



Saat Çizme Testinin İki Farklı Puanlama Sisteminin Türkiye Normları ve Geçerlik-Güvenirlik Çalışması

Validity, Reliability and Turkish Norm Values of the Clock Drawing Test for Two Different Scoring Systems

© Derya Durusu Emek-Savaş¹, © Deniz Yerlikaya², © Görsev G. Yener³

¹Dokuz Eylül Üniversitesi Edebiyat Fakültesi, Psikoloji Bölümü, İzmir, Türkiye

²Dokuz Eylül Üniversitesi Sağlık Bilimleri Enstitüsü, Sinirbilimler Anabilim Dalı, İzmir, Türkiye

³Dokuz Eylül Üniversitesi Tıp Fakültesi, Nöroloji Anabilim Dalı, İzmir, Türkiye

Öz

Amaç: Saat çizme testi (SÇT) ülkemizde çok sayıda merkezde uygulanmakta olup saatin dairesinin hazır olarak sunulduğu versiyonun Türkiye popülasyonuna ait norm değerleri bulunmamaktadır. Bu çalışmada, SÇT'nin uluslararası literatürde en yaygın kullanıma sahip skorlama yöntemlerinden Manos ve Wu ile Shulman yöntemleri ile standardizasyonunun yapılması ve geçerlik ve güvenilirliğinin test edilmesi amaçlanmıştır.

Gereç ve Yöntem: Çalışmanın norm belirleme aşamasına 50-92 yaş aralığındaki 244 sağlıklı birey dahil edilmiştir. Yaş, eğitim ve cinsiyet değişkenlerinin SÇT puanları üzerindeki etkisi incelenmiş, yaş ve eğitime göre norm değerleri oluşturulmuştur. Ayrıca, iki puanlama yönteminin test-tekrar test güvenilirliği, değerlendiriciler arası güvenilirliği, eş-zaman geçerliliği ve ölçüt geçerliliği test edilmiştir. Ölçüt geçerliliği için ROC eğrisi yöntemi kullanılarak, SÇT puanlarının sağlıklı bireyler, hafif kognitif bozukluk (HKB) olguları ve Alzheimer hastalarını ne düzeyde ayırt edebildiği incelenmiştir.

Bulgular: Manos ve Wu yönteminde sağlıklı bireylerin SÇT puanları üzerinde yaş ve eğitim etkisi, Shulman yönteminde ise sadece eğitim etkisi olduğu saptanmıştır. Her iki yöntemin de test-tekrar test (Manos ve Wu için, 0,81; Shulman için, 0,72) ve değerlendiriciler arası güvenilirliği (Manos ve Wu için, 0,98; Shulman için, 0,96) yüksek olup, birbirleriyle ve diğer kognitif tarama testleriyle yüksek düzeyde korelasyon gösterdiği bulunmuştur. ROC eğrisi altında kalan alan, sadece sağlıklı bireyler ile Alzheimer hastalarını ve sağlıklı bireyler ile HKB olgularını birbirinden ayırt etmede yüksek değerler göstermiştir.

Sonuç: Bu çalışmada, 50 yaş ve üzeri bireyler için ayrıntılı norm değerleri oluşturulmuştur. SÇT skorları yorumlanırken yaş ve eğitim düzeyi dikkate alınmalıdır. Her iki puanlama yönteminden elde edilen bulgular oldukça paralel olup, yüksek test-tekrar test ve değerlendiriciler arası güvenilirlik ile yüksek eş-zaman ve ölçüt geçerliklerine işaret etmektedir.

Anahtar Kelimeler: Saat çizme testi, normatif veri, geçerlik-güvenirlik, Alzheimer hastalığı, hafif kognitif bozukluk

Abstract

Objective: The clock drawing test (CDT) version with the pre-drawn circle has been widely used in research and clinical practice without standardized Turkish norms. The present study aimed to standardize CDT scores according to the most frequently used scoring methods in the literature (Manos-Wu and Shulman) and to estimate the validity and reliability of both methods.

Materials and Methods: The norm determination phase of the study was performed with 244 healthy individuals in the age range of 50-92 years. The effects of age, education, and sex on the CDT scores were examined and normative data stratified by age and education were derived. Test-retest reliability, inter-rater reliability, and concurrent validity and criterion validity of the two scoring methods were tested. Criterion validity was assessed using the ROC analysis to examine the extent to which the CDT scores could distinguish among healthy individuals, patients with mild cognitive impairment (MCI), and patients with Alzheimer's disease (AD).

Results: In the Manos-Wu method, age and education had significant effects on CDT scores of healthy individuals; however, in the Shulman method, only an education effect was found. Both methods had high test-retest reliability (Manos-Wu, 0.81; Shulman, 0.72) and inter-rater reliability (Manos-Wu, 0.98; Shulman, 0.96) and showed high correlations with each other and with other cognitive screening tests. The area under the ROC curve had high values only in discriminating between healthy individuals and AD, and healthy individuals and MCI.

Conclusion: In the present study, detailed normative data for individuals aged 50 and older were established. Age and education levels should be taken into account when interpreting CDT scores. Both scoring methods provided similar results in terms of high test-retest and inter-rater reliability, as well as high concurrent and criterion validity.

Keywords: Clock drawing test, normative data, validity-reliability, Alzheimer's disease, mild cognitive impairment

Yazışma Adresi/Address for Correspondence: Dr. Derya Durusu Emek-Savaş, Dokuz Eylül Üniversitesi Edebiyat Fakültesi, Psikoloji Bölümü, İzmir, Türkiye

Tel.: +90 536 950 20 50 E-posta: derya.emek@deu.edu.tr ORCID ID: orcid.org/0000-0001-7042-697X

Geliş Tarihi/Received: 21.09.2017 **Kabul Tarihi/Accepted:** 07.12.2017

©Telif Hakkı 2018 Türk Nöroloji Derneği
Türk Nöroloji Dergisi, Galenos Yayınevi tarafından basılmıştır.

Giriş

Nöropsikolojik değerlendirme, bilişsel işlevlerin değerlendirilmesinde altın standart olarak görülmektedir. Nöropsikolojik değerlendirmenin temelini oluşturan testler, bilişsel işlevleri etkileyen hastalıkların tanı, takip, tedavi planlama ve tedavi etkinliği değerlendirme süreçlerinde işlevsel uygulama alanı bulmaktadır. Bellek, dikkat, yanıt inhibisyonu, planlama, dil, soyut düşünme, konstrüksiyon (yapılandırma) ve praksi gibi pek çok bilişsel işlev nöropsikolojik testler ile değerlendirilebilmektedir (1). Ancak, yaş, cinsiyet ve eğitim gibi demografik değişkenler ile kültürel özelliklerin nöropsikolojik test skorlarını etkilediği bilinmektedir (2). Bu sebeple, kullanılan testlerin doğru yorumlanabilmesi için Türkçe'ye uyarlama ve standardizasyon çalışmalarının yapılmış olması büyük önem taşımaktadır.

Saat çizme testi (SÇT), 30 yılı aşkın süredir görsel-motor işlevler ile planlama, sıralama ve soyut düşünme becerilerini değerlendirmede kullanılan kolay, hızlı uygulanabilir, geçerlik ve güvenilirliği yüksek (3,4,5,6,7) bir kâğıt-kalem testidir.

SÇT'nin ilk kullanımından (8) bugüne, pek çok farklı uygulama ve puanlama biçimi geliştirilmiştir. Klasik bir SÇT uygulamasında kişiden bir saat çizmesi, rakamları düzgünce yerleştirilmesi ve son olarak da belirtilen zamanı (örneğin, 11:10, 15:00, 18:45) göstermesi istenmektedir. Uygulamadaki her bir basamak ayrı bir bilişsel işlev ile ilişkilidir. Bu sebeple, SÇT'nin uygulama biçimlerindeki farklılık, değerlendirilen bilişsel işleve odaklanma veya diğer işlevlerden birini göz ardı etme şeklinde sonuçlanabilmektedir. Literatürde SÇT'nin uygulama ve puanlama yöntemlerini karşılaştıran pek çok araştırma yer almaktadır. Farklı SÇT uygulamalarında saatin dairesi kişiye hazır olarak verilebilir (3,4,5,6,7,9) veya kişiden saatin dairesini çizmesi istenebilir (örneğin; 10). SÇT uygulamasında, kişiden bir (3,4,5,6) veya birkaç farklı zamanı (8) göstermesi istenebilir ya da saat gösterimi hiç değerlendirmeye alınmayabilir (6,9). Testin uygulama farklılıklarının yanı sıra literatürde pek çok farklı puanlama sistemi de yer almaktadır (11). Çalışmalarda en sık kullanılan puanlama yöntemlerinin başında Shulman, Manos ve Wu ve Sunderland gelmektedir. Shulman puanlama yönteminde, kişiye saatin dairesi hazır olarak verilir ve saatin 11:10'u göstermesi istenir. Puanlama 0-5 arasında yapılmakta olup, testten alınabilecek en yüksek puan 5'tir. Manos ve Wu yöntemi, toplam 10 puan üzerinden değerlendirme yapmaktadır ve uygulama açısından Shulman yöntemine benzer. Sunderland yönteminde ise saatin dairesini kişinin çizdiği ve bu aşamanın da puanlamaya dahil edildiği 10 puan üzerinden bir değerlendirme yapılır.

