

## DİL YETENEKLERİNİ ÖLÇEN ÜÇ TESTİN YAPI GEÇERLİKLERİ

Yrd.Doç.Dr. Satılmış Tekindal(\*)

### *Giriş*

Eđitimde bireyler haklarında deęer yargısına varmak için, onların özelliklerinin bilinmesi gerekmektedir. Çünkü bireyler arasında farklılıklar vardır. Bu farklılıklardan bazıları psikolojik yapılardır.

Psikolojik yapı, doğrudan gözlenebilir bazı davranışların tanımlanması ve sınıflandırılmasını ortaya koymak için geliştirilmiş bir fikir; biçimlendirilmiş bilimsel tasarım (imagination)'ın bir ürünü olarak tanımlanır (Crocker ve Algina 1986:231). Bireylerin psikolojik özelliklerine birer yapı olarak bakılabilir.

Psikolojik yapı doğrudan gözlenemez. Örneğin, "zeka" ve "yaratıcılık" gibi psikolojik yapıların iki seviyede tanımlanması gerekir. İlk önce, ilgilenilen yapının kavramsal olarak tanımı yapılır. Bu yapıdan hipotezler çıkarılır ve hipotezleri tanımlayan davranışlar ortaya konur. İkinci olarak, tanımlanan yapının ölçülmesinde kullanılacak prosedür tayin edilerek işlev olarak tanımlanır.

Yukarıda, zekanın psikolojik bir yapı olduğundan bahsedildi. Zeka kavramı, bireyler arasındaki psikolojik farklılıkların görülmesiyle ortaya çıkmış olabilir. Aynı sorularla karşı karşıya kalan bireylerin aynı şekilde davranmadıkları; aynı şeyleri öğrenme fırsatına sahip olan bireylerin birbirinden farklı bir şekilde ve nicelikte öğrendikleri bilinir. Gözlenebilen bu çeşit bireysel farklılıkların genişliği ve hemen her çeşit insan davranışlarında görülmesi, bu alanda çalışan bilim adamlarını "zeka" denilen bir kuram geliştirmeye, hatta zihin gücünün yapısı ve işleyişini gösteren "kuramsal yapılar" kurmaya götürmüştür. Bu kuramsal yapılardan hipotezler ve bunların davranışları çıkarılmış, bu davranışları ölçmek için de zeka testleri (yetenek testleri) geliştirilmiştir. Bu testlerin güvenilirlikleri ve geçerlikleri üzerine araştırmalar yapılmış ve değişik bilgiler elde edilmiştir (Turgut 1977:1).

\* Ondokuzmayıs Üniversitesi Amasya Eğitim Fakültesi Öğretim Üyesi.

Geçerlik kavramıyla, bir ölçme aracının ya da yöntemin onunla ölçülmek istenen deęişkeni ölçüp ölçmedięi, ölçebiliyorsa onu başka deęişkenlerden ne derecede arınık olarak ölçtüęü ifade edür (Turgut 1986:27). Test geçerlięi konusunda ileri sürülen bir çok geçerlik türü vardır, biz konumuz gereęi yapı geçerlięi üzerinde duracaęız. Yapı geçerlięini kanıtlamak için başvurulanan yollardan bazıları şunlardır:

#### *Çoklu Özellik-Çoklu Metod Matriks(\*)*

Bu yaklaşımda, aynı yapıyı farklı metodlarla ölçüp, bu ölçümler arasındaki korelasyonlar hesaplanır. Bu korelasyon katsayıları, birbirine yaklaşan (convergent) geçerlik katsayılarıdır. İkincisi, farklı yapıları aynı ölçme metodu ile ölçerek, bu ölçümler arasındaki korelasyonları ya da farklı yapılan farklı ölçme metodları ile ölçüp, bu ölçümler arasındaki korelasyonları hesaplama şeklidir. Bu korelasyon katsayıları ayırıcılık (discriminant) geçerlik katsayılarıdır. Son olarak da, aynı yapı, aynı metodla ölçülerek bu ölçümler arasındaki korelasyonlar hesaplanır, bu ise güvenilirlik katsayısıdır (Corcker ve Algina 1986: 232).

