

## Yükseköğretimde Öğretimi Değerlendirme Anketlerinde Öğrenci Algılarındaki Farklılıkların Puanlamaya Etkileri<sup>(\*)</sup>

Nurdan Kalaycı

*Bu çalışma, yükseköğretim kurumlarında kullanılan öğretim elemanlarını ve öğretimi değerlendirme anketlerindeki maddelere öğrenciler tarafından verilen puanların, puanlara verilen nedenlerin, nedenler arasındaki ilişkilerin analizinin yapılmasını ve öğrencilerin maddeleri okuduklarında, maddeler hakkındaki ilk yorumlarından, puanların birinin seçimine karar verme sürecinde etkili olan durumları saptamayı amaçlamıştır. Örnekleme Gazi Üniversitesi Gazi Eğitim Fakültesi İlköğretim Bölümü Sınıf Öğretmenliği, Okul Öncesi Eğitimi ve Sosyal Bilgiler Öğretmenliği Anabilim Dalı 3. sınıf öğrencilerinden oluşmuştur. Çalışmanın verileri aynı yıl içerisinde ardışık olarak iki farklı dönemde toplanmıştır. Bu iki ayrı dönemde yürütülen araştırmada aynı veri toplama aracı kullanılarak veriler içerik analizi ile çözümlenmiştir. Bulgular puanlar için verilen nedenlerin (yorumların) içerik analizi yapılması sonucu elde edilmiş ve her bir madde için matris oluşturulmuştur.*

*Ankette her madde için puan verme nedenlerinin farklılığı ve fazlalığı, öğrencilerin hatırlayamadıkları ve emin olmadıkları durumlarda bile puan verdiği, maddedeki binişikliğin, puanlamada kararsızlığa neden olduğu, öğrencilerin bilgileri olmadığını belirttikleri maddelere puan verdikleri, öğrenciler sınavlarda bekledikleri notu aldıklarında öğretim elemanına yüksek not verdiği, açık olarak kabul edilen maddelerde bile nedenler arasında zıtlıkların olduğu ve diğer araştırmalardan daha farklı hale etkisinin ortaya çıkması bu çalışmanın dikkat çekici sonuçlarıdır. Sonuçlar kötü yapılandırılmış problem çözme süreçleri temele alınarak yorumlanmıştır.*

**Anahtar sözcükler:** Yükseköğretimde öğretimi değerlendirme, öğretimi değerlendirme anketi, öğretimin öğrenciler tarafından değerlendirilmesi

---

<sup>(\*)</sup> Bu makalenin bir kısmı İngilizce olarak World Applied Science Journal'ın cilt 3, no:1 2008 sayısında yayınlanmıştır.

### **Students' Evaluation of Teaching in Higher Education: Effects of Students' Perception on Ratings**

*This research aims at identifying the situations influential over the ratings given by students to the items on questionnaires employed to evaluate instructors and teaching in higher education institutions as well as reasons for giving the ratings, the process of initial comments of the items on first reading and choosing the appropriate rate, and the current research also aims at analyzing the relations between the reasons. The research sample was composed of the junior students of Elementary Education Class Teaching, Pre-School Education and Social Sciences Teacher Education branches of Primary School Teaching Department of Gazi Teacher Training College of Gazi University. Research data were gathered in two successive semesters in the same academic year. The same tool of data collection was employed in both of these two semesters, and the content analysis of the data was conducted. Following a content analysis of the reasons for attributing rates, findings were obtained, and a matrix for each item was formed.*

*The fact that reasons for rating were various and too many, that students gave rates even when they were not sure, that overlapping items caused indecisiveness in rating, that they gave rates even to those items of which they had no knowledge, that they rated higher to those instructors who gave them the marks they had expected, that there were contrasts among the reasons for the items which are clearly agreed upon, and that a different halo effect emerged in this research was of importance in this research.*

**Keywords:** *Assessment of instruction in higher education, instructional assessment scales, assessment of instruction by students*

Yükseköğretim kurumlarında öğretim ve öğretim elemanın değerlendirme yöntemlerinden biri öğrenci değerlendirme anketleridir. Dünyada birçok üniversitede öğrenci değerlendirme anketinden elde edilen veriler, öğretim elemanının ve öğretiminin kalitesinin değerlendirilmesinde yaygın biçimde kullanılmaktadır (Seldin,1985; Abrami,1989a; Wagenaar,1995; Neath,1996; Hobson ve Tablot, 2001; Grant,1998; Platt,1993). Özellikle son yirmi yıl, bu alanda çok sayıda deneysel (ampirik) ve kavramsal (conceptual) çalışmaya tanık olmuştur. Öğretimi değerlendireme anketleri ilk olarak ABD’de Purdue Üniversite’sinde Remmers ve meslektaşları tarafından uygulanmaya başlamış ve ilk bilimsel çalışmalar da yine Remmers tarafından 1927, 1928, 1930 tarihlerinde yapılmıştır. Bu konudaki çalışmalar o tarihten günümüze yoğunlukla devam etmektedir. Bu çalışmalardan bazıları Avustralya’da (Bedgood ve Polard,1999; Haynes,2002; Marsh,1987; Marsh ve Bailey,1993; Neumann,2000; Wagner,1999; Wortington, 2002) ve 1990 yıllardan bu yana Avrupa’da yapılmaktadır (Husbands, 1996, 1997; Husbands ve Fosh, 1993; Shevlin, Banyard, Davies ve Griffiths, 2000).

Dünyada iki bini aşkın çalışmanın bulunduğu bu konu ile ilgili olarak ülkemizde bilimsel olarak bir çalışmaya rastlanmamıştır. Bu konudaki çalışmaların çoğu anketlerin geçerliliği güvenilirliği hakkında olmasına karşın anketlerin güvenilirliği ve geçerliği hala bir endişe konusudur ve anket sonuçların kesinliği hakkında şüpheler oldukça fazladır (Habson ve Talbot, 2001; Block 1998). Literatürdeki hata faktörleri konusunda yoğunlaşma için Block (1998) aşağıdaki açıklamayı yapmıştır. Eğitimde araştırma yöntemleriyle ilgili kitapları incelediğimizde bu kitaplarda anketlere ilişkin kısımlar görürüz. (örneğin Hopkins,1985/1993; Cohen ve Manion,1989/1994; Nunan,1992; McKernan,1996; McNiff ve diğerleri,1996; Breawell ve diğerleri,1995; Vaus,1996; Coolican,1994).

Tüm bu yayınlarda ankette yer alan maddelerin ifade ediliş tarzı ve formatı, ya da ölçek oluşturma kısmına odaklanma eğilimi görülür. Fakat bireylerin anket doldurma veya puanlama zamanı geldiğinde anketle nasıl bir etkileşime girecekleri konusu gibi ustalık isteyen konulara girmekten kaçınmaktadırlar (Block,1998,s. 403).

### **Öğrenci Değerlendirme Anketlerine İlişkin Araştırmaların Konularına Göre Dağılımı**

Öğrenci değerlendirme anketlerine ilişkin araştırmalar nitel ve nicel olmak üzere iki ana başlık altında ele alınabilir. Bu konudaki araştırmaların önemli bir kısmı nicel özellik taşımaktadır. Bu araştırmalar incelendiğinde

ölçme aracının geliştirilmesi ve geçerliği (Marsh, 1987), güvenilirlik konusundaki genellenebilirlik ve öğretimin etkililiğinin ölçülmesinde öğrenci puanlarının geçerliği (Kulik ve McKeachie, 1975; Cohen, 1981; Braskamp, Ory ve Pieper, 1981; Murray, 1983; Feldman, 1989a,b; Marsh ve Dunkin, 1992), anketlerin güvenilirliği (Feldman, 1987; Sixbury ve Cashin, 1995a-b), güvenilirliğin istikrarı (Braskamp ve Ory, 1994; Centra, 1993), öğrenci değerlendirme anketlerinde yaygın olarak bulunan hata faktörleri (Centra, 1993, Braskamp ve Ory, 1994), anketteki maddelerin özelliği ve sayısı gibi (Abrami ve d'Apollonia, 1991) konularını kapsadığı görülmektedir.

Öğrencilerin puan verirken hangi değişkenlerden etkilenebileceği konusunda, yapılan çalışmalar dört grupta toplanmıştır: Öğrenci, öğretim elemanı, ders, diğer değişkenler (idari, çevresel vb).

Bu çalışmada, yalnızca öğrenci değişkenleri ile ilgili araştırmalara yer verilmiştir. Öğrenci değişkenleri ile ilgili kontrol altında tutulabilen ve nicel olarak analiz edilen araştırmalara daha fazla rastlanmaktadır. Örneğin; öğrencinin yaşı (Centra, 1993), cinsiyeti (Feldman, 1977, 1993), sınıf düzeyi (McKeachie, 1979), dersin düzeyi, (Aleamoni, 1981; Braskamp ve Ory, 1994; Feldman, 1978), GPA [Grade-point average veya ratio bir çok üniversitede tarafından kullanılan, öğrencinin genel başarısını bulma amaçlı sistemdir] (Feldman, 1976), akademik alanı (Braskamp ve Ory, 1994; Cashin, 1990; Centra, 1993; Marsh ve Dunkin, 1992; Sixbury ve Cashin, 1995b), dersin zorluk derecesi (Centra, 1993; Marsh ve Dunkin, 1992), sınavlarda beklediği notları (Braskamp ve Ory, 1994; Howard ve Maxwell, 1980, 1982).

Bireysel farklılıklara göre değişebilen algı, güdülenme, kişilik özellikleri üzerinde yapılan araştırmalar diğerlerine göre daha az sayıdadır. Çünkü bu özelliklerin ölçülmesi diğer özelliklerin ölçülmesine göre daha zordur. Bu araştırmalar; öğrencinin konuya karşı ilgisi (Marsh ve Dunkin, 1992), algısı, güdülenme, öğrencilerin duygusal durumları (Kember ve Wong, 2000; Chen ve Hoshower, 2003; Robertson, 2004; LaForge, 2003), öğrencinin kişiliği (Abrami, Perry ve Leventhal, 1982), öğrenci puanlarındaki potansiyel hatalar gibi (Hoffman ve Kramer, 1980; Abrami ve Mizener, 1983) konuları kapsamaktadır.

Anketi dolduranların düşünce süreçleri ile çok az araştırma bulunmaktadır. Öğrencilerin, anket maddelerini nasıl yorumladığı konusunu Low (1988, 1991, 1996 ve 1999); Solas (1990); Alderson; (1992); Block (1998); Robertson (2004) gibi çok az araştırmacı çalışmıştır.

Literatürdeki bu boşluğu doldurmayı amaçlayan Alderson, anketlerde katılımcıların maddeleri nasıl yorumladıklarını anlamaya çalışan ilk araştırmacıdır.

macılardan birisidir. Alderson çalışmasının en önemli sonucunu şöyle ifade etmiştir:

Cevapların bir listeden seçim yapmak ya da bir ölçekten değerlendirme yapmak gibi kapalı olduğu anketlerde daha büyük gizli yorum problemleri vardır. Katılımcıların seçtikleri cevaplar kendi seçtikleri cevaplar olmayabilir. Yani farklı bir ölçüt verilse farklı cevaplar seçebilirler. Anketi cevaplayanlar seçtikleri cevaplarda anketi hazırlayanlarla aynı anlamı kastetmeyebilir. Fakat puanın kendisi incelendiğinde bu problemlerden hiç birisi göze çarpmayacaktır. Aynı şekilde değerlendirme ölçeklerinin kullanımında – ölçeklerde kullanılan puanların sözel açıklamalarına rağmen maddeyi hazırlayan ile cevaplayanın yükledikleri anlamlar aynı olmayabilir (Alderson, 1992,s. 4-5).

Anketlerin ifade ediliş tarzı, formatı ve ölçek oluşturma gibi çalışmaların bir istisnası da Low'un (1988, 1991, 1996, 1999) araştırmalarıdır. Alderson (1992) ve Low (1996), bu konudaki çalışmaların yetersiz olduğunu açıklamışlardır. Low' un ifade ettiği gibi;

Anket hazırlayıcılarının bilmesi gereken şey; kendilerinin bir maddede bilgi akışını sınırlamak ve retorik etki yaratmak amacıyla kullandıkları kelimelere deneklerin nasıl tepki verdikleridir.....Deneklerin verdikleri cevapları vermelerine nelerin yol açtığıнын bulunması önemlidir (Low, 1996, s.2).

Anket doldurma işi bir iletişim süreci olarak düşünülürse kaynaktan alıcıya ulaşan mesajın yani anketin okunmasından sonra yaşanabilecek olası etkilerin neler olabileceğinin saptanmasının önemini Alderson yukarıda net bir şekilde ifade etmiştir. Anket puanlamadaki önemli bir başka sorunu ise Alderson, aşağıda şöyle ifade etmiştir.

Hatta bir katılımcı ölçeğin uç noktalarını kullanmayı seçen başka bir katılımcının aksine orta puanlarda kalma eğilimi gösterebilir de; her ikisi de aynı maddeye aynı tavrı gösterebilir. Böylece bir kişinin seçtiği 3' ün anlamı bir başkasınınkinden farklı olabilir. Ama puanların istatistiksel analizinde, frekansların sayılıp ortalamaların hesaplanmasında benzer puanların aynı anlam taşıdığını varsayarsanız (Alderson 1992, s. 4–5).

Block (1998) ve Robertson (2004) ile birlikte bu çalışma, Alderson'un (1992) varsayımlarını destekler niteliktedir. Bu anlamda, her ne kadar titiz bir analiz yapılsa da yalnızca istatistiğe dayalı değerlendirmeler, formative ve summative amaçlar için istenen düzeyde sonuçlar sağlamayabilir.

Block, Alderson'un (1992) bulgularına ilgi duyarak benzer bir çalışma yapmış ve sonuç olarak şöyle bir açıklama getirmiştir:

Bu anketin ilk tasarımcılarının ifade tarzları ile neyi kastettiklerini bilmeyebilirim ama onların yorumlarının bu çalışmaya katılan öğrencilerin verdiği çok boyutlu yorumlardan daha dar çerçeveli olacağını biliyorum” (Block, 1998, s.424).

