Intra- and inter-rater accuracy and reliability of the ocular vergence test

Acurácia e confiabilidade intra e interavaliador do teste de vergência ocular

Vanessa Cristina Godoi de Paula [1*]
Tiago Del Antonio Tsunoda [1]
Paulo Fernandes Pires [1]
Rodrigo Antonio Carvalho Andraus [2]
Odivan Bukalowski Barbosa Alves [1]
Felipe Micheletti Bento [1]
Guilherme Luis Santana Luchesi [1]
Fabrício José Jassi [1]

[1] Universidade Estadual do Norte do Paraná (UENP), Jacarezinho, PR, Brazil
[2] Universidade Pitágoras Unopar Anhanguera (Unopar), Londrina, PR, Brazil

Date of first submission: June 30, 2024
Last received: April 13, 2025
Accepted: May 7, 2025
Associate editor: Ana Paula Cunha Loureiro

***Correspondence:** vanessa.cgodoi@hotmail.com

Abstract

Introduction: Convergence insufficiency is the inability to achieve and maintain adequate convergence without effort, associated with symptoms such as eye strain and headaches. Furthermore, the literature lacks information on the different tests used to assess convergence. **Objective:** To determine the accuracy and intra- and inter-rater reliability of the ocular vergence test**. Methods:** A cohort of 123 volunteers of both sexes, aged between 18 and 49 years, was recruited. The Convergence Insufficiency Symptom Survey was applied, and the near point of convergence was evaluated. A battery of tests was administered by three different evaluators across two days. The linear-weighted Cohen's Kappa coefficient was used to assess intra- and inter-rater reliability and agreement. **Results:** The analysis of agreement for the ocular vergence test revealed moderate inter-rater (Kappa = 0.43 to 0.59) and intrarater (Kappa = 0.40 to 0.75) reliability, coupled with low accuracy (AUC = 0.53) for the methodology employed. **Conclusion:** The ocular vergence test demonstrated moderate reliability but low accuracy.

Keywords: Ocular convergence. Ocular motility disorders. Data reliability. Test reproducibility.

Resumo

Introdução: *A insuficiência de convergência é a incapacidade de obter e manter uma adequada convergência sem esforço, apresentando associação de sintomas como a fadiga ocular e cefaleia. Além disso, a literatura carece de informações sobre os diferentes testes utilizados para a avaliação de convergência.* ***Objetivo:*** *Verificar a acurácia e a confiabilidade intra e interavaliadores do teste de vergência ocular.* ***Métodos:*** *Foram recrutados 123 voluntários de ambos os sexos, de 18 a 49 anos. Aplicou-se o questionário de Sintomas de Insufuciência de Convergência e realizou-se a avaliação do ponto próximo de convergência. A bateria de testes foi realizada por três avaliadores distintos em dois dias. O coeficiente Kappa de Cohen ponderado linear foi utilizado para determinar a confiabilidade/concordância intra e interavaliadores.* ***Resultados:*** *Na análise da concordância do teste de vergência ocular, obervou-se moderada confiabilidade intervaliadores (Kappa = 0.43 a 0.59) e intra-avaliadores (Kappa = 0.40 a 0.75), com baixa acurácia (AUC = 0,53) para a metodologia utilizada.* ***Conclusão:*** *O teste de vergência ocular mostrou confiabilidade moderada, com baixa acurácia.*

Palavras-chave: *Convergência ocular. Transtornos da motilidade ocular. Confiabilidade dos dados. Reprodutibilidade dos testes.*

Introduction

The visual receptor, the eye, functions as both an endo- and exteroceptor and is a key sensory receptor involved in musculoskeletal motor activity. It contributes to cervical spine positioning and eye and head movements by integrating afferent information from the vestibular, visual, and proprioceptive systems. This integration generates appropriate or adaptive compensatory postural musculoskeletal responses, triggered by the activation of the cervical musculature to support the execution of activities that require eye movement.[1]

The neural commands needed to produce timely responses are organized and sent to the appropriate brain centers to execute the intended action; that is, the sensory information received is processed and transmitted to the effector mechanism, where appropriate motor response strategies are determined.[2] In this context, the visual system contributes to maintaining natural balance within the limits of the base of support (BoS) by informing how to align the head and trunk. Disruptions in this mechanism and in the ability to maintain the head in a neutral position can lead to poor posture, overloading the spine and potentially resulting in consequences such as pain.[3]

One of the most common disorders in the oculomotor system affecting binocular vision is convergence insufficiency (CI), which occurs in over 8% of the population, with its incidence increasing with age.[4] Ocular convergence is the simultaneous adduction movement of the eyes, enabling the visualization of nearby objects. The near point of convergence (NPC) is the closest point at which the eyes can converge. Convergence insufficiency (CI) is the inability of the eyes to converge adequately to meet the convergence demand, leading to intermittent sensory fusion. Ocular fusion is responsible for maintaining proper binocular alignment relative to the fixation point distance, in order to project the image onto both retinas and produce a single fused image. Ocular phorias occur when a misalignment is observed upon disrupting binocular vision, and are classified according to the direction of the deviation.[4-7]

