Türk Psikoloji Dergisi, Haziran 2018, *33*(81), 62-78

Türkçe Kelimelerin Serbest Çağrışım Normları Üzerine Güvenirlik ve Geçerlik Çalışması

Ümit Akırmak
İstanbul Bilgi Üniversitesi

Mehmet Altan Orhon
İstanbul Bilgi Üniversitesi

Özet

Kelimelerin çağrışım yapıları, kelimeler üzerine işlem yapılan durumsal görevlerde performansı etkilemektedir (Nelson, Dyrdal ve Goodmon, 2005). Serbest çağrışım yöntemi kullanılarak kelimelerin anlamca güçlü olarak bağlı olduğu diğer kelimeler belirlenip kelimelerin çağrışım yapılarının anlamsal bellekteki organizasyonuna dair bilgiler elde edilebilir. Türkçe kelimelerin çağrışım normları Tekcan ve Göz (2005) tarafından toplanmış olsa da bu bulguların güvenirliği ve geçerliği özel olarak test edilmemiştir. Bu araştırmanın amacı var olan serbest çağrışım norm bulgularının güvenirlik ve geçerliğini belirlemektir. Güvenirliği ölçmek için yapılan birinci çalışmada, var olan normlardan 100 kelime seçilmiş ve bu kelimelerin uyandırdığı çağrışımlar hem klasik kalem-kağıt yöntemiyle hem de çevrimiçi olarak toplanılmıştır. Veriler çağrışım gücü, çağrışım seti büyüklüğü ve çağrışımın heterojenliği değerleri üzerinden karşılaştırılmıştır. Elde edilen bulgular normatif verilerin yüksek güvenirliğe sahip olduğunu göstermiştir. İkinci çalışma, var olan normların yordama geçerliğini ipucu ile hatırlama yönteminin kullanıldığı deneysel bir desende sınamıştır. İpucu kelimelerinin çağrışım seti büyüklüğü ve çağrışım gücü değerleri değişimlenerek (manipüle edilerek) doğru hatırlama oranına etkileri incelenmiştir. Alanyazındaki beklentiler çerçevesinde çağrışım gücü değerleri yüksek ipuçları düşük olanlara göre daha iyi bellek performansı ile sonuçlanmış ancak ipucu çağrışım setinin büyüklüğü beklenen etkiyi göstermemiştir. Bulgular, çağrışım normlarının büyük oranda güvenilir ve geçerli olduğunu göstermekle beraber normatif verilerdeki eksikliklerin ilave norm çalışmaları ile giderilmesi gerektiğinin önemine işaret etmektedir.

Anahtar kelimeler: Serbest çağrışım, ipucu ile hatırlama, çağrışımsal bellek, anısal bellek, heterojenlik

Abstract

Pre-existing semantic associations of words influence performance in various episodic tasks (Nelson, Dyrdal, & Goodmon, 2005). The free association task indexes strong associative relations among words. In turn, this makes it possible to map associative structure of the words in semantic memory. Although Tekcan and Goz (2005) collected free association norms of Turkish words in an earlier study, the reliability and validity of their findings have not been empirically tested with a specific study. The goal of the present study was to determine the reliability and validity of the existing free association norms. In Study 1, 100 words were randomly selected from the existing normative database and were normed again by using the standard free association method as well as an online version of the same task in order to assess the reliability. The obtained findings were compared in regard to associative strength, associative set size, and heterogeneity values. The results indicated that the existing norms have high reliability. Study 2 assessed the predictive validity of the existing norms in an extra-list cued recall experiment. The cue-to-target strength and cue set size were manipulated in order to examine their effects on correct recall. In line with the previous findings, high cue-to-target strength led to higher recall rates compared to low cue-to-target strength; however, the effects of cue set size were not in the expected direction. The present findings imply that the existing free association norms have high reliability and validity overall, but there is a crucial need to overcome the shortcomings of the Turkish free association norms by norming additional words.

Key words: Free association, cued recall, associative memory, episodic memory, heterogeneity

Yazışma Adresi: Dr. Ümit Akırmak, İstanbul Bilgi Üniversitesi, Sosyal ve Beşeri Bilimler Fakültesi, Psikoloji Bölümü, Eski Silahtarağa Elektrik Santralı Kazım Karabekir Cad. No:2/13 Eyüp / İstanbul
E-posta: umit.akirmak@bilgi.edu.tr
Gönderim Tarihi: 21.01.2016
Kabul Tarihi: 30.06.2017

Kelimeler ve kelimeler arasındaki anlam ilişkileri anlamsal belleğimizin bir parçasıdır. Kelime dağarcığımızın oldukça geniş olması ve aradığımız bir kelimeyi çoğu zaman oldukça hızlı bulmamız anlamsal belleğin organizasyonunun rastgele olmadığına ve belirli prensipler çerçevesinde organize olduğuna işaret etmektedir (Aitchison, 2003). Anlamsal belleğin organizasyonunu anlamaya yönelik yöntemlerden bir tanesi serbest çağrışım yöntemidir. Serbest çağrışım yönteminde katılımcıların gösterilen bir kelimeye karşı düşündükleri ilk kelimeyi söylemesi istenir (Deese, 1965; Nelson, McEvoy ve Schreiber, 2004). Bu sayede bir kavramın güçlü olarak çağrıştırdığı ve anlamca bağlı olduğu diğer kavramlar belirlenir (Nelson, McEvoy ve Dennis, 2000). Anlamsal belleği oluşturan kavramlar arasındaki karmaşık yapıyı ortaya çıkarmak özellikle kelimelerin kullanıldığı psikoloji araştırmalarındaki çıkarımların geçerliliği için büyük önem taşımaktadır. Örneğin, bir bellek araştırması sonucunda elde edilen bulguların deneysel yönlendirmelerden kaynaklandığını doğru bir şekilde çıkarsayabilmek için kullanılan kelimelerin çağrışım özelliklerinin karıştırıcı değişken olmaması, yani deneyde kullanılan durumlar arasında kontrol edilmesi gerekmektedir. Serbest çağrışım sonucunda elde edilen normatif bulgular, ağırlıklı olarak, kelimelerin anlam özelliklerinin kontrol edilmesi amacıyla kullanıldığı gibi, bellek performansını, yalancı bellek oranını, diller arasındaki farklılıkları, çeşitli davranışları ve kişiler arasındaki davranış ve tutum farklılıklarını incelemek için de kullanılmaktadır (Deese, 1965; McEvoy, Nelson ve Komatsu, 1999; Reich ve Goldman, 2005; Vikis-Freibergs ve Freibergs, 1976).

Bu çalışmanın öncelikli amacı daha önceden toplanılmış Türkçe kelimelerin serbest çağrışım normlarına (bkz., Türkçe Kelime Normları, Tekcan ve Göz, 2005) dair güvenirlik değerlerini incelemektir. Alanyazına bakıldığında farklı dillere ait kelime normlarının güvenirlik değerlerinin ek çalışmalar ile saptandığı görünmektedir (Brotsky ve Linton, 1967; De Deyne, Navarro ve Storms, 2013; Nelson ve ark., 2000). Fakat, Tekcan ve Göz'ün (2005) norm çalışmasının üzerinden uzun bir zaman geçmesine rağmen konu ile ilgili bir çalışma mevcut değildir.

Bu araştırmada güvenirlik çalışmasına ek olarak, Türkçe alanyazına bu konuda katkı sağlamak ve mevcut serbest çağrışım normlarının geçerliliğine dair veri üretmek amaçlarıyla deneysel bir çalışma da yapılmıştır. İpucu ile hatırlama yönteminin kullanıldığı bu deneyde, var olan normatif bulguların alanyazındaki beklentiler çerçevesinde bellek performansını ne derece tahmin edip edemeyeceği, yani yordama geçerliği araştırılmıştır. Yapılan bu iki çalışma Türkçe kelime normlarının güvenirlik ve geçerliği üzerine ilk verileri üretmiştir. Yapılan çalışmalara geçmeden önce serbest çağrışım yöntemi ayrıntılı olarak tanıtılmış, bu yöntemi matematiksel olarak modelleyen bir yaklaşımdan bahsedilmiş ve son olarak, serbest çağrışımdan elde edilen normatif bulguların ipucu ile hatırlama yönteminde nasıl kullanıldığı anlatılmıştır.

Serbest Çağrışım Yöntemi

Serbest çağrışım yönteminde birçok katılımcıya bir dizi kelime (hedef) gösterilir ve onlardan her bir kelimenin akıllarına getirdiği, anlamca güçlü olarak ilişkili olan ilk kelimeyi yazmaları istenir. Standart yöntemde katılımcılar yalnızca bir kelime ile cevap verirler. Yalnızca bir kelime ile cevap verilmesi ve birçok kelime ile cevap verilmemesi tepki zincirlemesini[1] (response chaining) ve hatırlama baskılamasını[2] (retrieval inhibition) engellemek içindir (Nelson, McEvoy ve Bajo, 1988). Bu yöntem kullanılarak her hedef kelime için o kelimenin çağrıştırdığı bir kelimeler dizisi belirlenir ve bu kelimeler anlamca bağlı oldukları hedef kelimenin çağrışım setini oluştururlar (Deese, 1965; Nelson ve ark., 2004). Çağrışım seti bir kelimenin anlamca bağlı olduğu diğer kelimelerin toplam sayısından oluşur. Genel olarak en az iki ve daha çok katılımcının verdiği cevaplar çağrışım setine dâhil edilir (Nelson ve ark., 2004). Hedef kelimeye karşın yalnız bir katılımcı tarafından üretilen cevaplar çağrışım setine dâhil edilmez çünkü çağrışım setini belirlemede çok güvenilir olmadıkları ama yine de ipucu kelimesinin anlamını yansıttıkları düşünülmektedir (De Deyne ve ark., 2013). Serbest çağrışım bir kelimenin anlamca güçlü olarak ilişkili olduğu diğer kelimeleri belirlemekte kullanılır fakat zayıf olarak ilişkili olduğu diğer kelimeleri belirleyemeyebilir.

Çağrışım setinin büyüklüğüne ek olarak, kelimeler arasındaki anlam bağının gücü de serbest çağrışım ile belirlenebilir. Anlam bağının gücü ipucu kelimesine karşı verilen cevapların toplam cevap sayısına oranlanması ile bulunur. Türkçe kelimelerin normatif verilerine bakıldığında TAKSİ hedef kelimesine 27 katılımcı ŞÖFÖR, 4 katılımcı ise PARA cevabını vermiştir (Tekcan ve Göz, 2005). 100 katılımcıdan toplanılan verilere göre TAKSİ kelimesinin ŞÖFÖR ile anlam ilişkisinin gücü .27 PARA ile .04 olarak hesaplanabilir. Yalnız, ŞÖFÖR cevabı PARA cevabına göre altı kat daha olasıdır demek mümkün değildir; çünkü serbest çağrışım yönteminin ölçüm skalasının sırasal olduğu ve oransal olmadığı önemle not edilmiştir (Nelson

1 Hedef kelimesine ilk verilen cevabın üzerine ek olarak verilecek ikinci, üçüncü, vesair cevaplar birbirlerinin çağrışımından olabilmektedir.

2 Hedef kelimeye karşın hatırlanan bir kelime, anlamca bağlı olduğu diğer kelimelerin hatırlanma olasılığını düşürecektir. Bu durumda, ilk kelime dışında verilecek tepkiler hedef kelimenin çağrışım setinin üyesi olma olasılığını düşürmektedir.

ve ark., 2000). Ölçülen, iki kelime arasındaki bağın ne derece olası olduğudur. Bu ölçüm göreceli bir değer verir ve bu değeri mutlak bir değer olarak almamak gerekir.

Kelimelerin çağrışım yapılarına dair önemli bir diğer özellik anlam ilişkisinin yönüdür. Alanyazındaki bulgular çağrışımın yönünün ve gücünün değişkenlik gösterdiğini bulmuştur (Nelson, McEvoy ve Pointer, 2003). Serbest çağrışımda bir kelimenin koşullu olarak diğer kelimeyi anımsatma olasılığı *ileri gücü* (forward strength) adını alır. Bu değer bir hedef kelimesini tepki kelimesine bağlayan güç değeridir (Maki, 2008; Nelson ve ark., 2004). *Geri gücü* (backward strength) ise hedef kelimesinin çağrıştırdığı bir kelimenin hedef kelimesini çağrıştırma olasılığıdır. Örneğin, A hedef kelimesi B, C, D kelimelerini anımsatmaktadır. Peki B, C, D kelimeleri de A kelimesini anımsatır mı? Bir başka deyişle, bir kelimenin çağrışım setindeki diğer kelimeler kendisini her zaman çağrıştırır mı? Bu sorunun cevabı görgül bir çalışma, yani hedef kelimesine verilen tepkilerin ayrı örneklemlerde normlanmasını gerektirmektedir. Kelimelerin anlam ilişkilerinde yönün önemli olması bir bakıma Tversky'nin (1977) bahsetmiş olduğu benzerlik derecelendirmelerine benzer. Tversky'nin bulgularına göre KIRMIZI'nın ÇİN'e olan benzerliği, ÇİN'in KIRMIZI'ya olan benzerliğine eşit değildir. Kelimeler arasındaki benzerlik, geçişken (transitivity) bir özellik değildir. Aynı şekilde, A kelimesinin B ve C'yi anımsatması, C ve B kelimelerinin de A'yı anımsatmasını veya aynı güçte anımsatmasını gerektirmez. Kelimelerin çağrışım yapılarına dair bulgular, kelimelerin çağrışımlarının geçişken olmadığını göstermiştir (De Deyne ve ark., 2013; Nelson ve ark., 2003). Var olan Türkçe normlardan konu ile ilgili bir örnek verilebilir: KAYISI kelimesinin en güçlü anımsattığı kelime MEYVE (33%)'dir, fakat MEYVE kelimesi KAYISI'yı normatif verilere göre anımsatmamaktadır. Yani KAYISI → MEYVE ileri gücü .33'tür, fakat MEYVE → KAYISI ileri gücü normatif verilere göre ölçülemeyecek kadar zayıf çıkmıştır[3] (.00). Aynı şekilde, ÇIĞLIK kelimesinin KORKU'yu (40%) anımsatma gücü, KORKU kelimesinin ÇIĞLIK'ı (5%) anımsatma gücünden yüksektir (Tekcan ve Göz, 2005).