#### *Faktör Analizi*

Bu yöntemde faktörler, özel ve genel faktörler olmak üzere iki ana gruba ayrılır. Bu faktörlerden özel faktörleri her birisi tanımı gereęi ayrı ayrı deęişkenlere, genel faktör ise ortak bir deęişkenlięi dayandırılır. Aynı yapıyı ölçtükleri farz edilen test veya alt testlerin deneysel olarak genel bir faktörle tanımlanıp tanımlanamayacaęına bakılır. Eđer, aynı şeyi ölçtüęü düşünölen bir kaç test ya da alt test farklı faktörlerde gruplanırsa, bunlar tarafından ölçölen yapımm geçerlięi hakkında sorular artabür (Crocker ve Algina 1986: 232).

#### *Genelleştirilebilirlik Teorisinden Yararlanma*

Yapıyı doğrulamanın diđer bir yolu ele alınan yapı ile ilgili olarak, bir birey için, farklı ölçme teknikleri ile elde edilen gözlemlerin her birisinin benzer olup olmadıęını araştırmaktır. Kane (1982), bu noktanın araşhılması ile ilgili olarak "genelleştirilebilirlik" (generalizability) teorisinin uygulanmasından elde edilen varyans kısımları (components)'nın analizini önerir. Bir yapıyı, deęişik madde formları (cümle tamamlama, eş anlamlı gibi) üe ölçmek mümkündür. Bu yapıyı ölçmek için, kullanılan maddeformu ne olursa olsun, eđer cevaplayıcıların görelili puanları benzer

\* The Multitrait-Multimethod Matrix.

olursa yapının oldukça önemli ya da "genelleşmiş" (generalized) olduğu söylenebilir. Madde formlarından hangisinin ölçülmek istenen özelliđi daha iyi ortaya çıkardığı görülmek istenirse, bir G Study ve buna bađlı olarak bir D study desenlenebilir (Crocker ve Alina 1986: 234).

Madde yapısında deđişiklik yapılarak oluşmrulan testlerin geçerlik!erinin çok az araştırıldığı söylenebilir.

Matematik problemlerini çözme gücünün ölçülmesi amacıyla, Adıgüzel (1985) tarafından hazırlanan çoktan seçmeli tesüerde, dođru çözüm sonucunu seçtiren alışılmış madde yapısı dođru çözüm yolunu seçtiren madde yapışma dönüşmmlmiştir. Bu durumda dođru çözüm sonucunu seçtiren maddelerden oluşan testin diđerine göre daha geçerli olduğu görülmüştür (Adıgüzel 1985: 79-87).

Turgut (1971) tarafından, kelime bilgisini ölçen iki cevaplı maddelerden oluşan test, bir cevaplı test formuna çevrilmiş ve önceki formdan elde edilen puanlar, sonraki formun cevaplarından daha geçerli bulunmuştur (Turgut 1971: 66).

Lord ve Wild (1985) tarafından, Graduate Record Examinations (GRE) testinin sözel bölümündeki dört madde tipi de aynı sözel yeteneđi ölçer sayılışıyla yola çıkılmış ve analogi maddelerinin toplam test puanlarına istikrarlı bir katkı getirdiđi, sözel yeteneđin ölçülmesinde en iyi maddelerin cümle tanımlama ve okuduđunu anlama maddeleri olduğu belirtilmiştir. Ayrıca zıt anlamlı kelimeler testinin sözel yetenek ölçümlerine en yüksek dođruluk katkısını sağladığı belirtilmektedir (Lord ve Wild 1985: 5-20).

