

Poor Concordance Between Remote Neurocognitive Testing and Face-to-Face Testing of Verbal Memory Processes in Young Healthy Volunteers

Murat Cenk Çelen¹ , Neda Nur Kayapınar² , Şenel Kayıkcı² ,
Yusuf Kaan Akbulut² , Melda Pelin Yargıç³ 

¹Ankara Medipol University, Faculty of Medicine, Department of Biophysics, Ankara, Turkey

²Necmettin Erbakan University, Meram School of Medicine, Konya, Turkey

³Ankara Medipol University, Faculty of Medicine, Department of Physiology, Ankara, Turkey

Murat Cenk ÇELEN

Neda Nur KAYAPINAR

Şenel KAYIKCI

Yusuf Kaan AKBULUT

Melda Pelin YARGIÇ

Correspondence: Murat Cenk Çelen

Ankara Medipol University, Faculty of Medicine, Department of Biophysics, Ankara, Turkey

Phone: +905310108705

E-mail: muratcenkcelen@gmail.com

Received: 22 November 2022

Accepted: 17 June 2023

ABSTRACT

Purpose: Recently published studies indicate increased interest in remote health services that have expanded greatly with digitalization and whether they can be implemented into neuropsychological testing. Remote neurocognitive testing has become a frequently used approach today. The current study examines the concordance between the face-to-face and remote Öktem Verbal Memory Processes Test results of healthy young participants.

Method: A total of 45 healthy young participants with a mean age of 23.41 were invited to participate in the study. Participants were 42.2% of male and 57.8% female. The Öktem Verbal Memory Processes Test was administered to participants. There was at least one week between face-to-face and remote testing. The results were compared with paired t-tests and Lin's Concordance Correlation analysis. Reliability was analyzed by examining the Intraclass Correlation Coefficient.

Results: For immediate memory score, ICC was 0.286 (95% CI: 0.016-0.524). ICC of total learning score was calculated as 0.665 (95% CI: 0.465-0.801). The highest learning score had an ICC of 0.868 (95% CI: 0.773-0.926). ICC of the total inconsistency score was 0.639 (95% CI: 0.414-0.79). Finally, the score for reaching the criterion displayed an ICC of 0.253 (95% CI: 0.079-0.535).

Conclusions: The findings demonstrate that there is poor concordance between the results obtained with remote and face-to-face application of the Öktem Verbal Memory Process Test. As a result, using it remotely is not recommended at this stage.

Keywords: Digital Neuropsychology, Memory, Remote Neurocognitive Testing, Technology

Genç Sağlıklı Gönüllülerin Sözel Bellek Süreçlerinin Uzaktan Nörobilişsel Test ve Yüz Yüze Testi Arasındaki Uyum Zayıftır

ÖZET

Amaç: Son zamanlarda yayınlanan araştırmalar, dijitalleşmeyle birlikte büyük ölçüde genişleyen uzaktan sağlık hizmetlerine olan ilginin arttığını ve bunların nöropsikolojik testlere uygulanıp uygulanamayacağını ilgi konusu olduğunu göstermektedir. Uzaktan nörobilişsel testler günümüzde sıklıkla kullanılan bir yaklaşım haline geldi. Mevcut çalışmanın amacı, sağlıklı genç katılımcıların yüz yüze ve uzaktan Öktem Sözel Bellek Süreçleri Testi sonuçları arasındaki uyumunu incelemektir.

Yöntem: Yaş ortalaması 23.41 olan 45 sağlıklı genç katılımcı çalışmaya davet edildi. Katılımcıların %42,2'si erkek, %57,8'i kadındı. Katılımcılara Öktem Sözel Bellek Süreçleri Testi uygulandı. Yüz yüze ve uzaktan test arasında en az bir hafta vardı. Sonuçlar, eşleştirilmiş t-testleri ve Lin'in Uyum Korelasyon analizi ile karşılaştırıldı. Güvenilirlik, Sınıf İçi Korelasyon Katsayısı incelenerek analiz edildi.