Serbest çağrışımda hedef kelimelere karşı verilen tepki kelimelerinin olasılıklarını modelleyen bir yaklaşım *göreceli güç hipotezi* (relative strength hypothesis) olarak adlandırılır (Maki, 2008; Nelson ve ark., 2000). Bu varsayıma göre hedef kelime yalnızca bir kelimeyi değil bir dizi kelimeyi çalışma yönergesiyle uyumlu bir şekilde aktif hale getirir ve bu kelimeler arasında en yüksek güç değerine sahip olan kelime serbest çağrışımda tepki olarak verilir. Hedef kelimenin aktif hale getirdiği kelime seti içerisinde bir kelimenin tepki olarak seçilmesi o kelimenin ortalama güç değerinin aktifleşen diğer kelimelerin güç oranlarına göre göreceli olarak daha yüksek olmasından yani diğer aktifleşen kavramlar arasındaki rekabeti kazanmasından dolayıdır (Nelson ve ark., 2000). Hedef kelimeye karşı verilen tepki, güç değerinin göstergesidir ama bunu salt bir güç değeri olarak ele almamak gerekir. Modele göre, bir hedef kelimesi ona bağlı olan kelimelerin güç dağılımlarını aktif hale getirir ve ilişkili her kelime için kendi dağılımından rastgele bir güç değeri örnekler. Kelimelere ait güç değerlerinin sıklık dağılımlarının normal olduğu varsayılır. Bazı kelimelerin ortalama güç değeri diğerlerinden daha yüksek olsa da, kelimelerin kendi güç dağılımlarından rastgele bir değer seçiliyor olmasından dolayı, ortalama olarak daha zayıf güç değerine sahip olan bir kelime de serbest çağrışımda tepki olarak verilebilir. Ayrıca, Nelson ve arkadaşları (2000) ve Maki (2008) tarafından bir tepki kriteri (0 değeri) de belirlenmiştir. Yani bir kelimenin serbest çağrışımda cevap olarak verilmesi hem diğer cevaplara göre daha güçlü olmasına hem de tepki kriterinin üzerinde olmasına bağlıdır. Verilen tepkilerdeki değişkenlik kişiler arası ve kişiler içi farklılıklardan kaynaklanmaktadır. Kişiler arası değişkenlik, kişilerin farklı hayat deneyimlerini; kişiler içi değişkenlik ise, yakında olmuş olayların birey üzerindeki etkilerini içermektedir. Bir tepkinin gücü, o tepkinin önceki deneyimlerde ne kadar örneklendiği veya kullanıldığının sayısı olarak da görülebilir. Serbest çağrışımda verilen tepkiler, kişi tarafından daha fazla deneyimlenmiş ve kullanım sıklığından ötürü bellekte daha fazla tekrarlanmış kavramlar gibi de düşünülebilir (Hintzman, 1976).

Serbest Çağrışım Normlarının İpucu ile Hatırlama Yönteminde Kullanımı

Kelimelerin çağrışım yapılarına dair bulgular çağrışım setlerinin, anlamsal ilişkinin yönünün ve gücünün değişkenlik gösterdiğini ve bu değişkenliğin ise belleği etkilediğini göstermiştir (Nelson, Fisher ve Akirmak, 2007; Nelson, Schreiber ve Xu, 1999). Normatif verilerden elde edilen kelimelerin çağrışım yapılarına dair bu bilgiler birçok bellek testindeki materyallerin seçiminde ve kontrolünde kullanıldığı gibi aynı zamanda farklı görev performanslarını (örneğin, ipucu ile hatırlama ve tanıma görevleri) tahmin etmek için de kullanılır (Nelson, Schreiber ve McEvoy, 1992; Nelson, Zhang ve McKinney, 2001). İpucu ile hatırlama bu bilgilerin sıklıkla kullanıldığı bir yöntemdir (Nelson, McKinney, Gee ve Janczura, 1998). İpucu ile hatırlama yöntemi belleğin günlük hayattaki genel kullanımını örnek alan bilişsel bir testtir. Kütüphaneden ödünç aldığımız bir kitabı iade etmeyi unuttuğumuzu kütüphane binasının önünden geçerken hatırlamak ipucu ile hatırlamaya bir örnek olarak gösterilebilir. Kütüphane binası unuttuğumuz bir şeyi hatırlatmakta ipucu görevi

3 Serbest çağrışım yönteminde eğer iki kelime arasında bağ bulunamamışsa, bu iki kelime arasında anlam ilişkisi olmadığını kesin olarak söyleyemeyiz.

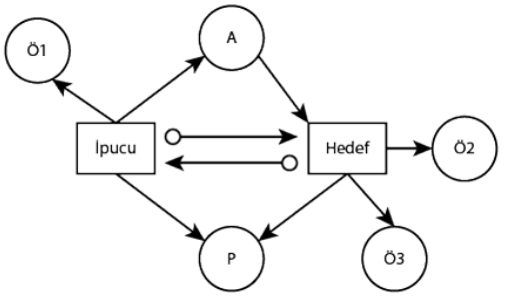
görmektedir. Deneysel yöntemde ise ilk önce katılımcılar bir kelime listesi çalışırlar. Listedeki kelimelere *hedef* denir ve katılımcılardan istenen şey, çalıştıkları hedef kelimeleri mümkün olduğunca iyi hatırlamalarıdır. Çalışma aşaması bittikten sonra test aşamasında, hedef kelimelerle anlamca ilişkili farklı bir grup kelime gösterilir. Bu kelimeler *ipuçlarıdır* ve katılımcıların ipucu kelimelerini kullanarak daha önce çalıştıkları hedef kelimeleri hatırlamaları istenir. Bu yöntemde, katılımcıların çalışma aşamasında gördüğü kelimeler, test aşamasında gösterilecek ipucu kelimeler ile serbest çağrışım verilerine dayanarak anlamca ilişkili olarak kurgulandığında, hangi hedef kelimelerin daha iyi hatırlanacağı önceden tahmin edilebilir. Örneğin; KÖPEK ipucu, KEDİ hedef kelimesini hatırlatmak için kullanılabilir. İpucu ve hedef kelimeleri birbirlerine anlamca bağlıdır ve bir ipucunun çalışma listesindeki yalnızca bir hedef kelimesi ile direkt ilişkili olmasına dikkat edilir. İpucu kelimeleri çalışma listesinde gösterilmez ve hedef kelimeler ile var olan anlam bağı daha önceden dil yolu ile edinilmiş bilgilere dayanır (Deese, 1965; Hintzman, 1976). Bu yöntem liste-dışı (extra-list) ipucu ile hatırlama olarak adlandırılır. İpuçları eğer çalışma aşamasında hedef kelime ile birlikte gösterilirse bu yöntem liste-içi (intra-list) ipucu ile hatırlama adını alır.

Hangi ipucu kelimelerinin daha iyi bellek performansı sağlayacağının tahmini serbest çağrışım yöntemiyle ölçülen kelimelerin birbirleriyle olan anlam ilişkilerine dayanır. Liste-dışı ipucu ile hatırlama yöntemi kullanılarak yapılan araştırmalar diğer değişkenler kontrol edildiğinde çağrışım seti küçük kelimelerin büyüklere göre daha iyi bir ipucu olduğunu, yani hatırlamayı olumlu etkilediğini göstermiştir (Nelson ve ark., 1999). Çağrışım setinin büyüklüğü hedef kelimenin hatırlanmasını da etkilemektedir (Nelson, McEvoy ve Schreiber, 1990) ancak kullanılan bellek testinin türü de bu etki üzerinde önemlidir. Serbest hatırlama (free recall) yönteminde çağrışım seti büyük hedef kelimeler iyi hatırlanırken, ipucu ile hatırlama yönteminde çağrışım seti küçük hedef kelimeler iyi hatırlanmaktadır (Nelson ve ark. 2007). İpucu ile hatırlama yönteminde çağrışım setinin etkisi büyük oranda hedef ve ipucu kelimeleri bir bütün olarak değerlendirildiğinde, her ikisinin de özgün çağrışımlarının az veya çok olmasına bağlıdır. Özgün çağrışımlar bir kelimenin çağrışım setinde bulunup diğerininkinde bulunmayan kelimelerdir. Dolayısıyla, hedef kelimelerin çağrışım setleri büyüdükçe özgün çağrışımlarının olasılığı da artmaktadır ki bu da ipucu ile hatırlama yönteminde hatırlama oranını azaltır[4]. Ayrıca, ipucunun direkt olarak hedef kelimeyi çağrıştırdığı ve bu çağrışımın güçlü olduğu durumlarda katılımcıların liste-dışı ipucu ile hatırlama performansları, bu çağrışımın güçsüz olduğu durumlara göre daha iyi olmaktadır (Nelson ve ark., 2007). Bu bağlamda, liste-dışı ipucu yöntemindeki bellek performansı anlamsal belleğin yapısından etkilenmektedir ve bu etki var olan serbest çağrışım normları ile tahmin ve kontrol edilebilir. Liste-dışı ipucu yöntemi anlamsal bellek organizasyonunun anlık hatırlama üzerine etkilerinin araştırılmasına imkan tanır (Nelson, Dyrdal ve Goodmon, 2005; Roediger, Jacoby ve McDermott, 1996; Tulving, 1985). Bu yöntem ile aynı zamanda işleme düzeyleri (levels of processing), bağlamın etkileri (context effects) ve çalışma süresi (study duration) gibi değişkenler yönlendirilerek (manipüle edilerek) anısal bellek araştırmaları yapmak da mümkündür.

Nelson ve arkadaşlarının, açık ve örtük temsilleri işleme modeli (processing implicit and explicit representations II - PIER II), liste-dışı ipucu ile hatırlama yöntemindeki bellek performansını kelimelerin çağrışım yapılarına dair bilgileri kullanarak matematiksel olarak modeller (Nelson, Schreiber ve McEvoy, 1992; Nelson ve Zhang, 2000). Bir kelime gördüğümüzde kelimenin anlamca ilişkili olduğu diğer tüm kelimelerin aynı anda (paralel olarak) aktif hale geldiği varsayılır (Goodmon ve Nelson, 2004; Nelson ve ark., 2007). Bu süreç otomatik olarak gerçekleşir ve kelimenin anlamını kavramamızda önemli bir rolü vardır. PIER II'nin genel varsayımı, farkında olunmadan (örtük olarak) aktifleşen kelimelere ait anlamsal bellek temsilinin anlık bellek performansını etkilediğidir (Nelson ve ark., 1998). Buna göre, bir kelimeyi gördüğümüz veya düşündüğümüz zaman yalnızca o kelime değil, anlamca ilişkili olduğu diğer tüm kelimelerin oluşturduğu yapı örtük olarak aktif hale gelmektedir. Bellek performansı kelimelerin hem açık (explicit) hem de örtük (implicit) temsillerine bağlıdır (Nelson, Bennett, Gee, Schreiber ve McKinney, 1993; Nelson ve ark., 1998). Bir kelimenin açık temsili kişinin kelimeyi hatırlamak için harcadığı kasıtlı ve dikkat gerektiren çabaların tümü sonucunda oluşur: tekrar türü, tekrar süresi, öğrenme niyeti gibi. Kelimenin çağrışım yapısı, yani anlamca bağlantılı olduğu diğer kavramlar ise o kelimenin örtük temsilini oluşturur. Kelimeler arasındaki ilişkileri içeren bu yapı dil deneyimi ile oluşur. Kişinin günlük hayatında hangi kavramları yan yana deneyimlediğine göre kavramsal yapı kişiler arasında değişkenlik göstermekle beraber anlamca ilişkili kavramların çoğunlukla aynı bağlamlarda kullanılmasından ötürü belirli bir tutarlılık da gösterir (Nelson ve ark., 2007; Stacy, 1997). Kelimelerin örtük temsilinin ne derece güçlü veya zayıf olduğu ise serbest çağrışım yöntemi ile elde edilen normatif bulgularla belirlenir. Normatif bulgular bir kelimenin bir diğeri için iyi bir

4 Serbest hatırlamada ise hedef kelimenin çağrışım setinin büyük olması, o kelimenin listedeki diğer kelimelere dolaylı olarak bağlı olma olasılığını artırmakta ve liste-içi ilişkileri güçlendirmektedir. Bu durum ise belleği olumlu etkilemektedir.

ipucu olup olmadığı hakkında bize bilgi verir. Serbest çağrışım ile toplanılan normatif veriler ipucu ile hatırlama görevinde katılımcıların bellek performansının 49%'unu tahmin etmektedir (Nelson ve Zhang, 2000). Bir başka deyişle, ipucu ile hatırlama yöntemindeki varyansın yarısı normatif verilerle, yani kelimelerin zihnimizdeki ilişkisel yapıları ile deney öncesinde tahmin edilebilir. Diğer yarısı ise, kelimenin açık temsili ve bağlam etkileri ile açıklanabilir.