Schrader (1984)'ın "Scholastic aptitude testin sözel madde tipleri ile ilgili üç çalışma" adlı araştırmasının bulgularından, okuduđunun anlama ve cümle tamamlama ile tanımlanan okuma puanları tanımlamasının hıtarlı olduğu ve cümle tamamlamasının okuduđunu anlama kadar iyi bir kelime bilgisi (vocabulary) ölçme arası olduğu sonucuna varılmıştır (Schrdder 1984: 32).

### *Problem Cümlesi*

Eş anlamlı sözcüğü bulduran maddelerden oluşan test (EAT), kavramlar arası bađmtıyı yoklatan maddelerden oluşan test (KAT) ve cümle tamamlatan maddelerden oluşan test (CTT) aynı yetenekler birleşimini mi ölçer?

### *Alt Problemler*

1. Test puanları arasındaki ilişkinin derecesi nedir, bu ilişki sıfırdan manidar derecede farklı mıdır?
- 2- Ortak faktör hangi büyüklüktedir? İkinci ve üçüncü dereceden faktörler var mıdır, varsa nelerdir?
- 3- Testlerin genelleştirilebilirlik teorisi (generalizability theory)'inden yararlanılarak hesaplanan geçerlik katsayısı nedir?

### *Yöntemi\*)*

#### *Evren ve Örneklem*

Bu araştırma, temel nitelikte bir araştırma olduğundan evren ve örneklem ilişkisi sadece araştırmanın yapıldığı okullar açısından göz önüne alınmıştır. Bu çalışmada her üç test tipinin karşılaştırılması için, Ondokuzmayıs Üniversitesi Amasya Eğitim Yüksekokulu birinci sınıf öğrencilerinin tamamı alınmış ve Meslek Yüksekokulu birinci sınıf öğrencilerinden de bir grup random olarak seçilmiştir. Toplam denek sayısı 249'dur.

#### *Çözümleme Yöntemleri*

Bu çalışmada, araştırmacı tarafından geliştirilen ve sözel yeleneğin ölçülmesinde kullanılan eş anlamlı kelimeler bulduran test (EAT), kavramlar arası bağıntıyı bulduran test (KAT) ve cümle tamamlatan maddelerden oluşan test (CTT)'in (Tekindal 1992: 27) yapı geçerlikleri üzerinde durulmuştur.(\*).

Araştırmanın birinci alt probleminin çözümü için, Peorson Moment çarpımı Korelasyon katsayısı (r), ikinci alt probleminin çözümünde faktör analizi (Sentroid metod), üçüncü alt problemin çözümünde genelleme **Dilirlik** katsayısını hesaplamak içinde genelleştirilebilirlik (generolizibility) teorisinde belirtilen **iki** boyutlu varyans analizi tekniğinin tek gözlemlili modeli (Crocker ve Algina 1986: 160 ve 234-235; Koksall 1985: 333) kullanılmıştır.

\* Bu yazının hacmini genişletmemek için, araştırmanın önemi, modeli, tanımlar, kısaltmalar, sayıtlar, sınırlamalar ve testlerin geliştirilmesi ile uygulanma şekline burada yer verilmemiştir.

\* Bu testler araştırmacının doktora tezi için geliştirdiği araçlardır.

Buna bađlı olarak da "G.Study" ve D-Study" (Crocker ve Algina 1986: 234-235) düzenlenmiştir. Pearson Momentler çarpımı korelasyon katsayısının sıfırdan manidar derecede farklı olup olmadığını test etmek için "t" testi (Akhun 1981: 5) uygulanmıştır.

### *Bulgular*

Sözel yeteneđin ölçülmesinde kullanılan eş anlamlı kelimeleri bulduran maddelerden oluşan test (EAT), kavramlar arası bađındıyı bulduran maddelerden oluşan test (KAT), ve cümle tamamlatan maddelerden oluşan test (CTT)'e ait bulgular ve bunlara dayalı yorumlar şöyledir:

Birinci alt probleme ilişkin olarak testlerin birbirleri ile olan ilişkisi pearson momentler çarpımı korelasyonu ile hesaplanmıştır.