SÇT'nin Cangoz ve ark. (12) tarafından 50 yaş ve üzeri popülasyonda yapılmış bir Türkiye norm ve geçerlik-güvenirlik çalışması bulunmaktadır. Söz konusu çalışmada, saatin dairesini kişinin çizmesinin istendiği dört puanlık bir SÇT puanlama yöntemi Türkiye örneklemini için standardize edilmiştir. Cangoz ve ark. (12), bu yöntemi demans olgularını ayırt etmede işlevsel oluşu ve kolay bir puanlama sistemi olması sebebiyle tercih ettiğini bildirmiştir. Ancak, Türkiye'de pek çok klinikte dairenin hazır olarak sunulduğu SÇT formu kullanılmaktadır ve bu uygulama biçimine ait Türkçe standardizasyonu yapılmış bir puanlama sistemi bulunmamaktadır. Birçok çalışmada SÇT puanları üzerinde yaş (13,14,15) ve eğitim (14,16,17) etkilerinin

gösterilmiş olması, özellikle klinik alanda norm verilerine duyulan ihtiyacı artırmıştır.

SÇT'nin özellikle demans sendromlarını ayırt etmede yüksek duyarlılık ve özgüllük değerlerine sahip olduğu bildirilmiştir (13,18). Shulman yönteminin duyarlılık ve özgüllüğünün sırasıyla %86 ve %72, Manos ve Wu yönteminin ise sırasıyla %76 ve %78 olduğu bildirilmiştir (18). Ülkemizde yapılmış bir çalışmada, üç farklı SÇT puanlama yönteminin demans tanısı olan hastalar ile demansı olmayan yaşlıları ayırt etme gücü incelenmiş ve Shulman yöntemi için yüksek duyarlılık ve özgüllük değerleri bildirilmiştir (19). Aynı çalışmada, SÇT'nin mini mental durum testi (MMDT) ile birlikte kullanımının duyarlılığı %100'e yükseldiği belirtilmektedir. Buna karşın, SÇT'nin ayırt edici geçerliğinin Alzheimer hastalığına (AH) bağlı hafif kognitif bozukluk (HKB) için daha düşük olduğu bulunmuştur (13).

Bu çalışmada, ülkemizde çok sayıda merkezde uygulanmakta olup Türkiye popülasyonuna ait norm değerleri bulunmayan SÇT versiyonunun, uluslararası literatürde en yaygın kullanıma sahip Shulman ile Manos ve Wu puanlama yöntemlerinin standardizasyonunun yapılması ve geçerlik ve güvenilirliğinin test edilmesi amaçlanmıştır. Ayrıca, bu puanlama yöntemlerinin HKB ve AH grupları için duyarlılık ve özgüllük değerleri belirlenmiş ve ayırt edicilik düzeyleri karşılaştırılmıştır.

Gereç ve Yöntem

Örneklem

Bu çalışmanın norm belirleme aşamasına, Dokuz Eylül Üniversitesi Sağlık Bilimleri Enstitüsü Sinirbilimler Anabilim Dalı'nda 2011-2017 yılları arasında boylamsal olarak takip edilmiş olan, klinik muayene, beyin görüntüleme ve nöropsikolojik değerlendirme sonuçlarına göre herhangi bir kognitif bozukluğu olmadığı belirlenen 50 yaş ve üzeri 244 sağlıklı birey dahil edilmiştir.

Uzman nöropsikologlar tarafından bir yıl ara ile rutin olarak uygulanan nöropsikolojik test bataryasında, (1) genel kognitif durum; MMDT (20,21), (2) sözel ve görsel bellek; Öktem Sözel Bellek Süreçleri Testi (22) ve WMS-R görsel üretim alt testi (23), (3) dikkat; WMS-R ileri ve geri sayı menzili testi (23), (4) yürütücü işlevler; WAIS-IV benzerlikler alt testi (24), sözel akıcılık testleri (semantik ve fonemik akıcılık) (25), SÇT ve Stroop testi (26), (5) görsel-mekansal işlevler; basit şekil kopyalama testleri ve (6) dil becerileri; Boston Adlandırma testi (27) ile değerlendirilmektedir.

Örneklem için eğitim düzeyi en az okur-yazar olarak belirlenmiştir. MMDT skorları 27'nin altında olan sağlıklı bireyler çalışmaya dahil edilmemiştir. Ayrıca, kognitif becerileri etkileyebilecek kontrol altında olmayan hastalık öyküsü, kafa travması öyküsü ve düzeltilmemiş görme bozukluğu varlığı dışlama kriteri olarak belirlenmiştir. Çalışmaya dahil edilen tüm bireyler, geçmiş çalışmalarda depresyon öyküsü ve depresif belirtiler açısından hem bir klinisyen tarafından hem de Yesavage Geriatrik Depresyon Ölçeği (GDÖ) (28) aracılığıyla değerlendirilmiştir. Yesavage GDÖ skoru 13 ve üzeri olan bireyler çalışmaya dahil edilmemiştir. Normatif verinin elde edildiği sağlıklı bireylerin demografik özellikleri Tablo 1'de sunulmuştur.

Geçerlik ve güvenilirlik analizlerinin gerçekleştirilmesi amacıyla, rutin olarak ortalama 12 ay arayla nörolojik muayene, nöropsikolojik testler ve yapısal manyetik rezonans görüntüleme

(MRG) ile takip edilen 128 amnestik HKB olgusu ve 109 Alzheimer hastası çalışmaya dahil edilmiştir. HKB tanısı için Albert ve ark. (29) tanı kriterleri, AH tanısı için McKhann ve ark. (30) tanı kriterleri kullanılmıştır. Çalışmanın çeşitli aşamalarında, farklı sayıda HKB ve Alzheimer olgusu yer almıştır.

Çalışmada yer alan sağlıklı bireyler, HKB olguları ve Alzheimer hastalarının geçmiş yıllara ait SÇT performansları, bu çalışma kapsamında puanlanarak analizlere dahil edilmiştir. Tüm katılımcılardan katılmış oldukları araştırmalar kapsamında yazılı bilgilendirilmiş onam alınmış olup, bu çalışma protokolü Dokuz Eylül Üniversitesi Etik Kurul Komitesi tarafından 13.07.2017 tarihinde 2017/18-01 karar numarası ile onaylanmıştır.

Materyaller

Bu çalışma kapsamında katılımcıların geçmiş yıllara ait SÇT performansları Manos ve Wu (5) ile Shulman (4) puanlama yöntemleri kullanılarak değerlendirilmiştir. Kullanılan SÇT versiyonunda, kağıt üzerinde 10 cm çapında boş bir daire verilmekte ve bireyden saatin rakamlarını yerleştirerek, saat 11'i 10 geçeyi göstermesi istenmektedir. Saat gösterimi ile ilgili hataların erken kognitif bozukluğu saptamada işlevsel olabileceği düşünülmektedir (31). Bu sebeple, yönerge verilirken saat gösterimi ile ilgili olarak akrep ve yelkovan sözcükleri, saatin kolları gibi terimlerin söylenmemesine özellikle dikkat edilmiştir. İhtiyaç duyulduğunda, hatırlatma amacıyla yönergenin tamamı veya bir kısmı aynı şekilde tekrar edilmiştir. Seçilen puanlama yöntemleri, saat dairesinin hazır olarak verildiği, saatin 11'i 10 geçeyi işaret ettiği ve kopyalama koşulunun bulunmadığı yöntemler olmaları sebebiyle tercih edilmiştir. Manos ve Wu ile Shulman puanlama yöntemleri alt başlıklarda ayrıntılı olarak tanıtılmıştır.