Bulgular: Anlık bellek puanı için ICC 0,286 (%95 GA: 0,016-0,524). Toplam öğrenme puanı ICC 0,665 (%95 GA: 0,465-0,801) olarak hesaplandı. En yüksek öğrenme puanının ICC değeri 0,868'dir (%95 GA: 0,773-0,926). Toplam tutarsızlık puanının ICC'si 0,639'du (%95 GA: 0,414-0,79). Son olarak, kriterlere ulaşma puanı 0,253'lük bir ICC gösterdi (%95 GA: 0,079-0,535).

Sonuç: Bulgular, Öktem Sözel Bellek Süreç Testi'nin uzaktan ve yüz yüze uygulanması ile elde edilen sonuçlar arasında zayıf bir uyum olduğunu göstermektedir. Sonuç olarak, bu aşamada uzaktan kullanılması önerilmez.

Anahtar Sözcükler: Dijital Nöropsikoloji, Bellek, Uzaktan Nörobilişsel Test, Teknoloji

Worldwide increase in digitalization during the 21st century has allowed development of new tools that enable neuropsychological assessment of cognition and behavior, sometimes referred to as digital neuropsychology (1). Developing alternatives to face-to-face practices in health care delivery has come to the fore due to the COVID-19 pandemic and with attempts to eliminate inequality in access to health services at the international level (2,3). Therefore, remote neurocognitive testing (RNCT) via video interview has gained momentum.

Memory consists of several components, including a primary act of learning which, via the continuity assumed vital and supplied by the storage mechanism, happens in the act of retrieval (4). Verbal memory is a relatively general concept that refers to memory for verbally delivered information. A variety of tasks, including learning word list tasks, story recall, and learning of sequences of paired words, can measure verbal memory ability (5). The Öktem Verbal Memory Processes Test is a valid and reliable test that was developed by the Istanbul University Neuropsychology Laboratory (Turkey) and examines verbal memory (6). This test can be applied to patient groups for diagnostic purposes, as well as to healthy individuals in neurocognitive studies. Memory tests are often administered to elderly people. When memory tests are used in the follow-up of the course of the disease, they can also be applied repeatedly at certain time intervals. This may require many additional hospital visits in a year. The frequent visits of elderly individuals to the hospital for neurocognitive evaluation is a challenge due to the high risk of infection during the pandemic and the difficulty and cost of access to the hospital for elderly individuals.

Studies have examined whether face-to-face and remote applications of many different scales affect the results for similar reasons (7,8). If there is no significant difference between the results obtained by the two application methods, it can be concluded that the scales/tests can be administered via remote access, creating a safe and easy alternative both in the provision of health services and in use for research purposes.

This study aims to examine the concordance between the face-to-face and remote Oktem Verbal Memory Processes Test results of healthy young individuals.

MATERIALS AND METHODS

This study has been approved by the Local Ethics Committee (2021/3369). Informed consent was obtained from all individual participants included in the study. The Öktem Verbal Memory Processes Test was administered to individuals who volunteered to participate in the study, once face-to-face and once remotely (6). It was ensured that there was at least one week and at most two weeks between the two tests. During the online tests, participants were asked to use their own smartphones which they use daily. Online tests were conducted in the format of videoconferencing via a WhatsApp videocall. All tests were administered by a single, experienced and trained researcher.

Healthy individuals between the ages of 18 and 50 were invited to the study. Being diagnosed with a neurological or psychiatric disease, using drugs that may affect memory functions, and having hearing loss at any level were exclusion criteria.

A randomized cross-over design was used to determine whether the participants would receive the face-to-face or videoconferencing testing first.

Masks were worn during face-to-face application due to the COVID-19 pandemic. Practitioners also wore masks during the test performed remotely in order to avoid an advantage due to lip reading which might provide additional cues.

