

# Tanıma Belleğinde Kriter Belirleme ve Tepki Yanlılığı Değişimlemeleri

Cavit Deniz Pala

Aslı Kılıç

Orta Doğu Teknik Üniversitesi

Orta Doğu Teknik Üniversitesi

## Özet

Tanıma belleği, geçmişte deneyimlenen uyarınları diğerlerinden ayırt etme becerisidir. Tipik bir tanıma belleği görevinde bir dizi uyarının katılımcıya sunulduğu çalışma aşamasının ardından test aşamasında sorulan maddelerin çalışma listesinde bulunup bulunmadığının ayırt edilmesi beklenir. Sinyal belirleme kuramına (Macmillan ve Creelman, 2005) göre test maddeleri bellek izleriyle kıyaslanarak bir bellek gücü değeri elde edilir. Bu bellek gücü değeri öznel karar kriterini aşılırsa madde tanınır. Test performansını belirleyen iki etken ise eski ve yeni maddeleri ayırt edebilme düzeyi olan duyarlılık ve sorulan maddelere “evet” yanıtı verme yatkınlığını ifade eden tepki yanlılığıdır. Alanyazında karar kriterinin çeşitli yöntemlerle deneysel olarak değiştirildiği çalışmalar mevcuttur. Bu yöntemlerden biri olan temel oran değişimlemesinde test listesindeki eski madde oranı değişimlenir ve kişilerin bu oranlara göre kriterlerini değiştirmesi beklenir. Temel oran değişimlemesi genellikle tek başına kriter değişimini sağlamaya yeterli olmamakta, ancak temel oranlar konusunda ön bilgi verilmesi, geribildirim gibi yardımcı ipuçlarıyla kriter değişimi sağlanabilmektedir. Bir diğer yöntemde ise eski ya da yeni maddelerin bellek güçleri değiştirilerek kriter değişiminin sağlanması amaçlanmaktadır. Bu yöntemlerin yanı sıra ödeme matrisleri, olasılıksal ipuçları, yanlı geribildirim kullanımı gibi değişimlemelerle de kriter değişimi elde edilmektedir. Başlangıçtaki kriter konumunun ve kriter değiştirmeye olan yatkınlığın kişiler arasında farklılıklar gösterdiği, bununla birlikte kişilerin kriterlerinin deneyler arasında tutarlı olduğu bilinmektedir.

**Anahtar kelimeler:** Tanıma belleği, sinyal belirleme kuramı, kriter, tepki yanlılığı

## Abstract

Recognition memory is the ability to distinguish previously experienced stimuli from new ones. In a typical recognition memory task, a series of items are presented to the participant, followed by a test phase in which they are expected to distinguish whether the probe item is on the study list. According to signal detection theory (Macmillan & Creelman, 2005), a probe item constitutes a memory strength value by being compared to the memory traces. The item is recognized if its memory strength exceeds the subjective decision criterion. Sensitivity, which is the extent to discriminate between the old and the new items, and response bias, the tendency to say “yes” to a probe item, are the two factors determining the performance. There are studies in the literature in which the response criterion is manipulated by several methods. In one of these methods, the proportion of old items in the test list is manipulated, and respondents are expected to change their criteria accordingly, a method called base rate manipulation. The base rate manipulation alone is usually limited in causing criterion changes; however, presenting it together with some additional clues, such as prior information about the base rates or feedback can provide robust criterion changes. In another method, the memory strength of the old or new items is manipulated to produce criterion shifts. Besides these methods, manipulations, such as payoff matrices, probabilistic cues or biased feedback resulted in a criterion shift. It is also reported that there are interpersonal differences in the initial criterion placement and the tendency to shift the criterion, but the criterion of the individuals is consistent across the experiments.

**Keywords:** Recognition memory, signal detection theory, criterion, response bias

**Yazışma Adresi:** Arş. Gör. Cavit Deniz Pala, Orta Doğu Teknik Üniversitesi, Fen Edebiyat Fakültesi, Psikoloji Bölümü, 06800 Çankaya / Ankara

**E-posta:** denizp@metu.edu.tr

**Gönderim Tarihi:** 09.10.2018

**Kabul Tarihi:** 05.02.2019

**Tablo 1.** Evet-Hayır Tanıma Belleği Görevinde Olası Sonuçlar Matrisi

Uyaran	Tepki	
	“Evet”	“Hayır”
Eski	İsabet	İska
Yeni	Yanlış Alarm	Doğru Ret

Bir kişiyle karşılaştığımızda onu tanıyıp tanımadığımızı o kişiye olan aşinalığımız ve o aşinalığı tanıma olarak değerlendirip değerlendirmedığımız belirler. Öyleyse birini tanımada başlıca rol oynayan iki etken, ona ne kadar aşına olduğumuz ve ne kadar aşına olursak onu tanıdığımızı kanaat getireceğimizi belirleyen “kriter”imizdir. Bu makalede tanıma belleğinde karar kriterine dair bulguların ve yorumların özetlenmesi ve değerlendirilmesi amaçlanmaktadır. Karar kriteri, tanıma belleğini açıklayan çeşitli modellerin (Cox ve Shiffrin, 2017; Dennis ve Humphreys, 2001; McClelland ve Chappell, 1998; Shiffrin ve Steyvers, 1997; Turner, Van Zandt ve Brown, 2011) bir bileşeni olması nedeniyle bellek süreçlerini açıklamada önemli bir çalışma konusudur. Bununla birlikte, tanıma belleği, çeşitli hatırlama belleği görevleri ve bellek dışı karar verme görevlerindeki karar süreçlerinin ilişkili olduğu araştırmalarla ortaya koyulmuştur (Kantner ve Lindsay, 2012, 2014). Söz konusu bilişsel süreçler arasındaki ilişkilerin doğasının açıklanmasında karar kriteri araştırmaları önemli bir yere sahiptir. Tanıma belleğinde kriter değişimleme konulu araştırmaların son yıllardaki artışı ve tartışmaların güncelliğine karşın Türkçe alanyazında bu konudaki çalışmaların derlendiği bir makale ya da kitap mevcut değildir. Bu gözden geçirme makalesinin amacı alanyazındaki bu ihtiyacı karşılamaktır. Makale, tanıma belleği görevlerinin kısaca tanıtılmasıyla başlayacak, daha sonra tanıma belleğinde performansın açıklanması ve ölçülmesinde kullanılan başlıca ölçüm modeli olan sinyal belirleme kuramı (SBK; Banks, 1970; Green ve Swets, 1966; Macmillan ve Creelman, 2005; Swets, Tanner ve Birdsall, 1961) açıklanacaktır. Sonraki bölümlerde ise tanıma belleğinde karar kriterinin deneysel olarak değişimlenmesi amacıyla yapılan deneyler, bu deneylerde elde edilen bulgular ve bulguların kriter temelli ve alternatif açıklamalarının detaylı olarak özetlenmesiyle söz konusu

çalışma alanının güncel durumunun ortaya konması hedeflenmektedir.

Tanıma belleği bir uyarının (kişi, nesne ya da durum) daha önce deneyimlenip deneyimlenmediğini ayırt etme becerisidir. Tipik bir tanıma belleği görevi olan “evet-hayır” madde tanıma (“yes-no” item recognition) çalışma, oyalama (distractor) ve test aşamalarından oluşur. Çalışma aşamasında katılımcılara bir dizi uyarın sunulur ve bunları daha sonraki test aşaması için öğrenmeye çalışması beklenir. Genellikle aritmetik işlemlerden oluşan oyalama görevinden sonra gelen test aşamasında, bir kısmı çalışma aşamasında gösterilen (eski), diğerleri ise yeni uyarılardan oluşan bir liste katılımcıya karışık sırayla sunulur ve ondan yeni ile eski uyarıların birbirinden ayırt etmesi beklenir. Evet-hayır testinin yanı sıra, yanıtların Likert tipi bir güven aralığı ölçeğinde verildiği güven kararı (confidence judgement) testi; birden fazla seçeneğin sunulup eski kelimenin seçilmesinin beklendiği zorunlu seçim (forced choice) testi; uyarıların tekil maddeler değil, birlikte sunulan madde çiftlerinden oluştuğu ve bu çiftlerin test aşamasında yeniden düzenlenmiş çiftlerden ayırt edilmesinin beklendiği ilişkisel tanıma (associative recognition) görevi; ayrık çalışma ve test aşamaları yerine, sunulan her uyarının daha önce görülüp görülmediğinin sorulduğu sürekli tanıma (continuous recognition) görevleri tanıma belleği görevlerinin başlıca varyantlarıdır.

Bir evet-hayır tanıma belleği testinde eski ve yeni maddelere verilen “evet” ve “hayır” yanıtları toplam dört sonuç olasılığı oluşturur (Tablo 1): Eski bir maddeyi tanımak “isabet” (hit) olarak adlandırılırken bu maddelere hayır yanıtı vermek bir “ıska”dır (miss). Yeni bir maddeyi hatalı olarak tanıdığını belirtmek “yanlış alarm” (false alarm) iken onu tanımamak ise “doğru ret” (correct rejection) olarak adlandırılır. Bellek performansı değerlendirilirken bu dört sonucun gerçekleşme sayılarından yararlanır.

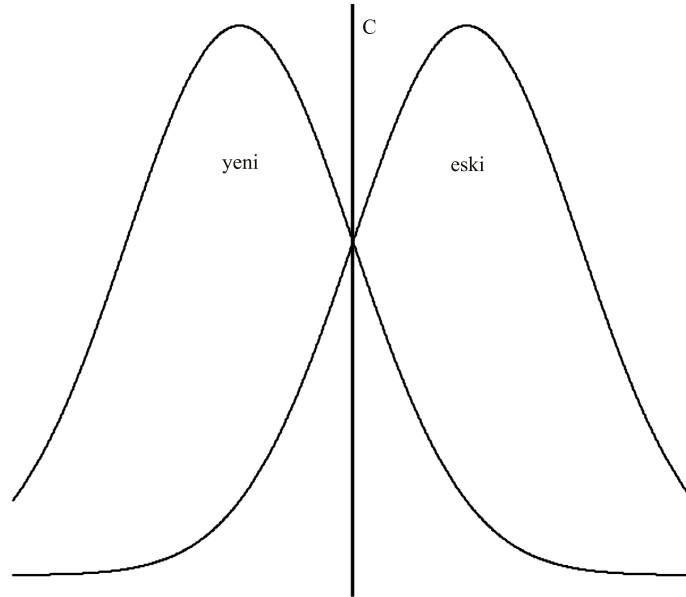
### Sinyal Belirleme Kuramı ve Kriter

Tanıma belleği performansını değerlendirmede kullanılan bir ölçüm modeli olan SBK'ye (Green ve Swets, 1966; Macmillan ve Creelman, 2005) göre bir tanıma görevi, yukarıdaki örneğimize benzer bir şekilde yerine getirilir. Test aşamasında sunulan her uyarının bir bellek gücü (memory strength), yani aşinalık (familiarity) değeri hesaplanır. Eski ve yeni maddelerin bellek güçlerinin ölçek üzerinde kısmen örtüşen iki dağılım oluşturduğu varsayılır (Şekil 1). Eski maddeler ortalama olarak daha güçlü olduklarından bu ölçekte daha yüksek değerler alırlar, ancak bazı eski maddeler yeni maddelerin bir kısmından daha düşük değerler aldığı için kusursuz bir ayrışma söz konusu değildir. Kişi tanıma kararını maddenin bellek gücü değerini kıyasladığı, öznel olarak belirlediği bir kriter değere göre verir: Maddenin gücü kriteri aşıyorsa o madde tanınır. En fazla sayıda doğru yanıt verilmek amaçlandığında optimum kriter değeri iki dağılımın kesiştiği noktadır. SBK'ye göre performansı belirleyen iki etken bulunur: duyarlılık (sensitivity) ve tepki yanlılığı (response bias).