Manos ve Wu Puanlama Yöntemi

Manos ve Wu (5) tarafından geliştirilen puanlama yöntemi, literatürde "Ten Point Clock Test" olarak yer almaktadır. Bu yöntemde, toplam 10 puan üzerinden bir değerlendirme yapılmakta olup, şeffaf bir daire şablonu, sekiz eşit dilime bölünerek değerlendirmeye yardımcı olarak kullanılmaktadır (Ek 1). Puanlamada 1, 2, 4, 5, 7, 8, 10 ve 11 rakamları eğer doğru saat dilimi içinde yer alıyorsa her bir doğru rakam için 1 puan verilir. Bir puan, kısa çizgi ile yapılmış 11 rakamını işaret eden akrep ve bir puan uzun çizgi ile yapılmış 2 rakamını işaret eden yelkovan için verilir. Eğer kolların uzunlukları hatalı ise puan verilmez. Puanlama formülü [8 (rakamlar) + 1 (akrep) + 1 (yelkovan)]=10 şeklinde ifade edilebilir.

SÇT performanslarının ayrıntılı incelenebilmesi için bireylerin toplam SÇT puanlarına ek olarak, 8 puan üzerinden rakam yerleştirme puanları ve 2 puan üzerinden saat gösterimi puanları da hesaplanmıştır.

Shulman Puanlama Yöntemi

Bu yöntemde SÇT performansı 0-5 aralığında puanlanmakta olup, testten alınabilecek en yüksek puan 5'tir. Beş puan, saat gösterimi ve rakam yerleşimleri kusursuzsa; 4 puan, saat gösterimi doğru ve minimal görsel-uzamsal hatalar varsa; 3 puan, görsel-uzamsal yerleşim iyi olmasına karşın saat gösterimi bozüksa; 2 puan, görsel-uzamsal yerleşim doğru saat gösterimine izin vermeyecek kadar bozüksa; 1 puan, görsel-uzamsal yerleşim ileri düzeyde bozüksa ve 0 puan, herhangi bir saat temsili bulunmuyorsa verilmektedir.

Shulman yönteminde herhangi bir alt puan türü bulunmadığı için, bireylerin yalnızca toplam puanları hesaplanmış ve değerlendirilmeye alınmıştır.

İstatistiksel Analiz

İstatistiksel analizler SPSS 20.0 programı ile gerçekleştirilmiştir. ROC eğrisi analizleri MedCalc 15.8 programı ile yapılmıştır.

Norm belirleme: Çalışmanın norm belirleme aşamasında öncelikle yaş, eğitim ve cinsiyet değişkenlerinin SÇT puanları üzerindeki etkisi adım adım (stepwise) çoklu regresyon analizi ile değerlendirilmiştir. Norm değerleri, SÇT skorları üzerinde etkisi saptanan yaş ve eğitim değişkenlerine göre oluşturulmuştur. Yaş ve eğitim grupları arasındaki istatistiksel farklar, iki faktörlü ANOVA ile incelenmiştir. 3x3 ANOVA desenine göre, grup içi faktörler yaş [(3 düzey): 50-59, 60-69 ve 70 yaş ve üzeri] x eğitim [(3 düzey): 0-5 yıl, 6-11 yıl ve 12 yıl ve üzeri] olmak üzere analiz edilmiştir. ANOVA, her iki yöntemin toplam puan ve alt puan türleri için tekrarlanmıştır. Post hoc analizlerde, yaş ve eğitim ana etkileri bağımsız örneklem t-testi ile, yaş x eğitim etkileşim etkisi ise eşleştirilmiş örneklem t-testi ile incelenmiştir.

Her iki puanlama yönteminin toplam puanları ve alt puanları için norm değerleri üç yaş grubu (50-59, 60-69 ve 70 yaş üzeri) ve üç eğitim grubuna (0-5 yıl, 6-11 yıl ve 12 yıl ve üzeri) göre oluşturulmuş ve her bir puan türü için gruplara ait ortalama ve standart sapma değerleri bildirilmiştir.

Test-tekrar test güvenilirliği: Test-tekrar test güvenilirliğinin değerlendirilmesi amacıyla, boylamsal olarak takip edilen sağlıklı bireyler arasından rastgele seçilen 50 katılımcının ortalama 12 ay

Tablo 1. Normatif örneklemin demografik özellikleri

	Toplam (n=244)		Yaş grupları					
	Aralık	Ortalama	50-59 yaş (n=58)		60-69 yaş (n=97)		+70 yaş (n=89)	
			Aralık	Ortalama	Aralık	Ortalama	Aralık	Ortalama
Yaş (yıl)	50-92	66,06±8,67	50-59	55,07±2,82	60-69	64,31±3,25	70-92	75,12±4,96
Eğitim (yıl)	0-23	12,11±4,32	7-21	13,62±2,62	0-23	11,87±4,57	0-21	11,40±4,73
Cinsiyet (E/K)	89/155		15/43		42/55		32/57	
El tercihi (Sağ/Sol)	239/5		58/0		94/3		87/2	
MMDT	27-30	29,04±1,09	27-30	29,22±0,90	27-30	29,05±1,07	27-30	28,92±1,19

Veri, ortalama ± standart sapma olarak sunulmuştur. E: Erkek, K: Kadın, MMDT: Mini mental durum testi

ara ile uygulanmış olan 1. ve 2. SÇT performansları güncel çalışma kapsamında puanlanmıştır. Birinci ve ikinci değerlendirmeler arasındaki uyum, sınıf içi korelasyon [Inter class correlation (ICC)] analizi ile değerlendirilmiştir. ICC analizlerinde, iki yönlü rastgele model (two-way mixed model) ile mutlak uyum test edilmiştir.

Değerlendiriciler arası güvenilirlik: Değerlendiriciler arası güvenirliliğin test edilmesi için 35 sağlıklı birey, 33 HKB olgusu ve 32 Alzheimer hastası olmak üzere toplam 100 katılımcının SÇT performansı ikinci bir değerlendirici tarafından bağımsız olarak puanlanmıştır. İki değerlendirici arasındaki mutlak uyum ICC analizi ile incelenmiştir. ICC analizlerinde, iki yönlü rastgele model (two-way mixed model) ile mutlak uyum test edilmiştir. Değerlendiriciler arası güvenirlilik sonuçları, sınıf içi korelasyon katsayıları tüm gruplar bir arada ve her bir grup ayrı ayrı olmak üzere hesaplanarak raporlanmıştır.

Eş-zaman geçerliliği: SÇT puanlarının (her iki puanlama yöntemi için) tüm katılımcılara eş zamanlı olarak uygulanmış olan MMDT puanları ile korelasyonu incelenmiştir. Korelasyon analizlerinde Pearson korelasyon katsayısı sağlıklı katılımcılar, HKB olguları ve Alzheimer hastalarının tümü ve her bir grup için ayrı ayrı hesaplanmıştır. Buna ek olarak, puanlama yöntemlerinin birbirleri ile korelasyonu incelenmiştir.

Ölçüt geçerliliği: Manos ve Wu ile Shulman yöntemlerine göre hesaplanan toplam puanların sağlıklı-HKB, sağlıklı-AH ve HKB-AH gruplarını birbirinden ayırt etme gücü ROC eğrisi yöntemi ile değerlendirilmiştir. SÇT skorlarının eğitim ve yaş değişkenlerinden etkilendiği bilindiği için, sağlıklı-HKB, sağlıklı-AH ve HKB-AH grupları yaş ve eğitim açısından eşleştirilmiş bireylerden oluşturularak karşılaştırılmıştır. Gruplar için kesme noktaları, duyarlılık ve özgüllük değerleri belirlenmiş, ölçüt geçerliliği test edilmiştir.

Bulgular

Norm Belirleme

Yaş, cinsiyet ve eğitim düzeylerinin SÇT skorları üzerindeki etkisi, adım adım (stepwise) regresyon analizi ile incelenmiştir. Öncelikle, Manos ve Wu puanlama sistemine göre toplam puan, saat gösterimi ve rakam yerleştirme skorları değerlendirilmiştir. Manos ve Wu yöntemine göre hesaplanan toplam SÇT puanlarını yordamada, yaş ve eğitim düzeylerinin %6,9 düzeyinde katkısı olduğu saptanmıştır [$F_{(2,244)}=9,111$, $p<0,001$]. Cinsiyet değişkeninin modele anlamlı bir katkısı saptanmamıştır. Manos ve Wu yöntemine göre, hesaplanan toplam SÇT puanları artan her yaş (yıl) ile 0,011 puan azalmakta ve eğitim düzeyindeki her bir yıllık artış ile 0,030 puan artmaktadır. Saat gösterimi skorlarının toplam varyansının %4,6'sını eğitim düzeyi açıklamakta [$F_{(1,245)}=11,848$, $p=0,001$] ve bu puan türü üzerinde yaş veya cinsiyetin modele anlamlı bir katkısı bulunmamaktadır. Rakam yerleştirme skorlarının toplam varyansının %3,1'ini ise yaş açıklamakta [$F_{(1,245)}=7,814$, $p=0,006$] ve bu puan türü üzerinde ise eğitim veya cinsiyetin katkısı bulunmamaktadır.