Implementation of Öktem Verbal Memory Processes Test (ÖVMPT) and Scoring

One of the word sets consisting of 15 words is read clearly at a constant speed (1 second per word), and then the participant is asked to repeat the words that he/she could remember to the practitioner. Then this is repeated for 10 trials. The words remembered in each round are recorded in order. Standard sentences are used to instruct the participant during test administration. The tests were carried out with different word sets during remote and face-to-face application. Word sets were shown to deliver similar results in previous studies by Öktem (6).

After the test is completed, various scores are calculated from the data. The number of words remembered in the first attempt constitutes the "immediate memory score". The sum of the words remembered in ten trials constitutes the "total learning score". Whichever of the ten repetitions

the participant was able to remember the largest number of words was recorded as the "highest learning score". If the participant succeeded in remembering all the words, the trial during which he/she succeeded in remembering all words was considered "reaching the criterion". For example, if the participant successfully repeated all 15 words in the third trial, reaching the criterion score was noted as 3. In addition, if participant is not able to remember a word that he/she had remembered twice in a row, or if he/she cannot remember a word that he/she had remembered three times in total even if not consecutively, one inconsistency point is given and the "total inconsistency score" of the person in the ten rounds is calculated.

Statistical Analysis

The Shapiro-Wilk test was used to evaluate whether the data show a normal distribution. Then, the results of the two applications were compared with paired t-test and Lin's Concordance Correlation analysis. Further, reliability was analyzed by examining the Intraclass Correlation Coefficient (ICC) for each score. Lin's concordance correlation coefficient was interpreted according to McBride's (2005) suggestions: <0.90: poor; 0.90 to 0.95: moderate; 0.95 to 0.99: substantial; >0.99 almost perfect. ICC levels were interpreted as follows: < 0.5 poor reliability, 0.5 to 0.75 moderate reliability, 0.75 to 0.9 good reliability, and > 0.90 excellent reliability (9). Alpha level was set at 0.05. Analyses were done using the R Studio program (10,11).

RESULTS

Participants (n=45) were between the ages of 18-45, with a mean age of 23.31 (SD: 6.47). Of the participants, 19 were male (42.2%) and 26 were females (57.8%). All participants had graduated at least from high school. 37 were university graduates (82.2%) and there 8 people who graduated from high school only (17.8%).

Results of the memory process tests are calculated in five different scores. Table 1 summarizes these scores, including the p values obtained from paired t tests.

Lin's Concordance Coefficients were evaluated as poor for all scores: ρ_c (Immediate Memory Score)=0.28, ρ_c (Total Learning Score)=0.66, ρ_c (Highest Learning Score)=0.86, ρ_c (Reaching the Criterion)=0.65, ρ_c (Total Inconsistency Score)=0.60.

Correlation plots of two methods (online vs. face-to-face) can be observed in Figure 1 for Immediate Memory score,

Total Learning score, Highest Learning score, Reaching The Criterion, and Total Inconsistency score.

Table 1. Results of the Verbal Memory Processing Test (n=45)

	Face-to-face (mean \pm SD)	Online (mean \pm SD)	p	(95% CI)
Immediate Memory Score	8.15 \pm 2.30	7.22 \pm 1.79	0.013	(-1.66; -0.20)
Total Learning Score	128.42 \pm 16.58	124.77 \pm 17.17	0.078	(-7.72; 0.43)
Highest Learning Score	14.57 \pm 1.13	14.6 \pm 1.11	0.799	(-0.15; 0.19)
Reaching the Criterion	5.11 \pm 2.26	5.67 \pm 2.33	0.272	(-1.45; 0.42)
Total Inconsistency Score	3.04 \pm 3.36	2.87 \pm 3.00	0.692	(-0.69; 1.03)