*Tablo-1*

*Test Puanlarının Birbirleri ile Olan Korelasyon Katsayıları (N= 249)*

	E A T	K A T	C T T
E A T	—		
K A T	0,35+	—	
C T T	0,43+	0,43 +	

+P < 0,01

Bu hesaplama sonucunda, EAT ile KAT arasındaki korelasyon 0.35, EAT ile CTT arasındaki korelasyon 0,43 ve KAT ile CTT arasındaki korelasyon da 0,43 bulunmuştur. Bu korelasyonlar 0,01 düzeyinde manidardır. Buna göre, testler birbirleri ile aynı yetenekleri yani dil yeteneđini ölçtüđü gibi, yine dille ilgili birbirinden farklı yetenekleri de ölçmektedir, denilebilir.

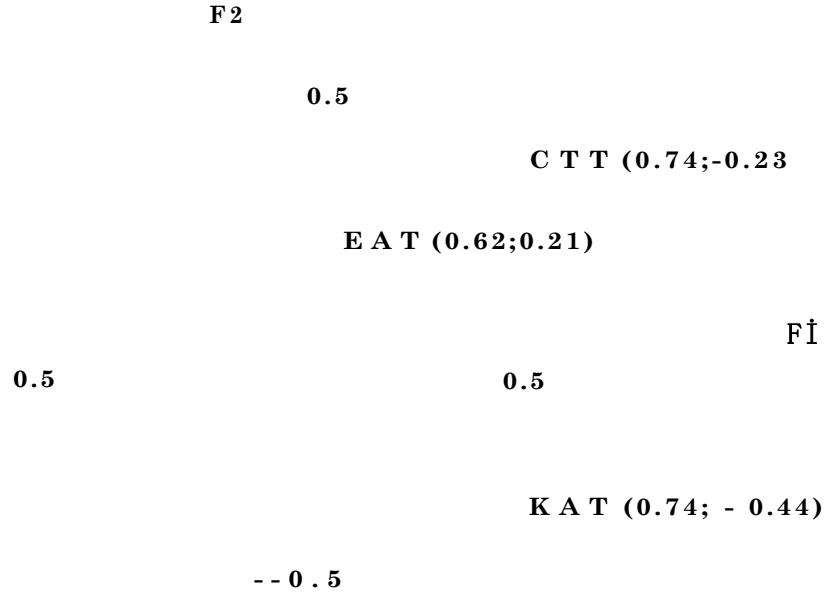
İkinci alt probleme ilişkin olarak, sentroid metodla faktör analizi yapılmıştır.

*Tablo-2*

*Testlerin Faktör Deđerleri (N= 249)*

	F1	F2	F3
EAT	0,62	0,21	0,29
KAT	0,71	-0,44	0,00
CTT	0,74	0,23	-0,29

Tablo 2'de görüldüğü gibi, birinci faktör yükleri yüksek, faktör yükleri arasındaki farklar düşük ve pozitifdir. Bu, EAT, KAT ve CTT testleri için birinci faktörün ortak olduğunu ifade eder. Bu ortak faktöre, dil yeteneği faktörü denilebilir. İkinci faktör de EAT ve CTT'in faktör yükleri düşük, KAT'in faktör yükü yüksek fakat negatif işaretlidir. O halde ikinci faktör iki kutupludur. Üçüncü faktör yorumlanmamıştır. Elde edilen faktörlerin grafik gösterimi Şekil 1'de görülmektedir.



#### Şekil 1' Faktör Analizi Sonuçlarının Grafik Gösterimi

Şekil 1'de görüldüğü gibi EAT ile CTT aynı bölgeye düşmüşlerdir. KAT ise ayrı bir bölgededir. Bu durum, KAT'in EAT ve CTT'ten farklı bir nitelik ölçtüğü şeklinde yorumlanabilir.