Regresyon analizleri, Shulman yöntemi ile hesaplanan toplam SÇT puanları için tekrarlanmış ve sadece eğitim düzeyinin modele anlamlı katkısı saptanmıştır. Eğitim düzeyi, Shulman toplam puanlarının %3,5'ini açıklamaktadır [$F_{(1,245)}=9,006$, $p=0,003$].

Yaş (50-59, 60-69 ve 70 yaş üzeri) ve eğitim (0-5 yıl, 6-11 yıl ve 12 yıl ve üzeri) değişkenlerinin düzeyleri arasındaki istatistiksel farklar, iki faktörlü ANOVA ile değerlendirilmiştir. Her iki yöntemte ait toplam puan ve Manos ve Wu yöntemine ait alt puan türleri üzerinde cinsiyet etkisi gözlenmediği için, ANOVA desenine cinsiyet değişkeni dahil edilmemiş, analizler yalnızca yaş ve eğitim düzeyleri ile oluşturulan 3x3 desene göre gerçekleştirilmiştir.

ANOVA analizinde, Manos ve Wu toplam puanları üzerinde yaş ana etkisi gözlenmiştir [$F_{(2,236)}=4,969$, $p=0,008$]. Post hoc analizlerde 70 yaş ve üzeri bireylerin SÇT puanlarının 50-59 yaş grubu ($p=0,003$) ve 60-69 yaş grubu ($p=0,006$) puanlarından anlamlı olarak daha düşük olduğu bulunmuştur. 50-59 ve 60-69 yaş grupları arasında ise fark saptanmamıştır ($p=0,633$).

Ayrıca, Manos ve Wu toplam puanları üzerinde eğitim ana etkisi gözlenmiştir [$F_{(2,236)}=4,253$, $p=0,015$]. Post hoc analizlerde, 12 yıl ve üzeri eğitime sahip bireylerin toplam puanlarının 0-5 yıl ($p=0,016$) ve 6-11 yıl ($p=0,026$) eğitilmiş bireylere kıyasla daha yüksek olduğu saptanmıştır. Eğitim düzeyi 0-5 yıl ve 6-11 yıl arasında olan bireylerin puanları arasında ise anlamlı fark gözlenmemiştir ($p=0,379$).

Manos ve Wu rakam yerleştirme skorları üzerinde yaş ana etkisi [$F_{(2,236)}=3,748$, $p=0,025$] saptanırken, eğitim ana etkisi [$F_{(2,236)}=2,415$, $p=0,092$] gözlenmemiştir. Post hoc analizlerde, 70 yaş ve üzeri bireylerin rakam yerleştirme puanlarının, 50-59 yaş grubu ($p=0,022$) ve 60-69 yaş grubu ($p=0,010$) puanlarından anlamlı olarak daha düşük olduğu bulunmuştur. 50-59 ve 60-69 yaş grupları arasında ise fark saptanmamıştır ($p=0,763$). Ayrıca, rakam yerleştirme puanları üzerinde yaş x eğitim etkileşim etkisi saptanmıştır [$F_{(3,236)}=4,020$, $p=0,008$]. 6-11 yıl eğitime sahip 70 yaş ve üzerindeki bireylerin rakam yerleştirme skorları, 50-59 yaş ($p=0,024$) ve 60-69 yaş ($p=0,030$) grubundaki bireylere kıyasla daha düşük bulunmuştur. Diğer eğitim ve yaş grupları arasında ise herhangi bir fark gözlenmemiştir ($p>0,05$).

Manos ve Wu saat gösterimi skorları üzerinde eğitim ana etkisi gözlenmiş [$F_{(2,236)}=3,809$, $p=0,024$]; ancak, yaş ana etkisi bulunmamıştır [$F_{(2,236)}=2,409$, $p=0,092$]. Eğitim düzeyi 0-5 yıl arasında olan bireylerin saat gösterimi puanları, 12 yıl ve üzeri eğitime sahip bireylerden anlamlı olarak daha düşük bulunmuştur ($p=0,030$). Diğer eğitim düzeyleri arasında ise fark gözlenmemiştir ($p>0,05$).

Shulman yöntemi ile hesaplanan toplam SÇT skorları üzerinde yaş [$F_{(2,236)}=3,248$, $p=0,041$] ve eğitim [$F_{(2,236)}=3,510$, $p=0,031$] ana etkileri bulunmuştur. Yetmiş yaş ve üzeri bireylerin toplam puanları, hem 50-59 yaş aralığındaki ($p=0,030$) hem de 60-69 yaş aralığındaki ($p=0,013$) bireylere kıyasla daha düşüktür. Ayrıca, 0-5 yıl arası eğitime sahip grubun puanları, 12 yıl ve üzeri eğitime sahip bireylerden anlamlı olarak daha düşüktür ($p=0,032$). Her iki yöntemte ait toplam puanların ANOVA bulguları Tablo 2'de özetlenmiştir.

Bu çalışmada, 50-59 yaş aralığında yer alan ve 0-5 yıl eğitime sahip popülasyona ulaşılmakta güçlük yaşanmış ve bu grupta yer alan 3 sağlıklı bireyin verileri analizlerden çıkarılmıştır. Bu durumun, Türkiye'de ilkökul eğitiminin zorunlu hale getirilmiş olması ile ilişkili olduğu düşünülmüştür. Bahsedilen yaş grubuna ait norm değerleri oluşturulmamış, 50-59 yaş aralığındaki bireylerin normları yalnızca 6-11 yıl ve 12 yıl ve üzeri eğitim için bildirilmiştir. Elde edilen norm verileri, Manos ve Wu ile Shulman yöntemleri için Tablo 3, 4, 5 ve 6'da sunulmuştur.

Test-Tekrar Test Güvenirliği

Manos ve Wu yöntemi ile hesaplanan toplam SÇT puanlarının test-tekrar test güvenirligi katsayısı 0,806 ve Shulman yöntemi ile hesaplanan SÇT puanlarının test-tekrar test güvenirligi katsayısı 0,720 olarak bulunmuştur. Ayrıntılı puan türlerine ait güvenirlilik katsayıları ve %95 güven aralıkları Tablo 7'de sunulmuştur.

Değerlendiriciler Arası Güvenirlik

Değerlendiriciler arası güvenirlilik sağlıklı bireyler, HKB olguları ve Alzheimer hastalarından oluşan toplam 100 katılımcının SÇT puanlarının iki bağımsız değerlendirici tarafından puanlanması ile hesaplanmıştır. Tüm katılımcılar bir arada değerlendirildiğinde, Manos ve Wu yöntemi ile hesaplanan toplam puanların değerlendiriciler arası anlaşma katsayısı 0,892 ve Shulman yöntemi ile hesaplanan toplam puanların değerlendiriciler arası anlaşma katsayısı 0,957 olarak saptanmıştır. Ayrıntılı puan türlerine ait güvenirlilik katsayıları ve %95 güven aralıkları Tablo 8'de sunulmuştur.

Eş-Zaman Geçerliliği

Korelasyon analizleri, sağlıklı katılımcılar, HKB olguları ve Alzheimer hastaları bir arada ve gruplar ayrı ayrı incelenerek gerçekleştirilmiştir. Tüm katılımcıların dahil edildiği analizlerde, MMDT puanları ile Manos ve Wu toplam puanları ($r=0,797$, $p<0,001$) ve Shulman toplam puanları arasında ($r=0,760$, $p<0,001$) yüksek düzeyde korelasyon saptanmıştır. Tüm puan türlerine ve gruplara ait korelasyon bulguları Tablo 9'da sunulmuştur. Ayrıca, Manos ve Wu yöntemi ve Shulman yöntemi ile hesaplanan toplam SÇT puanları arasında güçlü bir korelasyon gözlenmiştir ($r=0,929$, $p<0,001$).

Ölçüt Geçerliliği

Her iki puanlama yöntemine göre hesaplanan toplam puanların sağlıklı-HKB, sağlıklı-AH ve HKB-AH gruplarını birbirinden ayırt etme gücü ROC eğrisi yöntemi ile incelenmiştir. Yaş ve eğitim açısından eşleştirilen gruplara ait demografik bilgiler Tablo 10'da sunulmuştur.