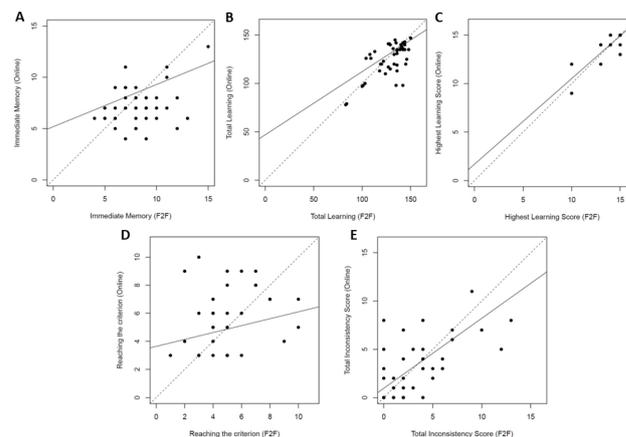


Figure 1. Correlation plots of memory scores obtained by two methods (face-to-face vs. online). A-Immediate Memory Score B-Total Learning Score C-Highest Learning Score D-Reaching the Criterion E-Total Inconsistency Score (F2F: Face-to-face)

Intraclass correlation coefficient estimates and their 95% confident intervals were calculated based on a single-rating, absolute-agreement, 2-way mixed-effects model. For immediate memory score, ICC was 0.286 (95% CI: 0.016-0.524). ICC of total learning score was calculated as 0.665 (95% CI: 0.465 -0.801).

The highest learning score had an ICC of 0.868 (95% CI: 0.773 -0.926). The ICC of the total inconsistency score was 0.639 (95% CI: 0.414 -0.79). Finally, the score for reaching the criterion displayed an ICC of 0.253 (95% CI: 0.079 -0.535). Consequently, immediate memory and reaching the criterion scores had poor reliability, total learning score and total inconsistency score had a moderate reliability, and only the highest learning score had a good reliability.

DISCUSSION

Many characteristics related to memory, such as immediate memory and the process of learning or acquiring knowledge, may be distinguished using verbal memory processes testing. It's particularly useful to help distinguish mild forms of cognitive impairment from dementia (12). The Öktem Verbal Memory Processes Test can identify various parameters associated with memory, such as immediate (sensory) memory, attention, and (short-term) working memory.

Remote neurocognitive testing has been implemented to evaluate immediate memory, semantic memory, working memory, immediate free recall, recognition, and delayed recall in various populations (13). The advantage of remote testing of verbal memory processes is that it increases the possibility of identifying minor cognitive changes over time in the natural context of a person's home setting (14). Therefore, the potential use of remote neurocognitive testing is exciting for clinicians and researchers. Sharing that motivation, we compared the agreement between the results obtained by online and face-to-face testing of verbal memory processes in this study. Although a significant difference was observed only between the immediate memory scores in the dependent variables *t* test; Lin's concordance correlation results, which is a superior analysis method in many respects in showing the compatibility of the two methods, showed that all the scores obtained in the verbal memory processes test were in poor concordance. Additionally, ICC values showed that most domains of the ÖVMPT displayed poor to moderate reliability between face-to-face and remote administration, whereas only the highest learning score had good reliability. These results show us that the use of ÖVMPT via videoconferencing does not give comparable results to face-to-face administration and the two methods cannot be used interchangeably. This reminds us once again that it is necessary to be cautious when switching to online administrations of traditional test batteries.