Duyarlılık, eski ve yeni test maddelerinin birbirlerinden ne kadar iyi ayırt edilebildiğini ifade eder. SBK'nin temel duyarlılık ölçüsü olan  $d'$  (d-üssü; Macmillan ve Creelman, 2005), eski ve yeni dağılımlarının arasındaki mesafeyi ifade eder. Bu mesafe,  $z$  normal kü-

mülatif dağılımın tersini,  $\bar{I}O$  isabet oranını, YAO yanlış alarm oranını ifade etmek üzere  $d' = z(\bar{I}O) - z(YAO)$  eşitliğiyle hesaplanır (Macmillan ve Creelman, 2005, s. 8). Buna göre isabet ve yanlış alarm oranlarının eşit olduğu bir şans performansı  $d' = 0$  sonucunu verir, duyarlılık yükseldikçe  $d'$  değeri de artar. Özetle, bellek testindeki duyarlılık eski ve yeni madde dağılımlarının ne kadar örtüştüğüne bağlı olarak belirlenir.  $d'$  ölçüsü eski ve yeni dağılımlarının eşit varyanslı ve normal oldukları varsayımına dayanır, ancak bu varsayımlar her zaman test edilememekte ya da karşılanamamaktadır. Bu nedenle dağılım varsayımlarına dayanmayan (parametrik olmayan) ölçüler geliştirilmiştir (ileri tartışma için bakınız Macmillan ve Creelman, 1996).  $A'$  parametrik olmayan bir duyarlılık ölçüsüdür ve  $A' = 0.5 + ([\bar{I}O + YAO][1 + \bar{I}O - YAO]) / (4\bar{I}O[1 - YAO])$  formülüyle hesaplanır (Sno-dgrass ve Corwin, 1988, s. 38).

Tepki yanlılığının temel ölçüsü olan  $C$  sembolüyle gösterilen "kriter konumu" (Macmillan ve Creelman, 2005) kriterin eski ve yeni dağılımlarının kesişim noktasına olan mesafesini ifade eder ve  $C = -.5[z(\bar{I}O) + z(-YAO)]$  formülüyle hesaplanır.  $C = 0$  değeri nötr kriteri gösterirken küçülen negatif değerler kriterin daha "yumuşak" (lenient) olduğu, yani kişinin "evet" demeye daha yatkın olduğu anlamına gelir, pozitif değerler ise "katı" (strict) bir kriterle sahip olduğunu, yani kişinin "hayır" demeye yatkın olduğunu ifade eder. Sık kulla-



Şekil 1. Sinyal belirleme kuramına göre yeni ve eski maddelerin bellek gücü ölçeğindeki dağılımları ile kriter konumu (C).

nılan alternatif bir tepki yanlılığı ölçüsü olan olabirlik oranı (likelihood ratio)  $\beta$  (beta) ve onun doğal logaritmik dönüşümüyle elde edilen  $\ln(\beta)$  değeridir. Bellek gücü ölçüğünde herhangi bir noktanın olabirlik oranı, o noktadaki bir test maddesinin eski ve yeni olma olasılıklarının oranlanmasıyla hesaplanır. Bu olasılıklar ise söz konusu noktadaki eski ve yeni dağılımlarının yükseklikleridir.  $\beta$ 'nın hesaplanmasının birçok yolu olmakla birlikte (Macmillan ve Creelman, 2005, s. 33-35; Stanislaw ve Todorow, 1999, s. 142) en basit ifadesi  $e^{cd'}$  formülüdür. Böylece  $\ln(\beta)$  da  $cd'$  değerine eşit olmaktadır (aynı zamanda  $\ln(\beta) = z(\text{YAO})^2 - z(\text{IO})^2 / 2$ ). Olabirlik oranı aynı zamanda tanıma testine ilişkin dışarıdan gelen ipuçlarının (temel oran, ödeme matrisi [payoff matrix] vb.) bellek gücü kanıtıyla bütünleştirilmesiyle oluşan kriter değerinin hesaplanmasında da kullanılabilir. (Örnekler için bakınız Healy ve Kubovy, 1978, s. 546-547; Jaeger, Lauris, Selmezy ve Dobbins, 2012, s. 102; Macmillan ve Creelman, 2005, s. 43; Selmezy ve Dobbins, 2013, s. 679-682).  $B''$  ise tepki yanlılığı için geliştirilen parametrik olmayan ölçüdür ve  $B'' = (\text{IO}[1 - \text{IO}] - \text{YAO}[1 - \text{YAO}]) / (\text{IO}[1 - \text{IO}] + \text{YAO}[1 - \text{YAO}])$  eşitliğiyle hesaplanır (Snodgrass ve Corwin, 1988, s. 38).

Alanyazında yanlış alarm oranının da, özellikle eski dağılımlarının konularının değiştiği çalışmalarda, sıklıkla tepki yanlılığı ölçütü olarak kullanıldığını görürüz (Bruno, Higham ve Perfect, 2009; Franks ve Hicks, 2016; Hicks ve Starns, 2014; Morrell, Gaitan ve Wixted, 2002; Verde ve Rotello, 2007). Bunun temel nedeni,  $C$  ölçüsüne göre kıyaslanan iki performansta yanlış alarm oranları eşit ise, isabet oranının yüksek olanın kriterinin daha yüksek bir değere sahip olmasıdır. Oysa bir görüşe göre, böyle bir kıyaslamada yeni dağılımların sabit olduğu varsayımından hareket edilerek, gerçek kriter değişiminin yanlış alarm oranlarındaki farklılaşmadan çıkarsanabileceği kabul edilir (Franks ve Hicks, 2016; Hicks ve Starns, 2014). Buna karşıt bir görüş olarak ayrıştırma (differentiation; Criss, 2006, 2010; Kılıç, Criss, Malmberg ve Shiffrin, 2017; McClelland ve Chappell, 1998; Shiffrin ve Steyvers, 1997) modellerinin yaklaşımına göre ise güçlü maddelerden oluşan çalışma listelerinde yeni maddelerin dağılımları eski maddelere göre daha solda yer alır, yani yeni kelimeler iyi öğrenilmiş kelimelerin arasında daha kolay ayırt edilir; dolayısıyla yanlış alarm oranları güçlü listelerde düşer. Bu durumda yalnızca yanlış alarm oranındaki değişimden kriter değişimi olduğu sonucu çıkarılamaz. Görüldüğü gibi, tepki yanlılığı ölçüsünün seçimi deney verilerinin yorumlanmasında önemli bir etkidir ve hangi koşullarda hangi kriter ölçüsünün en uygun olacağı üzerinde araştırmacılar arasında bir uzlaşma henüz sağlanamamıştır.

## Kriter Değişimleme Çalışmaları

Tanım belleğinde kriterin deneysel olarak değiştirilebilir olup olmadığı uzun yıllardır süregelen araştırma konularından biri olmuştur (Estes ve Maddox, 1995; Han ve Dobbins, 2008, 2009; Healy ve Kubovy, 1977, 1978; Hicks ve Starns, 2014; Rhodes ve Jacoby, 2007; Verde ve Rotello, 2007). Bu bölümde, söz konusu araştırmalarda en yaygın kullanılan yöntemler olan temel oran değişimlemeleri (Estes ve Maddox, 1995; Healy ve Kubovy, 1977, 1978; Rhodes ve Jacoby, 2007) ve bellek gücü değişimlemeleri (Hirshman, 1995; Stretch ve Wixted, 1998; Verde ve Rotello, 2007) ayrı alt başlıklar halinde incelenecek, daha sonra alternatif yöntemlere değinilecektir.

### Temel Oran Değişimlemeleri

Tanım belleğinde kriter kaymasını amaçlayan deneysel değişimleme yöntemleri içerisinde en yaygın kullanılanlardan biri, test aşamasındaki eski ve yeni maddelerin oranlarının, yani temel oranın değişimlenmesidir (Aminoff ve ark., 2012; Cox ve Dobbins, 2011; Estes ve Maddox, 1995; Franks ve Hicks, 2016; Healy ve Kubovy, 1977, 1978; Heit, Brockdorff ve Lamberts, 2003; Rhodes ve Jacoby, 2007). SBK'ye göre test maddelerinin çoğunlukla eski olduğu bir testte optimum performansın gösterilebilmesi için kriterin düşürülmesi (yumuşatılması); aynı şekilde, test maddelerinin çoğunluğu yeni olduğunda kişinin kriterini katılaştırarak hayır demeye yatkın olması beklenmektedir (Macmillan ve Creelman, 2005; Swets ve ark., 1961). Swets ve arkadaşları (1961) sinyal belirleme görevleriyle yaptıkları deneylerde temel oranın kriter üzerindeki etkisini raporlamış, tanım belleğinde de benzer bir etkinin görülüp görülmeyeceği araştırma konusu olmuştur. Bu konudaki ilk çalışmaların birinde Healy ve Jones (1975), uyarıların 3 basamaklı sayılar olduğu bir tanım belleği testi kullanmışlardır. Temel oranlarının %50 (eski madde oranı) ve %25 olarak değiştiği 40'ar maddelik test blokları katılımcılara rastgele sırayla sunulmuş, aynı zamanda her test bloğunun temel oran bilgisi ilgili blokun başında verilmiştir. Katılımcıların yanıtları dördü güven ölçüğünde alınmış, ayrıca test boyunca her maddeye verilen yanıtın doğruluğu konusunda geribildirim verilmiştir. Deneyde kriter ölçüsü olarak  $\beta$  ve  $\text{IO}/\text{YAO}$  hesaplanmış, 4 güven puanını birbirinden ayıran 3 kriterin yalnızca birinde temel orana göre bir farklılaşma gözlenmiştir. Healy ve Kubovy (1977) bu deneylerden yola çıkarak temel oran değişimlemesinin hem bir evet-hayır tanım belleği görevinde hem de bir sayının iki dağılımın hangisinden geldiğine karar verilen bir sayısal karar verme görevinde etkilerinin gözlemlendiği bir çalışma yürütmüşlerdir. İki görevde de temel oran %50 ve %25 olarak değiştirilmiş,

yine test aşamalarında temel oran ön bilgisi ve madde geribildirimini verilmiştir.  $\beta$  analizlerine göre sayısal karar verme testinde temel oranın kriter üzerinde etkisi gözlenirken tanıma belleği testinde bu etki bulunmamıştır. Bu deneyde sayısal karar verme görevinde duyarlılığın bellek görevine göre daha düşük olduğunu raporlayan Healy ve Kubovy duyarlılıktaki azalmanın daha çok kriter kaymasına yol açmış olabileceğini önermişlerdir (Benzer sonuçlar için bakınız Aminoff ve ark., 2012; Bruno ve ark., 2009; Estes ve Maddox, 1995; Hirshman, 1995; Selmeczy ve Dobbins, 2013).

Temel oran değişimlesinin kriter üzerinde etkisinin gözlenmediği bu çalışmaların yanı sıra bu etkinin raporlandığı araştırmalar da olmuştur. Healy ve Kubovy'nin (1978) test listeleri arasında temel oran değişimlesi yaptıkları deneyde katılımcılara testten önce temel oran ön bilgisi sunulmuş ve kriterin test listeleri arasında temele göre farklılaştığı bulunmuştur. Bu sonuç katılımcılara madde geribildirimini verilen ve verilmeyen koşullarda tekrarlanmıştır. Bir diğer çalışmada Estes ve Maddox (1995) çalışma-test ve sürekli tanıma görevlerinin karışımı bir görevle yaptıkları deneyde temel oran (%33 – %67) ile birlikte uyarın tipini (3 basamaklı sayılar, 3 harften oluşan uyarınlar, kelimeler) değişimlesilerdir. Çalışmanın diğer değişkeni ise geribildirim koşuludur. *C* analizlerine göre temele göre kriter kayması yalnızca geribildirim verilen sayı ve harf uyarınlarının kullanıldığı koşullarda gerçekleşmiş, diğer koşullarda gerçekleşmemiştir. Bu sonuçlar yorumlanırken dikkat edilmesi gereken detaylar, kelime tanıma performansının diğer koşullara göre daha iyi olması ve geribildirim ile birlikte performansa göre artan bir katılım ücreti verilmesinin etkilerinin bir arada gözlenmesidir. Daha sonra Aminoff ve arkadaşları (2012) ile Franks ve Hicks'in (2016, Deney 2) yaptığı araştırmalarda temel oran değişimlesi aynı test içinde 6 ile 9 madde arasında değişen bloklar arasında yapılmış, katılımcıların kriterlerini bloklar arasında temele göre değiştirebildikleri bulunmuştur.