Şekil 1. İpucu ile Hatırlama Deneyi Sırasında Örtük Aktifleşen Kelimelerin Birbirleriyle Olan İlişkisinin Temsili.
Not. Ö = Özgün, A = Aracı, P = Paylaşılan

İpucu ile hatırlama yöntemindeki bellek performansını tahmin edebilmek için hedef ve ipucunun birbiriyle direkt anlam bağları dışında dolaylı anlam bağlarının da incelenmesi gerekmektedir. İpucu ile hedef kelimelerinin ortak çağrıştırdığı kelimeler (paylaşılan) ve ipucunun hedef kelimeye bir başka kelime üzerinden (aracı) olan bağlantıları dolaylı bağları oluşturur (bkz., Şekil 1). Dolaylı anlam bağları her iki kelimenin örtük olarak aktif hale getirdiği kavramların incelenmesi ve bu kavramların kesişiminin analizinin yapılması ile mümkündür. PIER II kelimeler arasındaki direkt ve dolaylı ilişkileri matematiksel olarak modelleyerek ipucu ile hatırlama yöntemindeki bellek performansını yordar (Nelson ve ark., 1998). İpucundan hedefe (bir başka deyişle, ileri gücü) ve hedeften ipucuna (bir başka deyişle, geri gücü) direkt anlam bağının varlığı ve gücü bellek performansını olumlu etkiler. Bulgular, ileri gücünün bellek performansını tahmin etmede geri gücüne kıyasla önemli olduğunu göstermiştir (Nelson ve ark., 2007). Kelimeler arasındaki direkt bağların yanı sıra ipucu ve hedef kelimelerini birbirine dolaylı olarak bağlayan veya ortak olarak çağrıştırılan diğer kelimelerin sayısı ve gücü de bellek performansını iyileştirmektedir (bkz., Şekil 1, P ve A). Kelimelerin örtük olarak aktifleştirdikleri kavramlar hafızayı olumlu etkileyebildikleri gibi, aynı zamanda olumsuz etkileri de vardır. Hedef ve ipucunun çağrışım setleri düşünüldüğünde kendilerine özgün (unique associates), yani bir diğerinin çağrışım setinde bulunmayan kelimelerin çokluğu ve güç değerleri hafızayı olumsuz etkilemektedir (bkz., Şekil 1, Ö1, Ö2 ve Ö3; Nelson ve ark., 2007). Genel olarak, kelimelere ait özgün çağrışımların sayısı bir kelimenin çağrışım seti büyüdükçe fazlalaşır. Bunun sonucu olarak, hedef veya ipucu kelimelerinin çağrışım setleri büyüdükçe ipucu ile hatırlama yönteminde bellek performansı da düşer.

Alanyazındaki bu bulgu ve beklentilerden yola çıkarak Türkçe kelimelerin serbest çağrışım normları üzerine iki çalışma yapılmıştır. İlk çalışmada, var olan normlardan bir alt küme seçerek bu kelimelerle yeniden norm çalışması yapılmış ve elde edilen yeni verilerin kendi içinde ve var olan normatif veriler ile uyumuna bakılmıştır. İkinci çalışmada ise, var olan serbest çağrışım norm verilerinin bellek performansını PIER II'nin tahminleri doğrultusunda ne derece yordadığı incelenmiştir.

1. Çalışma

İlk çalışma Türkçe kelimelerin serbest çağrışım normlarının güvenirliği üzerine veri üretmeyi hedeflemiştir. Bu bağlamda Türkçe Kelime Normları kitabından rastgele seçilen kelimeler üzerine hem klasik kağıt-kalem yöntemiyle hem de internet üzerinden çevrimiçi olarak norm çalışması yürütülmüştür. Kelimelerin anlamları zaman içerisinde değişse de, daha önce farklı diller için yapılan norm çalışmaları serbest çağrışım yönteminin bir kelimenin çağrışım setinin büyüklüğünü ve bu setteki kelimelerin anımsanma olasılıklarını belirlemede geçen zaman aralığına rağmen yine de güvenilir olduğunu göstermiştir. Bu doğrultuda, var olan normların çağrışım seti ve güç değerleri bulguları ile mevcut çalışmada elde edilen bulguların uyumlu olması beklenmektedir.

Standart kağıt-kalem yöntemine ek olarak, norm çalışması internet üzerinden çevrimiçi olarak da yürütülmüştür. Amaç, çevrimiçi toplanılan verilerin klasik yöntem ile toplanılan verilerle nasıl uyuşacağını ve var olan normatif verilerle uyumunun nasıl olacağını incelemektir. İlerleyen teknoloji ile ortaya çıkan çevrimiçi veri toplama yöntemi klasik yönteme göre araştırmacılara özellikle katılımcılara erişimde kolaylık sağlamaktadır. Fakat yöntemsel olarak doğurduğu sonuçların bulgulara nasıl yansıyacağı ve bulgulardaki farklılıklar detaylı olarak araştırılmamıştır (bkz., De Deyne ve ark., 2013). Mevcut çalışmanın bir amacı da serbest çağrışım yöntemini çevrimiçi uygulamanın standart yönteme kıyasla nasıl benzerlikler ve farklılıklar üreteceğini saptamaktır.

Yöntem

Örneklem

105 İstanbul Bilgi Üniversitesi lisans öğrenci-

si (90.5% kadın) ders kredisi karşılığında kağıt-kalem yöntemiyle yapılan norm çalışmasına katılmışlardır. Katılımcıların yaş ortalaması 21.14'tir (S = 1.54). Çoğunluğu psikoloji bölümü (83.6%) olmakla birlikte çeşitli bölümlerden öğrenciler çalışmaya katılmışlardır. Aynı kelimelerin iki farklı yöntemle norm çalışması yapıldığı için farklı katılımcıları örnekleme gereği duyulmuştur. Kağıt-kalem yöntemi için üniversite ikinci ve üçüncü sınıf öğrencilerinden, çevrimiçi yöntemi için ise birinci sınıf psikoloji öğrencilerinden veri toplanılmıştır. Çevrimiçi yönteminde beş katılımcı anketi tamamlamadığı ve iki katılımcının da yarıdan fazla hedef kelimeye cevap vermediği için analizler 52 (87% kadın, $Ort._{Yaş}$ = 20.13, S = 1.13) katılımcı üzerinden hesaplanmıştır.

Veri Toplama Araçları ve İşlem

Tekcan ve Göz'ün (2005) kelime normlarından rastgele olarak seçilen kelimeler kullanılmıştır. Var olan normlardaki 600 kelime alfabetik olarak dizilmiş ve her kelimeye alfabetik sıralamasındaki numarası verilmiştir. Daha sonra 1 ve 600 arasında 100 tane rakam rastgele seçilmiş ve bu rakamlara karşılık gelen kelimeler mevcut çalışmaya dahil edilip norm çalışmasında kullanılmıştır[5]. Bu kelimelere hedef adı verilmiştir.

Seçilen 100 hedef kelimeden oluşan dört sayfalık bir kitapçık kağıt-kalem yöntemi için hazırlanmıştır. Kitapçığın ilk sayfasında demografik sorular (örn., yaş, cinsiyet, sınıf) ve serbest çağrışım yönergesi geri kalan sayfalarında ise hedef kelimeler iki sütun halinde sıralanmıştır. Hedef kelimelerin karşılarına gelecek kısım, katılımcıların cevaplarını yazabilmeleri için boşluk bırakılmıştır. Serbest çağrışım yönergesinde katılımcılardan her bir kelimenin anımsattığı, anlamca ilişkili veya güçlü bir şekilde bağlantılı olduğunu düşündükleri ilk kelimeyi hedef olarak verilen kelimelerin karşısındaki boşluğa yazmaları istenmiştir. Ayrıca, verilecek cevapların yalnızca bir kelimeden oluşması gerektiği vurgulanmıştır. Bu kritere uygun cevap vermeyenlerin çağrışımları veri analizinden çıkarılmıştır. Hedef kelimeler rastgele olarak sayfalara atanmış, kitapçık içerisindeki sayfalar rastgele bir şekilde karıştırılıp farklı kelime sıralarının olduğu kitapçıklar oluşturulmuş ve bunlar rastgele olarak katılımcılara dağıtılmıştır. Katılımcılar psikoloji laboratuvarında en fazla dörder kişilik gruplar halinde kitapçıkları doldurmuşlardır.

Normların çevrimiçi toplanılabilmesi için *Survey Monkey* adlı site üzerinden bir anket oluşturulmuştur. Anket hakkındaki bilgiler, birinci sınıflara zorunlu bir psikoloji dersinin çevrimiçi ders sayfası üzerinden duyurulmuştur. Katılımcılar, çevrimiçi anketin ilk sayfasındaki araştırma ile ilgili bilgileri de içeren bilgilendirilmiş onam formunu okuyup onay verdikten sonra ilk önce demografik bilgilerini girip sonrasında hedef kelimelere karşılık gelen çağrışımlarını hedef kelimelerin yanında bırakılan boşluğa yazarak anketi tamamlamışlardır. Çevrimiçi yöntemi mümkün olduğunca kağıt-kalem yöntemine benzetmek adına hedef kelimeler üç ayrı sayfada ve sayfa içerisinde iki sütun yer kaplayacak şekilde ekranda gösterilmiştir. Ayrıca, kelimeleri içeren toplam üç sayfanın sırası yine her katılımcı için rastgele olarak karıştırılıp sunulmuştur.

Bulgular

Kelime çağrışım verilerinin güvenirliğini incelemek için temel olarak iki değişken üzerinde analizler yapılmıştır. Kullanılan değişkenler hedef kelimelere karşılık verilen cevapların olasılık değerleri (çağrışım gücü) ve sayısıdır (çağrışım seti büyüklüğü). Tekcan ve Göz'ün (2005) normatif verileri (TG) ile kağıt-kalem (KK) ve çevrimiçi (Ç) yöntemleri ile yeni toplanılan veriler belirtilen değişkenler üzerinden karşılaştırılmıştır. Amaç, bu üç farklı veri setindeki benzerlik ve farklılıkları incelemektir. Bu değişkenlere ek olarak, De Deyne ve arkadaşlarının (2013) önerisi ile veri setlerinin heterojenliği de incelenmiştir. Heterojenlik, bir hedef kelimenin ürettiği tepki kelimelerinin çeşitliliğidir. Bu çeşitliliğin farklı veri setleri için benzer olması, uyumluluğun bir göstergesidir.

Hedef kelimenin çağrıştırdığı kelimeler ile olan ilişkisinin gücü o kelimeye verilen cevapların sıklığının toplanılan cevap sayısına bölünmesi ile yani olasılık değerlerinin hesaplanması ile bulunmuştur. Örneğin, KK yönteminde SOFRA hedef kelimesine 105 katılımcı bir çağrışımını yazmıştır. En güçlü ilişkinin 46 katılımcının kullandığı YEMEK kelimesi ile olduğu gözlemlenmiştir. Bu durumda, SOFRA'nın YEMEK kelimesini çağrıştırma gücü (veya olasılık değeri) .44 olarak (46/105) hesaplanabilir. Her bir hedef kelimenin çağrıştırdığı kelimeler ile olan güç ilişkisi bu şekilde hesaplanmıştır. Bazı katılımcıların bazı hedef kelimelere cevap vermemelerinden veya serbest çağrışım yönergesine uymamalarından kaynaklı veri kayıpları oluşmuştur ve bu sebeple toplam katılımcı sayısı hedef kelimelerde az da olsa farklılık göstermiştir. Çağrışım gücü TG veri setinde zaten hesaplandığı için yapılan analizlerde Türkçe Kelime Normları kitabında raporlanmış değerler kullanılmıştır.

Öncelikli olarak hedef kelimelere katılımcıların verdiği cevaplar en güçlü çağrışımın birinci sırayı alacağı şekilde çağrışım gücü değerlerine göre sıralanmıştır. Diğer bir deyişle, deneklerin hedef kelimelere vermiş oldukları cevaplar hedef kelimeler üzerinde gruplanmıştır. Sıralama yaparken cevapların güç değerleri göz önüne alınmış fakat cevabın aynı olup olmaması göz

5 Norm verilerine https://osf.io/mjp9z/ linki üzerinden ulaşılabilir.

Tablo 1. Her Veri Seti için En Güçlü Dört Çağrışımın Ortalama Olasılık Değerleri ve Aynı Sırada Verilen Cevapların Çağrışım Güçlerinin Korelasyonları

	Veri Seti											
	TG				KK				Ç			
	Sıra				Sıra				Sıra			
	1	2	3	4	1	2	3	4	1	2	3	4
Ort.	.29	.13	.08	.06	.25	.12	.08	.05	.26	.12	.07	.05
S	.14	.06	.03	.02	.14	.05	.03	.02	.14	.05	.03	.02
r^{TG}	—	—	—	—								
r^{KK}	.72	.46	.26	.47	—	—	—	—				
$r^{Ç}$	.70	.37	.26	.34	.89	.47	.51	.51	—	—	—	—

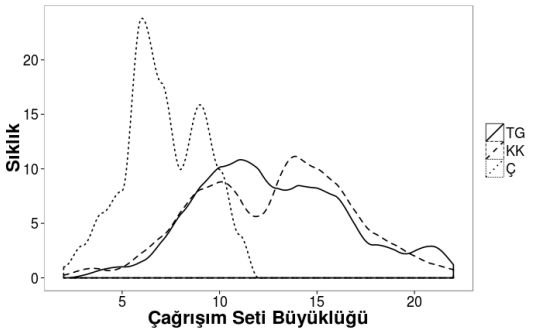
Not. Ort. = Ortalama, S = Standart sapma, TG = Tekcan ve Göz (2005), KK = Kağıt-kalem, Ç = Çevrimiçi, r = Pearson korelasyon katsayısı.