Ayrıca bu probleme açıklama getirilebilecek olan bir diğer çalışma ise Robertson'a aittir. Alderson (1992), Low(1996) ve Block'tan(1998) farklı olarak Robertson'da (2004) öğrencinin anketlere verdikleri puanların nedenlerini, problem çözme süreçlerini temele alarak açıklamıştır.

Bu çalışmanın, amaçlarından biri yukarıda bahsedilen konulardaki alan yazın eksikliğine katkı sağlamaktır. Çalışma anketlerdeki kapalı uçlu soruları öğrencilerin nasıl algıladıklarını, puanlamada kullandıkları kişisel ölçütleri öğrencilerin yorumlarından anlamayı amaçlamıştır.

Böylece sorulara ilişkin öğrenci algısının değişiminin incelenmesi geçerli ve güvenilir anketlerin hazırlanmasına bunun sonucu olarak formative ve summative\* değerlendirmelerin daha objektif olması açısından önemli olacaktır.

### **Problem Çözme Süreci Bakış Açısıyla Anket Doldurma İşi**

Bu çalışmada Robertson (2004), Newell ve Simon'ın (1972,), öğretim elemanı değerlendirme anketini doldurma sürecini, belirsizlikler içinde bir yargıya varma olarak kabul eden görüşleri temele alınmıştır. Öğrencilerin anketteki maddeler bazında alternatif çözümler üretmeleri ve ıraksak (divergent) düşünme tiplerini ortaya koymaları, anket doldurma işinin *kötü yapılandırılmış problem çözme* (ill-structured problem) özelliği taşıdığını ortaya koymaktadır. Kitchener'e (1983) göre farklı çözümlere yol açabilen çelişkili varsayım, kanıt ve fikirlerin olduğu problemler “kötü yapılandırılmış problem”lerdir.

Jonassen'in de (2000) vurguladığı gibi, kötü yapılandırılmış problem çözme sürecinde yer alan her bir öğrencinin; aşinalık, bulanıklık, işlem ba-

---

\* **Formative değerlendirme:** Öğretim elemanı değerlendirme anket sonuçları doğrultusunda, öğretim elemanına ve dersine ilişkin gerekli görülen düzeltmelerin veya güçlendirmelerin yapılması amacıyla işe koşulan değerlendirme.

**Summative değerlendirme:** Öğretim elemanı değerlendirme anket sonuçları doğrultusunda, öğretim elemanının terfi etmesi, sözleşmesinin uzatılması ve ödüllendirilmesi amacıyla işe koşulan değerlendirme.

samakları bilgisi, bilişsel stiller, özgüven, motivasyon, azim gibi bireysel farklılıklar çözümlerde farklılaşmaya yol açmaktadır.

Jonassen (2000)'in karmaşık problem tanımı göz önüne alındığında, anketlerdeki maddeleri puanlama işlemi, ilk bakışta karmaşık görünmese de aslında karmaşıktır. Çünkü maddelerle ilgili konu, işlemler, değişkenlerin sayısı, bu özellikler arası bağlantılılık derecesi ve işlevsel ilişkiler, puan verilecek madde ile ilgili alana özgü özellikler arasındaki ilişkilerin tamamı düşünülürse bu sürecin karmaşık olduğu görülür.

Kişiler, bu karmaşık karar verme sürecinde seçtikleri çözümü destekleyebilmek için neden ya da nedenler ileri sürmek zorundadır. Bu nedenler bilişsel nitelik taşıdığı gibi, bilişsel olmayan nitelik de taşıyabilir. Jonassen ve Tessmer (1996) kötü yapılandırılmış problem çözmede bilişsel ve metacognitif süreçlerin gerekli olduğunu; ancak bunların yetersizliğini vurgulayarak duygusal ve isteksel süreçlerin de gerekli olduğunu belirtmiştir.

Yukarıda belirtilen bu gerekçelerle analiz sonuçlarının, kötü yapılandırılmış problem çözme süreçleri temele alınarak açıklanmasına karar verilmiştir. Şu ana kadar açıklanmaya çalışılan olguları kısaca Low'un bir cümlesi ile özetlemek gerekirse,

*Belli soruların cevaplarının istatistiksel anlamda önemli hatalara yol açıp açmadığını bilmek kesinlikle önemli olsa da bundan daha önemlisi; cevap verenleri, verdikleri cevaba yönlendiren şeyin ne olduğunu bulmaktır (Low 1999, s.2).*

Bu amaca ulaşmak için kullanılan öğrenci yorumlarına dayalı analiz yapma yöntemi; öğretim elemanlarına, anketi değerlendirenlere ve hazırlayanlara standart bir değerlendirme anketindeki puanların ötesinde zengin bir bilgi sağlayacağı düşünüldüğü için önemli bulunmaktadır.

Bu çalışma, öğrencilerin anket sorularını puanlamalarındaki ölçütlerinin belirlenmesi, anketleri hazırlayanlara ve yorumlayanlara öğrencilerin etkili öğretimden ve etkili öğretim elemanından ne anladıklarını göstermesi açısından da önemlidir. McKeachie'nin (1997) iddia ettiği gibi değerlendirme anketine puan veren öğrencilerin zihinlerinde olup bitenler üzerine niteliksel bir çalışma, öğrencilerin anket maddelerini nasıl anladıklarını ve nasıl bir anlam yüklediklerini ortaya çıkaracaktır.

Bu çalışma, yükseköğretim kurumlarında kullanılan öğretim elemanlarını ve öğretimi değerlendirme anketlerindeki maddelere öğrenciler tarafından verilen puanların, puanlara verilen nedenlerin, nedenler arasındaki ilişkilerin analizinin yapılmasını ve öğrencilerin maddeleri okuduklarında,

maddeler hakkındaki ilk yorumlarından, puanların birinin seçimine karar verme sürecinde etkili olan durumları saptamayı amaçlamıştır.

Bu çalışma belirtilen amacı gerçekleştirmek için aşağıdaki sorulara cevap aramıştır.

1. Öğretim üyesi için verilen genel puan, form ve ankete verilen puanlar arasında bir fark var mıdır?
2. Ankette yer alan maddelerin ifadeleri öğrenci algılarını nasıl etkilemektedir?
3. Anketteki maddelerin ifadesinden bağımsız olarak öğrenci algıları nasıl bir değişim göstermektedir?
4. Ankette yer alan maddelere öğrenciler tarafından verilen puanlar ve puanlara verilen nedenler arasında nasıl bir ilişki vardır?

## **Yöntem**

Araştırma betimsel nitelikte bir çalışmadır. Bu çalışma iki ayrı aşamada tamamlanmıştır. Çalışmanın birinci ve ikinci araştırma sorusuna cevap aranırken ortaya çıkan sonuçlar, araştırmacıyı üçüncü ve dördüncü sorulara da cevap aramaya yönlendirmiştir. Bu nedenle veriler, birbirini izleyen iki farklı tarihte, örneklem sayısındaki farklılık dışında aynı bölüm, aynı öğrenciler ve aynı sınıflardan toplanmıştır. İki farklı tarihte toplanan verilerin birbiri ile olan ilişkilerinin bir bütünlük içinde aynı çalışmada verilmesinin nedeni, akademisyenleri daha yararlı yorumlara ve tartışmaya yönlendirmesi düşüncesidir.

Araştırmanın evreni Gazi Üniversitesi Gazi Eğitim Fakültesinde öğrenim gören öğrencileridir. Birinci ve ikinci araştırma soruları için Gazi Üniversitesi Gazi Eğitim Fakültesi İlköğretim bölümünden Sınıf, Okul Öncesi ve Sosyal Bilgiler Öğretmenliği Anabilim Dalı 3. sınıf öğrencilerinde araştırmaya istekli olarak katılan 192 öğrenci örneklem olarak alınmıştır. Örneklem grubunun, 3. sınıf öğrencilerinden seçilmesinin nedeni, 1. ve 2. sınıflardaki öğrencilerin etkili öğretimi değerlendirmede 3. ve 4. sınıflar kadar deneyimli olmadıkları düşüncesidir. 4.sınıflardan seçilmemesinin sebebi ise, örneklem için yeterli sayıda öğrenci bulunamamasıdır.

İkinci ve üçüncü araştırma soruları için ise, Gazi Üniversitesi Gazi Eğitim Fakültesi İlköğretim bölümünden Sınıf, Okul Öncesi ve Sosyal Bilgiler Öğretmenliği Anabilim Dalı 3. sınıf öğrencilerinden, araştırmaya istekli olarak katılan 138 öğrenci örneklem olarak alınmıştır. 3



sınıf öğrencilerinin amaca en uygun veri toplanacağı düşüncesi ile amaçlı olarak seçilmiştir(Purposive Sampling Approach).

Çalışmanın 1. ve 2. araştırma sorusuna ilişkin veriler 2004–2005 öğretim yılı güz döneminde; 3. ve 4. araştırma sorusuna ilişkin veriler ise 2004–2005 öğretim yılı bahar döneminde toplanmıştır.

#### *Birinci ve İkinci Araştırma Sorusu İçin Kullanılan Veri Toplama Araçları*

Araştırmanın bu bölümünde Gazi Üniversitesi Gazi Eğitim Fakültesi'nde öğretim elemanı değerlendirme amacıyla kullanılan veri toplama aracı kullanılmıştır. Bu veri toplama aracı iki farklı şekilde adlandırılmıştır. İlki, “Öğretim Elemanı Değerlendirme Formu”, ikincisi ise “Öğretim Elemanı Değerlendirme Anketi”dir. “Form” olarak adlandırılan araçla, öğrencilerin öğretim elemanına ilişkin puan vermeleri; “Anket” olarak adlandırılan araçla ise, verdikleri puanlara ilişkin neden belirtmeleri istenmiştir. Her iki araçta da, birbirinin aynı 20 kapalı uçlu soru yer almaktadır. Her iki araçta da Likert tipi 5’li dereceleme ölçeği bulunmaktadır (5 mükemmel, 4 çok iyi, 3 iyi, 2 orta, 1 zayıf). Öğrenciler güz döneminin son haftasında, ders saatinde ilk önce formu puanlamış ve daha sonra ankette hem puanlama yapmış hem de verdikleri puanların nedenlerini yazmışlardır. (Ek 1)

Öğrencilerden formu doldurmaya başlamadan önce, ilk olarak ders için 1–5 arasında “genel değerlendirme” puanı vererek formun üst kısmına yazmaları istenmiştir ve daha sonra diğer soruları puanlamışlardır.

Anketi doldururken daha önce forma verdikleri puanlara sadık kalarak ankete puan vermeleri istenmiştir. Ancak ankette herhangi bir puan değişikliği yapmak istediklerinde yapabilecekleri de belirtilmiştir. Öğrencilerden ankete vermiş oldukları puanların nedenlerini, her bir kapalı uçlu sorunun yanındaki boşluğa mutlaka yazmaları istenmiştir. Araştırmanın amacı, form ve anket doldurma işinden önce, öğrencilere açıklanmış, öğrencilerden anlamadıkları ve ders kapsamına uymayan kapalı uçlu soruları puanlamaları istenmiştir. Öğrenciler form ve ankete isimlerini yazmamışlardır.

#### *Üçüncü ve dördüncü araştırma sorusu için kullanılan veri toplama aracı*

1. ve 2. araştırma sorusuna ilişkin verilerin toplanmasında kullanılan “Öğretim Elemanı Değerlendirme Anketi” 3. ve 4. araştırma sorusu için de kullanılmıştır. İkinci kez aynı anketin aynı örneklem grubuna farklı zamanlarda uygulanmasının nedeni, öğrenci yorumlarının 3. ve 4. araştırma sorusuna uygun olarak analiz edilmesidir.

## Verileri Analizi

Birinci araştırma sorusu için form doldurulmadan önce ders için “genel” olarak verilen puanların ortalaması bulunmuştur (subjective module rating). Daha sonra formlardaki her bir kapalı uçlu soru için hesaplanan genel ortalamaların (20 madde için) tek bir ortalaması hesaplanmış (average module rating) bu puan “genel” olarak verilen puanlar (subjective module rating) ile karşılaştırılmıştır.

İkinci araştırma sorusu için ankette bulunan 20 kapalı uçlu soru için öğrenciler tarafından toplam 3517, 3. ve 4. araştırma sorusu için toplam 2574 yorum yazılmıştır. Nitel çözümleme, 2. 3. ve 4. araştırma sorusu için aynı biçimde yapılmıştır.

Bu çalışmada kullanılan “tema” terimi öğrencinin her soru bazında boşluklara yazdığı orijinal ifadelerini, “kategori” terimi ise öğrenci ifadelerinden çıkarılan fikirlerin ortak özelliklerine göre gruplandırılmayı ifade etmektedir. Bu verilerin analizinde klasik içerik analizi (Conventional Content Analysis) kullanılmıştır.

Van Manen (1990), bu tür bir analizde araştırmacı, öğrencilerin yaşadığı deneyimleri, yazılı yorumlarından bulup çıkarmaya çalışır, demiştir. Her soru düzeyinde öğrencilerin yazdığı yorumlar Latent Coding (meaning, idea) yaklaşımı ile analiz edilmiştir. Daha sonra Emergent Coding; tema olarak adlandırılan boşluklara yazılmış olan öğrenci yorumlarından, kategorize edilecek olan fikir veya anlamların bulunması tekniği temele alınmıştır. Diğer bir ifadeyle temalardaki fikirler, tüme varım yöntemi ile doğal olarak ortaya çıkmıştır.

Bu çalışmada, veriler “puan–neden analiz matrisi” olarak adlandırılan araç yoluyla sınıflandırılmış ve analiz edilmiştir.

Farklı tarihlerde toplanan araştırma verileri aynı işlem basamakları doğrultusunda analiz edilmiştir. Veriler sınıflara göre ayrılmıştır.

- İlk olarak nitel içerik analizi temalardaki fikirlerin sınıflanmasına yarayacak olan kategorilerin ve bunların çıkış frekansları ile birlikte sayılmasına karar verilmiştir.
- Daha sonra verilerin güvenilirliği için araştırmacı bir başka araştırmacı yardımı ile birbirinden bağımsız yorumların tümünü okumuşlardır.