Temel oran değişimlesinin testler arası ya da test içi bloklar arası yapıldığı araştırmalardan farklı olarak, Rhodes ve Jacoby (2007) deneylerinde aynı test blokunda maddeden maddeye kriter kayması gözlemişlerdir. Bu çalışmalarında Rhodes ve Jacoby, birinci deneylerinde ekranın iki ayrı konumunda gösterilen test maddelerinin temel oranlarını değişimlemiş, test aşamasında katılımcılara madde geribildirimini yanı sıra doğru yanıtların sayımını yapan bir sayaç sunulmuştur. Deneyde katılımcıların kriterlerini test maddesinin ekrandaki konumuna göre maddeden maddeye değiştirebildiği gözlenmiştir. İkinci deneyde ekranın farklı konumlarındaki maddeler için farklı tuşlarla yanıt vermeleri beklenen katılımcılarda kriter kayması gözlenmiş, ancak aynı tuşlarla yanıt

veren diğer gruptaki katılımcılarda kriter kayması gerçekleşmemiştir. Rhodes ve Jacoby üçüncü deneylerinde geribildirimini değişimlemiş, kriter kaymasının ancak geribildirim verildiğinde gerçekleştiği gözlenen Estes ve Maddox'un (1995) çalışmasının sonuçları burada da tekrarlanmıştır.

Yaratıcı bir yöntemle Heit ve arkadaşları (2003) temel oranı tepki sinyali tekniği (response signal technique; Reed, 1973) aracılığıyla değişimlemiştir. Bu teknikle katılımcılardan bir test maddesine yanıtlarını, o maddenin sunulmaya başlamasının ardından değişen sürelerde verilen bir sinyalden (örneğin, bir ses) hemen sonra verilmesi beklenmektedir. Heit ve arkadaşlarının deneylerinde sunum-sinyal arası gecikme süreleriyle (100ms. – 1500 ms. arasında değişen süreler) temel oranları (%10, %50 ve %90) değişimlenmiştir. Deneyde test boyunca maddeden maddeye sunum-sinyal gecikme sürelerinin temel oranlarına göre kriter kaymaları gözlenmiştir.

Temel oran değişimlesinin özel bir biçimi ise bu değişimlemenin uç noktalar arasında yapıldığı, yani test maddelerinin hepsinin eski (yalnızca-eski) ya da hepsinin yeni (yalnızca-yeni) maddelerden oluştuğu tanıma belleği görevleridir. Yalnızca-eski listelerde yanlış alarm, yalnızca-yeni listelerde ise isabet olmadığından dolayı bu deneylerde *C* hesaplanamaz. Bunun yerine yalnızca-eski listeler ile standart (%50 temel oranlı) listeler isabet oranlarına bakılarak kıyaslanırken yalnızca-yeni listeler de yine standart listeler ile yanlış alarm oranları ile kıyaslanır. Wallace, Sawyer ve Robertson (1978) yalnızca-eski ve standart test listelerinin kıyaslandığı deneylerinde katılımcıların isabet oranlarının test listesinde yeni madde bulunmasından etkilenmediğini gözlemişlerdir. Bu durum, katılımcılara test listesinin tamamının çalıştıkları kelimelerden olduğu bilgisinin sunulup sunulmamasından etkilenmemiştir. Yeni madde bulunmayan testlerde katılımcıların isabet oranlarını, dolayısıyla kriterlerini değiştirmemesi bulgusu daha sonraki araştırmalarda da tekrarlanmıştır (Cox ve Dobbins, 2011; Ley ve Long, 1987, 1988; Wallace, 1982). Cox ve Dobbins (2011) yalnızca-yeni test listesinin uygulandığı bir koşulu da standart ve yalnızca-eski listelere eklemiş, yalnızca-yeni ve yalnızca-eski test kriterlerinin standart teste göre bir farklılık göstermediğini ya da beklenenden ters yönde farklılaştığını gözlemiştir. Temel oranlar hakkında ön bilgi içeren deneylerde (Wallace ve ark., 1978, Deney 1; Cox ve Dobbins, 2011, Deney 1) katılımcıların bütün maddelere “evet” yanıtı vermeleri durumunda hatasız sonuç alabilecek olmalarına rağmen kriterlerinin sabit kalması dikkat çekicidir. Bu deneylerde katılımcıların yönergelerden yola çıkarak bütün test maddelerine “evet” yanıtı vermek yerine kriterlerini korumayı tercih etmiş olmaları muhtemeldir. Ön bilgi yalnızca eski maddelerin olduğu test listelerinde testi an-

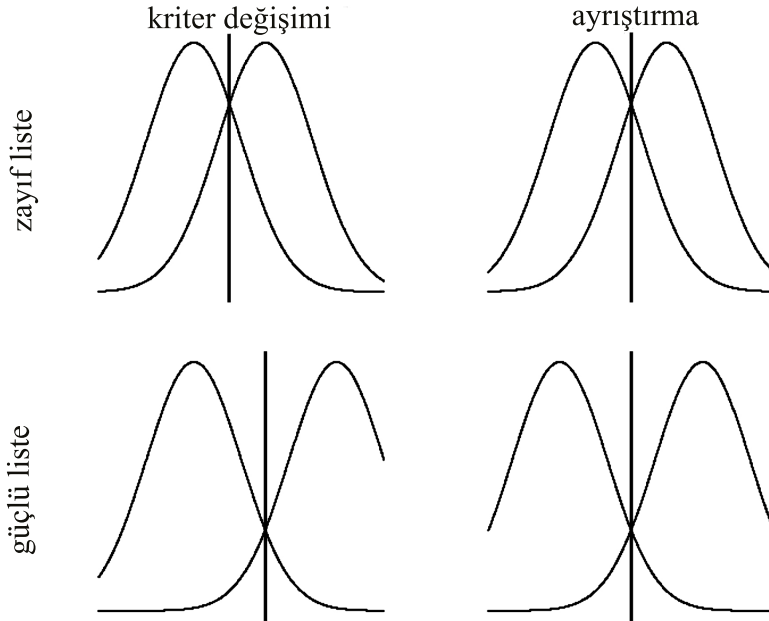
lamsız kılabilecek olması nedeniyle görmezden geliniyor, ancak eski ve yeni maddelerin bir arada bulunduğu temel oranlı listelerde bu bilgiden faydalanılıyor olabilir. Nitekim, katılımcıların performanslarına odaklanmalarını sağlayabilecek madde geribildirimlerinin sunulduğu yalnızca-eski ve yalnızca-yeni test listelerinde performans beklenen şekilde değişmektedir: Koop, Criss ve Malmberg (2015) standart, yalnızca-eski ve yalnızca-yeni listelerle yaptıkları tanıma belleği deneylerinde madde geribildirimini değiştirmiş, geribildirim olmayan durumlarda yalnızca-eski ve yalnızca-yeni listelerin isabet ve yanlış alarm oranları standart listeden farklılaşmamış, ancak geribildirim verilen koşulda beklenen kriter kayması gerçekleşmiştir.

Görülebileceği gibi, tanıma belleği görevlerinde test listelerinin temel oranlarının değişimlenmesiyle kriter kayması zaman zaman elde edilememiş (Cox ve Dobbins, 2011; Healy ve Kubovy, 1977) ancak uygun ve yeterli ipuçlarının yardımıyla temel orana bağlı kriter kaymasının gözlemlendiği araştırmalar da gerçekleştirilmiştir (Estes ve Maddox, 1995; Rhodes ve Jacoby, 2007). Buna göre, ön bilgi, geribildirim gibi yollarla desteklendiğinde temel oran değişimlenmesi, tanıma belleğinde karar kriteri konulu çalışmalarda kullanılan güvenilir bir

yöntemdir. Ancak bu çalışmalarda kullanılan yöntemler temel oran değişimlenmesinden ibaret değildir. Bir sonraki başlıkta kriter değişimlenmesinde kullanılan başlıca yöntemlerden bir diğeri, bellek gücü değişimlenmeleri incelenecektir.

### Bellek Gücü Değişimlenmeleri

Tepki yanlılığına ilişkin yapılan araştırmalarda yaygın olarak kullanılan bir diğer yöntemde eski ve yeni dağılımlarının konumları, yani bellek güçleri değişimlenerek kriter kaymasının sağlanması amaçlanır. Bellek gücü değişimlenmesinin yapıldığı "liste gücü" (list strength) paradigmasında (Ratcliff, Clark ve Shiffrin, 1990) çalışma listesindeki kelimelerin bellek güçleri çeşitli yöntemlerle zayıf ya da güçlü olacak şekilde değişimlenir. Maddelerin bellek güçlerinin değişimlenmesi, genellikle maddelerin sunulma süresinin (Bruno ve ark., 2009; Hirshman, 1995; Ratcliff ve ark., 1990, Deney 1-4; Ratcliff, Sheu ve Gronlund, 1992, Deney 1, Deney 3) ya da sunulma sayısının (Hicks ve Starns, 2014; Ratcliff ve ark., 1990, Deney 5-6; Ratcliff ve ark., 1992, Deney 2; Stretch ve Wixted, 1998) farklılaştırılmasıyla gerçekleştirilir. Örneğin, zayıf maddeler çalışma aşamasında birer kere sunulurken güçlü maddeler dörder kere



**Şekil 2.** Bellek gücü değişimlenmesinde kriter değişimi ve ayrıştırma açıklamalarına göre eski ve yeni maddelerin dağılımları ile kriter konumları. Sol sütun kriter değişimi, sağ sütun ayrıştırma mekanizmalarını tasvir etmektedir.

sunulabilir, ya da zayıf maddeler ekranda birer saniye süreyle görünürken güçlü maddeler ikişer saniye süreyle görünebilir. Güçlü ve zayıf maddeler ayrı listelerde sunulup test edildiğinde güçlü maddelerin zayıf maddelere göre isabet oranları daha yüksek, yanlış alarm oranları ise daha düşük olmaktadır. Bu bulguya “güç temelli ayna etkisi” (strength based mirror effect; Glanzer ve Adams, 1985; Ratcliff ve ark., 1990) denir. Güçlü ve zayıf test listelerindeki yeni maddelerin dağılımlarının eşit olduğu varsayıldığında bu etki, güçlü listeler için kriterin zayıf listelere göre daha yüksek olmasıyla açıklanmaktadır (Şekil 2, sol sütun; Hicks ve Starns, 2014; Starns, White ve Ratcliff, 2010; Stretch ve Wixted, 1998). Ancak güç temelli ayna etkisi ayırıştırma modelleri (Criss, 2006; Kılıç ve ark., 2017; McClelland ve Chappell, 1998; Shiffrin ve Steyvers, 1997) tarafından farklı bir mekanizmayla açıklanmaktadır. Buna göre güçlü maddeler bellek izleriyle zayıf maddelere göre daha iyi eşleşerek daha yüksek isabet oranları ortaya çıkarırken yeni maddeler güçlü maddelerle zayıf bir eşleşme oluşturarak daha düşük yanlış alarm oranı sonucunu verir, buna karşın zayıf maddeler bellekte daha kötü kodlanmış olduğundan yeni test maddeleriyle hatalı eşleşmeye daha müsaittirler (Şekil 2, sağ sütun).