Her iki yöntemin karşılaştırıldığı ROC eğrileri gruplar arası karşılaştırmaların her biri için hazırlanmıştır (Şekil 1). ROC eğrisi altında kalan alan (area under curve) ve %95 güven aralıkları her iki yöntem için Tablo 11'de gösterilmiştir. Manos ve Wu yönteminin sağlıklı-HKB ($p<0,0001$) ve sağlıklı-AH ($p<0,0001$) gruplarını ayırt etme gücü anlamlı olarak Shulman yönteminden daha yüksektir. HKB-AH gruplarını ayırt etme gücü açısından ise iki yöntem arasında anlamlı fark bulunmamaktadır ($p=0,721$).

ROC eğrisi analizlerinde hiçbir grup karşılaştırmasında %95 güven aralığı 0,5 değerini içermemektedir. Bu durum, puanlama yöntemlerinin ayırt etme güçlerinin şansa bağlı olmadığını ifade etmektedir.

Analizlerde, Manos ve Wu ile Shulman puanlama yöntemleri için kesme puanları ve karşılık gelen duyarlılık ve özgüllük değerleri hesaplanmıştır. Manos ve Wu puanlama yönteminde kesme noktası 9/10 olarak belirlendiğinde, testin HKB olgularını sağlıklı bireylerden ayırt etme duyarlılığının %80,8, özgüllüğünün %80,0; Alzheimer hastalarını sağlıklı bireylerden ayırt etme duyarlılığının %89,8, özgüllüğünün %76,8 olduğu gözlenmiştir. HKB ve AH olgularını ayırt etmede, kesme noktası 7/10 olarak belirlendiğinde, duyarlılık ve özgüllük değerleri sırasıyla %62,9 ve %61,9 olarak bulunmuştur.

Shulman puanlama yönteminde ise 4/5 kesme noktasında testin HKB olgularını sağlıklı bireylerden ayırt etme duyarlılığının

Tablo 2. Demografik değişkenlerin saat çizme testi toplam puanları üzerindeki etkisi

	Değişkenler	Ort.± SS	F	Post hoc
Manos ve Wu toplam puan	Yaş			
	A: 50-59	9,90±0,35	$F_{(2,236)}=4,969$, $p=0,008$	A=B, $p=0,633$
	B: 60-69	9,87±0,42		A>C, $p=0,003^*$
	C: >70	9,57±0,90		B>C, $p=0,006^*$
	Eğitim			
	A: 0-5	9,51±0,85	$F_{(2,236)}=4,253$, $p=0,015$	A=B, $p=0,379$
	B: 6-11	9,66±0,80		B<C, $p=0,026^*$
	C: >12	9,89±0,44		A<C, $p=0,016^*$
	Shulman toplam puan	Yaş		
A: 50-59		4,90±0,35	$F_{(2,236)}=4,488$, $p=0,012$	A=B, $p=0,858$
B: 60-69		4,91±0,35		A>C, $p=0,030^*$
C: >70		4,72±0,62		B>C, $p=0,013^*$
Eğitim				
A: 0-5		4,63±0,73	$F_{(2,236)}=3,188$, $p=0,043$	A=B, $p=0,222$
B: 6-11		4,80±0,50		B=C, $p=0,084$
C: >12		4,91±0,36		A<C, $p=0,032^*$

Veri, ortalama ± standart sapma olarak sunulmuştur. Ort.: Ortalama, SS: Standart sapma, * $p<0,05$

%81,5, özgüllüğünün %79,6; Alzheimer hastalarını sağlıklı bireylerden ayırt etme duyarlılığının %61,6, özgüllüğünün %83,2 olduğu saptanmıştır. HKB ve AH grupları için 3/5 olarak belirlenen

kesme noktası, düşük duyarlılık ve özgüllük değerleri göstermiştir (sırasıyla, %61,9 ve %59,8). Sağlıklı-HKB, sağlıklı-AH ve HKB-AH grupları için kesme puanları Tablo 12'de özetlenmiştir.

Tablo 3. Manos ve Wu yöntemi ile hesaplanan saat gösterimi puanları				
Eğitim düzeyi (yıl)	Yaş grupları (yıl)			Tüm yaş grupları
	50-59	60-69	70+	
0-5 yıl	-	1,80±0,56	1,55±0,83	1,66±0,73
N		(15)	(20)	(35)
6-11 yıl	1,94±0,24	1,84±0,45	1,77±0,59	1,84±0,47
N	(17)	(31)	(26)	(74)
+12 yıl	1,95±0,31	1,98±0,14	1,88±0,45	1,94±0,32
N	(41)	(51)	(43)	(135)
Tüm eğitim grupları	1,95±0,29	1,91±0,36	1,78±0,60	1,87±0,45
N	(58)	(97)	(89)	(244)

Veri, ortalama ± standart sapma olarak sunulmuştur

Tablo 5. Manos ve Wu yöntemi ile hesaplanan toplam saat çizme testi puanları				
Eğitim düzeyi (yıl)	Yaş grupları (yıl)			Tüm yaş grupları
	50-59	60-69	70+	
0-5 yıl	-	9,60±0,74	9,45±0,95	9,51±0,85
N		(15)	(20)	(35)
6-11 yıl	9,88±0,33	9,81±0,48	9,35±1,16	9,66±0,80
N	(17)	(31)	(26)	(74)
+12 yıl	9,90±0,37	9,98±0,14	9,77±0,65	9,89±0,44
N	(41)	(51)	(43)	(135)
Tüm eğitim grupları	9,90±0,36	9,87±0,42	9,57±0,90	9,77±0,65
N	(58)	(97)	(89)	(244)

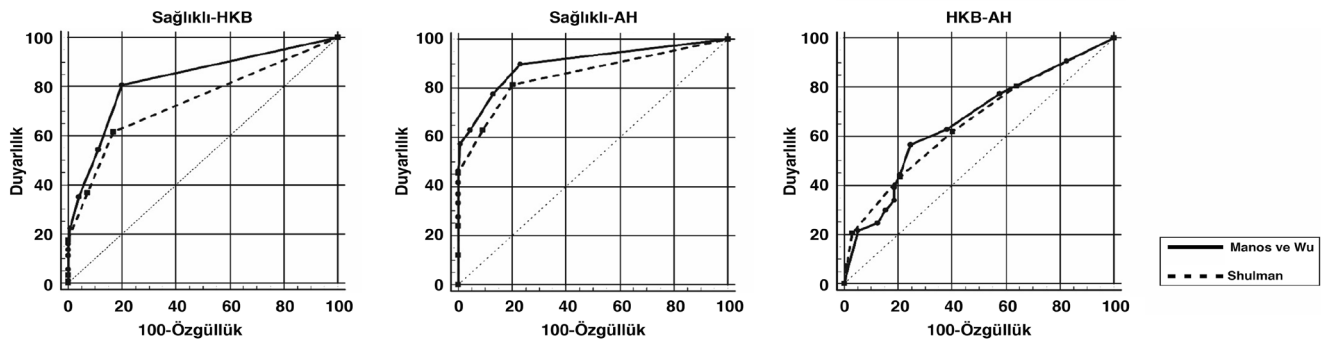
Veri, ortalama ± standart sapma olarak sunulmuştur

Tablo 4. Manos ve Wu yöntemi ile hesaplanan rakam yerleştirme puanları				
Eğitim düzeyi (yıl)	Yaş grupları (yıl)			Tüm yaş grupları
	50-59	60-69	70+	
0-5 yıl	-	7,80±0,41	7,90±0,31	7,86±0,36
N		(15)	(20)	(35)
6-11 yıl	7,94±0,24	7,97±0,18	7,58±0,81	7,82±0,53
N	(17)	(31)	(26)	(74)
+12 yıl	7,95±0,22	8,00±0,00	7,88±0,39	7,95±0,25
N	(41)	(51)	(43)	(135)
Tüm eğitim grupları	7,95±0,22	7,96±0,20	7,80±0,55	7,90±0,376
N	(58)	(97)	(89)	(244)

Veri, ortalama ± standart sapma olarak sunulmuştur

Tablo 6. Shulman yöntemi ile hesaplanan toplam saat çizme testi puanları				
Eğitim düzeyi (yıl)	Yaş grupları (yıl)			Tüm yaş grupları
	50-59	60-69	70+	
0-5 yıl	-	4,73±0,59	4,55±0,83	4,63±0,73
N		(15)	(20)	(35)
6-11 yıl	4,94±0,24	4,84±0,45	4,65±0,63	4,80±0,50
N	(17)	(31)	(26)	(74)
+12 yıl	4,88±0,40	5,00±0,00	4,84±0,49	4,91±0,36
N	(41)	(51)	(43)	(135)
Tüm eğitim grupları	4,90±0,36	4,91±0,36	4,72±0,62	4,84±0,48
N	(58)	(97)	(89)	(244)