- Ayrıca analizin güvenilirliğini arttırmak için örneklem içinden beş gönüllü öğrenci değişmeli olarak araştırmacılara kategorilerin oluşturulmasında yardımcı olmuştur.
- Her soru bazında, araştırmacılar tarafından birbirinden bağımsız olarak oluşturulan kategoriler karşılaştırılarak yeni kategoriler oluşturulmuştur. Bu aşamada oluşturulan, kategorilerin, temaların ana fikirlerini doğru şekilde temsil ettiklerini kontrol edebilmek için, 20 öğrenci anketi ve oluşturulan ortak kategoriler listesi Eğitim Fakültesindeki iki öğretim üyesine gösterilerek görüşleri alınmıştır.
- Lewis (1991-1992)'in açık uçlu maddeler analizi konusunda önerdiği matrise benzeyen “kategori-puan analiz matrisi” her bir kapalı uçlu soru için hazırlanmıştır. Bu kategorilendirme aşamasında incelenen her bir temanın yanına ilgili kategoriye ait işaret konulmuştur. Böylece herhangi bir karışıklık ve kontrol durumunda kategoriler ile temalar arasında kontrol sağlanabilmiştir.
- Araştırmacılar oluşturdukları her soru bazındaki kendi “kategori-puan analiz matris”lerini karşılaştırarak, her soru için ortak tek bir “kategori-puan analiz matrisi” oluşturmuşlardır (Bkz Tablo-1).

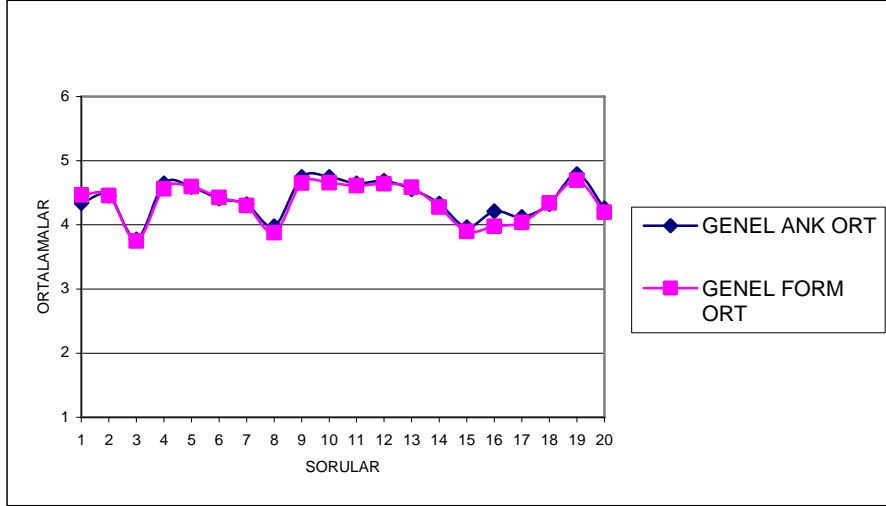
### **Bulgular ve Yorum**

Bu bölümde araştırma soruları doğrultusunda saptanan sonuçlara ve sonuçlara dayalı yorumlara yer verilmiştir.

#### *Öğretim Elemanları İçin Verilen Genel Puanlarla Form Ve Ankete Verilen Puanlar Arasındaki Farklar*

Formlar doldurulmadan önce öğretim elemanı için verilen “genel” puan ortalamaları (4.46) ile formlar doldurulduktan sonra tüm puan ortalamaları (4.35) arasındaki fark 0.11'dir. Daha sonra uygulanan anket puanlarının ortalamaları (4.29) anket doldurulmadan önce öğrenciler tarafından puanlanan form puanlarının ortalamalarının (4.35) arasındaki fark 0.06 bulunmuştur.

Grafik 1'de görüldüğü gibi anket puanlarının ortalamaları ile form puanlarının ortalamaları paralel göstermektedir. Ancak genel ortalamalar arasındaki fark hesaplandığında 0,04 oranında anket lehine bir fark görülmektedir. Bu fark oldukça önemsiz bir değerdir. Bu bölümde formlardan daha çok anket üzerinde durulmuştur. Bunun nedeni öğrenci yorumlarının ankette yer almasıdır.



Grafik 1. Form ve Anketlerin Puan Dağılımlarının Karşılaştırılması (%)

Formdaki puanlarla anketteki puanlar arasındaki bu fark öğrencilerin zihinsel modellerinin işleyiş süreci, kişisel, duygusal ve problem tabanlı hatıraları ve soyut kurallar gibi çok çeşitli hatıra ve bilgilerin gözlem ve değerlendirme sonucu olabilir. Ankette, ikinci kez maddeyi okuyan öğrenci bir de puanının nedenini yazdığına daha bilinçli bir şekilde zihinsel süreçlerini işletmekte ve üst bilişi kontrol altına almış olabilir ve sonuçtaki bu değişimin nedeni daha bilinçli düşünmesinden kaynaklanmış olabilir. Kötü yapılandırılmış problemler çözüm sürecinin seçimi ve uygulamasının kontrolü ve düzenlenmesine ihtiyaç duyarlar (Hong, 1998).

#### *Bazı Maddelerdeki Binişiklik, Puanlamada Kararsızlığa Neden Olmaktadır*

Bazı maddelere verilen nedenler irdelendiğinde, aynı maddede birden fazla öğretim elemanı davranışı/özelliği yer aldığı için, öğrencilerin maddenin farklı noktalarına odaklandıkları ve puanlarını odaklandıkları kısma göre verdikleri saptanmıştır. Örneğin; anketteki 20. madde “sınavları ve ödevleri zamanında değerlendirerek öğrencileri hata ve eksiklikler konusunda uyarır” şeklindedir. Madde incelendiğinde altı farklı davranışın ölçülmek istendiği görülmektedir. Aşağıda öğrenciler tarafından bu maddeye verilen bazı nedenlerden bazı örnekler sunulmuştur.

TABLO 1

*Öğretim Elemanı Dersi İlgi Çekici Hale Getiriyor mu? Maddesi İçin Hazırlanan Kategori- Puan Analiz Matrisi*

	A sınıfı					B sınıfı					C sınıfı					Genel toplam								
	1	2	3	4	5	T	1	2	3	4	5	T	1	2	3	4	5	T	1	2	3	4	5	T
zaman zaman konuyla ilgili bir fikri ya da güzel bir anıyı anlatır	0	0	0	3	21	24	0	0	0	0	8	8	0	0	0	0	6	6	0	0	0	3	35	38
derste ders hedeflerini dağıtmayacak fıkralar anlatıyor	0	0	0	0	5	5	0	0	0	0	3	3	0	0	0	0	2	2	0	0	0	0	10	10
daha hızlı anlatırsa daha iyi olur	0	0	0	1	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	2
materyaller kullanır, fotokopi verir	0	0	0	1	1	2	0	0	0	0	9	9	0	0	0	1	0	1	0	0	0	2	10	12
eğitsel örnekler verir.	0	0	2	1	4	7	0	0	0	0	5	5	0	0	0	0	0	0	0	0	0	0	0	0
yöntemleri derse uygun	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	9
materyaller kullanmıyor gerek yok zaten	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1
materyal kullanmaz ancak esprili ve sevecen biri	0	0	1	3	2	6	0	1	1	0	1	3	0	1	0	0	0	1	0	0	0	0	2	2
örnekler verir (Fenerbahçe ile ilgili) günlük yaşamla ilişki kurar	0	0	2	0	2	2	0	0	0	0	23	23	0	0	0	3	6	9	0	1	3	3	3	10
güler yüzlü ve sempatik	0	0	0	0	1	1	0	0	0	0	3	3	0	0	0	0	0	0	0	0	2	3	31	36
mimik ve jestleri çok iyi	0	0	0	0	1	1	0	0	0	0	2	2	0	0	0	0	1	1	0	0	0	0	4	4
asabiyet ve protokol yok	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	4	4
esprileri komik değil	0	0	0	0	1	1	0	0	0	2	1	3	0	0	0	0	0	0	0	0	0	0	1	1
kalbimizi kırmaz	0	0	0	0	0	0	0	0	0	0	3	3	0	0	0	0	0	0	0	0	0	2	2	4
sesi monoton değil, bu da ilgi çekicidir.	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	3	3
kendi motivasyonu yüksek	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	1
kendi anlattığı dersler çok iyi ama biz anlatırken derse katılmıyoruz	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	4	0	0	0	0	4	4
dersle ilgili ilginç örnekler verir ama dersle ilgili materyalleri biz hazırlıyoruz	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	1
hocamız derste bir tiyatrocunun gibi davranır, farklı rollere girer	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	0	0	0	0	2	2
doğal davranır ve espri yapar	0	0	0	1	2	3	0	0	0	1	8	9	0	0	1	1	0	2	0	0	1	3	10	14
sohbet eder gibi bilgi verir ancak bazen konudan sapma olur	0	0	0	0	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2
farklı fikir ve önerilere açıktır	0	0	0	0	2	2	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	2	3
şarkı ve şiir okur isimle hitap eder	0	0	0	0	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	3
ders çok monoton	0	0	0	0	0	0	0	0	1	0	0	1	0	4	4	3	0	11	0	4	5	3	0	12

“Sınavları zamanında okur, tek tek dönüt vermez, aklına gelirse verir”, “sınavları erken okur, dönüt verir”, “sınavları geç okur, dönüt verir/vermez”, “ödev verir, dönüt vermez”, “ödevleri zamanında okumaz ancak dönüt verir”, “bu derste ödev yok”, “sınavları zamanında okur ancak dönüt istenirse verir”. Bu maddede, puanlar için açıklanan nedenler belirgin biçimde birbirinden farklıdır. Burada önemli olan sonuç, öğrencilerin binişik maddelerde belirtilen birden fazla davranıştan yalnızca birine veya bir kısmına odaklanarak puan vermeleridir. Aynı problem 1.; 7.; 14.; 16. maddelerde de saptanmıştır.

Janossen’in (2000), belirttiği gibi kötü yapılandırılmış problemlerin çözüm sürecinde problemlerin açık olarak tanımlanması önemlidir. Tanımlamaların binişik olması, problem çözme sürecinde diğer bir faktör olan bireysel farklılıkların daha fazla ve belirgin şekilde ortaya çıkmasına neden olabilir. Hong (1998), problem çözümlerinin bir sorunun amaçları ve işlem alternatifleri açık olmadığına, kendi bilişsel çabalarını farklı derece ve yönlerde organize etmek ve yönlendirmek zorunda kaldığını açıklamıştır. Maddelerin puanlanmasında kullanılan ölçütlerin öğrenciden öğrenciye farklılık göstermesi maddenin geçerliğini azaltmaktadır.

Cashin (1995), “geçerliliği” anketindeki maddelerin öğretimin etkililiğinin bazı yönlerini/ boyutlarını ne dereceye kadar ölçtüğünün belirlenmesi olarak açıklamaktadır. Bu tanıma dayanarak, 1.; 7.; 14.; 16. ve 20. maddelerin amacına uygun ölçme yapılmadığından geçerlilikleri sorgulanmalıdır. Çünkü nedenler analiz edildiğinde, nedenlerin ölçülmek istenen boyutun maddenin bir bölümüne denk geldiği veya maddede ölçülmek istenen boyutun çok dışında konuları ölçtüğü saptanmıştır. Geçerlik konusunda problemler maddelere Amerikan ve İngiliz Üniversitelerinde kullanılan anketlerde de rastlanmaktadır. Örneğin, The American Üniversitesi’nin soru bankasındaki sorular incelenebilir.

Cashin (1995), öğretimi değerlendirme anketi için “güvenirliliği” anketindeki bir madde hakkında öğrenciler arası ortak görüş (puanlama) ya da bir puana yönelim olarak diğer bir deyişle “consistency” olarak görmektedir. Geçerliliği yukarıdaki tanımlar gereği düşük olması gereken anket maddelerinin, yani ne ölçtüğü tam olarak açık olmayan ve öğrenci tarafından çok farklı yorumlanan maddelerin, istatistikî metotlarla bulunan güvenilirlik puanının yüksek olması çelişkili görünmektedir.

Daha ileri bir analiz yapılması gerekirse; 7. maddede verilen nedenler incelendiğinde beş öğrenci “dersin amaçları her zaman bellidir, zamanında söyler”, altı öğrenci, “öğretim tekniği harika”, üç öğrenci “tahtayı kullanır, ders notları var, konu anlatması ve öğrenciyi anlaması ile öğretim amaçları-

nı yerinde kullanır”, beş öğrenci “örnek verir, öğretim yöntemi uygundur”, şeklinde neden belirtmiştir. Öğrencilerin çoğu belirttikleri bu nedenlere 5 puan vermiştir. Bu nedenlerin hepsi farklı ve hazırlayanlar tarafından maddenin ölçmeyi amaçladığı şeylerin dışındaki konuları öğrencide çağrıştırmış olabilir. Robertson’ın (2004) çalışmasında da farklı nedenlere aynı puanlar verilmiştir. Geçerlik problemi taşıyan bir maddenin istatistikî metotlarla güvenilirliğini hesaplamaya çalışmak ve yorumu bu sonuca dayalı yapmak, öğretim elemanının ve dersinin yanlış değerlendirilmesine yol açabilir.

#### *Şartlı İfade Edilen Maddelerin Bir Bölümüne Puan Verilmektedir*

Bazı maddelere verilen nedenlerden öğrenciler maddenin bir kısmına odaklanıp sadece o kısmına dayalı olarak puan vermişlerdir. Bu maddeler incelendiğinde maddelerin, şartlı cümleler olduğu saptanmıştır. Öğrenciler 9. maddede cümlenin şart kısmına, 10. maddede ise bağıl kısmına odaklanarak cevap vermişlerdir.

Örneğin 9. maddede, öğretim elemanının not kaygısı yaratmadan sınıfı yönetip yönetmediği sorulmuştur. Bu maddeye yedi öğrenci “not, öğrenmeden daha önemli değil”, On yedi öğrenci ”not kaygısı taşıyorum, dersi iyi öğreniyorum”, bir öğrenci “öğretim elemanı, bu soruyu yapanın vizesine 5 puan ekleyeceğim diyor, bu bende olumsuz etki yaratıyor”, bir öğrenci, “not kaygımız sınavlarda ortaya çıkıyor”, bir öğrenci ise “sözlü yapılmadığına göre not kaygısı da olmaz” şeklinde neden belirtip puanlama yapmışlardır. Öğrencilerin sorunun ilk bölümüne, yani notla ilgili uygulamalara ve bu konu ile ilgili duygularına odaklanarak puan verdikleri görülmektedir.