önüne alınmamıştır. Sıra, bir hedef kelimesine karşı verilen cevabın kaç katılımcı tarafından verildiğini yani sıklığını tanımlar. Bu sıralamanın sonucunda hedef kelimelere verilen ilk cevapların gücünün yüksek olduğu ve sistematik olarak güç değerlerinin ikinci, üçüncü ve dördüncü sırada verilen cevaplar için azaldığı gözlemlenmiştir (bkz. Tablo 1). Bu eğilimin üç veri seti için de söz konusu olup olmadığını saptamak için veri setlerinin kelimeler arası değiştiği ve sıranın kelimeler içerisinde değiştiği 3 (Veri setleri: TG, KK, Ç) x 4 (Sıra: bir, iki, üç, dört) karışık model faktörlü varyans analizi (mixed-model factorial ANOVA) kullanılmıştır. Aynı hedef kelimeler farklı katılımcılar tarafından cevaplandığı için, veri setleri katılımcılar arası bir değişkendir. Çağrışım setindeki kelimelerin güç sırası her bir hedef kelime içerisinde değişmektedir. Sıra değişkeni, hedef kelimelere verilen çağrışımların olasılıklarının en yüksekten düşüğe doğru dizilmiş halidir ve hangi cevapların daha fazla katılımcı tarafından verildiğini gösterir. Bu analizin sonuçları incelendiğinde, sıranın olasılık değerleri üzerine etkisi olduğu bulunmuştur[6], F (1.27, 375.83) = 417.33, p < .001, MSe = .006, η_p^2 = .58. Farkın nereden kaynaklandığını anlamak için yapılan LSD testinin sonucuna göre (.021) tüm karşılaştırmalar istatistiksel olarak anlamlı çıkmıştır. Veri setleri karşılaştırıldığında da istatistiksel olarak anlamlı bir fark bulunmuştur, F (2, 297) = 4.065, p = .018, MSe = .005, η_p^2 = .03. LSD test sonucuna göre (.01) TG verileri (*Ort.* = .143, *S* = .07), KK verilerine (*Ort.* = .128, *S* = .07) göre daha yüksek güç değerine sahip olurken Ç verileri (*Ort.* = .133, *S* = .07) diğer iki yöntemden farklı çıkmamıştır. Etki büyüklüğünün oldukça düşük (.03) olmasından dolayı veri setlerinin ortalama güç değerleri arasındaki fark önemsiz olarak yorumlanabilir. Sıralamanın etkisinin veri setlerine göre değişmediğini veri setleri ve sıralama arasındaki etkileşimin istatistiksel olarak anlamlı çıkmaması doğrulamıştır, F (2.53, 375.80) = 1.25, p > .05. Üç veri setinde de hedef kelimelere verilen cevapların olasılıklarının farklı sıralamalar için benzer olduğu gözlenmiştir.

Varyans analizine ek olarak eşit sıralarda verilen cevapların güç değerleri üç farklı veri seti için karşılaştırılmıştır. Bir hedef kelimesine tepki olarak üretilen aynı sıradaki cevapların farklı veri setlerindeki uyumuna Pearson korelasyon katsayısının hesaplanması yolu ile bakılmıştır. Bu karşılaştırma yalnızca en güçlü dört çağrışım üzerinden yapılmıştır. Tablo 1'in alt tarafında farklı veri setlerinin birbirleriyle olan korelasyon katsayıları her bir sıralama için görülebilir. Bulgular incelendiğinde korelasyon değerlerinin en güçlü çağrışımlar için oldukça yüksek olduğu, fakat diğer üretilen çağrışımlar için orta ve düşük seviyede olduğu saptanmıştır. En güçlü (birinci sıradaki) çağrışımların olasılık değerleri ele alındığında, var olan normlar ile kağıt-kalem yöntemiyle yeni toplanılan normlar arasındaki ilişkinin güçlü olduğu (r = .72) ve çevrimiçi toplanılan verilerin var olan verilerle yine aynı derecede uyumlu olduğu (r = .70) gözlenmiştir. En güçlü ilişkinin ise aynı dönemde toplanılan veri setlerinin en güçlü çağrışımları arasında, yani KK ve Ç arasında olduğu (r = .89) bulunmuştur.

Çağrışım gücü dışında çağrışım seti değerlerinin farklı veri setlerindeki uyumuna da bakılmıştır. Çağrışım setlerinin sıklığına dair dağılım Şekil 2'de görülebilir. Bulgulara göre, TG ve KK verileri normale yakın bir dağılım gösterirken çevrimiçi toplanılan verilerin pozitif kayış eğilimi vardır. Bu gözlemi dağılımın normalliğini test için yapılan Shapiro-Wilk testi doğrulamıştır, TG ve

6 Yapılan analiz sonucunda elde edilen Mauchly'nin küresellik testi sonucunun istatistiksel olarak anlamlı çıkmasından ötürü serbestlik dereceleri (*df*) "Greenhouse-Geisser" düzeltmesi ile rapor edilmiştir.



Şekil 2. Çağrışım Seti Büyüklüğünün Veri Setleri için Sıklık Dağılımı.
Not. TG = Tekcan ve Göz (2005), KK = Kağıt-kalem, Ç = Çevrimiçi.

KK için $p > .05$; Ç için $W = .96$ $p = .007$. Çağrışım seti büyüklüğünün veri setlerinde istatistiksel olarak farklılık gösterip göstermediğine kişilerarası tek yönlü varyans analizi ile bakılmıştır. ANOVA testinin normallik varsayımı sağlanamadığı için salt çağrışım seti değerlerinin logaritmik değerlere dönüştürülmesi uygun görülmüştür. Bu prosedür ile normallik varsayımı sağlanmıştır ve rapor edilen değerler logaritmik dönüşümden elde edilen değerler olmakla birlikte çağrışım seti ortalamaları logaritmik dönüşüm yapılmadan rapor edilmiştir. Varyans testinin sonucuna göre veri setleri arasında istatistiksel olarak anlamlı bir fark vardır, F (2, 297) = 101.67, $p < .001$, $MSe = .021$, $\eta_p^2 = .41$. Bu fark, çevrimiçi (*Ort.* = 7.13*, S* = 3.83) yöntemiyle elde edilen çağrışım seti değerlerinin kağıt-kalem (*Ort.* = 12.80, *S* = 3.90) ve var olan normlardan (*Ort.* = 12.76, *S* = 3.83) daha az olmasından kaynaklanmaktadır (*LSD* = .94). Yani, TG ve KK yöntemi için çağrışım setleri benzer bulunmuş, ancak çağrışım seti değerleri çevrimiçi toplanılan verilerde daha düşük gözlemlenmiştir. Çağrışım seti büyüklüğüne ek olarak veri setleri hedef kelimelere yalnız bir kişinin ürettiği ve genel toplam cevaplar üzerinden de ayrı tek yönlü varyans testleri ile karşılaştırılmıştır. Verilerin varyans analizinin normallik varsayımını sağlayabilmesi için tek cevap değerlerinin kare kökü, toplam cevapların ise logaritması alınmıştır. Rapor edilen ANOVA bulguları dönüşüm değerleri olmakla beraber, ortalamalar dönüşüm olmaksızın rapor edilmiştir. Hedef kelimelere yalnız bir kişinin verdiği cevapların sayısı tek yönlü varyans analizinin sonucunda veri setleri arasında istatistiksel olarak farklı çıkmıştır, F (2, 297) = 39.13, $p < .001$, $MSe = .67$, $\eta_p^2 = .21$. Buna göre, kağıt-kalem yönteminde (*Ort.* = 24.92*, S* = 7.92), çevrimiçi (*Ort.* = 16.95, *S* = 5.71) ve var olan normlardan (*Ort.* = 17.22, *S* = 7.19) daha fazla sayıda tek cevap verilmiştir (LSD = 1.96). Üretilen tüm cevaplar veri setleri için karşılaştırıldığında da benzer bir fark görülmüştür, F (2, 297) = 42.73, $p < .001$, $MSe = .023$, $\eta_p^2 = .22$. Genel olarak kağıt-kalem yönteminde (*Ort.* = 37.72*, S* = 9.94), çevrimiçi (*Ort.* = 24.08, *S* = 6.46) ve var olan normlara (*Ort.* = 29.95, *S* = 8.66) göre daha fazla cevap üretilmiş, en az sayıda cevap ise çevrimiçi yönteminde üretilmiştir (*LSD* = 2.37). Çevrimiçi yöntemindeki bu farklar kısmen verilerin diğer yöntemlere göre daha küçük bir örneklemden elde edilmiş olmasına bağlanabilir.

Son olarak, veri setlerinin heterojenliğine bakılmıştır. Belli bir veri grubu içerisinde herhangi bir çağrışım hedefine karşılık verilen yanıtların oluşturduğu kelime dağılımı ve sıklık açısından bu kelimelerin birbirine oranları, o çağrışım hedefinin heterojenlik yapısını tanımlar. Heterojenlik değerleri alternatif bir güvenirlik ölçütü olarak da kullanılabilir (De Deyne ve ark., 2013). Bir çağrışım setinin heterojenlik değeri (*H*), enformasyon teorisine ait olan Shannon entropi formülüne göre hesaplanır (Gray, 2011):

$$H = -\sum_{i=1}^{n} p_i log2\,(p_i)$$

Bu formülde *n,* yanıt olarak verilen farklı kelimelerin toplam sayısını; *p*, her bir kelimenin yanıtların toplamına olan oranını ifade eder. *H*'nin değeri, *n* ile logaritmik bir şekilde büyüme eğilimindedir; azami değeri de toplam yanıt sayısının logaritmasına eşittir. Hedef kelimelere aynı çağrışımlar ile cevap verildiği durumda heterojenlik 0 değerini, hedef kelimelere verilen cevaplar arasında hiçbir ortaklık olmadığı durumda ise heterojenlik maksimum değerini almaktadır. Tablo 2'de veri setlerinin heterojenlik değerleri ve her bir veri grubu için *n* ile *H* arasındaki korelasyon katsayısı belirtilmiş (bkz. n değerlerinin altında), ayrıca farklı gruplar arasındaki *H* değerlerinin korelasyonları da verilmiştir. Güvenirliğe denk bir ölçüt olarak kullanıldığında, gruplar arasındaki heterojenliklerin pozitif korelasyonu, farklı gruplardaki çağrışım setlerinin benzer şekilde heterojenlik gösterdiğine işaret eder. Gruplar arası korelasyonların hepsi pozitif ve $p < .001$ seviyesinde anlamlı bulunmuştur. Bu bulgulara göre, farklı zamanlarda, farklı örneklemlerden ve farklı yöntemlerle toplanılan serbest çağrışım normlarında katılımcılar hedef kelimelere benzer çeşitlilikte cevap vermişlerdir. Ayrıca, bulgular, serbest çağrışım yönteminin kelimelerin çağrışım yapılarına dair bilgileri tutarlı bir şekilde ölçtüğü yönündedir.

Veri setlerinin heterojenlik açısından farklılıkları da vardır. Heterojenlik üst sınırı örneklem büyük-

Tablo 2. Veri Setlerinin Korelasyon ve Heterojenlik Değerleri

	Veri Setleri				Heterojenlik		
	TG	KK	Ç	*n*	Üst Sınır	*Ort.*	*S*
TG	-			.89	6.64	3.79	.75
KK	.78	-		.92	6.70	4.22	.76
Ç	.73	.88	-	.95	5.70	3.83	.73

Not 1. Tablonun sol tarafında korelasyon değerleri, sağ tarafında ise betimsel istatistikler verilmiştir.
Not 2. TG = Tekcan ve Göz (2005), KK = Kağıt-kalem, Ç = Çevrimiçi, n = Örneklem büyüklüğü, H = Heterojenlik, Ort. = Ortalama, S = Standart sapma.

lüğünden etkilenmektedir ve veri setlerinin örneklem büyüklüğünün farklı olmasından dolayı analizlerde de farklılık göstermiştir. Salt heterojenlik değerlerini veri setleri için karşılaştırmanın bu farklılıktan dolayı anlamlı olmayacağından, her bir veri setinin kendi üst sınırına göre ortalanmış değerleri hesaplanmıştır. Üst sınıra göre ortalanmış heterojenlik değerleri karşılaştırıldığında heterojenliğin en fazla çevrimiçi verilerde (.67) olduğu, bunu kağıt-kalem yönteminin (.63) izlediği ve heterojenliğin var olan normlarda (.57) en düşük olduğu gözlenmiştir.

2. Çalışma

Türkçe Kelimelerin Çağrışım normlarında i) hedef kelimenin anımsattığı diğer kelimeler, ii) bu kelimelerin sayısı, yani hedef kelimenin çağrışım seti büyüklüğü, ve iii) hedef ve anımsattığı kelimeler arasındaki olasılık değerleri hakkında ayrıntılı bilgi verilmiştir (Tekcan ve Göz, 2005). Yapılan birinci çalışma ile serbest çağrışım yöntemi ile elde edilen var olan normatif bulguların güncelliğini koruduğu ve güvenilir olduğu sonucuna varılmıştır. İkinci çalışmanın yapılma amacı ise liste-dışı ipucu yönteminde Türkçe kelimelere dair normatif verileri kullanarak bellek performansının ne derece tahmin edilip edilemeyeceğini incelemek, böylece normların yordama geçerliği üzerine bulgular üretmektir.

Çağrışım seti büyüklüğü ve çağrışım olasılık değerleri bildiğimiz kadarıyla daha önce Türkçe materyallerin kullanıldığı hiçbir çalışmada özel olarak incelenmemiştir. Bu değişkenlerin incelenmesi sayesinde mevcut normların bellek performansını PIER II modelinin tahminleri doğrultusunda ne derece tahmin edip edemeyeceği araştırılmış ve normların yordama geçerliliği üzerine bulgular ikinci çalışma ile üretilmiştir. Ayrıca, elde edilen veriler sayesinde PIER II modelinin kültürler arası geçerliliği hakkında da bilgi sahibi olunmuştur. İkinci çalışmada çağrışım seti büyüklüğü ve çağrışım gücü değişkenleri liste-dışı ipucu ile hatırlama yönteminde yönlendirilerek katılımcıların hatırladıkları kelime sayısı ölçülmüştür. Temel olarak, i) ipucu kelimelerinin çağrıştırdığı diğer kelimelerin sayısının az veya çok olmasının ve ii) ipucu-hedef kelimelerinin anlam ilişkisinin zayıf veya güçlü olmasının bellek performansını nasıl etkileyeceği araştırılmıştır.