Veri, ortalama ± standart sapma olarak sunulmuştur



Şekil 1. İki farklı saat çizme testi puanlama yöntemi için ROC eğrileri
HKB: Hafif kognitif bozukluk, AH: Alzheimer hastalığı

Tartışma

SÇT dünyada olduğu gibi, Türkiye'de de klinisyenler ve araştırmacılar tarafından oldukça yaygın olarak kullanılan bir testtir; buna rağmen, testin ülkemizde en sık kullanılan uygulama biçiminin Türkiye popülasyonunda bir norm çalışmasının olmaması dikkat çekicidir. Bu çalışmanın amacı, 50 yaş ve üzeri Türkiye popülasyonunda Manos ve Wu (5) ile Shulman (4)

Tablo 7. İki uygulama yöntemine ait test-tekrar test anlaşma (agreement) katsayıları		
Manos ve Wu (5)	Anlaşma	%95 Güven aralığı
Rakam yerleştirme	0,950***	0,914-0,971
Saat gösterimi	0,330*	0,063-0,558
Toplam puan	0,806***	0,681-0,885
Shulman (4)		
Toplam puan	0,720**	0,554-0,831

Inter class correlation katsayısı: *Zayıf, **İyi, ***Mükemmel

Tablo 8. İki uygulama yöntemine ait değerlendiriciler arası anlaşma (agreement) katsayıları			
Puanlama yöntemleri	Gruplar	Güvenirlilik katsayısı	
Manos ve Wu (5)		Anlaşma	%95 Güven aralığı
Rakam yerleştirme	A	0,979	0,968-0,986
	B	0,892	0,797-0,944
	C	0,958	0,913-0,979
	D	0,983	0,966-0,992
Saat gösterimi	A	0,931	0,899-0,953
	B	0,460	0,165-0,683
	C	0,965	0,930-0,982
	D	0,910	0,825-0,955
Toplam puan	A	0,892	0,974-0,988
	B	0,825	0,683-0,907
	C	0,970	0,932-0,986
	D	0,984	0,968-0,992
Shulman (4)		Anlaşma	%95 Güven aralığı
Toplam puan	A	0,957	0,922-0,975
	B	0,790	0,625-0,888
	C	0,950	0,897-0,976
	D	0,948	0,814-0,980

A: Tüm katılımcılar (n=100), B: Sağlıklı bireyler (n=35), C: Hafif kognitif bozukluk olguları (n=33), D: Alzheimer hastaları (n=32)

puanlama yöntemlerine göre SÇT norm değerlerinin belirlenmesi ve konuyla ilgili klinisyen ve araştırmacıların kullanımına sunulmasıdır. Buna ek olarak kullanılan puanlama yöntemlerinin geçerlik ve güvenilirliği test edilmiştir.

Saat çizme görevi oldukça basit, anlaşılabilir ve uygulanabilir görünmekle birlikte, görevin başarıyla yerine getirilebilmesi için pek çok kognitif işlevin eş zamanlı etkileşimine ihtiyaç duyulmaktadır (4,32,33). Bu sebeple, farklı beyin bölgelerinin hasarı veya bu bölgeler ile ilişkili hastalıklarda farklı SÇT hata profilleri gözlenmektedir. Çalışmalar, sağ hemisfer hasarına bağlı olarak görsel-uzamsal hataların, sol hemisfer hasarına bağlı olarak ise saat gösterimi ile ilgili hataların meydana geldiğini bildirmektedir (33,34). Bu bilgilerden hareketle, Manos ve Wu puanlama yöntemi için, demografik değişkenlerin toplam puan üzerine etkisinin yanı sıra rakam yerleştirme ve saat gösterimi alt puan türleri üzerindeki etkileri de incelenmiştir. Bu çalışmada, Manos ve Wu toplam puanları üzerinde yaş ve eğitim etkisi bulunurken, rakam yerleştirme puanı üzerinde yalnızca eğitim etkisi ve saat gösterimi puanı üzerinde yalnızca yaş etkisi gözlenmiştir. Artan yaşla birlikte toplam puan ve saat gösterimi puanlarının azaldığı, buna karşın, eğitim düzeyi yükseldikçe toplam puan ve rakam yerleştirme puanlarının arttığı bulunmuştur. Shulman yöntemi ile hesaplanan SÇT puanları

Tablo 9. Saat çizme testi ile mini mental durum testi puanları arasındaki korelasyonlar			
	Tüm katılımcılar (n=484)	HKB olguları (n=128)	Alzheimer hastaları (n=109)
MMDT			
MW saat gösterimi	0,619	0,283	0,535
MW rakam yerleştirme	0,774	0,644	0,687
MW toplam puan	0,787	0,614	0,706
SH toplam puan	0,760	0,519	0,705

HKB: Hafif kognitif bozukluk, MW: Manos ve Wu puanlama yöntemi, SH: Shulman puanlama yöntemi, MMDT: Mini mental durum testi

Tablo 10. ROC analizleri için eşleştirilmiş grupların demografik özellikleri				
	Sağlıklı katılımcılar (n=125)	HKB olguları (n=125)	P	
Yaş	72,31±6,20	73,72±6,26	0,075	
Eğitim	10,66±4,74	10,09±4,93	0,347	
		Sağlıklı katılımcılar (n=108)	Alzheimer hastaları (n=108)	P
Yaş		70,39±8,61	70,57±8,76	0,876
Eğitim		9,48±4,31	8,55±4,39	0,115
		HKB olguları (n=97)	Alzheimer hastaları (n=97)	P
Yaş		73,36±6,46	72,51±7,17	0,384
Eğitim		9,11±4,43	8,70±4,39	0,516

Veri, ortalama ± standart sapma olarak sunulmuştur. HKB: Hafif kognitif bozukluk

üzerinde de yaş ve eğitim etkisi bulunduğu tespit edilmiştir. Hiçbir puan türü üzerinde cinsiyet etkisi gözlenmemiştir. Diğer bir deyişle, bu bulgular, sağlıklı bireylerde artan yaşın, daha çok soyut düşünme ve kavramsallaştırma becerilerini, düşük eğitim düzeyinin ise sıralama ve planlama becerilerini olumsuz yönde etkilediğine işaret etmektedir.

Her iki puanlama yöntemi için norm verileri, yaş ve eğitim düzeylerine göre oluşturulmuştur. Her iki puanlama yönteminde de 70 yaş ve üzeri bireylerin SÇT puanları diğer yaş gruplarından anlamlı olarak düşük bulunmuştur. Benzer bulgular, önceki çalışmalarda da bildirilmiş olup, yaş etkisinin, özellikle 70 yaş ve üzeri popülasyonda belirginleştiği rapor edilmiştir (13,14,15).

SÇT puanlarının düşük eğitim seviyesinden olumsuz etkilendiği çeşitli çalışmalarda bildirilmiştir (35,36). Bu çalışmada, 0-5 yıl eğitime sahip bireylerin SÇT puanlarının, lise ve üzeri eğitime sahip bireylerden daha düşük olduğu bulunmuştur. Ayrıca, Manos ve Wu yöntemi ile puanlanan SÇT toplam skorları, ortaokul/lise ve lise ve üzeri eğitim grupları arasında da farklılaşmaktadır. Bu bulgular doğrultusunda, eğitim etkisi özelinde Manos ve Wu yönteminin Shulman yöntemine kıyasla daha hassas olduğu söylenebilir. İki yöntem arasındaki bu farkın, Manos ve Wu yönteminde SÇT'nin daha geniş bir ölçüm aralığı üzerinden puanlanması ile ilişkili olabileceği düşünülmüştür.

SÇT'de hata tipinin (örneğin; rakamların yerleştirilmesi, saatin gösterilmesi, perseveratif hataların varlığı) demans sendromları için ayırt edici olabileceği düşünülmektedir. Parkinson hastalığı demansında sıklıkla perseveratif hata tipi gözlenirken, Alzheimer tipi demansta daha çok kavramsal hatalar (saat gösterimi ile ilişkili hatalar) bildirilmektedir (33,37). Bir çalışmada, eğitimsiz sağlıklı bireylerin de Alzheimer hastalarına benzer şekilde kavramsal hatalar yaptığı bulunmuştur (36). Bu bulgular, çalışmamızda saat gösterimi puanları üzerinde sadece eğitim etkisinin saptanmış olmasıyla uyumludur.