10. maddede derste demokratik ortam yaratarak soru sorma fırsatı verilip verilmediği sorulmuştur. 100 öğrenciden 27’si “öğrencilerin derse katılmalarını ve görüş bildirmelerini” sağlar, “sorulara istekle cevap verir”, nedenini belirtirken, 32 öğrenci “anlamadığımız yerleri sorabiliyoruz” diyerek, bu konu ile ilgili duygularına dayalı olarak yani, sorunun bağıl kısmına odaklanarak puan vermiştir.

Bu maddeler incelendiğinde şartlı cümleler olduğu saptanmıştır. Öğrenciler 9. maddede cümlenin şart kısmına, 10. maddede cümlenin bağıl kısmına odaklanarak cevap vermişlerdir. Hong’un (1998), açıkladığı gibi öğrenciler, üst bilişe bağlı olarak biliş bilgisi ile ilgili, alandan bağımsız kötü yapılandırılmış problem çözme stratejilerinden olan “decomposition” stratejisini kullanmış olabilirler. Çünkü her iki maddede öğrenciler problem tanımının bir kısmına yoğunlaşarak cevap vermişlerdir. Öğrencilerin puan vermeye yoğunlaştıkları madde parçacığı, onların bir dönem boyunca birebir yaşadık-

ları somut ve öncelikle kendilerini ilgilendiren, değerlendirebilecekleri veya daha yoğun bilgiye sahip oldukları kısım olan “not kaygısı” ve “soru sorma fırsatının verilip verilmemesi”ne ilişkindir. Polson ve Jeffries’in (1985), bir konu hakkında herhangi bir zihinsel şemaları olmayan problem çözücülerin, decomposition stratejisini kullanmada zorluk çekebileceklerini belirtmişlerdir. Çalışmada bu stratejiyi uygulamış gibi görünen öğrenciler, maddeyi parçalayıp bir kısmına odaklanmalarına karşın, maddenin tümüne cevap verdiklerinden decomposition stratejisini yanlış uygulamışlardır.

*Bazı Maddelere Verilen Nedenler Daha Geniş Bir Yelpaze Oluşturmaktadır*

Anketteki 2,4,6 ve 11. maddelere, diğer maddelere göre daha çeşitli yorumlar yapılmıştır. Bu maddelerin içerdiği bazı sözcüklerin muğlak olması (belirsiz, bulanık, muammalı, kapalı) bu durumun bir nedeni olabilir. Anketteki 2. maddede “derse hazırlıklı gelme”, 4. maddede “ders konusuna hakim olma”, 6. maddede “dersi ilgi çekici ve zevkli hale getirme”, 11. maddede ise, “başarılı iletişim kurma becerisi” sorulmaktadır. Özellikle “hazırlıklı olma” ve “hakim olma” sözcüğü geniş anlam içeren, göreceli sözcüklerdir. Bu durum bu sözcükleri kapalı kılmakta ve her öğrenci mesajı farklı algılamaktadır. Bu maddelerin kapsamında birçok davranış yer almaktadır. Ancak hangi davranışların ölçüt alınarak, değerlendirme yapılması gerektiği öğrenci için belirsizdir. Robertson’ın (2004) çalışmasında da bazı öğrenciler, “ders hakkında bilgiye ulaşıldı mı?” maddesinde geçen, göreceli anlam taşıyan “bilgi” sözcüğünü ders ve içerikle ilgili her türlü bilgi olarak anlamışlardır. Aynı biçimde Block’un(1998) araştırmasında da bir öğrenci ankette yer alan “dersin genel olarak değerlendirilmesi” maddesini, genel anlamda tüm derslerin değerlendirmesi olarak algılamıştır.

“ders konusuna hakimdir” maddesi için verilen puanların nedenlerinden bazıları şunlardır:

“Sorularımıza net cevap verir”, “günlük yaşamla bağ kurar bu durum daha iyi anlamamı sağlıyor”, “ders anlatımı akıcı”, “basitten karmaşığa doğru anlatır”, “yeni gelişmeleri takip eder” “kaynakları oldukça fazla”, “açıklamaların tam olarak yapıldığına inanmıyorum” “ders anlatırken materyali eksik”, “dersi öğrenci anlatıyor, bu durum yüzeysel anlamamıza neden oluyor”, “anlatımı bana uygun zaten alanını çok iyi biliyor”. Açıklanan nedenlerin büyük çoğunluğu eğitim literatüründe öğretim elemanının derse hakimiyetinin bir göstergesi olarak kabul edilebilir. Ancak “derse hakimiyet” kavramının kapsamı içinde hangi davranışların olması gerektiğine yönelik ortak bir görüş bulunmamaktadır.



Maddeler kötü tanımlandıkları için puanlamanın kesin şekilde bilinmeyen ya da yabancı problem çözme öğeleri taşıması (Wood, 1983, Kitchener, 1983), birden çok çözümü veya çözüm yolunun olması ya da çözümsüz olması olasılığı vardır.

Green ve diğerleri'nin (1998) araştırmasında öğrenci değerlendirme anketlerinin % 60.8'inde, bu araştırmadaki 2.; 4.; 6. ve 11. maddelere benzer en az bir tane madde bulunduğu ve bu maddelere öğrencilerin, alt yapıları, deneyimleri doğrultusunda puan verdikleri saptanmıştır.

Maddelere verilen nedenlerin geniş bir yelpazede olmasının diğer bir nedeni ise, öğrencilerin öğretim elemanı ve dersini değerlendirirken, kendi öğrenme stillerini ölçüt aldıklarını belirten Kember ve Wong'un (2000) görüşü ile açıklanabilir. Pasif öğrencilerin, öğretimin aktarım şeklinde yapıldığı dersleri tercih ettiğini, aktif öğrencilerin ise çeşitli yöntem ve tekniklerin kullanıldığı, öğrenci katılımının sağlandığı dersleri tercih ettiğini saptamıştır. Bu durumda öğrencilerin öğrenme stilleri ve tercihleri, öğretim elemana verilen puanı ve puanının verilme nedenleri farklılaştıracaktır.

#### *Bazı Maddeler Anlaşılmamış ve Yanlış Anlaşılmıştır*

Anketteki 7. maddede “öğretim elemanının, öğretim amaçlarını yerinde ve etkin kullanıp kullanmadığı” sorulmuştur. Bu maddenin anlaşılmadığı veya yanlış anlaşılmış olduğu aşağıda verilen nedenlerden görülebilir. “Notla tehdit etmez”, “disiplin ve samimiyeti bir arada bulunduran etkin bir öğretim elemanı”, “öğretmen olunca nelere dikkat edilmeli, bu konuda bilgi verdi” şeklinde nedenler irdelendiğinde, ilk neden başarılı bir öğretim elemanın özelliği; ikinci neden başarılı bir sınıf yönetimi; üçüncü neden ise ders içeriği hakkındadır. Bu açıklamaları yapan on altı öğrenci 5 puan vermiştir. Öğretim amaçlarının yerinde ve etkin kullanılmasının ölçütü olmayan davranışlar, yüksek puanla değerlendirilmiştir. Maddenin ifadesine dikkat edilirse “yerine ve etkin kullanma” ifadesi muğlak yada göreceli olduğundan farklı ölçüt geliştirmeye uygun olduğu görülmektedir.

Ayrıca yirmi dokuz öğrenci bu maddeyi anlamadıklarını belirtmelerine rağmen dört öğrenci 3, on beş öğrenci 4, on öğrenci ise 5 puan vermiştir. Hem puan verme nedenleri hem de puan dağılımı şaşırtıcıdır. Öğrenciler anlamadıklarını ifade ettikleri bir maddeye puan vererek öğretim üyesini değerlendirmişlerdir. Yirmi dokuz öğrenciden üçünün verdiği yorumlar puanlara nasıl karar verdiklerini göstermesi açısından önemlidir. Bir öğrenci “öğretim elemanı genelde başarılıdır, bu maddeyi anlamadım ancak ders anlatımı başarılı olduğu için puan verdim”, bir öğrenci “iletişimi gayet iyi

ama bu madde anlaşılır değil ancak bizi sevdiği için puanı yüksek verdim”, bir diğer öğrenci ise “genelde başarılı olduğu için puan vermek istedim” şeklinde açıklama yapmıştır.

Yukarıda sözü edilen öğrencilerin dışında on iki öğrenci hem not vermemiş hem de yorum yazmamış, yirmi öğrenci ise maddeyi anlamadıkları için puan vermediklerini belirtmişlerdir. Ancak anlamadan puan verenlerin oranı daha yüksektir. Robertson’ın (2004) çalışmasında yer alan bir madde, metodun problem çözmeye etkili biçimde uygulanıp uygulanmadığı sorulmuştur. Rapor yazma olarak anlaşılan bu maddeye öğrenciler “bu derste rapor yazmadık” sebebini göstererek en düşük puanı vermiştir. Robertson’ın çalışmasında, öğrenciler anlamadıkları madde için düşük, bu çalışmada ise yüksek puan vermişlerdir.

Tagomori ve Bishop (1995), iki yüz değerlendirme anketini inceleyerek maddelerin dörtte üçünün karışık, belli belirsiz veya öznel olduğunu saptamıştır. Böylece bu bulguların yazılı değerlendirme maddelerinin anlaşılabilirliği ve iyi öğretimin benimsenmiş yönlerine uygulanabilirliği konusunda ciddi kuşklar uyandırdığı sonucuna varmışlardır.

Bu çalışmada kullanılan ankette de bu özelliği taşıyan maddelerin bulunduğu, öğrencilerin verdiği nedenler yoluyla ortaya çıkarılmıştır. Maddeyi anlamamış öğrencilerden bazılarının öğretim elemanının onları etkileyen farklı özelliklerine göre puan verdiği anlaşılmıştır. Bu sonuçlar Gick’in (1986) öğrencilerin aşına olmadıkları problemlerle yüz yüze geldiklerinde, problem alanını kendi kişisel deneyimleri doğrultusunda yorumlayarak problemi çözmeye çalıştıkları görüşünü desteklemektedir.

#### *Maddeleri Beğenmeyen ve Eleştiren Öğrenciler Yine de Puan Vermişlerdir*

Bazı öğrenciler maddeleri beğenmediklerini belirtmelerine rağmen puan vermişlerdir. 3. maddede öğretim elemanının dersin başlama ve bitiş saatlerine titizlikle uyup uymadığı sorulmuştur. Bir öğrenci “bu maddeyi beğenmedim, anlamsız, derse giriş çıkış değil, ders işleme sürecinde yapılanlar önemlidir” nedenini belirterek 4 puan vermiştir.

10. maddede öğretim elemanının derste demokratik ortam yaratarak, öğrencilerin soru sormasına ve görüş bildirmesine fırsat verilip verilmediği sorulmuştur. Puan verme nedeni olarak soruyu beğenmediğini belirten üç öğrenciden biri 3, ikisi 4 puan vermiştir. Öğrencilerin bu puanları öğretim elemanının davranışına mı yoksa beğenmediklerini belirttikleri maddeye mi verdiklerini ayırt etmek zordur. Block’un makalesinde bu sonucun tersine

bir durum ortaya çıkmıştır. Maddeyi beğenmeyen öğrenci inatla bu maddeyi puanlamayacağını şu sözlerle açıklamıştır:

Sorun şu ki ben bu maddeyi öznel görüyorum, bu yüzden değerlendirme yapmayacağım. Benim için ilginç olabilecek bir şeyin bir başkası için daha az ilginç olacağını bilincindeyim çünkü birbirimizle konuşma esnasında birimizin laboratuvarında tek başına diğerimizin iki kişi çalışmayı sevdiğine şahit oldum (Block 1998, 417).

Öğrencilerden puanlama işi başlamadan önce beğenilmeyen veya anlaşılmayan maddeleri puanlamamaları istenmesine karşın onlar puanlama yapmışlardır. Bu öğrenciler, üst biliş altındaki biliş yönetimindeki değerlendirme sürecini uygulamış olabilirler. Öğrenciler değerlendirme sonunda soruyu beğenmediklerini belirtmişlerdir. Bu işlem bir değerlendirme işidir; ancak bunun sonucu olarak öğrencilerin puanlama yapmaması beklenirdi.

*Öğrenciler Bazı Maddelerin Kapsamında Yer Almayacak Özellikleri Neden Olarak Belirtmiştir*

Anketteki 4. maddede, öğretim üyesinin derse hakim olup olmadığı sorulmuştur. Öğrenciler tarafından verilen nedenler incelendiğinde, farklı dönemde toplanan verilerde benzerlik son derece yüksektir.

“Derse hâkimiyet” kavramına dâhil olabilecek öğretim üyesi özelliklerini öğrenciler şöyle açıklamıştır. En yüksek frekansa sahip olan nedenler şunlardır: “Sorulara net cevap verme”, “günlük yaşamla bağ kurma”, “somut örnekler verme” gibi. Buna karşın bazı öğrenciler, verdikleri puanın nedeni olarak bu maddenin kapsamı dışında yer alabilecek davranışları belirtmişlerdir. Örneğin: dört öğrenci “görüntüsü bize güven veriyor” şeklinde neden belirterek öğretim elemanına 5 puan vermiştir. Bu durum öğretim elemanının görünüşüne ilişkin öğrencide bıraktığı olumlu izlenimdir. Derse hâkimiyetinin bir göstergesi değildir.

Yukarıda açıklanan durum 5. ve 12. madde için de geçerlidir. Anketteki 5. maddede öğretim elemanının derste ses tonunu, dil ve ifadesini uygun olarak kullanıp kullanmadığı sorulmuştur. Bu maddeye üç öğrenci “bize son derece saygılı davranır” şeklinde neden belirterek 5 puan vermiştir. Bu üç öğrenciden biri “bize sen diye hitap etmez, bu bize olan saygısını gösteriyor” şeklinde daha ayrıntılı açıklama yapmıştır. Bu maddede öğretim elemanının öğrencilerle olan iletişim becerisi değil, dersi sunarken kullandığı ses tonu, dil ve ifadesi sorulmaktadır. Değerlendirme anketinde halihazırda öğretim elemanının iletişim becerisini ölçen bir başka madde bulunmaktadır.