Türkçe kelime normlarında 600 kelime ile ilgili çağrışım bulguları bulunmakla birlikte bu kelimelere verilen tepki kelimelerinin çağrışımları üzerine bulgular oldukça kısıtlıdır. Tepki kelimeleri üzerine normatif verilerin eksik veya hiç olmamasından dolayı bu kelimelerin çağrışım yapısı bilinmemektedir. Yani bir kelimenin hangi diğer kavramları anımsattığı bilinmekle beraber bu kavramların ipucu kelimesini geri anımsatıp anımsatmadığı bilinmemektedir. Bu da seçilecek bir kelime ikilisinde ya ileri gücünün ya da geri gücünün bilinebileceği anlamına gelmektedir. Bir başka deyişle, normatif çağrışım verisi bulunan kelimeler ya ipucu olarak alınıp ileri gücü değerleri deneyde kullanılabilir ya da hedef olarak alınıp geri gücü değerleri kullanılabilir. Ancak, her iki değeri de bilebilmek her iki kelimeye ait normatif veri mevcut değilse mümkün değildir. PIER II modelinde ipucunun hedef kelimeyi anımsatma olasılığı (ileri gücü) bellek performansını hedef kelimenin ipucunu anımsatma olasılığından (geri gücü) daha fazla etkilediği için (Nelson ve ark., 2007) ikinci çalışmada ipuçlarının normatif bulguları var olan kelimelerden, hedef kelimelerin ise bu kelimelerin çağrışım setlerinden seçilmesi uygun görülmüştür[7]. Bu doğrultuda Türkçe kelime normlarından ipucu kelimeleri seçilmiş ve sonrasında her bir ipucunun güçlü ve zayıf olarak anlamca ilişkili olduğu iki hedef kelime seçilmiştir. Beklenti, ipucu ve hedef arasındaki anlam ilişkisinin güçlü olduğu durumların zayıf olmasına göre daha iyi bir bellek performansı yaratmasıdır.

PIER II'ye göre çağrışım setinin bellek performansı üzerine etkisini kesin tahmin edebilmek için ipucu

7 Eğer normlanan kelimeler hedef olarak seçilmiş olsaydı, hedef ve ipucu arasındaki geri gücü değeri bilinecek, fakat ileri gücü yani ipucunun hedefi ne derece anımsattığı bilinemeyecekti.

ve hedef kelimeye özgü kelimelerin sayısının ve güç değerlerinin bilinmesi gerekmektedir. Her ne kadar bu bilgiye sahip olunmasa da ipucu kelimelerinin çağrışım seti büyüklüğünün bellek performansına olası etkisi bu çalışmada araştırılmıştır. Diğer değişkenler kontrol edildiğinde ipucu çağrışım seti bellek performansını etkilemektedir. Son olarak, iki değişken arasında önceki çalışmalarda etkileşim bulunmadığı için bu çalışmada da beklenmemektedir.

Yöntem

Deney Deseni ve Örneklem

İstanbul Bilgi Üniversitesi'nde lisans öğrencisi olan 64 kişi (% 95 kadın) ders kredisi karşılığında çalışmaya katılmışlardır. Katılımcıların yaş ortalaması 21.5'tir (*S* = 2.00; ranj = 19-30).

Çağrışım setinin (büyük veya küçük) denekler içi, çağrışım gücünün (yüksek veya zayıf) denekler arası manipüle edildiği 2 x 2 karışık faktörlü deney deseni kullanılmıştır. Bulguların seçilen kelimelerden bağımsız olduğunu gösterebilmek adına denekler arası değişken için aynı özelliklere sahip ikişer kelime listesi hazırlanmıştır. Hazırlanan dört kelime listesi için 16'şar katılımcıdan veri toplanılmıştır. Her listedeki ipuçlarının yarısı küçük, yarısı ise büyük çağrışım setine sahiptir. Test aşamasında kullanılan ipuçları yüksek ve zayıf çağrışım gücü grupları için aynı tutulmuş fakat seçilen, doğru cevap olan, hedef kelimeler değiştirilmiştir. Yani, aynı ipuçları kullanılmış fakat serbest çağrışım normlarından elde edilen olasılık değerleri göz önüne alınarak anlamca daha güçlü veya daha zayıf ilişkili hedef kelimeler denekler arasında farklılık göstermiştir.

Veri Toplama Araçları ve İşlem

Liste-dışı hatırlama yöntemini uygulamak için ipucu ve hedef olarak tanımlanan iki grup kelime Tekcan ve Göz'ün (2005) normatif veri tabanından seçilmiştir. Hedef kelimelerin birçoğu için çağrışım ve diğer norm bilgileri eksik olduğundan veya bulunmadığından dolayı rapor edilememişlerdir. Kontrol edilen ve yönlendirilen değişkenlerin tümü ipuçlarına aittir. Kelime listelerinin hazırlanmasında ipucu kelimeleri ile başlanmıştır. İpucu kelimeleri seçilirken somutluk, imgelem ve kelime sıklığı gibi özelliklerin mümkün olduğunca benzer olmasına, çağrışım seti değerlerinin ise değişkenlik göstermesine dikkat edilmiştir. Seçilen 48 ipucundan 24 tanesi (ipucu çağrışım seti küçük) ortalama 5.67 (*S* = 1.37) farklı kelimeyi anımsatmakta diğer 24 tanesi (ipucu çağrışım seti büyük) ise ortalama 16.67 (*S* = 1.01) farklı kelimeyi anımsatmaktadır. Somutluk (*Ort.* = 5.43, *S* = 1.47; *Ort.* = 4.99, *S* = 1.65), imgelem (*Ort.* = 5.05, *S* = 1.24; *Ort.* = 4.71, *S* = 1.31) ve kelime sıklığı (*Ort.* = 112.88, S = 252.70; *Ort.* = 131.67, *S* = 109.19) değerleri küçük ve büyük çağrışım setine sahip ipuçları için denk tutulmaya çalışılmıştır.

İpucu ile hedef kelime arasındaki anlam ilişkisinin gücü aynı ipucu için normatif veri tabanından seçilen farklı hedef kelimeler ile denekler arasında yönlendirilmiştir. Katılımcılar ipucu – hedef arasındaki ilişkinin güçlü veya zayıf olduğu durumlardan yalnızca birinde yer almıştır. Anlam ilişkisi güçlü olduğu durumda ipucunun hedef kelimeyi anımsatma olasılığı hem küçük çağrışım setine (*Ort.* = .25, *S* = .18) hem de büyük çağrışım setine (*Ort.* = .24, *S* = .06) sahip ipuçları için yüksek tutulmuştur. Anlam ilişkisi zayıf olduğu durumda ipucunun hedef kelimeyi anımsatma olasılığı, hem küçük çağrışım setine (*Ort.* = .02, *S* = .01) hem de büyük çağrışım setine (*Ort.* = .03, *S* = .02) sahip ipuçları için düşük tutulmuştur.

Deney "E-Prime 2.0 Professional" programı ile hazırlanıp bilgisayar ile katılımcılara uygulanmıştır. Tüm katılımcılar laboratuvar ortamında bireysel olarak test edilmişlerdir. Deney ile ilgili etik onay İstanbul Bilgi Üniversitesi etik kurulundan alınmıştır. Katılımcılar çalışmanın içeriğini bilgilendirilmiş onam formundan okuyup onay verdikten sonra deneye katılmışlardır.

Liste-dışı hatırlama yönteminde katılımcılardan gösterilecek bir kelime listesini daha sonraki bir bellek testi için çalışmaları istenmiş ancak testin doğası ve içeriği bu aşamada anlatılmamıştır. Yönergede, ekranda gösterilecek kelimelerin belirdikleri zaman sesli olarak okunması ve mümkün olduğunca fazla kelimeyi hatırlamaya çalışmaları istenmiştir. Bu aşamada eğer bir soruları yok ise yönergeyi anladıklarından emin olmak ve deneysel prosedürün hızına alıştırmak için yalnızca insan isimlerden oluşan kısa bir deneme listesi gösterilmiştir. İsimleri hatırlamaları gerekmemektedir ve bu kısım analizlerde kullanılmamıştır. Deneme aşaması bittikten sonra hedef kelimelerin çalışıldığı aşamaya geçilmiştir. Çalışma aşamasında 24 hedef kelime teker teker üç saniye aralıklarla ekranda gösterilmiş ve her katılımcı için kelimelerin sırası karıştırılmıştır. Kelimeler büyük harflerle yazılmış ve ekranın ortasında belirmişlerdir. Hedef kelimelerin tümü gösterildikten sonra, test yönergesi okunmuştur. Bu yönergede katılımcıların aynı şekilde bir başka kelime listesi göreceği ve gösterilecek ipucu kelimelerinin yardımıyla az önce görmüş oldukları hedef kelimeleri hatırlamaları gerektiği anlatılmıştır. İpucu kelimesi ekranda belirdiği zaman onu sesli olarak okumaları ve onunla anlamca ilişkili olduğunu düşündükleri ve çalışma listesinde gördükleri bir kelimeyi söylemeleri istenmiştir. İpuçları hedef kelimeler ile anlamca ilişkilidir ve çalışma listesinde gösterilmemişlerdir. Her bir hedef kelimesine karşılık anlamca ilişkili bir ipucu test aşamasında gösterilmiştir ve çalışma aşamasında olduğu gibi

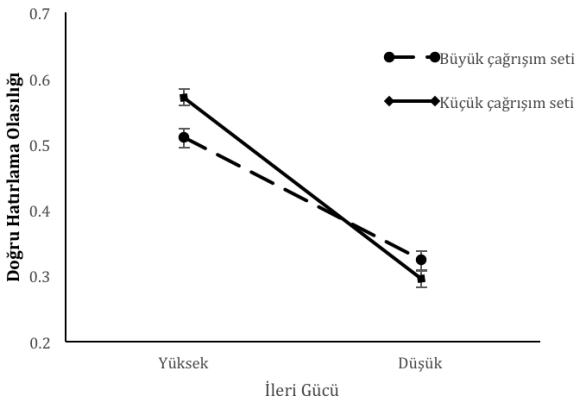
ipuçları da her katılımcı için farklı bir sırada gösterilmiştir. Hazırlanan kelime listelerinde bir ipucu kelimesinin anlamca yalnızca bir hedef kelimesi ile direkt ilişkili olmasına dikkat edilmiş ve bu bilgi katılımcılara da verilmiştir. Hedef kelimeyi hatırlayamadıkları durumlarda tahmin edebilecekleri de söylenmiştir. Test aşamasında zaman sınırlaması olmamış katılımcılar gösterilen ipucuna bir cevap verdikten sonra diğer ipucuna geçilmiştir. Cevaplar deneyi yapan asistan tarafından kayıt edilmiştir. Katılımcılar tüm ipuçlarına cevap verdikten sonra demografik bilgiler (yaş, cinsiyet, bölüm, sınıf) toplanmış ve sonrasında deney ile ilgili bilgilendirme yapılıp deney tamamlanmıştır.

Bulgular

İleri gücünün ve ipucu çağrışım seti büyüklüğünün bellek performansına etkisi Şekil 3'te görülebilir. Bulgular 2 (çağrışım seti büyüklüğü) x 2 (çağrışım gücü)'lik karışık model varyans analizi ile incelenmiştir. Doğru hatırlama olasılıkları kelime listelerinin ortalaması üzerinden hesaplanmıştır. Analiz sonuçlarına göre, hedef kelimeleri doğru hatırlama olasılığı ileri gücü yüksek olduğu durumda artmakta, düşük olduğunda ise azalmaktadır, $F(1, 62) = 48.22$, $p < .001$, $MSe = .036$, $\eta_p^2 = .44$. İpucu çağrışım setinin bellek performansına etkisi istatistiksel olarak anlamlı çıkmamıştır, ($F < 1$). Yani ipuçlarının daha az veya daha çok kelime çağrıştırması bellek performansını etkilememiştir. İleri gücü ve Çağrışım seti arasındaki etkileşim istatistiksel olarak anlamlı çıkmamıştır, $F(1, 61) = 3.74$, $p = .058$, $MSe = .018$. Ancak küçük çağrışım setine sahip ipuçları hedef kelimelere anlamca daha güçlü bağlantılı olduğu durumda (.57) büyük çağrışım setine sahip ipuçlarına (.51) kıyasla daha iyi hatırlanma eğilimi göstermiş, fakat bu avantaj, hedef kelimelere anlamca daha düşük güçte bağlantılı oldukları durumda (küçük ve büyük çağrışım setine sahip ipuçları için sırasıyla .29 ve .32) kaybolmuştur.

Varyans analizi bulgularında ipucu çağrışım seti büyüklüğünün etkisinin çıkmaması ve etkileşime dair bir eğilim çıkması alanyazındaki bulgulara dayalı öngörülerle çelişmiştir. Bu farklılığın nereden kaynaklandığını incelemek için ek analizler yapılmıştır. Elde edilen bulguların deney kapsamında kontrol edilen değişkenlerden birinin etkisi sonucunda olmuş olabileceği ihtimalini göz önüne alarak bir madde analizi yapılmıştır. Bu çerçevede her ipucu – hedef ikilisi için doğru hatırlama olasılıkları, yani hedef kelimesinin verilen ipucuna karşı hatırlama olasılığı hesaplanmış ve ipucu kelimelerinin diğer değerleri ile birlikte çoklu regresyon analizine sokulmuştur. İpuçlarına dair somutluk, imgelem, kelime sıklığı, çağrışım seti büyüklüğü ve ileri gücü değerlerinin hedef kelimelerin doğru hatırlanma olasılıklarını ne ölçüde yordadığı Tablo 3'ten görülebilir. Kelime düzeyinde yapılan bu analiz sonucunda, doğru hatırlama olasılığını yordayan tek değişkenin ileri gücü ($\beta = .45$, $p < .001$) olduğu saptanmıştır. Yani, deney içerisinde kontrol edilen değişkenlerin sonuçlara etkisi görülmemektedir.