Güç değişimlesinin yapıldığı çalışmalarda yaygın olarak kullanılan bir diğer yöntem ise, zayıf ve güçlü maddelerin bir arada sunulduğu (karışık) çalışma listelerindeki zayıf maddelerin yalnızca zayıf maddelerin sunulduğu listelerdeki (yalnızca-zayıf) maddelerle, karışık listelerdeki güçlü maddelerin ise yalnızca-güçlü maddelerle kıyaslandığı deneylerdir. Bu çalışmalarda duyarlılığın çalışma listesindeki diğer maddelerden etkilendiği gözlenmiştir (null list strength effect, Ratcliff ve ark., 1990; Ratcliff ve ark., 1992, Deneç 2; Starns ve ark., 2010). Hirshman (1995) liste gücü deneylerinde karışık listedeki zayıf maddeler için kriterin yalnızca zayıf listeye göre daha yüksek olacağı hipotezini öne sürmüş, deneyinde katılımcılara yalnızca zayıf ve karışık çalışma listelerini değişen sırayla sunmuştur. C analizinde elde ettiği bulgulara göre karışık listedeki zayıf maddeler için kriterin yalnızca zayıf liste kriterinden daha yüksek olduğunu, buna ek olarak, karışık listenin yanlış alarm oranının yalnızca zayıf listeden daha düşük olduğunu tespit etmiştir. Bu sonuçlarla paralel olarak, yalnızca-güçlü liste ile karışık listelerdeki güçlü maddeleri kıyasladığında da karışık listenin daha düşük bir kriteri olduğunu gözlemiştir. Hirshman ayrıca liste gücü etkisinin çalışıldığı geçmiş araştırmaların verilerini inceleyerek çoğunlukla beklenen yönde kriter değişiminin gözlendiğini raporlamıştır (1995, s. 306). Bir diğer deneyinde ise (Hirshman, 1995, Deneç 4) bu defa yalnızca-zayıf listeye kıyasladığı karışık listenin test aşamasında yalnızca zayıf maddeler test ederek iki test listesinin ortalama güçlerini eşit

tutmuş, yaptığı analizlerde aynı sonuçların tekrarlandığını raporlamıştır. Bu da, kriter konumunda test listelerinin değil çalışma listelerinin eski ve yeni dağılımlarının belirleyici olduğunu göstermiştir. Bu bulguya karşıt bir sonuç ise Hicks ve Starns’ın (2014) çalışmasında gözlenmiştir. Söz konusu çalışmada araştırmacılar güçlü ve zayıf maddelerden oluşan karışık çalışma listeleri kullanmış, test aşamasında ise güçlü maddelerin yeni maddelerle birlikte karışık halde sunulduğu güçlü-bloklar ile bu kez zayıf maddelerin yeni maddelerle birlikte bulunduğu zayıf-blokları arka arkaya sunmuşlardır. Güçlü ve zayıf test blokları farklı renklerde sunulduğunda test blokları arasında kriter kayması gözlenmiş, ancak renk ipucu olmadığında kriter kayması gerçekleşmemiştir. Bu sonuçlar, çalışma listeleri sabit kaldığı halde test bloklarının beklenen gücüne bağlı olarak aynı test içinde de bloklar arasında kriter kayması gerçekleşebildiğine yönelik bir kanıt olarak değerlendirilmiştir. İki çalışma birlikte değerlendirildiğinde, çalışma listeleri arasındaki farklılıkların kriter kaymasına yol açabileceği, bununla birlikte test listelerine bağlı olarak kriter kaymasının da mümkün olduğu, ancak bunun uygun ipuçları ile teşvik edilerek sağlanabildiği sonucuna varılabilir.

Bellek gücüne bağlı kriter kaymasının aynı test listesi içinde sağlanıp sağlanmadığını sınamak için yapılan bir diğer çalışmada Morrell ve arkadaşları (2002) güçlü ve zayıf maddelerin aynı listeler içinde çalışıldığı ve test edildiği bir dizi deney hazırlamışlardır. Deneylerinde çalışma aşamasında katılımcılara farklı anlamsal (semantik) kategorilerden (ör: meslekler ve hayvanlar) kelimeler gösterilmiş, bu kategorilerden birine ait kelimeler birer kere, diğer kategorideki kelimeler ise beşer kere sunulmuştur. Ayrıca katılımcılar bu değişimle konusunda yönergeyle açıkça bilgilendirilmiştir. Bu deneylerde güçlü maddelerin isabet oranı zayıf maddelerden fazlayken güçlü kategori ile zayıf kategoriden yeni kelimelerin yanlış alarm oranları birbirine eşit çıkmıştır. Anlamsal kategori ile birlikte uyaran tipinin de değişimlenerek ipucunun belirginleştirilmesi de bu sonucu değiştirmemiştir. Araştırmacılar bu sonucu aynı test içinde tek kriter belirlendiği şeklinde yorumlamışlardır. Bu deney sonuçları ve Hirshman’ın (1995) bulgularına paralel olarak Stretch ve Wixted (1998) de deneylerinde liste gücü paradigmasıyla listeler arası kriter farklılaşması gözlemekle birlikte aynı liste içinde farklı ipuçlarıyla birbirlerinden ayırıştırılan güçlü ve zayıf kelimeler için aynı kriterin kullanıldığını gözlemişlerdir.

Verde ve Rotello (2007), aynı test listesinde güçlü ve zayıf maddeler bloklar halinde sunulurken bloklar arasında kriterde bir değişim olup olmadığını test etmiştir. Deneylerinde karışık bir çalışma listesini takip eden test aşamasının ilk yarısında güçlü kelimeler, ikinci yarısında ise zayıf kelimeler 40’ar yeni kelimeyle karıştırıla-

rak sorulmuş, sonuçlara göre ilk bloktaki güçlü maddelerin isabet oranları sonraki bloktaki zayıf maddelerden daha yüksek olmasına rağmen iki bloktaki yeni maddelerin yanlış alarm oranları farklılık göstermemiştir. Bu sonuç, bellek gücü değişimlemesi sunulma süresi (500 ms. – 3 s.) ve sunulma sayısı (1 – 4) ile uygulandığında, güçlü ve zayıf kelimeler farklı ipuçları ile belirgin hale getirildiğinde ve zayıf ve güçlü maddelerin test bloku sıraları değiştirildiğinde de tekrarlanmıştır. Araştırmacılar, son olarak, test aşamasında madde geribildirim sunulduğu, ayrıca testin her çeyrek diliminden sonra doğru ve yanlış yanıtların toplam sayıları gösterildiğinde, yanlış alarm oranının güçlü maddelerin sunulduğu 1. bloktan zayıf maddelerin sunulduğu 2. bloka arttığını gözlemiştir (Deney 5). Bu deneylerde güçlü ve zayıf maddeler arasında kriter farklılaşması kategori ipuçları ile sağlanamazken bu farkın geribildirim ile sağlanabilmesi dikkat çekicidir.

Bunlara ek olarak, Verde ve Rotello (2007), test listesinin güçlü ve zayıf bloklarla başladığı durumları kıyaslamış ve ilk önce güçlü maddeler test edildiğinde katılımcıların katı bir kriter belirlediğini tespit etmişlerdir. Bu deneylere göre kriter ilk test maddelerinin bellek güçlerine göre belirlenirken daha sonra test boyunca sabit kalmıştır. Ancak, Starns ve arkadaşlarının (2010) deney sonuçları bu bulgularla çelişir görünmektedir: Söz konusu deneylerde katılımcılar, güçlü ve zayıf maddelerin bir arada sunulduğu çalışma aşamalarından sonra yalnızca zayıf ya da yalnızca güçlü test listelerini yanıtlamışlardır. Verde ve Rotello'nun çalışmasının aksine, yanıtladıkları ilk test listesinin zayıf ya da güçlü olması katılımcılar arasında kriter farkına yol açmamıştır. Bu deneyde ilk test listesinin zayıf ya da güçlü olması kıyaslanırken, Verde ve Rotello'nun deneyinde aynı testteki ilk blokların kıyaslandığını vurgulamak gerekir. Bununla birlikte, Hicks ve Starns (2014) bu iki deney arasındaki belirleyici farkın Verde ve Rotello'nun deneyinde güçlü ve zayıf test blokları 80'er maddeden oluşurken Starns ve arkadaşlarının (2010) deneyinde güçlü ve zayıf test listelerinin 24'er maddeden oluşması olduğunu iddia etmiş, kısa bloklar arasında kriter kaymasının daha zor olabileceği hipotezini ortaya atmışlardır. Bu hipotezi test etmek için, karışık çalışma listelerinin ardından güçlü ve zayıf maddelerin aynı test içinde bloklar halinde sunulduğu deneyler yürütmüşlerdir. Blokları renk ipucuyla birlikte sunduklarında bloklar arasında kriter farkı görülmüş, ancak renk ipucu olmadığında kriter sabit kalmıştır (Deney 1-2). Ayrıca bu deneylerde geribildirim verilip verilmemesinin kriter üzerinde bir etkisinin bulunmadığı raporlanmıştır. Deneylerde ayrıca blok uzunlukları değişimlenmiş, blok uzunluğu arttıkça bloklar arası kriter farkının da arttığı gözlenmiştir (Hicks ve Starns, 2014, Deney 3-4; benzer sonuçlar için bakınız Franks ve Hi-

cks, 2016). Son olarak, blok uzunluğu 40 maddeyken ilk test bloku zayıf ve güçlü olan katılımcılar arasında kriter farkı gözlenmiş, ancak bu fark 20 ve 10 maddeli bloklarda bulunmamıştır. Bu sonuç Verde ve Rotello'nun 80 maddeli bloklarla gözlediği, ancak Starns ve arkadaşlarının (2010) 24 maddeli bloklarla gözleyemediği ilk blok etkisiyle uyumlu ve onu açıklar niteliktedir.

Kriter değişimlemesinin bir diğer yöntemi, eski dağılımlarının sabit tutulması, ancak yeni dağılımlarının eski maddelere olan benzerliklerinin değişimlenmesidir. Benzer yeni maddeler bellek gücü dağılımında daha yüksek değerler alacağı için kriterin bu koşulda daha yüksek olması beklenmektedir. Benjamin ve Bawa (2004) belirli anlamsal kategorilerden kelimelerin ya da resimlerin gösterildiği çalışma listesinin ardından bu maddeler ile aynı kategoriden (benzer) ya da farklı kategorilerden yeni kelimelerin sunulduğu test listeleri hazırlamışlardır. Yaptıkları deneylerde, uyaran tipinden bağımsız olarak, benzer yeni kelimelerin kullanıldığı listelerde farklı yeni kelimelere göre daha yüksek bir kriter kullanıldığını belirtmişlerdir. Pendergrass, Olfman, Schmalstig, Seder ve Light (2011) kelime çiftlerinin çalışıldığı ilişkisel öğrenme görevlerinde çalışma aşamasında sunulmayan kelimelerden oluşan yeni kelime çiftlerinin (kolay test), ya da çalışmada farklı kelimelerle birlikte sunulan, yeniden düzenlenmiş kelime çiftlerinin (zor test) kullanıldığı testleri kıyaslamışlar, yeni çiftlerin eskilerine daha çok benzediği zor testte katılımcıların daha yüksek bir kriter belirlediklerini raporlamışlardır.

Alanyazında bellek gücü değişimlemelerinin etkileri ve deney sonuçlarının yorumlanmasında farklılıklar olduğu görülmektedir. Bu yöntemin kullanıldığı çalışmalarda yalnızca-güçlü ve yalnızca-zayıf çalışma listelerinin test aşamalarında kriterlerin farklılaşması sıklıkla tekrarlanan bir bulgudur, ancak karışık listelerde böyle bir kriter farklılaşmasının gerçekleşebilmesi için güçlü ve zayıf maddelerin ayrı bloklar halinde sunulması ve ipuçlarıyla bu farkın belirginleştirilmesi gibi yardımcı yollara başvurulması gerekmektedir. Bununla birlikte bellek gücü değişimlemelerinin isabet ve yanlış alarm oranlarında yol açtığı farklılaşmaların kriter kaymalarından dolayı değil, eski ve yeni dağılımlarının konumlarındaki kaymalardan dolayı olduğuna yönelik yorumlar da mevcuttur.