Anketteki 10.maddede demokratik ortam yaratarak derste soru sorma fırsatı verilip verilmediği sorulmuştur. Sekiz öğrenci “soru sorarız ancak ciddiye almaz”, iki öğrenci ise; “sorduğumuz sorulara istenilen şekilde açıklama yapmamıştır” şeklinde neden belirtmiştir. Öğretim elemanının sorulan soruyu ciddiye almaması veya cevabı geçiştirmesi davranışına odaklanarak puan verilmiştir.

Anketteki 12. maddede öğretim elemanının öğrencilere karşı eşitlik ilkesine uyup uymadığı sorulmuştur. Bu maddeye beş öğrenci “göz teması var, saygılıdır, güven verir”, iki öğrenci “selam verir”, iki öğrenci “gülümsemez”, beş öğrenci ise “esprili” açıklamasını yapmıştır.

Öğrencilerin belirttikleri bu nedenler incelendiğinde, öğretim elemanının eşitlik ilkesine uyup uymadığı açıklanmamaktadır. Belirtilen nedenler, anketteki 11. maddede sorulan iletişim becerileri kapsamında yer alan öğretim elemanının özelliklerine daha uygun görünmektedir.

Anketteki 11. ve 12. maddeye verilen nedenler benzerlik göstermektedir ve aralarındaki korelasyon 0.229 dur.

Öğrenciler maddenin kapsamında yer alabilecek özellikleri ölçmek yerine, öğretim elemanlarında kendilerini etkileyen en önemli duygusal özelliklere puan vermişlerdir. Sinnott (1989), problemin özünü anlamak veya çözümünü geliştirmek için duyguların ve konuyla ilgili olmayan düşüncelerin hızlandırıcı rol oynadığını belirtmiştir. Çünkü bu başlık altında incelenen nedenlerin çoğunluğu duygusal özellik taşımaktadır ve iletişim becerileri ile ilişkilidir. DeCanio (1986), öğrenciler için öğretim elemanının iletişim becerilerinin ve düzenlilik seviyelerinin en önemli özellikler olduğunu bulmuştur. DeCanio’un araştırma sonuçlarına göre, öğrencilerin öğretim elemanının iletişimi ile ilgili özelliklerinden etkilenecek puan vermeleri doğal görünmektedir. Bu bölümde belirtilen nedenlerin ortak noktası ise öğrencilerin öğretim sürecinde öğretim elemanı tarafından “dikkate ve ciddiye alınma” isteği olarak belirlenmiştir. Bu isteğin ortaya çıkış nedeni ise Türk Eğitim Sisteminin her kademesinde kuramsal olarak öğrenci merkezli öğretimden söz edilmesine rağmen, uygulamada öğretmen merkezli öğretim uygulamaları olabilir.

#### *Hale Etkisi (Halo Effect) Görülmüştür*

Bir insanın sahip olduğu bir olumlu ya da olumsuz özelliğinin, onunla ilgili genel bir yargının oluşmasına ve diğer özelliklerinin bu çerçevede değerlendirilmesine yol açması. Örneğin. Dürüstlüğü ile bilinen birisinin bu özel-

liğinden yola çıkılarak aynı zamanda zeki veya çalışkan olduğu yargısına varılması gibi. Psikoloji alanında “ayla etkisi olarak da kullanılmaktadır.

Nedenler irdelendiğinde bazı maddelerde hale etkisi göze çarpmaktadır. Bu durum yine ardışık zamanlarda toplanan verilerin ikisinde de ortaya çıkmıştır “Sorduğumuz sorulara net cevap verir”, “elinde bol miktarda kaynak var”, “günlük yaşamla bağ kurar”, “ders anlatımı ve aktarımı çok iyi”, “alan bilgisi çok iyi” gibi nedenler 2. ve 4. maddede puanlama nedenleri olarak açıklanmıştır. Buradaki hale etkisinin nedeni, öğrencilerin kendilerince önemli nedenleri sıklıkla farklı maddelerde ölçüt almaları olabilir.

Robertson’ın (2004) çalışmasında öğrenciler tarafından belirtilen bir neden, yedi ayrı soruda görülmüştür. Bu çalışmada ise beş neden iki ayrı maddede görülmüştür. Marsh, hale etkisinin iyi yapılanmış testlerde önemsiz olduğuna inanmaktadır (Marsh, 1984; Marsh ve Roche, 1997). Hale etkisi 10, 11 ve 12. “demokratik ortam yaratma” “Başarılı iletişim kurma”, “ve “davranışlarda eşitlik ilkesi” maddelerine verilen nedenler incelendiğinde öğretim elemanının “öğrencilere eşit davranma” özelliği ortaktır. Öğrenciler bu nedeni herhangi bir biçimde üç madde ile ilişkilendirmişlerdir. Öğretim elemanının “güler yüzlü olma” özelliği ise 6.11.12.13 ve 14. maddelere verilen puanların nedeni olarak belirtilmiştir.

Bu araştırmada bulunan sonuçlar, “tek bir konunun, bütün olarak dersin puanlanmasında önemli bir etkiye sahip olabileceğine ilişkin kuvvetli kanıt vardır” yorumunu yapan Robertson’ı (2004) desteklemektedir. “Güler yüzlü olma” ve “eşit davranma” davranışları öğrencilerin zihninde iz bırakan öğretim elemanı özelliği olmasından dolayı hale etkisi ortaya çıkmış olabilir.

Roberson (2004), hale etkisinin maddelerin kastettiği yapının veya konunun daha detaylı ve açık belirtilmesiyle azaltılacağını belirtmiştir. Bu görüşe farklı bir yaklaşımla katılan Low (1999), maddelerin yazımında, öğrencilerin yorum farklılıklarını en aza indirgeyecek şekilde ifade edilmesinin önemini vurgulamıştır. Bu ise maddenin ölçmek istediği davranışın açık, net olarak ifade edilmesi anlamına gelebilir. Buna karşın, Cashin (1995) makro düzeyde etkili öğretim ile ilgili daha genel maddeler sorulabilir şeklinde açıklama yaptıktan sonra bu konudaki endişelerini şöyle dile getirmiştir:

“Bir grup tarafından kabul gören boyutlar bir diğer grup tarafından kabul görmeyecektir. Çünkü etkin öğretimin” herkes tarafından kabul edilmiş bir tanımı yoktur ve her şeyi kapsayıcı tek bir ölçüt de bulunmamaktadır”.

Cashin’in (1995) eksikliğini belirttiği ortak bir “etkili öğretim tanımı” yapılacak olursa; bu öneri veya tanım Johnson’ın (2000) araştırmasında belirttiği endişeyi ortaya çıkabilir;

“Bu tür sınırlayıcı tanımlar yapıldığında tek tip öğretim elemanı yaratılacak ve bunun sonucunda etkili öğretimde en önemli niteliklerden biri olan öğretim elemanın ortaya koyacağı veya sahip olduğu kişisel farklılıklar, pozitif özellikler yok sayılacak veya ölçülmeyecektir. Bu anlamda böyle bir yaklaşımda gerçek “etkili öğretimden sapılarak tek tip öğretim elemanı ve etkin öğretim ödüllendirilecektir”

Tüm bu görüşler ışığında hale etkisinden kaçınmak için maddelerin çok özel bir davranışı ölçecek kadar daraltılması, sorunları azaltmayabilir. Bu durumda anket maddelerinin belli düzeyde genel ifade edilmesi kaçınılmazdır. O halde hale etkisini ve bunu ortaya çıkaran nedenleri yararlı veriler olarak kabul ederek bu durumu avantaj haline dönüştürebiliriz. Çünkü hale etkisini ortaya çıkaran sebepler, aslında öğrenci gözünde etkili öğretimin en “önemli” ölçütlerinin göstergesi olabilir. Hale etkisinin görüldüğü maddeler sadece istatistikî analiz açısından incelenirse, anketin sağlığı açısından bu etki olumsuz kabul edilebilir. Maddeler ve ilgili nedenler, ancak niteliksel analizle incelenirse ortaya çıkan sonuçlar “etkili öğretimin” sağlanması açısından çok faydalı ve pozitif etkilere sahip olabilir. Çünkü bu nedenler anket maddelerinde sıklıkla ortaya çıkan ölçütlerdir ve öğrenciler tarafından puanlamada etkili olarak kullanılmaktadır. Hale etkisini yaratan nedenler, pozitif ve negatif ifadedi olarak ayrılacak olursa, öğrenciler açısından pozitif olan nedenler devam etmesi istenen veya daha fazla arzu edilen öğretim elemanı davranışlarına; negatif olanlar ise, ortadan kaldırılmasını istedikleri ve rahatsız oldukları öğretim elemanı davranışlarına işaret etmektedir. Bu anlamda anketi değerlendirenler tarafından hale etkisini ortaya çıkaran nedenler dikkatle incelenirse, faydalı sonuçlar elde edilebilir.

*Açık Olarak Kabul Edilebilecek Maddelerde Bile Nedenler Arasında Zıtlıklar Görülmüştür*

Aynı sınıftaki öğrenciler, aynı öğretim elemanı için zıt açıklamalar yapmıştır. Bu durum ankette yer alan birçok maddede gözlenmiştir. Özellikle 3. maddede, öğretim elemanının dersin başlama ve bitiş saatlerine titizlikle uyup uymadığı sorulmuştur. Bu maddenin öğretim elemanının derse giriş ve çıkış saatleri açısından öğrenciden öğrenciye değişen, yoruma açık bir maddede olmadığı varsayılmıştı. Ancak öğrencilerin verdikleri birbirine zıt nedenlerden ve verdikleri farklı puanlardan bu varsayımın doğru olmadığı saptanmıştır. Örneğin “dersin başlangıç ve bitiş saatlerine özen gösterir”, beş öğrenci 5 bir öğrenci ise 4 puan vermiştir. Öğrenci yorumları yeniden incelendiğinde de bunun sebebi anlaşılammıştır. “Dekan yardımcısı olmasından dolayı başlangıç saatine uymaz”, açıklamasını yapan öğrencilerin ver-

dikleri puanlar da değişim göstermiştir. Yedi öğrenciden ikisi 5, ikisi 4, biri 3, biri ise 2 puan vermiştir. 2 puan veren öğrenci diğerlerinden farklı olarak;

“Yönetim görevi ayrı bir iş, öğretim ise başka bir iş, bunu dengelemeli” yorumunu yazmıştır. Olasılıkla 5 puan verenler diğerlerine oranla öğretim elemanın hatasına daha fazla tolerans göstermiş olabilirler. Bu durum anketlerin subjektifliği hakkındaki endişeleri destekler niteliktedir. Block’un (1998) araştırmasındaki benzer bir madde için öğrenciler derse geç kalan öğretim elemanına tam puan vererek kısa süreli gecikmelerin olabileceğini belirtmişlerdir. 4 puan veren öğrencilerden biri şu açıklamayı yapmıştır;

“Bazen bende geç kalıyorum, sınıftan bazı arkadaşlar da, bazen bu olabilir. Dersin içinin dolu geçmesi daha önemlidir”

Ayrıca nedenler ve puanlar incelendiğinde öğrenciler, dersin başlangıcı kadar, bitiş saatine gösterilen titizliğin de önemli olduğunu belirterek, “her zaman tam saatinde çıkmayız, dersin bitiş saatlerine uymaz” şeklinde neden belirtmişlerdir. Diğer nedenlerden bazı örnekler ise aşağıda verilmiştir. “Başlama saatine uymaz”, “konu ne zaman biterse dersi bitirir”, “genellikle uyar”, “tenefüsleri unuttur”, “blok ders yapar”, “öğrencinin durumuna göre erken bırakır”, “başlama bitiş saatlerine uymaz”.

Voss ve Post (1988) ve Jonassen’e (1997) göre, kötü yapılandırılmış problem çözme süreçlerinde problemin tanımlanması adımı çok önemlidir. Bu görüşe göre 3. maddenin öğrenciler tarafından farklı ve yanlış anlaşılmasına yol açmayacak şekilde açık ve anlaşılır olarak tanımlandığı düşünülmüştür. Bu denli açık ve net olduğu düşünülen bir maddede bile aynı sınıfın, aynı öğretim elemanı için çok farklı nedenler belirtmiş olması, anket doldurma işinin kişiler arası farklılıklardan oldukça etkilendiğini göstermesi açısından önemlidir.

Net olduğu düşünülen bir maddede bile, böyle farklı cevapların olması, Jonassen’in (1997), belirttiği gibi problem çözenler problemin ortaya konusuna dayalı olarak bir çözümü belirleyebilmek seçmek veya sentezlemek amacıyla problemin zihinsel modelini oluştururlar. Bu zihinsel modeli oluştururken modelin parçacıklarından biri; Jonassen’e göre deneyim, Sinnot’a (1989) göre hatıralar (memories) olarak ifade edilmiştir.

Bu hatıra veya deneyimler öğrenciden öğrenciye algının farklılığı nedeni ile farklı hiyerarşik önem sırasında saklanırlar. Bu madde ile ilgili farklı hiyerarşik sıradaki hatıralara veya deneyimlerden en baskın olanı veya olanları neden ya da nedenler olarak ortaya çıkmış olabilir. Farklı yorumlamaya neden olmayacak gibi görünen bu maddeye verilen nedenlerdeki zıtlık, her öğrenciye göre baskın olan hatıraların farklılığından kaynaklanmış olabilir.

*Dereceleme Ölçeğinde İlgili Çekici Farklılaşma Belirlenmiştir*

Clouder (1998)'e göre, öğrencilerin değerlendirme yapabilme yetenekleri, kompozisyon yazma yetenekleri kadar farklılık gösterir. Bu çalışmada da öğrencilerin dereceleme ölçeğini kullanmaları farklılık göstermiştir. Aynı sınıftaki öğrenciler, aynı öğretim elemanı için, aynı madde ile ilişkili olarak aynı nedenleri belirtmelerine rağmen farklı puan vermişlerdir. Örneğin 6. maddede öğretim elemanının dersi ilgi çekici ve zevkli hale getirip getirmediği sorulmuştur. On iki öğrenci “dersin çok monoton olduğu” nedenini belirtmiştir. Bu öğrencilerden dördü 2, beşi 3 puan verirken, üç öğrenci ise 4 puan vermiştir.