Tablo 3'teki bir diğer sonuç ise, önceki varyans analizi bulguları ile paralel olarak çağrışım seti büyüklüğünün bellek üzerine etkisinin olmamasıdır. Alanyazındaki bulgularla çelişen bu durumun, kelime listeleri arasındaki farklılıklardan, yani seçilen kelimelerden, kaynaklanabileceğini düşünerek, tüm listeleri tek bir analize tabi tutmak yerine benzer durumları ölçen listeler iki ayrı varyans analizi ile incelenmiştir. Bir başka de-



Şekil 3. Doğru Hatırlama Olasılığının Çağrışım Seti Büyüklüğü ve İleri Gücü ile İlişkisi.
Not. Hata çubukları ortalamanın standart hatasını göstermektedir.

Tablo 3. Kelime Özelliklerinin Doğru Hatırlama Olasılığını Yordadığı Çoklu Regresyon Analizi Bulguları

	Doğru Hatırlama Olasılığı		
Değişken	*B*	*SH*	*β*
İmgelem	.03	.04	.15
Somutluk	-.02	.03	-.14
Kelime Sıklığı	.00	.00	-.15
Çağrışım Seti	.00	.00	-.00
İleri Gücü	.80	.17	.45*
Düzeltilmiş R^2		.19	
F		5.47*	
sd		5.90	

Not 1. $*p < .001$.
Not 2. B = Standardlaştırılmamış regresyon katsayısı, β = Standardlaştırılmış regresyon katsayısı, SH = Standart hata, sd = Serbestlik dereceleri.

yişle, kelime listelerinden birinci ve ikinci listelerin ileri gücü yüksek, üçüncü ve dördüncü listelerin ise ileri gücü düşüktür. Ancak deney içerisinde kontrol edilemeyen değişkenlerden dolayı (örneğin, hedef kelimelerin çağrışım özellikleri) listeler arasında farklı sonuçlar elde edilmiş olma ihtimali vardır. Aynı ipucuna ve farklı ileri gücüne sahip olan listeler gruplanmış (1 ile 3; 2 ile 4) ve karışık model varyans analizi uygulanmıştır. Birinci ve üçüncü listenin varyans analiz sonuçlarına göre çağrışım seti küçük ipuçları (.43), büyük ipuçlarına (.34) göre daha iyi bellek performansı vermişlerdir, F (1, 30) = 10.62, p = .003, MSe = .013, η_p^2 = .26. İleri gücü yüksek ipuçları (.47), ileri gücü düşük ipuçlarına (.29) göre daha iyi bellek performansıyla sonuçlanmıştır, F (1, 30) = 14.43, p = .001, MSe = .034, η_p^2 = .33. İki değişken arasındaki etkileşim ise istatistiksel olarak anlamlı çıkmamıştır, (F < 1). Bulgular alanyazındaki bulgularla eşleşmektedir. Buna karşın, ikinci ve dördüncü listelerin varyans analiz sonuçlarına göre çağrışım seti küçük ipuçları (.44), büyük ipuçlarına (.50) göre daha kötü hatırlanmakla beraber bu fark istatistiksel olarak anlamlı değildir, F(1, 30) = 3.42, p = .074, MSe = .017. İleri gücü yüksek ipuçları (.61), ileri gücü düşük ipuçlarına (.32) göre daha iyi hatırlanmıştır, F(1, 30) = 46.49, p < .001, MSe = .029, η_p^2 = .61. İki değişken arasındaki etkileşim ise oldukça güçlü bir eğilim göstermiştir, F(1, 30) = 4.04, p = .053, MSe = .017. Bu eğilime göre büyük çağrışım setine sahip ipuçları (.39), küçüklere (.26) göre düşük ileri gücü durumunda daha iyi hatırlanmış, fakat yüksek ileri gücü durumunda bu fark kaybolmuştur (büyük ve küçük çağrışım setine sahip ipuçları için sırasıyla .61 ve .62). Küçük çağrışım setine sahip ipuçlarının doğru hatırlanması farklı listeler arasında benzer olmakla birlikte (liste 1 ve 3 ile liste 2 ve 4 için sırasıyla .43 ve .44), büyük çağrışım setine sahip ipuçlarının doğru hatırlanması listeler arasında büyük farklılık göstermiştir (liste 1 ve 3 ile liste 2 ve 4 için sırasıyla .34 ve .50). Ayrıca, yüksek ileri gücü durumunda da listeler arası farklılıklar söz konusudur (.47 ve .61). Bu farklılıklar, büyük oranda, deney kapsamında kontrol edilemeyen ve listeler arasında farklılaşmış değişkenlere (örneğin, hedef kelimesinin özellikleri, hedef kelimesinin ipucu kelimesine olan direkt ve dolaylı bağlantıları gibi) bağlanabilir.

Tartışma

Bir kelimenin anımsattığı anlamca ilişkili diğer kelimeler ve bu kelimelerin birbirleriyle olan anlam ilişkileri kelimelerin çağrışım yapılarını tanımlar. Bu çağrışım yapıları hakkındaki bilgiler serbest çağrışım yönteminin birçok kişiye uygulanması, yani normatif bir çalışma ile elde edilebilir. Türkçe kelimelere ait çağrışım normları önceki bir çalışmada (Tekcan ve Göz, 2005) toplanılmış olsa da, bulguların güvenirlik ve geçerliği incelenememiştir. Kelimelerin anlam ve yapı özellikleri kelimelerin kullanıldığı bilimsel çalışmalarda katılımcıların tepkilerini etkileyebileceğinden, normatif bulguların geçerliğinin ve güvenirliğinin bilinmesi oldukça önemlidir. Mevcut çalışma, Türkçe kelimelerin çağrışım normlarının güvenirlik ve geçerliğini inceleyerek, alanyazına katkı sağlamıştır. Var olan normların güvenirliği, önceki norm çalışmasında kullanılan kelimelerden rastgele olarak bir alt küme oluşturulması ve bu kelimelere dair serbest çağrışım verilerinin hem kağıt-kalem hem de çevrimiçi yöntemleri kullanılarak toplanılıp elde edilen bulguların karşılaştırılması ile değerlendirilmiştir. Geçerliğin de-

ğerlendirilmesi için ipucu ile hatırlama yöntemi kullanılmış ve alanyazındaki beklentiler çerçevesinde normatif verilerin bu görevdeki doğru hatırlama oranını ne derece yordadığına bakılmıştır. Türkçe kelimelerin çağrışım verilerinin yüksek test-tekrar test ve iç tutarlılığı gösterdiği ve ileri gücünün (bir başka deyişle, ipucunun hedef kelimeyi anımsatma olasılığının) güvenilir olarak ve alanyazındaki beklentiler çerçevesinde bellek performansını yordadığı bulunmuştur.

Çağrışım yapıları kişiler içinde ve kişiler arasında çeşitli nedenlerden dolayı farklılık gösterse de, bu yapılar zaman içerisindeki ölçümlerde tutarlılık göstermektedirler (De Deyne ve Storms, 2008; Nelson ve ark., 2000). Birinci çalışmada, seçilen kelimelere ait var olan normatif bulgular, aynı kelimeler için yeniden toplanılan bulgular ile karşılaştırılmıştır. Ayrıca, yöntemsel farklılıkları incelemek için, veriler hem klasik kağıt-kalem yöntemiyle hem de çevrimiçi olarak toplanılmış ve var olan normlar ile benzerlik ve farklılıklarına bakılmıştır. Karşılaştırmalar, kelime yapılarına ait üç farklı boyut üzerinden yapılmıştır. Bunlar, çağrışım gücü, çağrışım seti büyüklüğü ve kelimelere verilen cevapların çeşitliliği (heterojenliği) olarak ifade edilebilir. Bulgular, bu üç boyutta veri setlerinin sistematik olarak farklı olmadığını ve tutarlılığa sahip olduğunu göstermiştir. Diğer bir ifadeyle, serbest çağrışım yöntemi ile farklı bir zamanda, farklı örneklemlerde ve farklı yöntemlerle elde edilen normatif bulgular birbirleriyle uyum göstermiştir. Nelson ve arkadaşlarının (2000) farklı yöntemler kullanılarak elde ettiği serbest çağrışım güç değerleri arasındaki tutarlılık değeri (.89) mevcut çalışmada elde edilen kağıt-kalem ve çevrimiçi yöntemleri arasındaki tutarlılık değeri ile aynıdır (.89). Bu iç tutarlılık, serbest çağrışım yönteminin kelimelerin çağrışım yapılarını ölçmede oldukça güvenilir bir yöntem olduğunu göstermektedir. Ayrıca, güç değerlerinde, yeni toplanılan veriler ve var olan bulguların arasında da yüksek tutarlılık bulunmuştur. Güç değerlerine ek olarak, çağrışım seti değerleri de yeni toplanılan verilerde benzerlik göstermiştir. En büyük fark olarak, çevrimiçi yöntemin genel olarak daha az sayıda çağrışım üretmesi ve elde edilen çağrışımların sıklık dağılımının normal olmaması gösterilebilir. Bu bulgu, büyük oranda bu yöntemden elde edilen verilerin diğerlerine göre daha küçük bir örneklemden gelmesiyle açıklanabilir. Önceki çalışmalar, çağrışım setinin toplanılan veri sayısının büyüklüğünden etkilendiğini göstermiştir (De Deyne ve ark., 2013; De Deyne ve Storms, 2008). Mevcut çalışmada, çevrimiçi yöntemdeki örneklem küçüklüğüne rağmen diğer veri setleri ile büyük oranda benzer bulgular elde edilmiştir. Bu bulgu, serbest çağrışım verilerini çevrimiçi çalışmalar ile toplamanın mümkün olduğunu ve klasik yönteme göre benzer bir sonuç verdiğini desteklemektedir (benzer bir öneri için, De Deyne ve ark., 2013). Son olarak, farklı veri setlerinde elde edilen tepkilerin heterojenliğine bakılmıştır. Veri setlerine ait heterojenlik değerlerinin karşılaştırılması ile bir kelimeye karşı verilen tepkilerin çeşitliliğinin farklı veri setlerindeki benzerliği veya tutarlılığı hakkında bilgi sahibi olunabilir. Elde edilen bulgular, kelimelere verilen çağrışımların benzer çeşitliliğe sahip olduğu yönündedir. Ancak, ortalama heterojenlik değerlerinde farklılıklar gözlenmiştir. Bu değerlerdeki farklılığın veri toplama işlemlerindeki farklılıklarla paralel olması dikkati çekmiştir. Var olan normlar 10 ila 40 kişilik gruplar üzerinden toplanmışken (Tekcan ve Göz, 2005), kağıt-kalem ile yeni toplanılan normlar en fazla 4'er kişilik gruplardan toplanmıştır. Çevrimiçi yöntemle toplanılan verilerde katılımcıların kaç kişilik bir ortamda olduğu bilinmemekte, ancak yalnız oldukları tahmin edilmektedir. Böyle bir etki öngörülmediği ve konuyla ilgili deneysel bir yönlendirme yapılmadığı için kesin bir sonuç çıkarsamak mümkün olmamakla birlikte, elde edilen bulgulara göre, heterojenlik değerleri katılımcıların test edildiği grup büyüdükçe küçülme eğilimindedir. Yani, deney ortamının ne derece kalabalık olduğu potansiyel olarak kişilerin cevaplarındaki çeşitliliği etkileyebilir.

Birinci çalışmanın bulguları hedef kelimelerin çağrışım seti büyüklüğü ile çağrışımlarının güç ve çeşitlilik değerlerinin yüksek tutarlığa sahip olduğunu göstermiştir. Tutarlılık ile ilgili yapılan istatistiksel analizlerde katılımcıların verdiği cevapların aynı olup olmaması hesaba katılmamıştır. Veri setleri hedef kelimelere verilen cevaplar doğrultusunda incelendiğinde bazı farkları barındırmaktadır. Bu farklar bireysel farklılıkların yanı sıra toplumdaki değişimlerin ve güncel olayların etkisinden de kaynaklanabilir. Örneğin, Tekcan ve Göz'ün (2005) veri setinde bulunan HASAR → DEPREM ilişkisi yeni toplanılan normlarda büyük ölçüde yok olurken, baskın ilişki olarak HASAR → ARABA yerini almıştır. Niceliksel açıdan düşünüldüğünde, bunun 1999 yılı İstanbul depreminden birkaç sene sonra Boğaziçi Üniversitesi'nde yapılan Tekcan ve Göz (2005) araştırmasının tarihsel ve mekânsal bağlamını yansıttığı düşünülebilir. Göze çarpan başka bir örnek, çevrimiçi veri setinde bulunan, 2010–2014 arasında İstanbul'un valiliğini yapmış Hüseyin Avni Mutlu'nun adını anan VALİ→MUTLU ilişkisidir. Dolayısıyla, araştırmaların gerçekleştirildiği zaman ve mekâna özel faktörlerin katılımcıların hedef kelimelere verdiği çağrışımları etkilediği görülmektedir. Kelimelerin çağrışımlarındaki bu farklar göreceli güç hipotezi (Maki, 2008; Nelson ve ark., 2000) ile uyumludur. Serbest çağrışım yönteminde, bir kelimenin bir başka kelimeyi anımsatma güç değeri salt bir güç değeri olarak farz edilirse; serbest çağrışım görevinin zaman, mekan ve uygulama yönteminden bağımsız olarak aynı çağrışım ve güç dağılımlarını üretmesi beklenir. Buna karşın

göreceli güç hipotezi serbest çağrışım verilerinin kişisel deneyimlerden ve güncel koşullardan dinamik bir şekilde etkilenmesini ve bunun yol açtığı değişkenliği doğal olarak açıklar. Göreceli güç hipotezine göre, hedef kelime bir dizi kelimeyi ve onların güç dağılımlarını aktif hale getirmektedir. Hangi kelimenin tepki olarak verileceği aktifleşen kelimelerin güç dağılımlarından rastgele bir değer seçilmesi ve tüm değerlerin birbirleriyle karşılaştırılıp en güçlü olanının rekabeti kazandığı bir süreç sonucunda belirlenir. Bu yaklaşıma göre, hedef kelimenin bir başka kelimeyi çağrıştırması gücün göstergesidir ancak salt, değişmeyen bir güç değeri değildir (Nelson ve ark., 2000). Bireysel farklılıklar, yakında olmuş deneyimler ve kültürel değişimler kelimelerin çağrışımlarını ve güçlerini değiştirebilmektedir.