Many researchers have studied the agreement between online vs. remote neurocognitive testing via a number of test batteries and reported positive results. One study compared the remotely conducted Montreal Cognitive Assessment (MoCA) scores of patients with Parkinson's Disease and obstructive sleep apnea to those which were done in person. They concluded that two methods were in good agreement. However, this study examined only seven participants remotely (15). A similar study with seventeen participants examined the reliability of the MoCA using telehealth in a rural setting with veterans and reported ICC levels >0.98 (16). Another study tested the feasibility and reliability of the Wechsler Adult Intelligence Scales - 4th Edition (WAIS-IV) when administered via home-based videoconferencing (17). Similar to our study, they have selected a nonclinical cohort. The results have demonstrated that WAIS-IV can be effectively administered through videoconferencing. Additionally, results of a recent study have shown that video-conference calls to complete cognitive assessments during the COVID-19 pandemic was feasible and acceptable (18). However, it should be noted that this study reports only the subjective experience of clinicians and service-users, but lacks any objective comparisons between videoconferencing and face-to-face assessments. Additionally, one study has even shown that the type of the platform used, such as the browser, an iPad or an iPhone, can also significantly affect the results of neurocognitive tests when administered remotely (21). This finding shows that the type of platform used in remote applications is also an important variable and should be taken into account in research.

Our study results, in contrast to many positive reports, have shown that online and face-to-face administration of ÖVMPT provide different results. However, similar to our study findings, there are other studies in the literature that show that reliability is very low when some neurocognitive tests applied with traditional paper and pencil are applied remotely. One important example would be the study by Zeghari et. Al, that compared face-to-face and online administration of several neurocognitive test batteries on 50 participants (aged 55 and older). Results showed that Mini Mental State Test total score ICC was only 0.371, The Free and Cued Selective Reminding Test total recall score, delayed recall score and recognition score ICC values were all below 0.5, and verbal fluency test semantic and phonological z-score ICC values were below 0.5 (19). It is well established that, ICC values below 0.5 indicate poor reliability. However, when a similar study was conducted with a larger cohort (202 adult subjects, including 83 with cognitive impairment and 119 healthy

controls) by Cullum et al, ICC for the total score of MMSE was found 0.905. This once again shows the importance of the high number of participants in the studies in the results (20) and the limited number of participants is an important limitation in our study. Nevertheless, the number of participants in our study exceed that of many similar studies.

In conclusion, remote neurocognitive assessment is becoming more and more widespread, which benefits from the point of view of convenience and elimination of inequalities in access to health. In this context, the question of whether traditional assessment tools give valid results when applied with remote access has also been a common research question. Although remote use has yielded valid results for different test batteries in the literature, as a result of this research, it has been determined that there is poor concordance between the results obtained with remote and face-to-face application of the Öktem Verbal Memory Process Test. Therefore, using this test with a remote application is not recommended at this stage. A similar study can be repeated by increasing the number of participants.

DECLERATIONS

Funding

None.

Conflicts of Interest/Competing Interests

None.

Ethics Approval

Necmettin Erbakan University Non-Pharmaceutical and Medical Device Research Ethics Committee (2021/3369).

Availability of Data and Material

Available on request

Authors' Contributions

All authors (MCC, NK, YKA, SK, MPY) contributed to conceptualisation, data collection, writing the manuscript. MPY has also done the statistical analysis.

ACKNOWLEDGEMENT

All participants of the study are thanked for their time and attention.