Öğrenciler aynı ders ve aynı öğretim elemanı için, farklı nedenlere aynı puanı vermiştir. Örneğin, 14. maddede “öğretim elemanının görünüş ve davranışlarıyla model oluşturup oluşturmadığı sorulmuştur. Altı öğrenci, “örnek alınabilecek iyi bir model”, “model almam”, “model alınabilir”, “cömert, esprili”, nedenlerini belirterek 5 puan vermiştir.

Aynı maddeye bir başka sınıfta, “örnek alınabilecek bir öğretim elemanı”, “anlatım tarzı model alınabilir”, giyimine dikkat etmeli”, “model alınabilir, cömert, esprili, bilgili”, “bazı davranışları hariç evet”, “özellikle görgü kurallarını uygular davranışları iyi” nedenlerini veren on beş öğrenci 4 puan vermiştir.

*Anketin Tüm Maddelerindeki Puanlama Nedenlerinde Farklılık ve Fazlalık Saptanmıştır*

Ankette her madde için puan verme nedenlerinin farklılığı ve fazlalığı dikkat çekici sonuçlardan biridir. Her madde için saptanan neden en az on beştir. Bu sayı bazı maddelerde yirmi beşe kadar yükselmiştir.

Bu sonuca bir örnek verilecek olursak 6. maddede öğretim elemanının dersi ilgi çekici ve zevkli hale getirip getirmediği sorulmuştur. Aşağıda, öğrencilerin bu madde için verdikleri puanlar ve yorumlardan çıkan nedenler “puan–neden analiz matrisi” ile verilmiştir.

Yukarıdaki nedenler incelendiğinde, öğrencilerin puanlama ölçütlerinin çok farklı olduğu görülmektedir. Bu sonuç daha çok öğretimin ve öğretim elemanının etkinliğinin öğrenci tarafından değerlendirilmesindeki geleneksel fonksiyonlarından biri olan formative değerlendirme için önem taşımaktadır.

Öğretim elemanın yalnızca dereceleme tipi ölçek kullanılarak yapılan puanlama sonuçları ile elde edemeyeceği veriyi, puanlara verilen nedenler-



den elde edebileceğini göstermesi açısından bu tür nedenlerin yazıldığı anketler önemlidir.

Bu çalışmada ortaya çıkan anket maddelerine verilen nedenlerin ve ilgili puanların geniş bir yelpazede olması durumu, üstelik bu durumun da hemen hemen her maddede görülmesi Block (1998), Low (1988, 1996 ve 1999) ve Robertson (2004)'in araştırma sonuçları ile benzerlik göstermektedir. Çok az çalışma, öğretim elemanlarının etkililiğinin değerlendirilmesinde öğrenci katılımının davranışsal boyutunu incelediğinden (Chen ve Hoshower 2003), bu çalışmanın sonuçları bu açıdan literatüre katkı getirebilir.

Birçok bireysel farklılığın bireylerin problem çözme yeteneklerini etkilediği bilinmektedir. Bireyler bilişsel stillerinde ve denetimlerinde farklılaşır ve bu durum bireylerin bilgiyi işleme ve muhakeme etme yollarını denetleyen düşünme kalıplarını gösterir (Jonassen ve Grabowski, 1993). Ayrıca Kitchener (1983) epidemik varsayımların kişiden kişiye değiştiğini ileri sürmektedir. Bu görüşe göre bireyler kötü tanımlanmış problemler için farklı sonuçlara ulaşırlar. Çünkü bilgiye dair farklı inançları vardır. Her iki görüş doğrultusunda öğrencilerin belirttiği nedenlerde bu denli fazlalığın ve farklılığın ortaya çıkması doğal görünmektedir. Cashin (1990) "Idea Paper 22" isimli çalışmasındaki 24. önerisinde, frekans dağılımlarının öğretim elemanı için değerli veriler sağladığını açıklamıştır. Ardından öğretim elemanlarına, değerlemelerin iki uça öbeklenip, dağılım çift doruklu olma eğiliminde ise belki yaptıkları şeye devam etmelerini ve diğerleri için bir şeyler eklenilmesi gerektiğini önermiştir. Ancak öğretim elemanı, elde ettiği sayısal sonuçlar doğrultusunda öğretimine yenilik katmaya karar verdiğinde hangi davranışları katacağını ancak anketlere verilen puanların nedenlerinin belirtilmesi ile saptayabilir. Cashin'in (1990), 26 ve 32 nolu önerilerinde belirttiği açık uçlu soruların gelişim amaçlı kullanılması görüşüne ek olarak puanlamaların nedenlerinin yazılmasının da gelişim amaçlı kullanılabilir.

#### *Öğrenciler Hatırlayamadıkları ve Emin Olmadıkları Durumlarda Bile Puan Verdiler*

Anket uygulanmadan önce öğrencilerden hatırlayamadıkları veya emin olmadıkları durumlarda maddelere puan vermemeleri istenmişti. Bu açıklamaya rağmen öğrenciler puan vermiştir. Örneğin;

Anketteki 1. maddede öğretim elemanının dönem başında işlenecek konuları ve amaçları açıklayıp açıklamadığı sorulmuştur. Bu maddeye altı öğrencinin "hatırlayamadım, belirtildi veya belirtilmedi, üzerinden zaman geçti" şeklinde neden belirtmelerine rağmen puan verdikleri saptanmıştır.

Anketteki 19. madde, sınavların objektif değerlendirilip değerlendirilmediğini ölçmeyi amaçlamaktadır. Yedi öğrenci “sınava giremediğini” belirtmiş; ancak yine de puan vermiştir. Sınava girmeyen öğrencinin bu puanı hangi ölçüte göre vermiş olduğu belirsizdir.

Birçok araştırma konusu, öğretim elemanının kontrolü dışındaki değişkenlerin değerlendirmeyi etkileyip etkilemediği ile ilgilidir. Öğrencilerin değerlendirme sistemini algılamaları ve puanlamaya katılım konusundaki güdülenmelerinin puanlamayı etkilediği (Chen ve Hoshower, 2003), öğrenci mevcudu düşük olan sınıflarda puanların yüksek çıktığı (Centra ve Creech, 1976), bazı disiplinleri öğretenlerin, diğer bazı disiplinleri öğretenlere oranla daha yüksek puan aldıkları (Cashin, 1990), seçmeli derslerin, zorunlulardan daha yüksek puan aldığı (Feldman, 1978), zor derslerin daha yüksek puan aldığı (Cashin, 1996), öğretim elemanının seksi olarak algılanmasının puanları yükselttiği (Felton ve diğerleri, 2004) saptanmıştır.

Yukarıdaki çalışmalara paralel olarak, bu çalışmada da öğretim elemanının dışında olan, fakat öğrenciden kaynaklanan sebepler ölçüt alınarak puanlama yapıldığı ve öğretim elemanına düşük puan verildiği saptanmıştır. Hoffman ve Kramer (1980) puanların öğretim elemanının tamamıyla kontrolü dışındaki özelliklerden etkileneceği endişelerini belirtmişler ve bu problemi çözümlerin en iyi yolunun öğrenci yorumlarına dayalı araştırmalar olduğunu belirtmişlerdir.

Öğretim elemanının kontrolü dışındaki etkenler, öğrencilerin verdikleri puanları yükseltmekte veya düşürmektedir. Bu sonuçlar; Wetzstein, Broder ve Wilson (1984), Gramlich ve Greenlee, (1993) değerlendirme puanlarının, tek başına öğretim elemanı performansının değerlendirilmesinde güvenilir bir yol olmadığı görüşlerini desteklemektedir.

#### *Öğrenciler Bilgileri Olmadığını Belirttikleri Maddelere Puan Verdiler*

Bazı öğrenciler anketteki 8., 10., 12. ve 20. maddeler hakkında fikirlerinin olmadığını veya maddelerin yeterli olmadığını belirtmelerine rağmen, maddelere puan vermişlerdir. 8. maddede dört öğrenci “bu konu hakkında fikrim yok”, 10. maddede üç öğrenci “bunu hiç düşünmedim biz rahatız”, 12. maddede beş öğrenci, “kesin bir fikrim yok, ayırım yapar gibi ama fazla bilmiyorum”, 20. maddede altı öğrenci “bilmiyorum” şeklinde puan verme nedeni belirterek maddelere 2-5 arasında puan vermişlerdir. Konu hakkında fikirlerinin olmadığını belirten öğrencilerin form ve anket doldurulmadan önce yapılan açıklama doğrultusunda puan vermemeleri beklenirdi.

Eğer öğrenciler değerlendirme anketinde yer alan maddeleri puanlar-ken, bilgileri olmadığı için bir kararsızlık yaşıyor ve yine de karar vermeleri gerektiğini düşünüyorlarsa bu engeli aşmak ve sonuca ulaşmak için kendine ait bir problem çözme süreci işletmiş olabilir. Whetley'de (1984) problem çözmeyi, bireyin ne yapacağını bilmediği anda yaptığı şey olarak açıklamış-tır.

Öğrencilerin bilgileri olmadığını belirttikleri maddeyi puanlama neden-lerinden biri Robertson'ın (2004), belirttiği gibi öğrencilerden anketi ta-mamlamalarının istenmesiyle birlikte ve öğrencilerin tüm maddeleri puan-lamalıym şeklindeki açık veya gizli bir varsayımla hareket etmesi olabilir.

Bu varsayımla hareket eden öğrencilerin kararlarını bilişsel olmayan duygu ve tutumlar veya değerlendirme becerileri de (justification skills) yönlendirmiş olabilir. Jonassen'de (1997), kişiler aşına olmadıkları bir prob-lemle yüz yüze geldiklerinde, problem alanı hakkındaki kişisel bilgilerini, duygu ve tutumlarını tüm sürece kattığını belirtmiştir.

Robertson'ın (2004) eğer anket maddesi ilk okumada akılda bir cevap oluşturmuyorsa, öğrenciler çaresiz bir durumla yüz yüze gelmiş ve bundan kurtulmak için sorunun özünü değil öğretim elemanına ilişkin zihnindeki madde ile ilgisiz bir bilgi veya bilgi parçacığına göre bir işlem yapmış olabi-rlirler görüşü de bu durumun gerekçesi olabilir.

## **Tartışma ve Öneriler**

Elde edilen sonuçlar, nedenlerin yazılması ile ortaya çıkartılmıştır. An-ketteki puanlara karar verme eylemi zihinden ilk gelen şeylere dayanır.

Üniversitenin resmi anketi olan ve bu araştırmada kullanılan ankette yer alan maddelerdeki ifade tarzları bu anketi tasarlayanların neleri kastettikle-ri anlaşılamayabilir ancak tasarlayanların kastettikleri madde ifadeleri, öğ-rencilerin algıladıklarından daha dar çerçevede olduğu anlaşılmıştır. Mesajı (anket) hazırlayıp gönderenin bu denli farklı ve fazla miktarda davranışın ölçülmesini planlamış olma ihtimali düşük görünmektedir.

Maddenin ifade ediliş biçiminin ve seçilen kelimelerin puanlama süre-cinde önemli olduğu da ortaya çıkmıştır. Maddenin tanımında kullanılan ve kaçınılmaz olan, genel ifadeler, maddelerin yorumlanmasında, öğrenciler arası bireysel farklılıkların olması gerektiğinden daha etkin hale getirmekte ve bu durum öznel değerlendirmelere sebep olmaktadır.

Belli puanların anlam ve önemi öğrenciler ve maddeler arasında farklı-laşmaktadır. Bu farklılaşma önemlidir ve üniversitelere kullanılan anket-

lerin maddeleri yeniden yazılarak ve geçerlik güvenirlik çalışması yapılarak daha tutarlı bir anketler elde etmek mümkündür. Ayrıca bu sonuçlar tüm anket hazırlayanları ve araştırma tekniklerinde veri toplama araçlarının hazırlanması konularında çalışma yapanlar için önemlidir.

Öğretim elemanının kontrolü dışındaki öğrenci tabanlı sebepler, anketlerin puanlanmasında öğretim elemanının lehine veya aleyhine bir değişim göstermektedir. Öğrencilerin değerlendirme yeteneklerinde, öğrenme stillerinde ve tercihlerindeki farklılıklar bu değişimin sebebi olabilir. Böylece öğrenciler maddeyi hazırlayandan farklı olarak anladıklarında veya maddeyi anlamadıklarında, maddenin ölçmek istediği şeyi değil, kendi kişisel deneyimleri ve bilgisinin uyandırdığı şeyi ölçüt olarak puan vermektedirler.

Öğrencilerin beğenmedikleri, eleştirdikleri veya hakkında bilgiye sahip olmadıkları maddelerin puanlanmasında maddenin ölçmeyi amaçladığından farklı bir durumu ölçtükleri saptanmıştır. Bunu yaparken bilişsel bilgiden çok; bilişsel olmayan duygu, tutum ve değerlendirme becerilerine dayanarak öğretim elemanın lehinde veya aleyhinde puan vermiş oldukları, belirttikleri nedenlerden anlaşılmıştır.

Puanlama nedenlerinin fazlalığı ve farklılığı ile dereceleme ölçeğindeki farklılaşmalar dikkat çekici sonuçlardandır. Bu durum, öğrencilerin öğretimi değerlendirme yeteneklerinin birbirinden çok farklı olabileceği gerçeğini ortaya çıkarmaktadır.

Öğrenci değerlendirmelerin de geçerliğin denetlenmesinin çok zor olduğunu, ortada birden fazla etkin öğretim ölçütü bulunduğunu belirten Kemp ve Kumar'ın (1990), endişesinin haklılığı aynı sınıftaki öğrencilerin aynı öğretim elemanı için aynı maddede farklı nedenlere aynı puanları vermeleri sonucu ile desteklenmektedir. Farklı nedenler ise, öğrencilerin bir maddede farklı şeyleri ölçüt olarak puan verdiklerinin göstergesidir. Bu durum söz konusu anket maddesinin geçerliği hakkında ciddi şüpheler uyandırmaktadır. Benzer şekilde bir madde ile ilgili olarak, aynı sınıftaki öğrenciler, aynı öğretim elemanı için aynı nedene farklı puan vermişlerdir. Bu ise bir madde ile ilgili tek bir nedene atanan değerlerin öğrenciden öğrenciye değiştiğinin göstergesidir.