Serbest çağrışım yöntemi ile önceden elde edilen çağrışım yapısı bulguları, ipucu ile hatırlama yönteminde hedef kelimenin verilen ipucu sayesinde ne derece iyi hatırlanıp hatırlanamayacağını büyük oranda belirler. Örneğin, ileri gücü (Nelson ve ark., 2007) ve ipucu çağrışım seti büyüklüğü (Nelson ve ark., 1998) bellek performansını olumlu etkilemektedir. Var olan normların bellek performansını ne derece yordadığı geçerliğin bir ölçütü olarak alınıp ikinci çalışmada sınanmıştır. Bu doğrultuda yapılan ipucu ile hatırlama deneyinde ipucu kelimeleri daha önce norm çalışması yapılmış kelimelerden (Tekcan ve Göz, 2005) seçilip hedef kelimeler ise bu kelimelere karşı katılımcıların verdiği yine aynı veri tabanında bulunan cevaplardan seçilmiştir. İpucu kelimelerinin diğer özellikleri kontrol edilip ipucunun hedef kelimeyi anımsatma olasılığı ve ipucu çağrışım seti büyüklüğü değerleri manipüle edilmiştir. Bulgular alanyazındaki tahminlerle kısmen uyumlu çıkmıştır. Aynı ipucuna karşı seçilen iki farklı hedef kelimenin doğru hatırlanma oranları karşılaştırıldığında, kelimeler arasındaki bağ normatif verilerce güçlü olduğu durumda güçsüz olduğu duruma göre daha iyi bir bellek performansı elde edilmiş ve etki büyüklüğü de oldukça yüksek çıkmıştır. PIER II'nin tahminleri, farklı bir dilde farklı araçlar kullanan bir çalışma üzerinde doğrulanmıştır. Bu durum, modelin kültürler arası genellenebilirliğini göstermektedir. PIER II'nin beklentileriyle uyuşmayan bulgular da elde edilmiştir. Alanyazındaki benzer çalışmalardan farklı olarak, ipucu çağrışım setinin etkisi istatistiksel olarak anlamlı çıkmamış; ileri gücü ile çağrışım setlerinin büyüklüğü arasındaki etkileşim anlamlı çıkmıştır. Çağrışım setlerinin büyüklüğüne dair olan etkiler büyük oranda hedef ve ipucu kelimelerinin bir bütün olarak incelenmesi, ortak ve özgün çağrışımların değerlendirilmesine dayalıdır (Nelson ve ark., 1999). Yalnız ipucu kelimelerine ait normatif çağrışım verileri bulunmasından ve hedef kelimelerin çağrışım yapılarına dair bilgilerin bulunmamasından ötürü mevcut çalışma için kelime seçilirken bu durumu göz önüne almak mümkün olmamıştır. Yapılan ek analizlerin sonucunda kullanılan kelime listelerinin birinde ve özellikle ipucu çağrışım setinin büyük olduğu durumda bir karıştırıcı değişkenin etkisi olduğu tahmin edilmektedir. Çoklu regresyon analizinin sonucu, bu karıştırıcı değişkenin deney kapsamında kontrol edilen bir değişken olmadığını göstermiştir.

Bu durum mevcut normların kullanımına dair bir eksiklik olarak nitelendirilebilir. Eksiklik, norm verisi toplanılmış kelimelerin çağrışım setlerini oluşturan diğer kelimelerin çağrışım yapılarının bilinmemesinden kaynaklanmaktadır. Tepki kelimelerine dair bilgilerin olmaması ipucu ve hedef kelimelerinin bir bütün olarak bilimsel araştırmalarda kullanımını kısıtlamakta hatta sonuçları öngörülmeyecek şekilde etkileyebilmektedir. Alanyazına bakıldığında, tepki kelimeleri üzerine farklı örneklemler kullanılarak yapılan ilave çalışmalarla normatif çağrışım verisi toplanıldığı görülmektedir (De Deyne ve ark. 2013, Nelson ve ark., 2004). İkinci çalışmadan elde edilen bulgular mevcut normların kullanımının kısıtlılıklarının araştırmaların sonuçlarını doğrudan etkileyebileceğine işaret etmektedir. Var olan normatif kelime bulguları, bugüne kadar yapılan psikolojik araştırmalarda öncelikli olarak kelimelerin imgelem, somutluk, sıklık ve çağrışım sayısı gibi özelliklerinin farklı deneysel yöntemlerde kontrol edilmesine yönelik kullanılmıştır (Babayiğit ve Stainthorp, 2007; Baran, Cangoz ve Salman, 2015; Booth, Mackintosh, Mobini, Oztop ve Nunn, 2014; Kılıç ve Öztekin, 2014; Tabakçı ve Karakelle, 2012) ve alanyazında da kelime normlarının en sık rastlanan kullanım amacı budur. Alanyazın incelendiğinde Türkçe kelimelere dair serbest çağrışım verilerinden yayınlanmış yalnızca az sayıda araştırmada (Boduroğlu, Tekcan ve Kapucu, 2014; Coşkun, 2009) kelime özelliklerini kontrol etmek amacı dışında normatif bulguların yönlendirilmesi ile doğrudan faydalanıldığı görünmektedir. Türkçe kelimelerin çağrışım yapılarına dair normatif bulguların psikoloji araştırmalarında oldukça seyrek kullanılmış olması dikkat çekicidir. Bu durumun en büyük nedenlerinden biri olarak benzerleri ile karşılaştırıldığında Türkçe normların (Tekcan ve Göz, 2005) 600 kelime ile oldukça kısıtlı kalmış olması gösterilebilir. Örneğin, Nelson ve arkadaşları (2004) yaklaşık 5000, De Deyne ve arkadaşları (2013) ise yaklaşık 12500 kelimenin normatif verilerini toplamışlardır.

Türkçe serbest çağrışım normlarının yeterli olmaması yapılabilecek araştırma olanaklarını da kısıtlamaktadır. Çağrışım normlarının kullanım alanlarını genişletmek için var olan normatif bulgulardan yola çıkarak ilave norm çalışmaları yapılmasını gerekli görüyoruz. Bu sayede, normatif bulguların çeşitli psikolojik çalışmalarda sağlıklı olarak kullanılması mümkün olacaktır. İpucu ile hatırlama yöntemi dışında, kelimelerin

çağrışım yapılarına dair normatif bilgiler yalancı bellek (*false memory*) oranını tahmin etmek için de kullanılmıştır. Deese–Roediger–McDermott (DRM) paradigması (Roediger ve McDermott, 1995) kullanıldığında, çalışma listesinde gösterilmeyen kritik kelime eğer oldukça bağlantılı bir çağrışım yapısına sahip ise yanlış hatırlama oranı çağrışım yapısı daha az bağlantılı olan kritik kelimelere göre daha fazla olmaktadır (McEvoy ve ark., 1999). Serbest çağrışım normlarının bir başka kullanım alanı da popülasyonların çağrışımları arasındaki farklılıkları belirlemektir (Coronges, Stacy ve Valente, 2007; Stacy, 1997). Örneğin, alkole dair beklentilerin deneyimlerle oluştuğu ve bu deneyimlerin ilişkisel bellekteki çağrışım yapılarındaki farklılığa sebep verdiği varsayılmaktadır (Dunn ve Goldman, 2000). Serbest çağrışım kullanılarak elde edilen bulgular, daha sık alkol kullanan ve kullanım dozu daha yüksek olanların alkolün daha heyecan uyandıran ve sosyal sonuçlarını belirten çağrışımlara sahip olduğunu, alkol kullanım sıklığı ve dozu düşük olanların ise alkolün yatıştırma ve sakinleştirici özelliklerini tercih ettiklerini göstermektedir (Reich ve Goldman, 2005). Bahsedilen çalışmaların yapılabilmesi, genişletilmesi ve bunlara ilave farklı çalışma desenlerinin oluşturulması Türkçe kelimelerin çağrışım normlarının bütünlüklü olarak tamamlanması ile desteklenebilir.

Kelimelerin çağrışım yapıları çeşitli görevlerde çalışmaların sonuçlarını etkilemektedir. Bu çalışma, mevcut kelime çağrışım normlarını detaylı olarak inceleyerek güvenirlik ve geçerliğe dair ilk bulguları elde etmiştir. Çalışmanın bir sınırlılığı, çevrimiçi yöntemde diğer yöntemlere göre daha az katılımcının yer alması ve bu katılımcıların yalnızca üniversite birinci sınıf öğrencilerinden oluşmasıdır. Bu bağlamda elde edilen sonuçların kısmen örneklem sayılarının ve yaşlarının farklı olmasından kaynaklandığı söz konusu olabilir. Ayrıca, ipucu ile hatırlama yönteminde bir kelime listesinin benzer koşullarda hazırlanan diğer kelime listesine göre farklı sonuç vermesi listedeki hedef kelimelerin ipucu ile olan anlam ilişkilerindeki farklılıklara bağlanmıştır. Bu farklılıkları kesin olarak bilebilmek eldeki verilerle mümkün değildir ve çalışmada ölçülmeyen ve hedef kelimelerinden bağımsız farklı bir değişkenden ötürü olması düşük de olsa ihtimal dahilindedir. Son olarak, geçerliliğe dair çıkarımlar çağrışım gücü değerlerinden yapılmıştır, ancak çağrışım setlerinin büyüklüğüne dair yordama geçerliği normların yapısından da kaynaklanan eksikliklerden dolayı bilinmemektedir. Bu kısıtlılıklara rağmen elde edilen sonuçlar, mevcut serbest çağrışım verilerinin büyük oranda güvenilir ve geçerli olduğunu ancak ilave norm çalışmaları ile zenginleştirilmesi gerektiğini göstermektedir.

Kaynaklar

Aitchison, J. (2003). *Words in the mind: An introduction to the mental lexicon (3. baskı)*. Malden, MA: Blackwell Pub.

Babayiğit, S. ve Stainthorp, R. (2007). Preliterate phonological awareness and early literacy skills in Turkish. *Journal of Research in Reading*, *30*(4), 394–413. doi:10.1111/j.1467-9817.2007.00350.x

Baran, Z., Cangöz, B. ve Salman, F. (2015). Duygusal bağlam eşik altı ve eşiküstü hazırlamayı etkiler mi? *Türk Psikiyatri Dergisi*, *26*.

Boduroglu, A., Tekcan, A. İ. ve Kapucu, A. (2014). The relationship between executive functions, episodic feeling-of-knowing and confidence judgements. *Journal of Cognitive Psychology*, *26*(3), 333–345. doi:10.1080/20445911.2014.891596

Booth, R. W., Mackintosh, B., Mobini, S., Oztop, P. ve Nunn, S. (2014). Cognitive bias modification of attention is less effective under working memory load. *Cognitive Therapy And Research*, *38*(6), 634–639. doi:10.1007/s10608-014-9628-6

Brotsky, S. J. ve Linton, M. L. (1967). The test-retest reliability of free association norms. *Psychonomic Science*, *8*(10), 425–426.

Coronges, K. A., Stacy, A. W. ve Valente, T. W. (2007). Structural comparison of cognitive associative networks in two populations. *Journal of Applied Social Psychology*, *37*(9), 2097–2129. doi:10.1111/j.1559-1816.2007.00253.x

Coşkun, H. (2009). Beyin fırtınası sürecinde çağrışım alıştırmalarının düşünce üretimine etkisi (The effects of associative exercises on the idea generation during brainstorming). *Türk Psikoloji Dergisi*, *24*(64), 34–44.

De Deyne, S., Navarro, D. J. ve Storms, G. (2013). Better explanations of lexical and semantic cognition using networks derived from continued rather than single-word associations. *Behavior Research Methods*, *45*(2), 480–498. doi: 10.3758/s13428-012-0260-7

De Deyne, S. ve Storms, G. (2008). Word associations: norms for 1,424 Dutch words in a continuous task. *Behavior Research Methods*, *40*(1), 198–205. doi: 10.3758/BRM.40.1.198

Deese, J. (1965). *The structure of associations in language and thought*. Baltimore, MD: Johns Hopkins University Press, Derry.

Dunn, M. E. ve Goldman, M. S. (2000). Validation of multidimensional scaling-based modeling of alcohol expectancies in memory: Age and drinking-related differences in expectancies of children assessed as first associates. *Alcoholism, Clinical and*

Experimental Research, *24*(11), 1639–1646. doi: 10.1111/j.1530-0277.2000.tb01965.x

Goodmon, L. B. ve Nelson, D. L. (2004). Strengthening the activation of unconsciously activated memories. *Memory & Cognition*, *32*(5), 804–818. doi: 10.3758/BF03195870

Gray, R. M. (2011). *Entropy and information theory*. Springer Science & Business Media.