REFERENCES

- Germine L, Reinecke K, Chaytor NS. Digital neuropsychology: Challenges and opportunities at the intersection of science and software. *Clin Neuropsychol* 2019;33:271–86.
- Monaghesh E, Hajizadeh A. The role of telehealth during COVID-19 outbreak: a systematic review based on current evidence. *BMC Public Health* 2020;20:1–9.
- Wosik J, Fudim M, Cameron B, Gellad ZF, Cho A, Phinney D, et al. Telehealth transformation: COVID-19 and the rise of virtual care. *J Am Med Informatics Assoc* 2020;27:957–62.
- Klein SB. What memory is. *Wiley Interdiscip Rev Cogn Sci* 2015;6:1–38. <https://doi.org/10.1002/WCS.1333>.
- Tatsumi IF, Watanabe M. Verbal Memory. *Encycl Neurosci* 2009;4:176–8. https://doi.org/10.1007/978-3-540-29678-2_6266.
- Öktem Ö. Öktem Verbal Memory Processes Test (ÖKTEM-VMPT) Manuel book. Ankara: Turkish Psychologist Society Publish; 2011.
- Pietkiewicz IJ, Helka AM, Tomalski R. Validity and reliability of the Polish online and pen-and-paper versions of the Somatoform Dissociation Questionnaires (SDQ-20 and PSDQ-5). *Eur J Trauma Dissociation* 2019;3:23–31.
- de Brito Silva B, de Brito A, Monteiro EP, Mondelo GP, Remor E. Evidence of validity for the online version of the Assessment of Adherence to Antiretroviral Therapy Questionnaire. *SAGE Open* 2019;9:2158244019877201.
- Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016;15:155–63.
- Steichen TJ, Cox NJ. A note on the concordance correlation coefficient. *Stata J* 2002;2:183–9.
- McBride GB. A proposal for strength-of-agreement criteria for Lin's concordance correlation coefficient. *NIWA Client Rep HAM2005-062* 2005;45:307–10.
- Şair A, Şair YB, Ayar S. Differentiating Mild Forms of Cognitive Impairment and Dementia: Where Other Tests Fail, Verbal Memory Assessment May Prove Critical. *Meandros Med Dent J* 2021;22:187.
- Moore RC, Swendsen J, Depp CA. Applications for self-administered mobile cognitive assessments in clinical research: A systematic review. *Int J Methods Psychiatr Res* 2017;26:e1562.
- Resnick HE, Lathan CE. From battlefield to home: a mobile platform for assessing brain health. *Mhealth* 2016;2.
- Lajoie AC, Crane J, Robinson AR, Lafontaine AL, Benedetti A, Kimoff RJ, et al. Feasibility of remote neurocognitive assessment: pandemic adaptations for a clinical trial, the Cognition and Obstructive Sleep Apnea in Parkinson's Disease, Effect of Positive Airway Pressure Therapy (COPE-PAP) study. *Trials* 2021;22:1–9. <https://doi.org/10.1186/S13063-021-05879-1/TABLES/3>.
- DeYoung N, Shenal B V. The reliability of the Montreal Cognitive Assessment using telehealth in a rural setting with veterans. *J Telemed Telecare* 2019;25:197–203. <https://doi.org/10.1177/1357633X17752030>.
- Mahon S, Webb J, Snell D, Theadom A. Feasibility of administering the WAIS-IV using a home-based telehealth videoconferencing model. <https://doi.org/10.1080/1385404620211985172> 2021;36:558–70. <https://doi.org/10.1080/13854046.2021.1985172>.
- Sumpter R, Camsey E, Meldrum S, Alford M, Campbell I, Bois C, et al. Remote neuropsychological assessment: Acceptability and feasibility of direct-to-home teleneuropsychology methodology during the COVID-19 pandemic. <https://doi.org/10.1080/138540462022056922> 2022;1–16. <https://doi.org/10.1080/13854046.2022.2056922>.

19. Zeghari R, Guerchouche R, Tran-Duc M, Bremond F, Langel K, Ramakers I, et al. Feasibility Study of an Internet-Based Platform for Tele-Neuropsychological Assessment of Elderly in Remote Areas. *Diagnostics* 2022, Vol 12, Page 925 2022;12:925. <https://doi.org/10.3390/DIAGNOSTICS12040925>.
20. Cullum CM, Hynan LS, Grosch M, Parikh M, Weiner MF. Teleneuropsychology: Evidence for video teleconference-based neuropsychological assessment. *J Int Neuropsychol Soc* 2014;20:1028–33.
21. Ye S, Sun K, Phi H, Hosseini Ghomi R, Ghomi RH. Validation of Remote Testing using BrainCheck, a Computerized Neurocognitive Test. *MedRxiv* 2020:2020.06.01.20119289. <https://doi.org/10.1101/2020.06.01.20119289>.