### **Diğer Yöntemler**

Kriter değişimlemesinde kullanılan yöntemler temel oran ve bellek gücü değişimlemelerinden ibaret değildir. Bu yöntemlerin yanı sıra kişilerin tepki yanlılıklarını çeşitli açık ya da örtük yönlendirmelerle değişimlemeyi amaçlayan başka yöntemler de geliştirilmiştir. Bunlardan biri olan ödeme matrisi (Curran, DeBuse ve Leynes, 2007; Healy ve Kubovy, 1978; Van Zandt, 2000)



yönteminde test edilen kişiler test performanslarına göre bir ödül (genellikle para) kazanırlar ve test maddelerine verdikleri “evet” ya da “hayır” yanıtlarının kazanç-kayıp dengesi farklılaştırılır. Böylece katılımcılara “evet” ya da “hayır” yanıtının daha cazip hale getirilmesiyle kriter değişimi sağlanır. Örneğin, Healy ve Kubovy'nin (1978) deneyinde ödeme matrisi koşulunda katılımcılar isabetleri için 1 puan kazanırken yanlış alarmları için 3 puan kaybedecek, bununla birlikte doğru retleri için 3 puan kazanırken ıskaları için 1 puan kaybedecektir. Bu koşullarda “hayır” yanıtının olası kazancı olası kaybindan daha çok iken evet yanıtı için bu durum tersidir. Healy ve Kubovy deneylerinde bu ödeme matrisi (katı matris) ile doğru ve yanlış yanıtların eşit düzeyde puan kazanıp kaybettiği dengeli bir ödeme matrisini (nötr matris) birbirini takip eden testlerde değişen sırayla sunmuş,  $\beta$ ,  $\log \beta$  ve doğru yanıt oranı ile yapılan kriter analizlerine göre katı matrisin nötr matrise göre daha yüksek bir kritere yol açtığını raporlamışlardır. Ayrıca iki ödeme matrisi arasındaki kriter farkının, aynı deneydeki standart temel oran koşullarındaki kriter farkına eşit düzeyde olduğu görülmüştür. Curran ve arkadaşları (2007) araştırmalarında benzer bir ödeme matrisi değişimlemesi yapmışlardır. Buna göre “katı” blokta doğru retler 5, isabetler 1 puan kazandırırken “yumuşak” blokta bu puanlar tam tersi şekildedir. *C* analizine göre 3 ayrı deneyde katı koşulda kriter yumuşak koşula göre daha yüksek bulunmuştur (Benzer sonuçlar için bakınız Snodgrass ve Corwin, 1988).

Selmez ve Dobbins (2013) ise kriter değişimini sağlamak için çalışmalarında, her bir test maddesinin “muhtemelen eski” ya da “muhtemelen yeni” olduğunu belirten bir ipucu sundukları “olasılıksal mnemonic ipucu” (probabilistic mnemonic cue) yöntemini tercih etmiştir. Birinci deneylerinde katılımcılara %75 olasılıkla doğru olduğu belirtilen bu ipucunun yanı sıra madde geribildirim ve bellek gücü de değişimlenmiştir. Katılımcılar olasılıksal ipuçları yardımıyla performanslarını geliştirmişler, aynı zamanda “muhtemelen yeni” ipuçlu maddeler için “muhtemelen eski” maddelere göre daha düşük kriter kullanmışlardır. Dahası, duyarlılığın daha düşük olduğu koşulda kriterler arası fark güçlü koşula göre daha fazla olmuştur; yani katılımcılar duyarlılıkları düşük olduğunda olasılıksal ipuçlarından daha çok yararlanmışlardır. Bu deneyde ayrıca, madde geribildiriminin duyarlılık ya da kriter üzerinde bir etkisi olmadığı raporlanmıştır. İpuçlarının doğruluk olasılıklarının (%65 – %85) değişimlendiği ikinci deneyde ise daha güvenilir olan %85 koşulunda olasılıksal ipucunun duyarlılığı ve kriter farkının büyüklüğünü daha çok artırdığı görülmüştür.

Han ve Dobbins (2008) alanyazını gözden geçirerek kriter değişimlemesinin katılımcılara yönerge, açık

ipuçları gibi araçlarla görünür hale getirildiği çalışmalarda etkili olduğunu ifade etmiş, bu yöntemlerden farklı olarak kriter kaymasını katılımcıların farkındalığı olmadan sağlamayı hedefledikleri bir yöntem geliştirmişlerdir. Bu yöntemde katılımcılara test sırasında yanlı (biased) geribildirim sunulan iki koşul vardır. “Yumuşak” koşulun standart madde geribildiriminden tek farkı yanlış alarmlara “doğru” yanıtı verilmesi, böylece kişilerin “evet” deme eğiliminin artırılmasının hedeflenmesidir. “Katı” koşul ise standart geribildirimden ıskalara “doğru” yanıtı verilmesiyle ayrılır, böylece kişiler “hayır” yanıtına yönlendirilir. Han ve Dobbins, katılımcıları sırayla katı (K), normal (N) ve yumuşak (Y) geribildirim verilen 3 test aşamasından oluşan KNY koşulu ile ters geribildirim sırası uygulanan YNK koşulundan birine atamışlardır. Bu deneyde 1. ve 2. test bloklarında KNY grubunun kriteri YNK grubuna göre daha katı olmuş, ancak bu fark 3. testte kaybolmuştur. Sonraki deneylerinde test listelerinin bu defa KYY ya da YKK şeklinde sıralandığı iki koşul oluşturulmuş, 1. testlerde KYY grubunun kriteri YKK grubundan anlamlı düzeyde yüksek değilken 2. ve 3. testlerde YKK grubu KYY grubuna kıyasla daha yüksek kritere sahip olmuştur (Han ve Dobbins, 2008). Han ve Dobbins sonraki araştırmalarının (2009) ilk deneyinde bu defa katı ve yumuşak koşullarında sırasıyla ıskala ve yanlış alarmların %70'ine “doğru” geribildirimini vererek, ya da bu yanıtlara geribildirim vermeyerek değişimlemeyi daha da örtük hale getirmeyi hedeflemişlerdir. Katılımcılar, iki yumuşak geribildirimli testin ardından iki geribildirimsiz teste tabi tutuldukları “yumuşak” ve bu koşuldaki ilk iki testte katı geribildirim almalarıyla farklılaşan “katı” koşullarından birine atanmışlardır. Bu deneylerde katı koşulunun daha yüksek bir kritere sahip olduğunu ve yanlı geribildirim aşamasında belirlenen kriterin sonraki testlerde de korunduğunu gözlemişlerdir. Bu sonuç katılımcıların katı ya da yumuşak test aşamalarında belirledikleri kriterleri koruduklarına işaret etmektedir. Çalışmada bütün katılımcılar test sonrası anketinde kendilerine yanlı geribildirim verildiğinin farkında olmadıklarını belirtmişlerdir. Han ve Dobbins (2009) böylece katılımcıların farkındalığını gerektirmeden kriter kaymasını sağlayan bir yöntemi alanyazına kazandırmışlardır.

Son olarak, alanyazında çeşitli uyaran özelliklerinin de isabet ya da yanlış alarmlarda değişime yol açabildiğini gösteren birçok çalışma vardır (kelime frekansı, Glanzer ve Adams, 1985; Glanzer, Adams, Iverson ve Kim, 1993; kelime ayırt ediciliği [distinctiveness], Dobbins ve Kroll, 2005; Singer, Fazaluddin ve Andrew, 2011; kelimenin duygusal değeri [emotionality], Kapucu, Rotello, Ready ve Seidl, 2008; uyaran tipi, Scimeca, McDonough ve Gallo, 2011; tanıma görevi tipi, Hockley ve Niewiadomski, 2007); ancak söz konusu uyaran

özelliklerinin kriter kayması ya da başka bir mekanizma aracılığıyla mı bu farklılaşmalara yol açtığı belirsizdir. Örneğin, normatif frekansı (kullanım sıklığı) düşük olan kelimelerin yüksek olanlara kıyasla daha yüksek isabet oranı ve daha düşük yanlış alarm oranına sahip olduğu raporlanmıştır (Glanzer ve Adams, 1985; Glanzer ve ark., 1993); bu bulgu kelime frekansı ayna etkisi (word frequency mirror effect) olarak bilinmektedir. Tıpkı güç temelli ayna etkisi gibi kelime frekansına bağlı gelişen ayna etkisinin altında da bellek gücü dağılımları ya da karar süreçlerine bağlı bir kriter kayması mekanizmasının, ya da ayırıştırma mekanizmasının yattığı görüşleri tartışılmaya devam edilmektedir (Criss, 2010; Glanzer ve ark., 1993; Hirshman, 1995; Hirshman ve Arndt, 1997; Kılıç ve ark., 2017).

### **Kriter Değişimindeki Etkenler**

Alanyazında tanıma belleğinde kriterin incelendiği birçok araştırmada yapılan değişimlemelerin etkili ya da etkisiz olduğu deneyler raporlanmıştır. Bazıları diğerleriyle çelişir görünen bu çalışmalarda temel yöntemlere (temel oran, bellek gücü vb.) ek olarak çeşitli yardımcı yöntemler de uygulanmış ve sonuçlar üzerinde etkili oldukları tespit edilmiştir.

Morrell ve arkadaşları (2002), Stretch ve Wixted (1998) ve Verde ve Rotello'nun (2007) çalışmalarında zayıf ve güçlü maddeler arasındaki farkın renk ya da semantik kategori gibi ipuçları aracılığıyla görünür kılınması amaçlanmıştır. Bu deneylerde ipuçları ve bunlarla birlikte değişimlemeyi yönergede açıkça belirtmek kriterde beklenen kaymaya yol açmamış, ancak madde geribildirimini sunulmasıyla kriter kayması sağlanmıştır. Ancak, Hicks ve Starns'ın (2014) araştırmasında kriter kaymasının gerçekleşmesi geribildirim değil renk ipucuna bağlı olmuştur. Benzer bir sonuç Starns ve arkadaşlarının (2010) 3. deneyinde görülmüştür: Bu deneyde karışık ve yalnızca-zayıf test listelerindeki zayıf maddelerin karar kriterleri aynı bulunmuş, bu sonuç madde geribildirimini verilen ve verilmeyen koşullarda değişmemiştir (Benzer sonuçlar için bakınız Healy ve Kubovy, 1978; Selmeczy ve Dobbins, 2013). Bu çalışmaların ortak yönü, geribildirim deneysel değişimlemenin farkındalığına yönelik diğer ipuçlarına ek bir katkısı olmadığı yöntemlerin kullanılmış olması olabilir. Nitekim, diğer ipuçlarıyla kriter kaymasının sağlanmadığı bazı çalışmalarda geribildirim sunulmasıyla beklenen fark elde edilebilmiştir (Estes ve Maddox, 1995; Rhodes ve Jacoby, 2007; Verde ve Rotello, 2007). Geribildirim etkisine yönelik çalışmaların bulgularındaki bu farklılıkların, bahsedilenler haricinde, geribildirim birlikte kullanıldığı diğer ipuçlarıyla etkileşimi, kriter değişimi kıyaslamasının test listeleri ya da bloklar arası yapılması, liste

ya da blok uzunlukları gibi birçok etkenle ilişkisi için ileri incelemelere ihtiyaç duyulmaktadır. Örneğin, renk ve kategori ipuçlarının test blokları arası kriter kaymasını sağlamasının ancak bloklar yeterince uzunsu mümkün olduğu son yıllardaki çalışmalar sonucunda saptanmıştır (Hicks ve Starns, 2014; Starns ve ark., 2010; Verde ve Rotello, 2007). Bu tür çalışmalar alanyazındaki çelişen sonuçların nedenlerini açıklamada önemli bir role sahip olacaktır.

İpucu tipinin kriter değişimini sağlamadaki belirleyici etkisini vurgulayan bir çalışma ise Singer (2009) tarafından gerçekleştirilmiştir. Singer çalışmasında anlamsal olarak farklı kategorilere ait (kuşlar, kimyasal elementler, meslekler vd.) kelimeleri kullanmış, kategorileri bir zayıf ve bir güçlü olacak şekilde çiftlere ayırarak sunmuş ve test etmiştir. Çalışma aşamasında sunulan her kelimenin olumluluk (pleasantness) düzeyinin puanlandığı anlamsal yönelme görevi (orienting task) kullanılmıştır. Test aşamasında ise zayıf ve güçlü kategoriden kelimeler karışık sırayla sorulmuş, zayıf kategorideki kelimeler için belirlenen kriter güçlü kategoriye göre daha yumuşak olmuştur. Bu bulgu, deneyin kategori çiftlerinin ayrı bloklar halinde (Deney 1) ya da bir arada karışık sırayla sunulduğu (Deney 2) iki varyasyonunda da raporlanmıştır. Singer'ın bu deneydeki bir diğer önemli bulgusu, yönelim görevinin kriter değişimi üzerindeki etkisidir. Anlamsal yönelim görevi kullanılan deneylerde maddeden maddeye kriter kayması gözleyen Singer, çalışma aşamasında bu görevin uygulandığı koşulla standart bir çalışma aşamasının uygulandığı koşulu kıyaslamış ve standart koşulda kriter kayması olmadığını ortaya koymuştur. Kelimelerin somutluğunun (concreteness; Hirshman ve Arndt, 1997) ya da olumluluğunun (Shiffrin, Huber ve Marinelli, 1995) oylandığı diğer çalışmalarda da liste içi kriter değişiminin gözlenmesinden yola çıkan Singer, yönelim görevinin kriter değişimi için kısmen belirleyici olduğu sonucuna varmıştır (2009). Singer ayrıca kriter değişiminin sağlanması için kişinin mevcut test maddesinden bilgi alabilmesi gerektiği çıkarımını yapmıştır. Ancak Han ve Dobbins'in (2008, 2009) yanlı geribildirim yoluyla kriter kaymasının gerçekleşmesini sağladığı araştırmalarında katılımcıların söz konusu değişimlemenin farkında olmamaları Singer'ın bu çıkarımını yanlışlar niteliktedir.