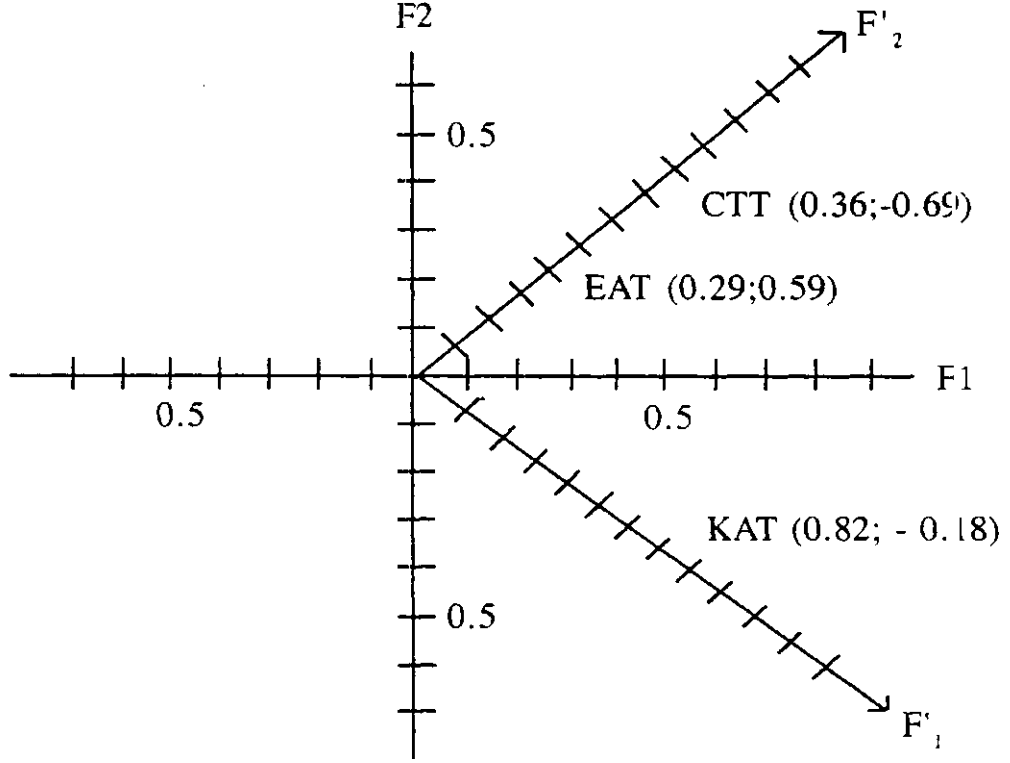
Döndürülmüş eksenlere göre yeni faktör yükleri de şöyledir.

Tablo-3

#### Döndürülmüş Eksenlere Göre Testlerin Faktör Değerleri

	F1	F2	F3
EAT	0,29	0,59	0,29
KAT	0,82	0,18	0,00
CTT	0,36	0,69	-0,29

Tablo 3'te görüldüğü gibi, dönüştürme durumunda, EAT ve CTT'in birinci faktör yükleri düşük, KAT'inki ise oldukça yüksektir. İkinci faktör yüklerinde ise tersi bir durum vardır. Üçüncü faktör önemsizdir.



Şekil 2: Döndürülmüş Eksenlere Göre Yeni Faktör Yükleri

EAT kelimelerin anlamım bulabilme ağırlıklı, CTT anlamlı cümle kurabilme ağırlıklıdır. Bu nedenle, bu iki testin daha çok yük aldığı F/2 faktörü dil yeteneğinin Türkçe'yi anlama ve kullanma davranışlarını ölçtüğü söylenebilir. KAT'de dil yeteneğine dayanmakla birlikte, usavırma (reasoning) ağırlıklıdır; öyleyse F/1 faktörü bir tür genel yetenek faktörüdür., denilebilir.

Üçüncü alt ilişkin olarak da genelleştirilebilirlik (generalizability) teorisi ile ilgili olan "G-study ve D-Study" den yararlanılmış ve genelleştirilebilirlik katsayısı hesaplanmıştır. İki yönlü tek gözlemlili varyans analizi tekniği ile hesaplanan sonuçlar Tablo 3'de gösterilmiş olup, F istatistikleri satırlar için 3,03, sütunlar için 17,60 bulunmuştur.

