmasında**
 - ayrımı,
 - gereksinimi,
 - Ve rolünü :

açıklar

▪

Tamamlayıcı yaklaşımlar

- **Klasik ölçüm teorisi**
 - örn. Güvenirlilik, duyarlık, standart hata, ANOVA
- **Geleneksel psikometri –
Klasik Test Teorisi**
 - e.g. EFA, α , SEM
- **Modern psikometri / Eğitim metrikleri**
 - Especially IRT
- **kalitatif yöntemler**
- **Klinimetrik –**
 - Genel mantık ve klinik ilişkililik.
- **Ve Ekonometri gibi diğer yaklaşımlar**



Özet – 1

- **Tüm bu yaklaşımların güçlü yönleri ve zayıflıkları vardır.**
- **Hepsi soru formu gelişimine katkıda bulunurlar.**
- **Egemen yaklaşımın seçimi aracın amaçları ve aynı zamanda içerilen itemlerin doğasına (özellikle nedensel vs. indikatör) dayanmalıdır**
- **Ortaya çıkan soru formunun karakteri benimsenen yaklaşımlar ile belirgin değişecektir.**

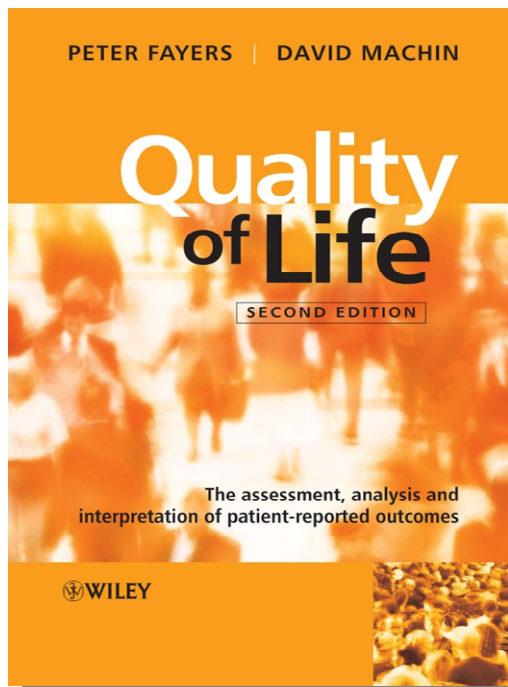


Özet – 2

- Araç geliřtirmenin en önemli fazı ilk başlataki kavramsal modeldir.
- İlk bakış ve içerik geçerliğine ilk baştaki *kalitatif* yaklaşım da yaşamsaldır.
- Eğer bir araç iyi oluşturulmuş ve düzenlenmiş ise kantitatif yöntemler onay (geçerlik) sağlar, ve itemlerin ince ‘ayarını’ mümkün kılar.



PETER FAYERS / DAVID MACHIN
**Quality of Life: The assessment, analysis
and interpretation of patient-reported outcomes**
Second Edition



Reviews of the First Edition – Winner of the
first prize in the Basis of Medicine
Category of the BMA Medical Book
Competition 2001:

*“This book is highly recommended to clinicians who are
actively involved in the planning, analysis and
publication of QoL research.”* CLINICAL
ONCOLOGY

*“This book is highly recommended
reading.”* QUALITY OF LIFE RESEARCH

Buy online at www.wiley.com

Also available online. For further information, visit: www.interscience.wiley.com/onlinebooks

 WILEY