Bu çalışmada, sağlıklı bireylerin SÇT performansları her iki yöntem için test-tekrar test güvenilirliğinin değerlendirilmesi amacıyla incelenmiştir. İlk ve ikinci uygulamalar arasındaki uyumun, Manos ve Wu yöntemi için mükemmel düzeyde, Shulman

yöntemi için iyi düzeyde olduğu bulunmuştur. Çalışmamızda, Shulman yönteminin test-tekrar test güvenilirliğinin Manos ve Wu yöntemine göre daha düşük bulunması, yöntemin yarı kantitatif olması ve subjektif değerlendirme içermesiyle ilişkili olabilir. Fuzikawa ve ark.(38)Shulman yöntemi ile iyi test performanslarının puanlanmasının (performansa 4 veya 5 puan verme kararının) düşük performansları puanlamaya göre daha zorlayıcı olduğunu ve test-tekrar test güvenilirliğini düşürebileceğini belirtmiştir. Manos ve Wu'nun (5) çalışmasında, iki gün ara ile uygulanmış SÇT'nin test-tekrar test güvenilirliği 0,94 olarak bildirilmiştir. Çalışmamızda, SÇT ortalama 12 ay arayla uygulanmış ve test-tekrar test güvenilirliği 0,81 olarak bulunmuştur. Bu farklılığın, çalışmamızda iki uygulama arasındaki zaman aralığının uzun olmasıyla ilişkili olduğu düşünülmüştür. Sağlıklı bireylerin 12 ay sonra tekrarlanan tüm nöropsikolojik test bulguları (bellek, dikkat, yürütücü işlevler vb.), yaş ve eğitim normları ile uyumlu olup ilk değerlendirmeleri ile belirgin bir farklılık göstermemiştir. Yine de test-tekrar test güvenilirliğinin yorumlanmasında, uzun aralıkların güvenilirliği düşürebileceği göz önünde bulundurulmalıdır. Ancak, Manos ve Wu (5) yöntemi için her iki çalışmada da bildirilen değerler oldukça yüksek güvenilirlik katsayılarına işaret etmektedir.

Değerlendiriciler arası güvenilirliğin, her iki puanlama yöntemi için de mükemmel düzeyde olduğu bulunmuştur. Ancak, Manos ve Wu yönteminde sağlıklı bireylerin saat gösterimi puanlarında değerlendiriciler arası güvenilirlik kabul edilebilir düzeyde olmakla birlikte, diğer tüm puan türlerinden daha düşüktür. Bu yöntemde, saatin doğru gösterilmesinin yanı sıra saat kollarının doğru uzunlukta olması puanlama açısından önem taşımaktadır. Akrep, yelkovanı kıyasla daha kısa çizilmediği sürece, 11 rakamının işaret etse dahi puan verilmemektedir. Test performansının değerlendirilmesi sırasında zaman zaman kolların uzunluklarını ayırt etmekte güçlük yaşanmıştır. Bu durumun saat gösteriminin doğru, ancak kol uzunluklarının net olmadığı durumlarda puanlamayı etkileyebileceği düşünülmüştür.

Çalışmamızda, iki farklı puanlama yöntemiyle elde edilen SÇT puanları ile genel kognitif durumu değerlendiren MMDT puanları arasında güçlü bir ilişki saptanmıştır. Bu bulgu, Shulman ve ark.'nın

Tablo 11. İki farklı saat çizme testi puanlama yöntemi için ROC eğrisi altında kalan alan, standart hata ve %95 güven aralıkları

	Manos ve Wu			Shulman		
	AUC	Standart hata	%95 Güven aralığı	AUC	Standart hata	%95 Güven aralığı
Sağlıklı-HKB	0,83	0,03	0,773-0,871	0,74	0,03	0,680-0,792
Sağlıklı-AH	0,90	0,02	0,853-0,937	0,85	0,02	0,799-0,897
HKB-AH	0,66	0,04	0,591-0,729	0,66	0,04	0,585-0,723

HKB: Hafif kognitif bozukluk, AH: Alzheimer hastalığı, AUC: Eğri altında kalan alan (area under curve)

Tablo 12. Saat çizme testi puanlama yöntemlerinin kesme puanları ile karşılık gelen duyarlılık ve özgüllük değerleri

	Manos ve Wu			Shulman		
	Kesme puanı	Duyarlılık (%)	Özgüllük (%)	Kesme puanı	Duyarlılık (%)	Özgüllük (%)
Sağlıklı-HKB	9	%80,8	%80,0	4	%61,6	%83,2
Sağlıklı-AH	9	%89,8	%76,8	4	%81,5	%79,6
HKB-AH	7	%62,9	%61,9	3	%61,9	%59,8

HKB: Hafif kognitif bozukluk, AH: Alzheimer hastalığı

(3) çalışmasıyla uyumlu olup, puanlama yöntemlerinin eş-zaman geçerliğinin yüksek olduğunu göstermektedir. Ayrıca, iki puanlama yönteminin geçerliği, yöntemlerin arasındaki korelasyonun yüksek düzeyde bir uyum göstermesiyle de kanıtlanmıştır.

SÇT, ilk olarak hemisferik ihmal olgularını değerlendirmek amacıyla kullanılmaya başlanmış (18,39,40); ancak, son 20 yılda yapılan pek çok çalışma, SÇT'nin demans sendromlarını saptamada işlevsel bir tarama testi olabileceğini göstermiştir. Bu çalışmada da her iki skorlama yönteminin de Alzheimer hastaları ve HKB olgularını sağlıklı bireylerden ayırt etme gücünün oldukça yüksek olduğu bulunmuştur. Ancak, Manos ve Wu yönteminin HKB olgularını sağlıklı bireylerden ayırt etmede duyarlılık ve özgüllüğü Shulman yönteminden daha yüksektir. Her iki yöntem ile de HKB olgularını Alzheimer hastalarından ayırt etmede güvenilir sonuçlar elde edilememiştir. Bu bulgular, klinik kullanımda SÇT'nin HKB ve AH için bir tarama testi olarak kullanılabilmesine; ancak, HKB ve AH olgularını birbirinden ayırt etmede tek başına yetersiz kalabileceğine işaret etmektedir. Literatürde, yaşlı bireylerde demans varlığını tespit etmede SÇT'nin sözel akıcılık testleri ile birlikte kullanımının, diğer testler ile birlikte kullanımına göre en yüksek faydayı sağladığı bildirilmiştir (4).

Sonuç olarak, bireyin SÇT performansının norm değerlerinin altında yer almasının, kognitif etkilenmenin kaynağını araştırmak üzere daha ayrıntılı tetkiklerin yapılması için yol gösterici olabileceği düşünülmüştür. AH tanısı için tek başına yeterli olmasa dahi, SÇT performansının hastalığın erken evrelerinden itibaren etkilenilebileceği yönündeki literatür bulguları bu çalışmada desteklenmiştir. Kognitif işlevleri etkileyen nöropsikiyatrik hastalıkların boylamsal takibinde ise, SÇT'nin diğer nöropsikolojik testlerle birlikte kullanımı önerilmektedir.

Sonuç

SÇT'nin farklı uygulama ve değerlendirme biçimlerinin varlığı, her bir yöntem için ayrı norm değerlerine ihtiyaç doğurmaktadır. Ülkemizde klinik değerlendirme ve araştırma amacıyla en sık uygulanan SÇT versiyonu, norm değerleri olmaksızın kullanılmaktadır. Bu çalışmada, SÇT'nin Türkiye popülasyonu için yaş ve eğitim düzeylerine göre norm değerleri, uluslararası çalışmalarda en sık kullanılan Manos ve Wu ile Shulman puanlama yöntemleri için oluşturulmuştur. Geçerlik ve güvenilirlik analizlerinde, her iki puanlama yönteminden elde edilen sonuçlar oldukça paralel olup, yöntemlerin yüksek test-tekrar test ve değerlendiriciler arası güvenilirlik ile yüksek eş-zaman ve ölçüt geçerliklerine sahip olduğunu göstermektedir.

Bu çalışmayla birlikte, 30 yılı aşkın süredir dünyada ve ülkemizde görsel-mekansal beceriler ile yürütücü işlevleri değerlendirmede yaygın olarak kullanılan, uygulama ve puanlaması kolay, hızlı ve yatak başı muayeneye uygun bir tarama testi olan SÇT'nin Türkiye norm değerleri konuyla ilgili klinisyen ve araştırmacıların kullanımına sunulmuştur.

Etik

Etik Kurul Onayı: Bu çalışma protokolü Dokuz Eylül Üniversitesi Etik Kurul Komitesi tarafından 13.07.2017 tarihinde 2017/18-01 karar numarası ile onaylanmıştır.