Tablo-4

*Testlerin Puanlarına İlişkin n = l için İki Boyutlu Varyans Analizi  
Tablosu  
(N=249)*

Varyans Kaynağı	SD	KT	KO	F
Satırlar	248	11355,53	45,78	3,03
Sünunlar	2	530,70	265,35	17,60
Kalan	496	7476,62	15,07	
Toplam	746	19362,85		

Sd: 248,496 a: 0,01 Fr = 1,32

Sd: 2,496 a: 0,01 Fc = 4,66

Bu değerler manidardır. Genelleştirilebilirlik katsayısı (&) 0,57' dir EAT, KAT ve CTT aynı yapıyı (sözel yeteneği) ölçmek üzere geliştirilmiş farklı maddelerden oluşan testlerdir. Bu bakımdan, genelleştirilebilirlik katsayısı birbirine yaklaşan (convergent) geçerlik katsayısı olarak düşünülebilir (Crocker ve Algina 1986: 235). Bu, ortalama birbirine yaklaşan (the average convergent) geçerlik katsayısı, üç testin birlikte olan geçerliği olarak yorumlanabilir. O halde, testler ortak bir yeteneği yani dil yeteneğini ölçtüğü gibi, farklı yetenekleri de ölçüyor, denilebilir. Bunlar, yine dile dayalı olarak okuduğunu anlama, dili kullanma ve usavurma davranışları olabilir.

Özet olarak, bu araştırmada eş anlamlı kelimeleri bulduran maddelerden oluşan test (EAT), kavramlar arası bağıntıyı bulduran test (KAT) ve cümle tamamlatan maddelerden oluşan test (CTT)'in yapı geçerlikleri çalışmıştır.

EAT ile KAT, EAT ile CTT ve KAT ile CTT arasındaki Pearson Momentler Çarpımı korelasyon katsayıları sırasıyla 0,35, 0,43 ve 0,43 bulunmuştur. Bu korelasyon katsayıları 0,01 düzeyinde N= 248 için manidardır. Testlerin genel dil yeteneği yanında yine dille ilgili bazı yetenekleri de ölçtüğü söylenebilir.

EAT, KAT ve CTT' in puanlarına uygulanan faktör analizi sonuçlarına göre, anılan testlere ait üç faktör bulunmuş ve genel faktör dil yeteneği faktörü olarak yorumlanmıştır. İkinci faktörde EAT ve CTT' in faktör yükleri düşük, KAT' in faktör yükü yüksek fakat negatif işaretlidir. Bu faktör iki kutupludur. Üçüncü faktör yorumlanmamıştır.



Dönüştürme durumunda, KAT için F/1 yüksek, EAT ve CTT için ise düşüktür. F/2 de tersi bir durum vardır. F/3 önemsizdir. Bu durumda EAT ve CTT'nin dili kullanma ve anlama davranışlarını, KAT'in ise yine dile dayanan usavurma davranışını ölçtüğü düşünülebilir. F/1 faktörü bir tür genel yetenek faktörüdür, denilebilir.

Genelleştirilebilirlik teorisinden yararlanılarak hesaplanan ve geçerlik katsayısı olarak yorumlanan genelleştirilebilirlik katsayısı 0,57'dur. Buna göre, bu testler ortak bir yeteneđi, yani dil yeteneđini ölçüyor denilebilir.