Bu durum ile ilgili diğer ilginç sonuç ise, çok net olduğu ve tek bir cevabı olduğu düşünülen bir maddede bile, aynı sınıftaki öğrencilerin aynı öğretim elemanı için bu denli birbiri ile çelişen ve geniş yelpazede neden belirtmeleri ve bu nedenlere de farklı puan vermiş olmalarıdır. Bu sonuçlar, anket doldurma sürecinde bireysel farklılıkların ne derece etkili olduğunun ve anketlerde güçlü bir geçerlik sorununun varlığının göstergesidir. Kemp ve Kumar'da (2000), öğrenci değerlendirmelerinin geçerliğini denetlemenin gerçekten zor bir iş olduğunu açıklamışlardır.

Analiz sonucunda maddelere verilen nedenlerde hale etkisi görülmüştür. Hale etkisi yaratan nedenler birden fazla maddede etkin ölçüt olarak kullanıldıklarından, bu durum, öğrenciler tarafından hangi boyutların etkili öğretim kapsamında yer aldığını göstermesi açısından önemlidir. Ancak anketlere verilen puanlar üzerinde sadece istatistiksel analizle araştırma yapanlar “hale Etkisi”nin olumsuz olduğunu düşünerek bu durumdan kaçınmak için maddelerin çok özel davranışları ölçecek şekilde daraltılmasını önermişlerdir (Low 1996). Anketler üzerinde nitel araştırma yapanlar ise hale etkisinden kaçınmak için soruların bu denli daraltılmasının, etkili öğretim açısından tek tip öğretim ve öğretim elemanı yaratacağını belirterek, bu durumun olumsuz etkilerinin olabileceğini vurgulamışlardır (Johnson, 2000). Ayrıca herkes tarafından kabul görmüş “etkili öğretimin” bir tanımı olmadığından maddeleri, bu derece daraltmak da mümkün görünmemektedir. Bu şartlar altında anketlerde “hale etkisi” az ya da çok mutlaka görülecektir. O halde bu sonuç avantajla dönüştürülebilir. Anketlerde hale etkisini ortaya çıkaran neden ya da nedenler belirlenerek, bunlar etkili öğretimin tanımlanmasında kullanılabilir.

Yukarıdaki analizler sonucunda, sadece anketlerden gelen puanlamalar yoluyla toplanan verilerin, formative ve summative değerlendirmeler için sağlıklı olamayacağı açıktır. Damron,(1995); Haskell, (1997a,1997b,1997c, 1997d), Martinson da (2000), anketlerin summative amaçlı kullanımına karşı çıkmaktadırlar. Bu durum yurt dışındaki üniversiteler için geçerlidir. Çünkü ülkemizde bu anketler zaten summative amaçlı kullanılmamaktadır.

Anketlere puan verilirken nedenlerinin de belirtilmesinin özellikle formative değerlendirmeleri daha işlevsel kılacağı sonucuna varmışlardır. Bunun yanı sıra değerlendirmelerde veriler çeşitli araçlarla toplanmalı ve bunların ortak sonucuna dayanarak formative ve summative değerlendirmeler yapılmalıdır.

Bu çalışmada kullanılan anket puanlamadaki neden yazma biçimine benzer bir işlem kullanılarak anketlerin oluşturulması ve geliştirilmesi öneriler arasındadır. Değerlendirme sürecinin önemli bir ögesi olan öğrencilerin, yalnızca puanları yerine, “orijinal fikirlerinin ” de öğretim sürecine katılmasını sağlamak öğretimi ve öğretim elemanının kalitesini artırabilir. Üniversitelerimiz kullanılan bu anketlerin geçerliğine katkı sağlayabilir.

Üniversitelerimizde kullanılan anket maddelerinin değerlendirilmesi ve kullanılan bu anketlerin sonuçlarının nasıl değerlendirildiğinin araştırmaları ivedilikle yapılmalıdır.

### **Summary**

Student evaluation of teaching (SET) by questionnaire is one of the methods of evaluating instruction and instructors in higher education institutions. The data obtained from students' evaluation questionnaires have been widely used in universities across the world so as to evaluate the quality of instructors and instruction (Seldin, 1985; Abrami, 1989a; Wagenaar, 1995; Neath, 1996; Hobson and Tablot, 2001; Grant, 1998; Platt, 1993).

Whereas over 2000 studies concerning the issue have been available across the globe, no studies on this issue have been encountered in Turkey. The topic of how students interpret questionnaire items was studied by very few researchers – namely Low (1988, 1991, 1996, and 1999), Solas, (1990), Alderson (1992), Block (1998), Robertson (2004).

Considering the work of filling in a questionnaire as a process of communication, it is of critical importance to identify what are the probable effects of the message coming from the source to the receiver are, that is to say, the effects following the reading of the questionnaire. Hence this study aims at revealing the way students perceive the closed-ended questions in the questionnaire and their individual scale in scoring the survey items through students' comments.

Thus, examination of the change in students' perception of the questions will result in designing questionnaires which have attained validity and reliability, and in turn, this will lead to more objective formative and summative evaluations. This research is based on Robertson (2004), Newell and Simon's (1972) views which consider badly structured problem-solving as reaching an ambiguous statement. Since the work of filling in the questionnaire is regarded as ill-structured problem-solving duty, the results are described on this basis.

Answers are sought for the following questions:

1. Are there any differences between the general ratings assigned to instructors, to the form and the questionnaire?
2. Do the statements in the questionnaire items affect students' perception?
3. How do students' perceptions vary independent of the statements in questionnaire items?
4. What kind of correlations holds between the scores assigned by students to questionnaire items and the reasons for the ratings?

## **Method**

This is a descriptive study. This research was conducted in two phases. The results emerging while seeking answers to the first and second questions led to finding answers to the third and fourth questions. Consequently, the data were collected at two subsequent times in the same department, for the same students and for the same classes; however, with a difference in the size of the sample. The research was the students of Gazi Teacher Training College of Gazi University in Ankara, Turkey. For the first and second research questions, the sample was composed of the 192 volunteering junior students of Elementary Education Class Teaching, Pre-School Education and Social Sciences Teacher Education branches of Primary School Teaching Department of Gazi School of Education of Gazi University.

As for the third and fourth research questions, 138 students were used to form the sample (through Purposive Sampling Approach).

The data concerning the first and second research questions were obtained in the fall semester of 2004- 2005 academic year whereas the ones concerning the third and fourth questions were obtained in the spring semester of the same academic year.

### *Data Collection Tool for the First and Second Research Questions*

In this part of the research, the data collection tool used in instructor evaluation at Gazi Teacher Training College was employed for the purposes of data collection. The tool had 20 items and Likert type 5-pointed grading scale (namely, 5: excellent, 4: very good, 3: good, 2: average, 1: poor).

### *Data Collection Tool for the Third and Fourth Research Questions*

“Teacher Evaluation questionnaire”, which had been utilized for the first and second questions, was also used for the collection of data concerning the third and fourth questions. The reason for applying the same survey to a different sample for the second time was to perform an appropriate analysis of students’ comments in accordance with the third and fourth questions. A space was provided next to each item so that students could write their reasons for their rating.

### *Qualitative Analysis*

Students wrote 3517 comments for the 20 closed-ended questions concerning the second research question whereas they wrote 2574 comments

for the questions concerning the third and fourth research questions. Qualitative analysis was performed for the first research question. Quantitative analysis was conducted in a similar way for the second, third, and fourth research questions.

Conventional content analysis was utilized for the analysis of the data. The data were classified and analyzed through a tool called “rating-reason analysis matrix” in the current research.

### *Findings and Interpretations*

An insignificant difference was found for the overall ratings assigned to a course, form, and questionnaire, regarding the first research question.

Concerning the second research question, overlaps in some questions were found to cause indecisiveness in rating. Some statements with a conditional clause were rated, and reasons given for some items were various; which demonstrated that some of the items were not sufficiently comprehended or misunderstood. It was further found that even those students who did not like the items or who criticized them also rated the questionnaire.

As for the third question of the research, students mentioned some features out of the content of the items as their reasons; and thus Halo effect was found. Contrasts were noticed even for the items which are considered to be clear.

With regard to the fourth research question, an interesting differentiation was identified in the grading scale. Differences and excessiveness was found in the reasons for scoring all the survey items. Students assigned scores even when they were not able to remember appropriately or were not certain of their answers, and thus it was found that they had assigned ratings even to the items of which they had no knowledge.

### *Discussion and Suggestions*

The results obtained were reached by writing the reasons. The action of assigning a rating to a questionnaire item is based on the first thing coming into mind. If the reasons had not been required, evaluators of the questionnaire could not have determined the criterion on which scores were given.

The way of expressing an idea is not understood appropriately in the questionnaire, which is the official questionnaire of the university and which has also been used in our research. It was found that the reference of the items were in a narrower framework than the students had perceived. The



likelihood of planning to measure so wide range of behaviors by the questionnaire designer is considered low.

Student- based reasons which are out of the instructors' control can display changes in favor of or against the instructor in rating the questionnaires. Differences in students' evaluation skills, learning styles and preferences may stand as the reasons for such changes. In consequence, when students understand an item differently from what the item designer meant or when they do not understand it at all, they tend to base their rating the item on their own experiences or knowledge rather than on the element that is to be measured by the item.

### **Kaynakça / References**

- Abrami, P.C & D. Mizener. (1983). "Does the attitude similarity of college professors and their students produce 'bias' in course evaluations?" *American Educational Research Journal*, 20, 123-136.
- Abrami, P. C. (1989). How should we use student ratings to evaluate teaching? *Research in Higher Education*. 30, 221-227.
- Abrami, P. C. (1989). SEEQing the truth about student ratings of instruction. *Educational Researcher*. 18 (1), 43-45.
- Abrami, P. C., Perry, R. P., & Leventhal, L. (1982). The relationship between student personality characteristics, teacher ratings and student achievement. *Journal of Educational Psychology*, 74 (1), 111-125.
- Abrami, P.C. & d'Apollonia, S. (1991). Multidimensional students' evaluations of teaching effectiveness; the generalizability of "n = 1" research: comment on marsh (1991). *Journal of Educational Psychology*. 83, 411-415.
- Alderson, J. C. (1992). Validating questionnaires. centre for research in language education. Working Papers No. 15. University of Lancaster.
- Aleamoni, L. M. (1981). Student ratings of instruction. In J. Millman (Ed.), *Handbook of teacher evaluation*. (pp. 110-145). Beverly Hills, California: Sage.
- Bedgood, R.E., & Polard, R.J. (1999). Uses and misuses of student opinion surveys in eight Australian Universities. *Australian Journal of Education*, 43 (2), 129-156.
- Block, D. (1998). Exploring interpretations of questionnaire items. *System*, 26 (3), 403-425.
- Braskamp, L. A., & Ory, J. C. (1994). Assessing faculty work: enhancing individual and institutional performance. San Francisco: Jossey-Bass.

- Braskamp, L. A., Ory, J. C. & Pieper, D. M. (1981). Student written comments: dimensions of instructional quality, *Journal of Educational Psychology*, 73, 65-70.
- Cashin, W.E. (1990). Student rating of teaching: Recommendations for use. IDEA Paper no 22. Kansas State University, Center for Faculty Evaluation & Development.
- Cashin, W. E. (1996). Developing an effective faculty evaluation system, IDEA Paper no 33. Kansas State University, Center for Faculty Evaluation & Development.
- Cashin, W. E. (1995). Student Ratings of Teaching: A summary of The Research. IDEA Paper no. Kansas State University, Center for Faculty Evaluation & Development.
- Centra, J. A. (1993). Reflective faculty evaluation: enhancing teaching and determining faculty effectiveness. San Francisco: Jossey-Bass.
- Centra, J. A. & Creech, F. R. (1976). The relationship between student, teachers, and course characteristics and student ratings of teacher effectiveness. Project Report 76- 1, Educational Testing Service, Princeton, NJ.
- Chen, Y., Hoshower, L. B. (2003). Student Evaluation of Teaching Effectiveness: an assessment of student perception and motivation, *Assessment & Evaluation in Higher Education*, 28 (1), 71-88.
- Clouder, L. (1988). Getting The 'Right Answer': Student Evaluation As A Reflection On Intellectual Development, *Teaching in Higher Education*, 3 (2), 185-196.
- Cohen, P. A. (1981). Student ratings of instruction and student achievement: a meta-analysis of multisection validity studies. *Review of Educational Research*, 51, 281-309.
- Damron, J. C. (1995). The three faces of teaching evaluation, Unpublished Manuscript, Douglas College, New Westminster, British Columbia.
- DeCanio, S. J. (1986). Student evaluations of teaching – a multinomial logit approach, *Journal of Economic Education*. 17, 165-175.
- Feldman, K. A. (1976). The superior college teacher from the student's view, *Research in Higher Education*, 5, 243-288.
- Feldman, K. A. (1977). Consistency and variability among college students in rating their teachers and courses: a review and analysis. *Research in Higher Education*, 6 (3), 223-274.
- Feldman, K. A. (1978). "Course characteristics and college students ratings of their teachers: what we know and what we don't". *Research in Higher Education*, 9, 199-242.