Hintzman, D. L. (1976). *Repetition and memory*. G. H. Bower (Ed.), *The psychology of learning and motivation* içinde (47–91). New York: Academic Press.

Kılıç, A. ve Öztekin, İ., (2014). Retrieval dynamics of the strength based mirror effect in recognition memory. *Journal of Memory and Language*, 76, 158–173. doi: 10.1016/j.jml.2014.06.009

Maki, W. S. (2008). A database of associative strengths from the strength-sampling model: A theory-based supplement to the Nelson, McEvoy, and Schreiber word association norms. *Behavior Research Methods*, *40*(1), 232–235. doi: 10.3758/BRM.40.1.232

McEvoy, C. L., Nelson, D. L. ve Komatsu, T. (1999). What is the connection between true and false memories? The differential roles of interitem associations in recall and recognition. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *25*(5), 1177–1194. doi: 10.1037/0278-7393.25.5.1177

Nelson, D. L., Bennett, D. J., Gee, N. R., Schreiber, T. ve McKinney, V. M. (1993). Implicit memory: Effects of network size and interconnectivity on cued recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(4), 747–764.

Nelson, D. L., Dyrdal, G. M. ve Goodmon, L. B. (2005). What is preexisting strength? Predicting free association probabilities, similarity ratings, and cued recall probabilities. *Psychonomic Bulletin & Review*, *12*(4), 711–719. doi: 10.3758/BF03196762

Nelson, D. L., Fisher, S. L. ve Akirmak, U. (2007). How implicitly activated and explicitly acquired knowledge contribute to the effectiveness of retrieval cues. *Memory & Cognition*, *35*(8), 1892–1904.

Nelson, D. L., McEvoy, C. L. ve Bajo, M. T. (1988). Lexical and semantic search in cued recall, fragment completion, perceptual identification, and recognition. *The American Journal of Psychology*, *101*(4), 465–80. http://www.ncbi.nlm.nih.gov/pubmed/3232723 adresinden alınmıştır.

Nelson, D. L., McEvoy, C. L. ve Dennis, S. (2000). What is free association and what does it measure? *Memory & Cognition*, *28*(6), 887–899. doi: 10.3758/BF03209337

Nelson, D. L., McEvoy, C. L. ve Pointer, L. (2003). Spreading activation or spooky action at a distance? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(1), 42–52. doi: 10.1037/0278-7393.29.1.42

Nelson, D. L., McEvoy, C. L. ve Schreiber, T. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers: A Journal of the Psychonomic Society, Inc*, *36*(3), 402–407. doi: 10.3758/BF03195588

Nelson, D. L., McEvoy, C. L. ve Schreiber, T. A. (1990). Encoding context and retrieval conditions as determinants of the effects of natural category size. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 34–41.

Nelson, D. L., Schreiber, T. A. ve McEvoy, C. L. (1992). Processing implicit and explicit representations. *Psychological Review*, *99*, 322–348.

Nelson, D. L., Schreiber, T. A. ve Xu, J. (1999). Cue set size effects: sampling activated associates or cross-target interference? *Memory & Cognition*, *27*(3), 465–477. doi: 10.3758/BF03211541

Nelson, D. L. ve Zhang, N. (2000). The ties that bind what is known to the recall of what is new. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *27*(5), 1147–1159. doi: 10.1037/0278-7393.27.5.1147

Nelson, D. L., Zhang, N. ve McKinney, V. M. (2001). The ties that bind what is known to the recognition of what is new. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *27*(5), 1147–1159. doi: 10.1037/0278-7393.27.5.1147

Reich, R. R. ve Goldman, M. S. (2005). Exploring the alcohol expectancy memory network: The utility of free associates. *Psychology of Addictive Behaviors*, *19*(3), 317–325. doi: 10.1037/0893-164X.19.3.317

Roediger, H. L., Jacoby, J. D. ve McDermott, K. B. (1996). Misinformation Effects in Recall: Creating False Memories through Repeated Retrieval. *Journal of Memory and Language*, *35*(2), 300–318. doi: 10.1006/jmla.1996.0017

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*, 803–814.

Stacy, A. W. (1997). Memory activation and expectancy as prospective predictors of alcohol and marijuana use. *Journal of Abnormal Psychology*, *106*, 61–73.

Vikis-Freibergs, V. ve Freibergs, I. (1976). Free association norms in French and English: Inter-linguistic and intra-linguistic comparisons. *Canadian Journal Of Psychology/Revue Canadienne De Psychologie*, *30*(3), 123–133. doi:10.1037/h0082054

Tabakçı, H. Ş. ve Karakelle, S. (2010). Öğrenilmiş Çaresizliğin Bilme Hissi Kararı Üzerindeki Etkisinin Gelişimsel Olarak İncelenmesi. *Psikoloji Çalışmaları Dergisi*. http://www.journals.istanbul.edu.tr/iupcd/article/view/1023018497 adresinden alınmıştır.

Tekcan, A. İ. ve Göz, İ. (2005). *Türkçe Kelime Normları (Turkish Word Norms).* İstanbul: Boğaziçi Üniversitesi Yayınevi.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist*, *26*, 1–12.

Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327–352.

Turkish Journal of Psychology, June 2018, *33*(81), 79-81

Summary
A Study on the Reliability and Validity of the Turkish Free Association Norms

Ümit Akırmak
İstanbul Bilgi University

Mehmet Altan Orhon
İstanbul Bilgi University

Free association helps us in mapping the associative structure of words in semantic memory by indexing the strong associations of a given word (Deese, 1965; Nelson, McEvoy, & Schreiber, 2004). Understanding the associative relations among the words is crucial especially for studies that make use of words as stimuli since the characteristics of the words have been shown to directly influence the results of the experiments (Nelson, Fisher, & Akirmak, 2007; Nelson, Schreiber, & Xu, 1999). For this reason, pre-existing associative knowledge about the words is usually utilized to control various word attributes across experimental conditions. In addition to controlling for experimental conditions, associative knowledge is also used to predict memory performance in various tasks (Deese, 1965; Nelson, Fisher, & Akirmak, 2007), false memory rates in the Deese–Roediger–McDermott (DRM) paradigm (McEvoy, Nelson, & Komatsu, 1999), as well as differences in behaviors and attitudes of different populations (Reich & Goldman, 2005; Stacy, 1997). As it was shown, the free association task has broad applicability in psychological research and is still being widely used in different types of research.

In the free association task, participants are provided with target words and are asked to produce the first word that comes to their mind in association with each target word. A prior study by Tekcan and Goz (2005) collected free association norms for 600 Turkish words; however, the reliability and validity of the responses they obtained have not been tested by an additional study. When the studies on free association norms are examined in the literature (e.g., De Deyne, Navarro, & Storms, 2013; Nelson, McEvoy, & Schreiber, 2004), it can be seen that norming procedures often involve reliability and validity analyses through new studies. The purpose of the present study was to examine the reliability and validity of the existing free association norms for the Turkish words (Tekcan & Goz, 2005). Two studies were conducted to this end as part of the present study. These studies focused on associative information, which was indexed by free association, such as associative strength, set size, and heterogeneity. The present study first reviewed these concepts and theoretical models that were specifically aimed to explain free association and cued recall tasks. The relative strength hypothesis (Maki, 2008; Nelson, McEvoy, & Dennis, 2000) for predicting response probabilities in a free association task and "Processing Implicit and Explicit Representations II" (PIER II; Nelson, McKinney, Gee, & Janczura, 1998) for predicting memory performance in cued recall task were explained in detail. The specific information on studies as well as the findings were presented.

Study 1

The first study aimed to assess test-retest and internal reliability of the associative structure of the Turkish words as measured by free association. For this purpose, 100 words were randomly chosen from the existing free association normative database (Tekcan & Goz, 2005) and were used as materials in the present study. The standard free association task was applied to 105 undergraduate students. In addition to the standard method, an online version of the task was designed and was applied to 59 undergraduate students through an online survey site namely *Survey Monkey*. The analyses were mainly focused on two variables of interest: associative strength and associative set size. To assess the reliability of the existing association norms, the newly obtained values of strength and set size through standard and online methods were compared with the existing values. In addition, as it was suggested by De Deyne et al. (2013), heterogeneity values for the three databases were computed and compared as an alternative index of reliability. The results indicated that for all three databases, the associative strength values decreased consistently with the rank of a given associate. In addition, the first associates showed a high test-retest

Address for Correspondence: Dr. Ümit Akırmak, İstanbul Bilgi University, Faculty of Social Sciences and Humanities, Department of Psychology, Eski Silahtarağa Elektrik Santralı Kazım Karabekir Cad. No:2/13 Eyüp / İstanbul
E-mail: umit.akirmak@bilgi.edu.tr

correlation of associative strength (r = .72) and high internal consistency as shown by comparison of the standard task with the online task (r = .89). Finally, the three databases also showed a strong positive correlation of heterogeneity values, indicating that they produced similarly different associative responses for the target words. Thus, these findings implied that free association reliably indexes associative word knowledge. More importantly, existing Turkish free association norms were found to be reliable as indicated by similar pattern of responses obtained from a different sample, measured at a different time and by utilizing a different method.

Study 2

The purpose of the second study was to examine whether the existing normative free association database could be used to predict memory performance in a new task, based on the expectations from similar studies in the literature. Specifically, according to PIER II, cues that have stronger forward connections to their targets should result in better correct cued recall as compared to cues that have weak or no forward connections when all other variables are controlled (Nelson et al., 2007). The same advantage is also present for the cues that produce fewer associates in free association as compared to the cues that produce many associates (Nelson et al., 1998). These predictions were tested in a cued recall experiment designed to vary some of the word characteristics that were measured by free association. In the cued recall task, participants first studied a list of words (targets) and then were asked to recall these words with the help of associatively related words (cues) that were not shown during the experiment. Associative strength was manipulated between subjects and cue set size was manipulated within subjects in this cued recall experiment. For this purpose, 48 cue words and their corresponding targets were chosen from the normative database (Tekcan & Goz, 2005) and were allocated into four different word lists. There were two word lists in each between-subjects condition in order to examine the generalizability of the findings across chosen materials. Sixty-four undergraduates participated in the study, in exchange for course credit. Because of the structure and the extent of the existing word norms, only various word attributes for the cue words could be controlled in the present study. These attributes included word frequency, concreteness, and imagery. The results of the experiment indicated that stronger associative strength between the cue–target pair resulted in higher probability of cued recall as compared to weaker associative strength between the cue–target pair. This result replicated earlier findings on associative strength and added to the generalizability of PIER II by confirming its predictions with material from a different language's free association norms. However, the manipulation of the cue set size did not result in the expected advantage for the cues having a smaller associative set size. A close inspection of the findings revealed that one set of lists confirmed PIER II's predictions on cue set size effects while the other set contradicted it, implying that there might be a confound in the construction of the study lists. Further analyses revealed that this effect was not due to confound variables controlled within the study. Although this result seems to contradict PIER II's predictions in general, the apparent contradiction can be explained by uncontrolled word attributes belonging to the target words. Specifically, according to PIER II, cue set size effects result from an intersection of the cue and target's associative structures. Thus, in order to fully predict the cue set size effects, knowledge regarding the target's associative structure must be present; this was not the case for the present study. Overall, the findings implied that Turkish word association norms possess predictive validity, especially when associative strength was considered. However, more research on associative set size is needed to be able to entangle how the associative structure of the target influences correct recall. In order to be able to assess the contribution of the associative structure of the target words on cued recall and to map the associative structure of these words, it is needed to conduct additional norming studies on the responses of the normed words in the normative database.

General Discussion

The purpose of the present study was to examine the reliability and validity of Turkish word association norms. Study 1 evaluated the test-retest reliability of the three databases (existing norms, newly collected norms through standard method, and newly collected norms through an online survey) by comparing associative strength, set size, and heterogeneity values. The findings of Study 1 indicated high test-retest reliability. Although there exist differences in the particular associates produced, the overall pattern of associates provided in response to a target word was highly similar. The present findings showed for the first time that Turkish word associations (Tekcan & Goz, 2005) are reliable despite the time elapsed and different population utilized. The validity of the normative database was examined by designing a cued recall experiment based on the predictions of a mathematical model of cued recall task, namely PIER II (Nelson et al., 1998). The results of Study 2 partially supported the validity of the existing norms by showing that associative strength reliably predicted cued recall memory performance. The results in regard to the cue

set size failed to provide full support mainly because of the insufficient knowledge for the associative structure of the target words utilized in the second study. In other words, cue set size effects are based on the intersection of cue and target's associates, and thus the knowledge of both cue and target and their associative structures are needed in order to reliably demonstrate their effects. In the absence of such information, it is not possible to accurately predict the recall advantage for the cues with smaller associative sets as compared to cues with larger associative sets. The findings obtained from one set of lists confirmed the cue set size effects, while the results obtained from another set of lists failed to provide a support. Overall, the results of the cued recall study was promising in that it confirmed the predictive validity of the associative strength values indexed by free association norms collected about a decade ago. In addition, present findings supported the cross-cultural validity of PIER II in predicting cued recall performance based on the word attributes measured by free association norms from a different language.

Word association norms are used in various tasks and for various purposes. We believe collecting further word association norms will be a productive endeavor as long as the norms serve as reliable predictors of task performance. As an example, the seemingly contradictory results of Study 2 might be derived from the insufficient knowledge on word associations. Thus, extending the normative database for word associations may enable researchers, who are working with words as stimuli, to apply experimental control more reliably in their studies. In addition, an extended database for word associations may enable researchers to make use of different research methodologies and designs, which, currently, can not be utilized without the potential confounds mentioned in the present study as lack of associative knowledge.