### *Sonuç ve Öneriler*

Bu araştırmadan elde edilen bulgulara göre şu sonuçlara varılmıştır:

Eş anlamlı kelimeleri bulduran (EAT), kavramlar arası bağıntıyı yoklatan (KAT) ve cümle tamamlatan maddelerden oluşan test (CTT) lerin üç yoldan yapı geçerlikleri hesaplanmıştır. EAT, KAT VE CTT faktör analizine tabii tutulmuş ve üç faktör bulunmuştur. Bu faktörlerden birincisi yani genel faktör dil yeteneđi faktörü, ikincisi dili kullanma, anlam ve usavurma ile ilgili bir faktördür, denilebilir. Üçüncü faktör yorumlanmamıştır, İkinci olarak, anılan testler arasındaki Pearson Momentler çarpımı korelasyonları hesaplanmıştır. EAT ile KAT, EAT ile CTT ve KAT ile CTT e r arasmdaki korelasyonlar sıfırdan manidar derecede büyük bulunmuştur. Bu testlerin genel dil yeteneđi yanında kendilerine özgü dil ile ilgili bazı yetenekleri de ölçtüğü sonucuna varılmıştır.

Son olarak da, genelleştirilebilirlik teorisinden yararlanılarak hesaplanan ve birbirine yaklaşan (convergent) geçerlik katsayısı olarak yorumlanan generalizability katsayısı bulunmuştur. Üç test ile ölçülen yetenek bileşiminin yüzde 57'lik bir kısmı ortaktır. Bu genel dil yeteneđi olabilir. Geri kalan kısmı da, dili kullanma, usavurma ve anlama ile ilgili olabilir.

Sonuç olarak, faktör analizi sonuçları, testlerin birbirleri ile olan korelasyonları ve genelleştirilebilirlik teorisi sonuçlarından elde edilen geçerlik katsayısı, bu testlerin belli bir genel dil yeteneđini ölçtüğünü gösteriyor denilebilir. Ayrıca testlerin kendilerine özgü dille ilgili bazı yetenekleri de ölçtüğü söylenebilir.

*Yukarıdaki sonuçlara göre şu öneriler geliştirilmiştir:*

1. Sözel yetenek testlerinde kullanılan diğer madde tipleri ile bu çalışmada kullanılan madde tipleri karşılaştırılabilir.
2. Örnekleme aynı seviyedeki çok sayıda farklı gruptan oluşturularak anılan testlerin yapı geçerliğine bakılabilir.

**KAYNAKÇA**

**A K H U N , İlhan**

- 1981 İstatistik Formüller ve Tablolar.**  
(İkinci Baskı), Ankara.

**ADIGÜZEL, Muzaffer**

- 1985 "Madde Yapısının Tesat Güvenirliğine ve Geçerliğine Etkisi".**  
(Yayınlanmamış Doktora Tezi)  
Ankara: Hacettepe Üniversitesi.

**CROCKER, Linda ve James Algina**

- 1986 Introduction to Classical and Modern Test Theory**  
New York. CBS Colege Publishing

**KOKSAL, Bilge Aloba**

- 1985 İstatistik Analiz Metodları**  
(Üçüncü Baskı),  
İstanbul: Çağlayan Kitabevi

**LORD, M.Frederik and Cherly L. VVild**

- 1985 Conhtribution of Verbal Item Types in the G R E**  
**General Test to Accuarcy of Measurement of the**  
**Verbal Scores**  
Princeton: Educational Testing Service.

**SCHRADER, B. VVilliam**

- 1984 Three Studies of SAT-Verbal Item Types.**  
New York: College Board Report No: 84-7 ETS  
RR No: 84-33.

**TURGUT, M.Fuat**

- 1971 Şans Başarısının Test Puanlarına Etkisi.**  
Ankara: Orta Dođu Teknik Üniiversitesi,  
Fen ve Edebiyat Fakültesi Yayın No:22
- 1977 Zekanın Mahiyeti ve Ölçülmesi.**  
(Çođaltma),  
Ankara: H.Ü. Eğitim Bölümü.
- 1986 Eğitimde Ölçme ve Deđerlendirme Metodları.**  
(Dördüncü Baskı),  
Ankara: Saydamı Matbaası.

**TEKİNDAL, Satılmış**

- "Madde Çeşidinin Test Güvenirliğine ve Geçerliğine Etkisi".**  
(Yayınlanmamış Doktora Tezi),  
Ankara: Hacettepe Üniversitesi