- Feldman, K. A. (1987). Research productivity and scholarly accomplishment of college teachers as related to their instructional effectiveness: a review and exploration. *Research in Higher Education*, 26 (3), 227-298.
- Feldman, K. A. (1989a). The Association between student ratings of specific instructional dimensions and student achievement: refining and extending the synthesis of data from multisection validity studies. *Research in Higher Education*, 30 (6), 583-645.
- Feldman, K. A. (1989b). Instructional effectiveness of college teachers as judged by teachers themselves, current and former students, colleagues, administrators and external (neutral) observers. *Research in Higher Education*, 30, 583-645.
- Feldman, K. A. (1993). College students' views of male and female college teachers: part ii- evidence from students' evaluations of their classroom teachers. *Research in Higher Education*, 34 (2), 151-211.
- Felton, J., Mitchell, J., & Stinson, M. (2004) Web-based student evaluation of professors: the relations between perceived quality, easiness and sexiness, *Assessment and Evaluation in Higher Education*, 29 (1), 91-108.
- Gick, M. L. (1986). Problem Solving Strategies. *Educational Psychologist*, 21 (1&2), 99-120.
- Glaser, B. G., & Strauss, A. L. (1967). *The Discovery of Ground Theory*. Chicago: Aldine.
- Gramlich, E. M. and Greenlee, G. A. (1993). Measuring Teaching Performance, *Journal of Economic Education*, 24, 3-13.
- Grant, H. (1998). Academic contests: Merit pay in Canadian universities, *Relations Industrially / Industrial Relations* 53 (4), 647-664.
- Green, B. P., Caleron, T. G. & Reider B. P. (1998). A Content analysis of teaching evaluation instruments used in accounting departments, *Issues in Accounting Education*, 13 (1), 15-30.
- Haskell, R. E. (1997a). Academic freedom, tenure, and student evaluations of faculty: galloping polls in the 21st century, *Education Policy Analysis Archives* 5 (6), February 12.
- Haskell, R. E. (1997b). Academic Freedom, Promotion, Reappointment, Tenure, and The Administrative Use of Student Evaluation of Faculty (SEF): (Part II) Views from court, *Education Policy Analysis Archives* 5 (6), August 25.
- Haskell, R. E. (1997c). Academic Freedom, Promotion, Reappointment, Tenure, and The Administrative Use of Student Evaluation of Faculty (SEF): (Part III) Analysis and implications of views from the court in relation to accuracy and psychometric validity, *Education Policy Analysis Archives* 5 (6), August 25.

- Haskell, R. E. (1997d). Academic Freedom, Promotion, Reappointment, Tenure, and The Administrative Use of Student Evaluation of Faculty (SEF): (Part IV) Analysis and implications of views from the court in relation to academic freedom, standards, and quality of instruction, *Education Policy Analysis Archives* 5 (6), November 25.
- Haynes, B. (2002). The use of student evaluations as an indicator of teaching quality in higher education. Paper presented at the Conference of Australians Association for Research in Education, Brisbane.
- Hobson, S. M. & Talbot, D. M. (2001). Understanding student evaluations, *College Teaching*, 49 (1), 26–31.
- Hoffman, F. E. & Kremer, L. (1980). Attitudes toward higher education and course evaluation, *Journal of Educational Psychology*, 72 (5), 610–617.
- Hong, S. N. (1998). The Relationship between Well- Structured and Ill-Structured Problem Solving in Multimedia Simulation. Unpublished Doctoral Thesis. The Pennsylvania State University, The Graduate School College of Education.
- Howard, G. S., & Maxwell, S. E. (1980). Correlation between student satisfaction and grades: a case of mistaken causation? *Journal of Educational Psychology*, 72, 810-820.
- Howard, G. S., & Maxwell, S. E. (1982). Do grades contaminate student evaluations of instruction? *Research in Higher Education*, 16,175-188.
- Husbands, C. T. (1996). Variations in students' evaluations of teachers' lecturing and small group teaching: A study at the London school of economics and political science. *Higher Education*, 33(1), 51-70.
- Husbands, C. T. & Fosh, P. (1993). student evaluations of teaching in higher education: Experiences from four European countries and some implications of the practice. *Assessment and Evaluation in Higher Education*, 18(2), 95-114.
- Johnson, R. (2000). The authority of the student evaluation questionnaire. *Teaching in Higher Education*; 5 (4), 419-434.
- Jonassen, D. H. & Grabowski, B. L. (1993). Handbook of Individual Differences, Learning and Instruction. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Jonassen, D. H. & Tessmer, M. (1996/1997). An outcomes-based taxonomy for instructional systems design, evaluation, and research. *Training Research Journal*, 2, 11-46.
- Jonassen, D. H. (1997). Instructional design model for well-structured and ill-structured problem-solving learning outcomes. *Educational Technology: Research and Development*. 45 (1), 65-95.

- Jonassen, D. H. (2000a). Integrating Problem Solving Into Instructional Design. in R.A. Reiser & J. Dempsey (Eds.), *Trends and issues ion instructional design and technology*. Upper Saddle River, NJ: Prentice-Hall.
- Jonassen, D. H. (2000b). Towards a design theory of problem solving. *Educational Technology Research and Development* 48 (4), 63-85.
- Kember, D. & Wong A, (2000). Implication for evaluation from a study of students' perceptions of good and poor teaching, *Higher Education*, 40 (1), 69-97.
- Kemp, B. and G.Kumar (1990). Student evaluations: are we using them correctly, *Journal of Education for Business*, 65, 106-111.
- Kitchener, K. S. (1983). Cognition, metacognition, and epidemic cognition: a tree-level model of cognitive processing, *Human Development*, 4, 222-232.
- Kolitch, E & Dean, A. V. (1999). Students' ratings of instruction in the usa: hidden assumptions and missing conceptions about 'good' teaching, *Studies in Higher Education*. 24 (1), 27-42.
- Kulik, J. & McKeachie, W. (1975). The Evaluation of Teachers in Higher Education. Ed. Kerlinger, F. (Ed). *Review of Research in Education*. Itasca: IL: F.E. Peacock.
- LaForge, M. C. (2003). Student mood and teaching evaluation ratings. *Journal of The Academy of Business Education*, 4.
- Lechtreck, R. (1990). College Faculty Evaluation By Students - an Opportunity For Bias. *College Student Journal*, 24 (September), 297-299.
- Lewis. K (1991). Making sense of student written comments, *New Direction for Teaching and Learning*, 87, 25-32.
- Low, G. (1988). The semantics of questionnaires rating scales' *Evaluation and Research in Education*, 2, 69-79.
- Low, G. (1991). Talking to Questionnaires Pragmatic Models in Questionnaire Design. In B Heaton, P Adams and P. Howarth (eds). *Sociocultural aspect of english for academic purposes*. London, Macmillan.
- Low, G. (1996). Intensifiers and hedges questionnaire items and the lexical invisibility hypothesis, *Applied Linguistics*, 17 (1), 1-36.
- Low, G. (1999). What respondents do with questionnaires: accounting for incongruity and fluidity. *Applied Linguistics*, 20 (4), 503-533.
- Marsh, H. W. (1984). Students' evaluations of university teaching: dimensionality, reliability, validity, potential biases, and utility. *Journal of Educational Psychology*, 76, 707-754.
- Marsh, H. W. (1987). Students' evaluations of university teaching: research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, 11 (3), 253-388.

- Marsh, H. W., & Bailey, M. (1993). Multidimensional students' evaluations of teaching effectiveness. *Journal of Higher Education*, 64(1), 1-18.
- Marsh, H. W., & Dunkin, M. (1992). Students' evaluations of university teaching: a multi-dimensional perspective. In J. C. Smart (Ed.), *Higher Education: Handbook on Theory and Research*, New York: Agathon Press, pp. 143-234.
- Marsh, H. W. & Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective, *American Psychologist*, 52 (11), 1187-1197.
- Martinson, D. L. (2000). Student evaluations of teaching and their short term validity. *Journalism and Mass Communication Educator*, 54 (4), 77-82.
- McKeachie, W. (1997). Students ratings: the validity of use, *American Psychologist*, 52 (11), 1218-1225.
- McKeachie, W. J. (1979). Student ratings of faculty: a reprise. *Academe: Bulletin of the AAUP*, 65(6), 384-397.
- Murray, H. G. (1983). Low-inference classroom teaching behaviors and student ratings of college teaching effectiveness. *Journal of Educational Psychology*, 75 (1), 138-149.
- Neumann, R.(2000). Communicating student evaluations of teaching *Assessment and Evaluation in Higher Education*, 25 (2), 121-134.
- Newell, A. & Simon, H. A. (1972). *Human Problem Solving* (Englewood Cliffs, NJ; Prentice Hall).
- Platt, M. (1993). What student evaluations teach, *Perspectives in Political Science* 22 (1), 29-40.
- Polson, P. & Jeffries, R.(1985). Instruction in problem- solving skills: an analysis of four approaches. In J. W. Segal, S. F. Chipmen, & R. Glaser (eds.), *Thinking and learning skills*. 1,417-455. Hillsdale, NJ: Lawrence Erlbaum.
- Remmers, H.H. (1927). The Purdue rating scale for instructors. *Educational Administration and Supervision*, 6,399-406.
- Remmers, H.H. (1928).The relationship between students' marks and students' attitudes toward instructors. *School and Society*, 28,759-60.
- Remmers, H.H. (1930). To what extent do grades influence student ratings of instructors? *Journal of Educational Psychology*, 21, 314-16.
- Remmers, H.H., & Brandenburg, G.C. (1927). "Experimental data on the Purdue rating scale for instructors." *Educational Administration and Supervision*, 13, 519-27.
- Remmers, H.H., Martin F.D., & Elliot D.N. (1949). Are students ratings of their instructors related to their grades? *Purdue University Studies in Higher Education*. 44,17-26.

- Robertson, I. (2004). Student perceptions of student perception of module questionnaires: questionnaire completion as problem solving, *Assessment and Evaluation in Higher Education*, 29 (6), 662-669.
- Seldin, P. (1985). Current practices in evaluating business school faculty. Pleasantville, NY: Center for Applies Research, Lubin School of Business Administration, Pace University.
- Shevlin, M., Banyard, P., Davies, M., & Griffiths, M. (2000). The validity of student evaluation of teaching in higher education: Love me love my lectures? *Assessment and Evaluation in Higher Education*, 25(4), 397- 405.
- Sinnott, J. D. (1989). A model for solution of ill-structured problems: implications for everyday and abstract problem solving. In J. D. Sinnott (Ed.), *Everyday problem solving: Theory and applications*, 72-99. New York: Praeger.
- Sixbury, G. R., & Cashin, W. E. (1995a). *IDEA technical report no. 9: Description of database for the IDEA Diagnostic Form*. Manhattan: Kansas State University, Center for Faculty Evaluation and Development.
- Sixbury, G. R., & Cashin, W. E. (1995b). *IDEA technical report no. 10: Comparative data by academic field*. Manhattan: Kansas State University, Center for Faculty Evaluation and Development.
- Solas, J. (1990). Effective teaching as construed by social work students. *Journal of Social Work Education*, 26, 145-154.
- Tagomori, H. & Bishop, L. (1995). Student evaluation of teaching: flaws in the instruments, *Thought and Action*, 11, 63-78.
- Van Manen, M. (1990). *Researching lived experience: human science for an action based pedagogy*. London, Ont: The Althouse Press.
- Voss, J. F., & Post, T. A. (1988). On the solving of ill-structured problems. In M.T.H. Chi, R. Glaser and M.J Farr (Eds), *The nature of expertise* (261-285). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Wagenaar, T. A. (1995). Student evaluation of teaching: some cautions and suggestions. *Teaching Sociology*, 23 (1), 64-8.
- Wagner, Z.M.(1999). Using student journals for course evaluation. *Assessment and Evaluation in Higher Education*, 24(3), 261-272.
- Wetzstein, M. E., Broder J. M., & Wilson G. (1984). Bayesian Inference and student evaluations of teachers and courses. *Journal of Economic Education* 15 (Winter): 40-45.
- Whetley, G. H. (1984). Problem solving in school mathematics, MEPS Technical Report no. 84.01 (West Lafayette, in, Purdue University School of Mathematics and Science Center).

*Nurdan Kalaycı*

- Wigington, H., Tollefson, N., & Rodriguez, E. (1989). Students' ratings of instructors revisited: interactions among class and instructor variables. *Research in Higher Education*, 30 (3), 331-44.
- Worthington, A.C. (2002). The impact of student perceptions and characteristics on teaching evaluations: a case study in finance education. *Assessment and Evaluation in Higher Education*, 27(1), 49-64.
- Wood, P. K. (1983). Aspects of problem solving: an inquiry systems perspective of decision theory. *Human Development*.

**İletişim/Communication:**

Yard. Doç. Dr. Nurdan Kalaycı  
Eğitim Bilimleri Bölümü  
Gazi Eğitim Fakültesi  
Gazi Üniversitesi  
kalayci@gazi.edu.tr

Alındığı tarih/received: 24/11/2007  
Düzeltilme/revision: 04/04/2008  
İkinci düzeltme/second revision: 22/07/2008  
Kabul/approved: 23/07/2008  
kalayci@gazi.edu.tr



Yükseköğretimde Öğretimi Değerlendirme

**Ek:1:**

Gazi Üniversitesinde öğretim ve öğretim elemanını değerlendirme formunda yer alan maddeler

	Öğretim Elemanı ve Dersi Değerlendirme Maddeleri	5: mükemmel, 4: çok iyi, 3: iyi, 2: orta, 1: zayıf).					Nedenler
		5	4	3	2	1	
1.	Dönem başında işlenecek konuları ve dersin amaçlarını açıkça belirtir.						
2.	Derse hazırlıklı gelir.						
3.	Dersin başlama ve bitiş saatlerine uyar.						
4.	Ders konusuna hakimdir.						
5.	Derste ses tonunu, dil ve ifadelerini uygun olarak kullanır						
6.	Dersi ilgi çekici ve zevkli hale getirir.						
7.	Öğretim amaçlarını yerinde ve etkin kullanır.						
8.	Derste ders dışı konulara gereğinden fazla önem vermez.						
9.	Not kaygısı yaratmadan sınıf yönetimini sağlar.						
10.	Derste demokratik ortam yaratarak, öğrencilerin soru sorma ve görüş bildirmelerine fırsat verir.						
11.	Öğrencilerle iletişim kurmada başarılıdır.						
12.	Öğrencilere karşı davranışlarında eşitlik ilkesine uyar.						
13.	Kişisel problemlerini derse ve öğrenciye yansıtmaz.						
14.	Görünüş ve davranışları ile model oluşturur.						
15.	Öğrencileri araştırma ve inceleme gibi aktivitelere yöneltir						
16.	Her dersin sonunda konuyu özetler ve önemli noktaları vurgular						
17.	Ders dışında, gerekli durumlarda öğrenciye zaman ayırır.						
18.	Sınav soruları dersin amaç ve içeriğini yansıtır						
19.	Sınavların değerlendirilmesinde objektif davranır						
20.	Sınavları ve ödevleri zamanında değerlendirerek öğrencileri hata ve eksikleri konusunda uyarır						