Kriter kaymasını sağladığı saptanan bir diğer ipucu farklı deneysel koşullar için farklı yanıt tuşları kullanmaktır (Franks ve Hicks, 2016; Rhodes ve Jacoby, 2007; Starns ve Olchowski, 2015). Rhodes ve Jacoby'nin (2007) çalışmasında 4-tuş koşulunda çoğunluğu eski ve çoğunluğu yeni maddelere klavyenin farklı tuşlarıyla yanıt verilmiş, 2-tuş koşulunda ise bütün test maddeleri için aynı tuşlar kullanılmıştır. 4-tuş koşulunda temel oranlar arası kriter farkı 2-tuş koşulundakinden

daha büyük olmuş, ayrıca 2-tuş koşulunda temel oranların kriter farkları deney blokları boyunca sabit kalırken farklı tuş koşulunda kriter farkı test boyunca artmıştır. Stars ve Olchowski (2015) ise bu yöntemi zayıf ve güçlü kelimelere evet yanıtının farklı tuşlarla verildiği 3 tuşlu versiyonuyla kullanmıştır. Bu çalışmanın test aşamasında güçlü ve zayıf kelimeler, karışık sırayla, farklı renklerde ve ekranın farklı taraflarında sunulmuştur. Bu iki koşula eski kelimelerle birlikte sunulacak yeni kelimeler atanmış, güçlü ipuçlarıyla sunulan yeni kelimelerin yanlış alarm oranı zayıflara göre daha düşük olmuştur. Ancak, aynı deneyin 2 tuşlu versiyonunda bu fark gözlenmemiştir (Benzer sonuçlar için bakınız Franks ve Hicks, 2016).

Kriter değişimine yardımcı etkenlerin yanı sıra bazı etkenlerin ise kriter değişiminin önüne geçmekte olduğu ya da kişiler arasında farklılaşma yaratarak onu görünmez kıldığı raporlanmıştır. Kişilerin kriter değişimi olmadığı durumlardaki kriter konumlarının büyük farklılıklar göstermesi (Aminoff ve ark., 2012; Kantner ve Lindsay, 2012) başlıca karıştırıcı etkenlerden biridir. Kriter konumu açısından kişiler arasında görülen dikkate değer farkların olası nedenlerini inceleyen çalışmalarda araştırmacılar, kriter konumu ile çeşitli örneklemeler (Alzheimer'lı kişiler, Beth, Budson, Waring ve Ally, 2009; demanslı kişiler, Woodard, Axelrod, Mordecai ve Shannon, 2004) ve kişilik özellikleri (içedönüklük, Gillespie ve Eysenck, 1980; olumsuz kişilik, eğlence arayan kişilik, Aminoff ve ark., 2012) arasındaki ilişkiyi ortaya koymuştur. Bununla birlikte, kişilerin kriter değiştirme düzeyleri açısından farklı zamanlar ve görevler arasındaki tutarlılıkları gözlenmiştir (Aminoff ve ark., 2012; Kantner ve Lindsay, 2012, 2014). Kantner ve Lindsay'ın çalışmasında (2012) katılımcılar, birbirlerinden 10 dakika ya da 2 gün farkla ayrılan 2 standart tanıma belleği görevine girmişlerdir. Katılımcıların iki testteki C kriter puanları arasında hem 10 dakika koşulunda ( $r(37) = .67$ , Deney 1) hem de 2 gün koşulunda ( $r(42) = .73$ , Deney 2) yüksek korelasyon bulunmuştur. Aynı sonuç birbirini takip eden tanıma belleği görevlerindeki uyaran tiplerinin kelime ve resimler olarak değiştiği deneylerinde de gözlenmiştir (2012; benzer sonuçlar için bakınız Kantner ve Lindsay, 2014). Franks ve Hicks (2016) güçlü ve zayıf maddelerin aynı liste içinde test edildiği çalışmalarında katılımcıların birbirini takip eden iki testteki kriter değişim düzeyleri arasında korelasyon olduğunu gözlemiş, aynı zamanda güçlü ve zayıf test bloklarının büyüklüğü (bloksuz – 10 – 40) arttıkça bu korelasyonun azaldığını raporlamışlardır. Franks ve Hicks'e (2016) göre güçlü ve zayıf blokların büyüklüğü ne kadar az olursa, kişilerin test içinde kriterlerini kaydırmaları için o kadar çok fırsatı olmakta, bu sayede daha güvenilir sonuçlar alınabilmektedir.

Selmeczy ve Dobbins'in (2013) çalışmasına göre kriter belirlemede dışsal ipuçlarından yararlanma becerisi kişiden kişiye göre farklılık göstermekte ve üstbilişsel izleme (metacognitive monitoring) becerisi tarafından yordanabilmektedir. Ayrıca kişilerin farklı güvenilirlik düzeyindeki (%65 – %85) ipuçlarından yararlanma becerileri tutarlı ve bir değerinin yordayıcısıdır.

Kişilerin farklı değişimleme görevlerindeki kriterleri arasında tutarlılık olup olmadığı da inceleme konusu olmuştur. Franks ve Hicks'in (2016) kişilerin güç ve temel oran değişimlemesiyle gerçekleştirdiği kriter kaymalarının büyüklükleri arasında bir korelasyon olup olmadığını inceledikleri çalışmalarında katılımcılar, bir güç bir de temel oran değişimlemeli iki testten oluşan bir seansa iki gün aralıkla ikişer kere katılmışlardır. Toplam dört bellek testinin hepsinde deneysel koşullar arasında kriter kayması gözlenmiştir. Bu kriter farklarının büyüklüğü aynı tip testler arasında korelasyon göstermiş, ancak güç değişimlemeli testler ile temel oran değişimlemeli testler arasında korelasyon bulunamamıştır. Ayrıca kişilerin kriter konumları farklı uyaran tipleriyle yapılan tanıma belleği görevleri arasında da tutarlıdır: Kantner ve Lindsay (2014) çalışmalarında kişilerin kelime ve insan yüzleri uyaranlarıyla yaptıkları deneylerdeki kriterlerinin uyaran tipleri arasında tutarlılık gösterdiği görülmüştür (Benzer bir sonuç için bakınız Aminoff ve ark., 2012).

Son olarak, tanıma belleğinin diğer bilişsel görevlerle ilişkisinin incelendiği çalışmalarda tanıma belleği kriter değişim düzeyi ile Deese/Roediger-McDermott görevi (Roediger ve McDermott, 1995) kritik yanlış alarm oranı (Kantner ve Lindsay, 2012), görgü tanığı görevi yanlış teşhis oranı (Kantner ve Lindsay, 2014) ve yap-yapma (go-no-go; Donders, 1969) görevi doğru yanıt oranı (Kantner ve Lindsay, 2014) arasında ilişki bulunmuştur. Bu çalışmalarda elde edilen bulguları açıklamak amacıyla söz konusu görevlerdeki karar verme süreçleri arasında nasıl bir bağlantı olduğuna ilişkin bir görüş sunulmamıştır. Bu süreçlerin ve aralarındaki ilişkilerin açıklanması karar kriteri alanyazınında geleceğin çalışma konularından biri olmaya adaydır.

## Sonuç

Tanıma belleğinin temel ölçüm modeli olan SBK, performansı maddelerin bellek gücüne, yani aşinalık değerine, ve tanıma kararının verildiği aşinalık eşliğini ifade eden karar kriterine bağlı olarak açıklar. Söz konusu karar kriteri, kişisel farklılıklara ya da deneysel değişimlemelere bağlı olarak çeşitlilik gösterir ve bu farklılıklar tepki yanlılığı olarak adlandırılır. Alanyazında karar kriteri temel oran, bellek gücü, ödeme matrisleri gibi çeşitli yollarla değişimlenmiştir. Bulgulara göre temel oran de-

ğişimlemeleri ön bilgi, geribildirim gibi ipuçlarının yardımıyla belirgin kılındığında karar kriteri üzerinde etkili olmaktadır (Estes ve Maddox, 1995; Rhodes ve Jacoby, 2007). Bellek gücü değişimlemelerinin etkili olmaları için ise söz konusu yardımcı yöntemlere ek olarak, renk, anlamsal kategori gibi ipuçlarının yardımına da başvurulması gerekli olmuştur (Hicks ve Starns, 2014; Singer 2009; Verde ve Rotello, 2007). Ancak bellek gücü değişimlemesi ile elde edilen isabet ve yanlış alarm oranlarındaki farklılaşmanın kriter kayması değil, ayırıştırma adı verilen bir süreç nedeniyle gerçekleştiğine yönelik yorumlar da mevcuttur (Criss, 2006, 2010; Kılıç ve ark., 2017). Karar kriteri tanıma belleği modellerinde önemli bir bileşeni oluşturmaktadır (Cox ve Shiffrin, 2017; Dennis ve Humphreys, 2001; McClelland ve Chappell, 1998; Shiffrin ve Steyvers, 1997; Turner ve ark., 2011). Bu nedenle, karar kriteri çalışmalarında elde edilen bulguların kriter kayması ya da ayırıştırma süreçleriyle açıklanmaları tanıma belleğinin işleyişine ilişkin modellemelerde önemli farkları beraberinde getirmektedir. Bu iki alternatif açıklamanın yeni çalışmalarla sınanması ve güncellenmesi tanıma belleğinin açıklanmasına yönelik değerli katkılar sunacaktır.

Alanyazındaki bir diğer eksiklik, kriter değişimleme çalışmalarındaki çelişen bulguların açıklanmasına yardımcı olacak yeterli veriye sahip olunmamasıdır. Söz konusu çalışmalarda, değişimlemenin kriter kaymasına yol açmasını belirleyen temel etkenin yardımcı ipuçları olduğu görülmektedir, ancak bu ipuçlarının hangi koşullarda tutarlı bir etkiye sahip oldukları ya da etki mekanizmaları açık değildir. Örneğin, bazı çalışmalarda kriter değişiminin sağlanması ancak geribildirim varlığına bağlı olarak gerçekleşmekteyken (Estes ve Maddox, 1995; Rhodes ve Jacoby, 2007), diğerlerinde geribildirim bir etkisi gözlenmemiştir (Selmecezy ve Dobbins, 2013; Starns ve ark., 2010). Rhodes ve Jacoby (2007) geribildirim etkisini temel oran değişimlemesinin farkındalığını artırması üzerinden açıklarken Han ve Dobbins'in (2008, 2009) araştırmalarında yanlı geribildirim karar kriterinde farklılaşmaya yol açmış, ancak katılımcılar söz konusu değişimlemenin farkında olmadıklarını raporlamışlardır. Han ve Dobbins yanlı geribildirim yöntemlerine "artan pekiştirme öğrenmesi" (incremental reinforcement learning) adını vererek geribildirim evet ve hayır yanıtlarının sonuçlarında bir ayrışma yaratarak tepki yanlılığına yol açtığı yorumunda bulunmuşlardır (2009). Bu yorum farklılıkları dikkate alındığında geribildirim mnemonik etkileri ve üstbilişsel karar mekanizmaları üzerindeki etkilerinin detaylı bir biçimde araştırılması ve modellenmesi, tanıma belleği alanyazınının gelecek araştırma konularından biri olacaktır.