Hasta Onayı: Tüm katılımcılardan katılmış oldukları araştırmalar kapsamında yazılı bilgilendirilmiş onam alınmıştır.

Hakem Değerlendirmesi: Editörler kurulu ve editörler kurulu dışında olan kişiler tarafından değerlendirilmiştir.

Yazarlık Katkıları

Konsept: D.D.E.S., **Dizayn:** D.D.E.S., G.G.Y., **Veri Toplama veya İşleme:** D.Y., **Analiz veya Yorumlama:** D.D.E.S., D.Y., **Literatür Arama:** D.D.E.S., D.Y., **Yazan:** D.D.E.S., D.Y., G.G.Y.

Çıkar Çatışması: Yazarlar bu makale ile ilgili olarak herhangi bir çıkar çatışması bildirmemiştir.

Finansal Destek: Bu çalışma için hiçbir kurum ya da kişiden finansal destek alınmamıştır.

Kaynaklar

- Öktem Ö. Nöropsikolojik testler ve nöropsikolojik değerlendirme. *Türk Psikoloji Dergisi* 1994;9:33-44.
- Strauss E, Sherman EM, Spreen O. A compendium of neuropsychological tests: Administration, norms, and commentary. Oxford University Press, USA, 2006.
- Shulman KI, Shedletsky R, Silver IL. The challenge of time: clock-drawing and cognitive function in the elderly. *Int J Geriatr Psychiatr* 1986;1:135-140.
- Shulman KI. Clock-drawing: is it the ideal cognitive screening test? *Int J Geriatr Psychiatry* 2000;15:548-561.
- Manos PJ, Wu R. The ten point clock test: a quick screen and grading method for cognitive impairment in medical and surgical patients. *Int J Psych Med* 1994;24:229-244.
- Watson YI, Arfken CL, Birge SJ. Clock completion: an objective screening test for dementia. *J Am Geriatr Soc* 1993;41:1235-1240.
- Tuokko H, Hadjstavropoulos T, Miller JA, Beattie BL. The clock test: a sensitive measure to differentiate normal elderly from those with Alzheimer's disease. *J Am Geriatr Soc* 1992;40:579-584.
- Goodglass H, Kaplan E, Barresi B. Boston diagnostic aphasia examination (BDAA). Philadelphia: Lea and Febiger, 1983.
- Wolf-Klein GP, Silverstone FA, Levy AP, Brod MS. Screening for Alzheimer's disease by clock drawing. *J Am Geriatr Soc* 1989;37:730-736.
- Sunderland T, Hill JL, Mellow AM, Lawlor BA, Gundersheimer J, Newhouse PA, Grafman JH. Clock drawing in Alzheimer's disease. A novel measure of dementia severity. *J Am Geriatr Soc* 1989;37:725-729.
- Agrell B, Dehlin O. The clock-drawing test. *Age and Ageing* 1998;27:399-403.
- Cangoz B, Karakoc E, Seleker K. The norm determination and validity-reliability studies of clock drawing test on Turkish adults and elderly (ages 50 and over). *Turk J Geriatr* 2006;9:136-142.
- Mazancova AE, Nikolai T, Stepankova H, Kopecek M, Bezdicek O. The Reliability of Clock Drawing Test Scoring Systems Modeled on the Normative Data in Healthy Aging and Nonamnestic Mild Cognitive Impairment. *Assessment* 2016;pii:1073191116632586.
- Bozikas VP, Giakoulidou A, Hatzigeorgiadou M, Karavatos A, Kosmidis MH. Do age and education contribute to performance on the clock drawing test? Normative data for the Greek population. *J Clin Exp Neuropsychol* 2008;30:199-203.
- Sugawara N, Yasui-Furukori N, Umeda T, Sato Y, Kaneda A, Tsuchimine S, Saito M, Nakaji S, Kaneko S. Clock drawing performance in a community-dwelling population: Normative data for Japanese subjects. *Aging Ment Health* 2010;14:587-592.
- Hubbard EJ, Santini V, Blankevoort CG, Volkens KM, Barrup MS, Byerly L, Chaisson C, Jefferson AL, Kaplan E, Green RC, Stern RA. Clock drawing performance in cognitively normal elderly. *Arch Clin Neuropsychol* 2008;23:295-327.
- Lourenço RA, Ribeiro-Filho ST, Moreira Ide F, Paradelo EM, Miranda AS. The Clock Drawing Test: performance among elderly with low educational level. *Rev Bras Psiquiatr* 2008;30:309-315.
- Mainland BJ, Shulman KI. Clock drawing test. *Cognitive Screening Instruments*. Springer London 2013:79-109.
- Can SS, Özel Kızıl ET, Varlı M, Turan E, Atlı T. Demans hastalarında üç farklı saat çizme testinin Türkçe uyarlamalarının psikometrik özellikleri. *Nöropsikiyatri Arşivi* 2010;47:91-95.

20. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189-198.
21. Güngen C, Ertan T, Eker E, Yaşar R, Engin F. Reliability and validity of the standardized Mini Mental State Examination in the diagnosis of mild dementia in Turkish population. *Turk Psikiyatri Derg* 2002;13:273-281.
22. Öktem Ö. Sözel Bellek Süreçleri Testi. Bir ön çalışma. *Nöropsikoloji Arşivi* 1992;29:196-206.
23. Wechsler D. Wechsler memory scale-revised (WMS-R). Psychological Corporation 1987.
24. Wechsler D. Wechsler adult intelligence scale-Fourth Edition (WAIS-IV). San Antonio, TX: NCS Pearson, 22, 2008;498
25. Tumaç A. Normal deneklerde frontal hasarlara duyarlı bazı testlerde performansa yaş ve eğitimin etkisi. İstanbul Üniversitesi Sosyal Bilimler Enstitüsü Psikoloji Bölümü, Yüksek Lisans Tezi, İstanbul, 1997.
26. Stroop JR. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology* 1935;18:643.
27. Kaplan E, Goodglass H, Weintraub S. Boston naming test. Pro-ed, 2001.
28. Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, Leirer VO. Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res* 1983;17:37-49.
29. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, Gamst A, Holtzman DM, Jagust WJ, Petersen RC, Snyder PJ, Carrillo MC, Thies B, Phelps CH. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 2011;7:270-729.
30. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, Klunk WE, Koroshetz WJ, Manly JJ, Mayeux R, Mohs RC, Morris JC, Rossor MN, Scheltens P, Carrillo MC, Thies B, Weintraub S, Phelps CH. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 2011;7:263-269.
31. Teixeira Fabricio A, Aprahamian I, Sanches Yassuda M. Qualitative analysis of the Clock Drawing Test by educational level and cognitive profile. *Arq Neuropsiquiatr* 2014;72:289-295.
32. Freedman, M. Clock drawing: A neuropsychological analysis. Oxford University Press, USA. 1994.
33. Caffarra P, Gardini S, Zonato F, Concaro L, Dieci F, Copelli S, Freedman M, Stracciari A, Venneri A. Italian norms for the Freedman version of the Clock Drawing Test. *J Clin Exp Neuropsychol* 2011;33:982-988.
34. Tranel D, Rudrauf D, Vianna EP, Damasio H. Does the Clock Drawing Test have focal neuroanatomical correlates? *Neuropsychology* 2008;22:553-562.
35. von Gunten A, Ostos-Wiecheteck M, Brull J, Vaudaux-Pisquem I, Cattin S, Duc R. Clock-drawing test performance in the normal elderly and its dependence on age and education. *Eur Neurol* 2008;60:73-78.
36. Kim H, Chey J. Effects of education, literacy, and dementia on the Clock Drawing Test performance. *J Int Neuropsychol Soc* 2010;16:1138-1146.
37. Lee AY, Kim JS, Choi BH, Sohn EH. Characteristics of clock drawing test (CDT) errors by the dementia type: quantitative and qualitative analyses. *Arch Gerontol Geriatr* 2009;48:58-60.
38. Fuzikawa C, Lima-Costa ME, Uchoa E, Barreto SM, Shulman K; Bambuí Health and Ageing Study. A population based study on the intra and inter-rater reliability of the clock drawing test in Brazil: the Bambuí Health and Ageing Study. *Int J Geriatr Psychiatry* 2003;18:450-456.
39. Critchley M. The Parietal Lobes. Hafner: New York, NY, 1953.
40. Hazan E, Frankenburg F, Brenkel M, Shulman K. The test of time: a history of clock drawing. *Int J Geriatr Psychiatry* 2018;33:e22-e30.

Ek 1. Manos ve Wu (5) puanlama yöntemi için kullanılabilir 10 cm çaplı saat şablonu