Son olarak, tanıma belleğinde karar kriteri ile Deese/Roediger-McDermott görevi, görgü tanığı belleği

ve çeşitli bellek dışı karar verme görevlerindeki karar kriterleri arasındaki ilişkileri inceleyen çeşitli çalışmalar yapılmıştır (Kantner ve Lindsay, 2012, 2014); ancak bu çalışmalarda ortaya konan bulgulardan yola çıkan, farklı karar verme süreçleri arasında ne tür bağlantılar olduğuna yönelik açıklamalar mevcut değildir. Farklı karar verme süreçleri arasındaki bağlantıları ortaya koyacak yeni çalışmalar bu süreçlerin açıklanmasına yardımcı olacak, tanıma belleğinde kriter konusunun araştırıldığı gündün güne genişleyen bir alanyazının bilişsel psikolojide daha geniş bir bağlamda anlam kazanmasına katkıda bulunacaktır.

## Kaynaklar

- Aminoff, E. M., Clewett, D., Freeman, S., Frithsen, A., Tipper, C., Johnson, A., ... Miller, M. B. (2012). Individual differences in shifting decision criterion: A recognition memory study. *Memory & Cognition, 40*(7), 1016-1030. doi: 10.3758/s13421-012-0204-6
- Banks, W. P. (1970). Signal detection theory and human memory. *Psychological Bulletin, 74*(2), 81-99. doi:10.1037/h0029531
- Benjamin, A. S. ve Bawa, S. (2004). Distractor plausibility and criterion placement in recognition. *Journal of Memory and Language, 51*(2), 159-172. doi:10.1016/j.jml.2004.04.001
- Beth, E. H., Budson, A. E., Waring, J. D. ve Ally, B. A. (2009). Response bias for picture recognition in patients with Alzheimer's disease. *Cognitive and Behavioral Neurology: Official Journal of the Society for Behavioral and Cognitive Neurology, 22*(4), 229-235. doi:10.1097/WNN.0b013e3181b7f3b1
- Bruno, D., Higham, P. A. ve Perfect, T. J. (2009). Global subjective memorability and the strength-based mirror effect in recognition memory. *Memory & Cognition, 37*(6), 807-818. doi:10.3758/MC.37.6.807
- Cox, J. C. ve Dobbins, I. G. (2011). The striking similarities between standard, distractor-free, and target-free recognition. *Memory & Cognition, 39*(6), 925-940. doi:10.3758/s13421-011-0090-3
- Cox, G. E. ve Shiffrin, R. M. (2017). A dynamic approach to recognition memory. *Psychological Review, 124*(6), 795-860. doi:10.1037/rev0000076
- Criss, A. H. (2006). The consequences of differentiation in episodic memory: Similarity and the strength based mirror effect. *Journal of Memory and Language, 55*(4), 461-478. doi:10.1016/j.jml.2006.08.003
- Criss, A. H. (2010). Differentiation and response bias in episodic memory: Evidence from reaction time distributions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(2), 484-499. doi:10.1037/a0018435
- Curran, T., DeBuse, C. ve Leynes, P. A. (2007). Conflict and criterion setting in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(1), 2-17. doi:10.1037/0278-7393.33.1.2
- Dennis, S. ve Humphreys, M. S. (2001). A context noise model of episodic word recognition. *Psychological Review, 108*(2), 452-478. doi:10.1037/0033-295X.108.2.452
- Dobbins, I. G. ve Kroll, N. E. (2005). Distinctiveness and the recognition mirror effect: Evidence for an item-based criterion placement heuristic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(6), 1186-1198. doi: 0.1037/0278-7393.31.6.1186
- Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica, 30*, 412-431. doi:10.1016/0001-6918(69)90065-1
- Estes, W. K. ve Maddox, W. T. (1995). Interactions of stimulus attributes, base rates, and feedback in recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(5), 1075-1095. doi:10.1037/0278-7393.21.5.1075
- Franks, B. A. ve Hicks, J. L. (2016). The reliability of criterion shifting in recognition memory is task dependent. *Memory & Cognition, 44*(8), 1215-1227. doi:10.3758/s13421-016-0633-8
- Gillespie, C. R. ve Eysenck, M. W. (1980). Effects of introversion-extraversion on continuous recognition memory. *Bulletin of the Psychonomic Society, 15*(4), 233-235. doi:10.3758/BF03334517
- Glanzer, M. ve Adams, J. K. (1985). The mirror effect in recognition memory. *Memory & Cognition, 13*(1), 8-20. doi:10.3758/BF03198438
- Glanzer, M., Adams, J. K., Iverson, G. J. ve Kim, K. (1993). The regularities of recognition memory. *Psychological Review, 100*(3), 546-567. doi:10.1037/0033-295X.100.3.546
- Green, D. M. ve Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Han, S. ve Dobbins, I. G. (2008). Examining recognition criterion rigidity during testing using a biased-feedback technique: Evidence for adaptive criterion learning. *Memory & Cognition, 36*(4), 703-715. doi:10.3758/MC.36.4.703
- Han, S. ve Dobbins, I. G. (2009). Regulating recognition decisions through incremental reinforcement learning. *Psychonomic Bulletin & Review, 16*(3), 469-474. doi:10.3758/PBR.16.3.469
- Healy, A. F. ve Jones, C. (1975). Can subjects maintain a constant criterion in a memory task? *Memory & Cognition, 3*(3), 233-238. doi:10.3758/BF03212903
- Healy, A. F. ve Kubovy, M. (1977). A comparison of recognition memory to numerical decision: How prior probabilities affect cutoff location. *Memory & Cognition, 5*(1), 3-9. doi: 10.3758/BF03209184
- Healy, A. F. ve Kubovy, M. (1978). The effects of payoffs and prior probabilities on indices of performance and cutoff location in recognition memory. *Memory & Cognition, 6*(5), 544-553. doi:10.3758/BF03198243
- Heit, E., Brockdorff, N. ve Lamberts, K. (2003). Adaptive changes of response criterion in recognition memory. *Psychonomic Bulletin & Review, 10*(3), 718-723. doi:10.3758/BF03196537

- Hicks, J. L. ve Starns, J. J. (2014). Strength cues and blocking at test promote reliable within-list criterion shifts in recognition memory. *Memory & Cognition*, 42(5), 742-754. doi:10.3758/s13421-014-0397-y
- Hirshman, E. (1995). Decision processes in recognition memory: Criterion shifts and the list-strength paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(2), 302-313. doi:10.1037/0278-7393.21.2.302
- Hirshman, E. ve Arndt, J. (1997). Discriminating alternative conceptions of false recognition: The cases of word concreteness and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(6), 1306-1323. doi:10.1037/0278-7393.23.6.1306
- Hockley, W. E. ve Niewiadomski, M. W. (2007). Strength-based mirror effects in item and associative recognition: Evidence for within-list criterion changes. *Memory & Cognition*, 35(4), 679-688. doi:10.3758/BF03193306
- Jaeger, A., Lauris, P., Selmeczy, D. ve Dobbins, I. G. (2012). The costs and benefits of memory conformity. *Memory & Cognition*, 40(1), 101-112. doi:10.3758/s13421-011-0130-z
- Kantner, J. ve Lindsay, D. S. (2012). Response bias in recognition memory as a cognitive trait. *Memory & Cognition*, 40(8), 1163-1177. doi:10.3758/s13421-012-0226-0
- Kantner, J. ve Lindsay, D. S. (2014). Cross-situational consistency in recognition memory response bias. *Psychonomic Bulletin & Review*, 21(5), 1272-1280. doi:10.3758/s13423-014-0608-3
- Kapucu, A., Rotello, C. M., Ready, R. E. ve Seidl, K. N. (2008). Response bias in "remembering" emotional stimuli: A new perspective on age differences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(3), 703. doi:10.1037/0278-7393.34.3.703
- Kılıç, A., Criss, A. H., Malmberg, K. J. ve Shiffrin, R. M. (2017). Models that allow us to perceive the world more accurately also allow us to remember past events more accurately via differentiation. *Cognitive Psychology*, 92, 65-86. doi:10.1016/j.cogpsych.2016.11.005
- Koop, G. J., Criss, A. H. ve Malmberg, K. J. (2015). The role of mnemonic processes in pure-target and pure-foil recognition memory. *Psychonomic Bulletin & Review*, 22(2), 509-516. doi:10.3758/s13423-014-0703-5
- Ley, R. ve Long, K. (1987). A distractor-free test of recognition and false recognition. *Bulletin of the Psychonomic Society*, 25(6), 411-414. doi:10.3758/BF03334727
- Ley, R. ve Long, K. (1988). Distractor similarity effects in tests of discrimination recognition and distractor-free recognition. *Bulletin of the Psychonomic Society*, 26(5), 407-409. doi:10.3758/BF03334898
- Macmillan, N. A. ve Creelman, C. D. (1996). Triangles in ROC space: History and theory of "nonparametric" measures of sensitivity and response bias. *Psychonomic Bulletin & Review*, 3(2), 164-170. doi:10.3758/BF03212415
- Macmillan, N. A. ve Creelman, C. D. (2005). *Detection theory: A user's guide* (2. Baskı). New York: Erlbaum.
- McClelland, J. L. ve Chappell, M. (1998). Familiarity breeds differentiation: A subjective-likelihood approach to the effects of experience in recognition memory. *Psychological Review*, 105(4), 724-760. doi:10.1037/0033-295X.105.4.734-760
- Morrell, H. E., Gaitan, S. ve Wixted, J. T. (2002). On the nature of the decision axis in signal-detection-based models of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(6), 1095-1110. doi:10.1037/0278-7393.28.6.1095
- Pendergrass, R., Olfman, D., Schmalstig, M., Seder, K. ve Light, L. L. (2011). Age, criterion flexibility, and associative recognition. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 67(1), 36-42. doi:10.1093/geronb/gbr071
- Ratcliff, R., Clark, S. E. ve Shiffrin, R. M. (1990). List-strength effect: I. Data and discussion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(2), 163-178. doi:10.1037/0278-7393.16.2.163
- Ratcliff, R., Sheu, C. F. ve Gronlund, S. D. (1992). Testing global memory models using ROC curves. *Psychological Review*, 99(3), 518-535. doi:10.1037/0033-295X.99.3.518
- Reed, A. V. (1973). Speed-accuracy trade-off in recognition memory. *Science*, 181(4099), 574-576. doi:10.1126/science.181.4099.574
- Rhodes, M. G. ve Jacoby, L. L. (2007). On the dynamic nature of response criterion in recognition memory: Effects of base rate, awareness, and feedback. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(2), 305-320. doi:10.1037/0278-7393.33.2.305
- Roediger, H. L. ve McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(4), 803-814. doi:10.1037/0278-7393.21.4.803
- Scimeca, J. M., McDonough, I. M. ve Gallo, D. A. (2011). Quality trumps quantity at reducing memory errors:

- Implications for retrieval monitoring and mirror effects. *Journal of Memory and Language*, 65(4), 363-377. doi:10.1016/j.jml.2011.04.008
- Selmeçy, D. ve Dobbins, I. G. (2013). Metacognitive awareness and adaptive recognition biases. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(3), 678-690. doi:10.1037/a0029469
- Shiffrin, R. M. ve Steyvers, M. (1997). A model for recognition memory: REM—retrieving effectively from memory. *Psychonomic Bulletin & Review*, 4(2), 145-166. doi:10.3758/BF03209391
- Shiffrin, R. M., Huber, D. E. ve Marinelli, K. (1995). Effects of category length and strength on familiarity in recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(2), 267-287. doi:10.1037//0278-7393.21.2.267
- Singer, M. (2009). Strength-based criterion shifts in recognition memory. *Memory & Cognition*, 37(7), 976-984. doi:10.3758/MC.37.7.976
- Singer, M., Fazaluddin, A. ve Andrew, K. N. (2011). Distinctiveness and repetition in item recognition. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, 65(3), 200-2007. doi:10.1037/a0023704
- Snodgrass, J. G. ve Corwin, J. (1988). Pragmatics of measuring recognition memory: applications to dementia and amnesia. *Journal of Experimental Psychology: General*, 117(1), 34-50. doi:10.1037/0096-3445.117.1.34
- Stanislaw, H. ve Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, 31(1), 137-149. doi:10.3758/BF03207704
- Starns, J. J. ve Olchowski, J. E. (2015). Shifting the criterion is not the difficult part of trial-by-trial criterion shifts in recognition memory. *Memory & Cognition*, 43(1), 49-59. doi:10.3758/s13421-014-0433-y
- Starns, J. J., White, C. N. ve Ratcliff, R. (2010). A direct test of the differentiation mechanism: REM, BCDMEM, and the strength-based mirror effect in recognition memory. *Journal of Memory and Language*, 63(1), 18-34. doi:10.1016/j.jml.2010.03.004
- Stretch, V. ve Wixted, J. T. (1998). On the difference between strength-based and frequency-based mirror effects in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(6), 1379-1396. doi:10.1037/0278-7393.24.6.1379
- Swets, J. A., Tanner Jr, W. P. ve Birdsall, T. G. (1961). Decision processes in perception. *Psychological Review*, 68(5), 301-340. doi:10.1037/h0040547
- Turner, B. M., Van Zandt, T. ve Brown, S. (2011). A dynamic stimulus-driven model of signal detection. *Psychological Review*, 118(4), 583-613. doi:10.1037/a0025191
- Van Zandt, T. (2000). ROC curves and confidence judgments in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(3), 582-600. doi:10.1037/0278-7393.26.3.582
- Verde, M. F. ve Rotello, C. M. (2007). Memory strength and the decision process in recognition memory. *Memory & Cognition*, 35(2), 254-262. doi:10.3758/BF03193446
- Wallace, W. P. (1982). Distractor-free recognition tests of memory. *The American Journal of Psychology*, 95(3), 421-440. doi:10.2307/1422134
- Wallace, W. P., Sawyer, T. J. ve Robertson, L. C. (1978). Distractors in recall, distractor-free recognition, and the word-frequency effect. *The American Journal of Psychology*, 91(2), 295-304. doi:10.2307/1421539
- Woodard, J. L., Axelrod, B. N., Mordecai, K. L. ve Shannon, K. D. (2004). Value of signal detection theory indexes for Wechsler Memory Scale-III recognition measures. *Journal of Clinical and Experimental Neuropsychology*, 26(4), 577-586. doi:10.1080/13803390490496614

## *Summary*

# Criterion Setting in Recognition Memory and The Manipulations of Response Bias

Cavit Deniz Pala

Middle East Technical University

Aslı Kılıç

Middle East Technical University

Recognition memory is the ability to distinguish whether a stimulus (person, object or situation) is previously experienced. There are many variants of recognition tasks; namely, yes-no recognition, confidence judgement, forced choice, associative recognition, and continuous recognition. Correct recognition of an old item is called a “hit”, when rejecting an old item is a “miss”. “False alarm” is endorsement of a new item, and, finally, rejecting a new item is called “correct rejection”.

### Signal Detection Theory and The Criterion

According to Signal Detection Theory (SDT; Banks, 1970; Green and Swets, 1966; Macmillan and Creelman, 2005; Swets, Tanner and Birdsall, 1961), memory strength, or familiarity, is calculated for each probe item. Old and new items form two overlapping distributions on memory strength scale. Recognition decision is made by using a decision criterion to which memory strength of test items is compared. Sensitivity and response bias are the factors determining the performance according to SDT.

Sensitivity is the extent to which old items are distinguished from the new items.  $d'$ , which is the main sensitivity measure of SDT, is calculated via the following equation:  $d' = z(\text{HR}) - z(\text{FAR})$ , where  $z$  is the inverse of the normal cumulative distribution, HR is hit rate and FAR is false alarm rate (Macmillan and Creelman, 2005, p. 8).  $d' = 0$  is chance performance while increasing positive values indicating better sensitivity.

Main response bias measure of the SDT, Criterion location  $C$ , is calculated using the equation  $C = -.5 [z(\text{HR}) + z(\text{FAR})]$  (Macmillan and Creelman, 2005, p. 29). A value of 0 indicates neutral criterion while increasing negative values signify more lenient criteria, tendency to say “yes” to a probe; and, positive values mean stricter criteria. Likelihood ratio  $\beta$ ,  $\ln(\beta)$ , and  $B''$  are other common response bias measures. FAR is also

used as a response bias measure in the studies in which the locations of old distributions are manipulated. In these studies, new distributions are assumed to be fixed (Franks and Hicks, 2016; Hicks and Starns, 2014). But differentiation models (Criss, 2006, 2010; Kılıç, Criss, Malmberg and Shiffrin, 2017; McClelland and Chappell, 1998; Shiffrin and Steyvers, 1997) suggest that new items are better distinguished among well-learned study items, compared to poorly-learned items; thus the new distributions of strong (well-learned) lists are located more to the left than that of weak lists.

### Studies with Criterion Manipulation

#### Base Rate Manipulations

Base rate (the proportion of old items in a test list) manipulation is one of the most common methods to induce a criterion shift (Aminoff et al. 2012; Cox and Dobbins, 2011; Estes and Maddox, 1995; Franks and Hicks, 2016; Healy and Kubovy, 1977, 1978; Heit, Brockdorff and Lamberts, 2003; Rhodes and Jacoby, 2007). Participants tend to say “yes” in mostly-old test list; however, adopt stricter criteria for mostly-new test lists to achieve optimum performance according to SDT (Macmillan and Creelman, 2005; Swets et al., 1961). Swets et al. (1961) reported the effect of base rate on criterion in a signal detection task. Examining such effects in recognition memory tasks, Healy and Jones (1975) manipulated the base rates of test blocks (25% - 50%) in random order, with the prior information given to the participants about the base rates, and found no reliable criterion shifts between the test blocks. Healy and Kubovy (1977) also reported no changes in criterion in a recognition memory task but they observed criterion shifts in a numerical decision task.

In contrast with these results, there are studies in the literature that reported criterion shifts due to base rate manipulations. Estes and Maddox (1995) observed



criterion shifts when base rate manipulation is supported with item feedback for 3-digit-number and 3-letter stimuli; but failed to observe such shifts when feedback is absent or with word stimuli. Later, Aminoff et al. (2012) and Franks and Hicks (2016, Experiment 2) found that subjects were able to shift their criteria in the test blocks ranged 6-to-9 items within the same test phases. Moreover, Rhodes and Jacoby (2007) managed to induce item-by-item criterion shifts by assigning different base rates to the items appearing on different sides of the screen. They also noted that criterion shifts depend on the presence of feedback.

A special case of base rate manipulations is when test list consists of only old items (pure-old) or only new items (pure-new). Surprisingly, there are many studies in the literature that found null effects of pure list manipulations on hit or false alarm rates (Cox and Dobbins, 2011; Ley and Long, 1987, 1988; Wallace, 1982; Wallace, Sawyer and Robertson 1978), regardless of the presence of prior information. A possible explanation for null results is that participants focused on their subjective mnemonic evidence rather than the composition of test lists. As a matter of fact, Koop, Criss and Malmberg (2015) found criterion shifts due to base rate manipulations with pure lists when feedback is present, but replicated the previous results in the absence of feedback.

### Manipulations of Memory Strength

In another method, locations of old and new distributions (i.e. memory strength) are manipulated to induce criterion shift. When weak and strong items are studied and tested in separate lists, hit rates are higher and false alarms are lower for strong lists than weak lists (strength based mirror effect; Glanzer and Adams, 1985; Ratcliff, Clark and Shiffrin, 1990). It is suggested that this pattern stems from criterion shifts, more specifically, stricter criterion is set for strong lists (Hicks and Starns, 2014; Starns, White and Ratcliff, 2010; Stretch and Wixted, 1998). However, differentiation models (Criss, 2006; Kılıç et al., 2017; McClelland and Chappell, 1998; Shiffrin and Steyvers, 1997) suggest that new probes reveal poor match with strong items, result in fewer false alarms.

Another finding called null list strength effect (Ratcliff et al. 1990; Ratcliff, Sheu and Gronlund, 1992, Experiment 2; Starns et al., 2010) is that sensitivity is not affected by the strength of other items in study list. However, Hirshman (1995) found that criterion is more lenient for weak lists than weak items in mixed lists (consisting of strong and weak items), and criterion is stricter for strong lists than strong items in mixed lists. That means composition of test lists affects criterion.

Morrell, Gaitan and Wixted (2002) investigated the possibility of criterion shifts within a test list via strength manipulation, and reported that the criterion was stable during test phase. Stretch and Wixted (1998), and Verde and Rotello (2007, Experiments 1-4) replicated this result; however, Verde and Rotello (2007, Experiment 5) were able to induce such criterion shifts when subjects are given feedback during test phases. Other studies examining criterion shifts within test lists revealed that as the size of weak or strong test blocks increase, criterion shifts become more likely and their size tends to be larger (Hicks and Starns, 2014; Starns et al., 2010; Verde and Rotello, 2007).

In some studies, strength of new distributions is manipulated rather than old distributions. Adopting this approach, Benjamin and Bawa (2004) used new words from the same categories with old words in a test list, but the other test list included unrelated new words. Stricter criterion was set for the former test list.

### Other Methods

Methods used in criterion manipulations are not limited to base rate and strength manipulations. One of the other methods is using payoff matrices by which the gain and loss balances of “yes” and “no” answers are manipulated. When “yes” answer becomes more advantageous, subjects tend to say “yes” more often; likewise, when gains of “no” answer outweigh gains of “yes” answer, their criteria become more lenient (Curran, DeBuse and Leynes, 2007; Healy and Kubovy, 1978; Van Zandt, 2000).

Selmecezy and Dobbins (2013) used “probabilistic mnemonic cues” that give the participants the probability of the following probe being old. Subjects adopt stricter criterion following the probably-old cues rather than the probably-new cues. As the sensitivity decrease, they tend to rely more on probabilistic cues.

Han and Dobbins (2008) managed to induce criterion shifts by using a covert manipulation, biased feedback. In this method, participants are given “correct” feedback to their misses (strict condition) or false alarms (lenient condition). Participants were not aware of the manipulation; yet, their criteria shifted in accord with the conditions.

Finally, there are many studies in the literature reporting changes in hit and false alarms due to several stimulus properties, such as word frequency (Glanzer and Adams, 1985), or stimulus type (Scimeca, McDonough, and Gallo, 2011); but it remains unclear and is an ongoing debate among the scientists whether this changes stem from criterion shifts or any other mechanisms (e.g. differentiation).

### **Factors in Criterion Shifts**

Literature has conflicting findings regarding the effectiveness of manipulations on response bias. Studies have revealed that additional clues, such as feedback or prior information, are needed to induce reliable criterion shifts. For instance, some studies reported that criterion shifts were observed only when feedback was provided to the participants (Estes and Maddox, 1995; Morrell et al., 2002; Rhodes and Jacoby, 2007; Verde and Rotello, 2007); yet, other studies failed to establish such decisive role of feedback (Hicks and Starns, 2014; Selmezy and Dobbins, 2013). It is observed that additional clues, such as using different font colors (Hicks and Starns, 2014), semantic categories (Singer, 2009), or response keys (Rhodes and Jacoby, 2007) for different conditions are effective in revealing criterion shifts.

Unlike the auxiliary clues, some factors hinder criterion shifts. For instance, initial criterion placement varies substantially among the subjects (Aminoff et al., 2012; Kantner and Lindsay, 2012); besides, the degree to which they shift their criterion is consistent across different time and tasks (Aminoff et al., 2012; Franks and Hicks, 2016; Kantner and Lindsay, 2012, 2014). Furthermore, the ability to benefit from external cues varies between individuals (Selmezy and Dobbins, 2013). Finally, it is reported that the degree of criterion shifts is related to critical false alarm rate in Deese/Roediger-McDermott task (Kantner and Lindsay, 2012; Roediger and McDermott, 1995), false identifications in eyewitness task (Kantner and Lindsay, 2014), and hit rates in go-no-go task (Donders, 1969; Kantner and Lindsay, 2014).