CI can be assessed using the ocular vergence test. However, recent literature lacks information on the combined reliability and accuracy of these tests. In clinical settings, two methods are commonly used to assess ocular vergence: 1) the NPC test, which determines the visual system's ability to simultaneously maintain a single image on both foveae (the central region of the retina where images are formed and transmitted to the brain). The shift in gaze toward a near target is classified as convergence, and toward a distant target as divergence; and 2) the Convergence Insufficiency Symptom Survey (CISS), which is considered the gold standard instrument for this analysis.[8]

The present study aims to assist clinicians and researchers in diagnosing and identifying alterations in the ocular receptor by providing reproducible information on tests already used in clinical and research settings. These alterations can manifest as musculoskeletal consequences, postural adaptations, possible pain conditions, and other associated symptoms that may, over time, lead to poor posture and chronic pain, thereby supporting the development of treatment and prevention strategies. Finally, the hypothesis of this study is that the convergence test is reliable and accurate for assessing alterations in the ocular receptor.

Methods

This is a cross-sectional observational study, approved by the Research Ethics Committee of the State University of Northern Paraná (UENP), under protocol no. 5.501.335 (CAAE 57521821.2.0000.8123). The sample comprised volunteers of both sexes, aged 18 to 49 years, recruited voluntarily through social media advertisements,[9-10] and referred to the Physical Therapy Clinic at the UENP Health Sciences Center, where they were informed about the assessment protocol.

Volunteers who consented to the assessment protocol and signed the informed consent form were included in the study. Exclusion criteria included the presence of neurological symptoms and/or psychiatric disorders, the use of medications that could alter sensory perception, and those with ocular tropia. The sample size calculation, based on the 8% prevalence of convergence insufficiency reported by Dornbos et al.,[4] considered an acceptable inter-rater Kappa value ≥ 0.3, a power of 80%, and an alpha of 0.05, indicating an ideal sample size of 113 individuals. A total of 123 individuals were included in this study.

The research followed the recommendations of the Guidelines for Reporting Reliability and Agreement Studies (GRRAS)[11] and was conducted at the UENP clinical physical therapy outpatient clinic in the city of Jacarezinho, Paraná state, between June and August 2022. To ensure privacy and prevent any potential exposure or embarrassment for the participants, all procedures were performed in a private room. Three posturology specialists, trained and calibrated in administering visual tests and with an average of three years of experience, conducted all assessments.

Assessment protocol

After signing the informed consent form, volunteers underwent a battery of tests, beginning with assessment of the Near Point of Convergence (NPC) to measure the maximum capacity for ocular convergence using a real object (a pen). After a 5-minute rest period (with a black mask covering their eyes), the volunteers completed the CISS questionnaire to identify any ocular convergence alterations. The test battery was administered by three different evaluators. The volunteer remained in the evaluation room while the evaluators entered individually in a randomized order determined by a draw.

The battery was conducted on two occasions, with a one-week interval between sessions. All evaluators applied the complete test battery on both occasions. A random draw defined the order of the evaluators, who did not have access to each other's assessments between sessions. This procedure was implemented to analyze the intra- and inter-rater reliability of the test.

Instruments and data collection techniques

The ocular convergence test, which measures the maximum capacity of ocular convergence under conditions of sustained fusion during near vision,[8] was used to assess NPC. With the volunteer standing, they were instructed to fix their gaze on a real object – a pen placed 40 cm away at eye level – and then asked if the object appeared doubled. If the response was affirmative, the pen was moved away until the volunteer saw only one object. If the response was negative, the pen was moved closer until the volunteer indicated that it appeared doubled (or if the examiner observed that one eye lost fixation or the reflexes were misaligned). At this point, the distance was measured and recorded in centimeters using a millimeter ruler. To interpret the values obtained,8 the following crirteria were used: between 3.0 and 4.0 cm = normal; between 4.1 and 6.9 cm = sufficient; 7.0 cm or more = insufficient.

To assess symptoms of convergence insufficiency, the 15-item CISS validated for Portuguese was administered (Table 1). Responses to each item ranged from "never" to "always" on a five-point Likert scale. This standardized, approved, validated, and reliable tool yields a final score by summing all responses, which is used to detect changes in ocular vergence, according to the following criteria: 0 to 10 = normal binocular vision; 11 to 36 = suspected convergence insufficiency; 37 to 60 = convergence insufficiency.

Data analysis

Data from the three evaluators were used for analysis. Weighted linear Cohen's Kappa was used to determine intra- and inter-rater reliability. Inter-rater agreement was calculated for three pairs of therapists. Intra- and inte-rater agreement rates were analyzed by calculating the proportion of agreement, exact agreement, and agreement expected by chance,[12,13] and confidence intervals [95%] calculated for the values.

A Kappa coefficient of 1.0 indicates perfect agreement beyond chance. Values greater than 0.80 are considered excellent; between 0.60 and 0.80, substantial; 0.40 to 0.60, moderate; 0.20 to 0.40, fair; and less than 0.20, poor.[14] According to Akobeng,[15] an area under the curve (AUC) of 0.5 represents chance.

The AUC values for classifying accuracy were as follows: > 0.5 to ≤ 0.7 = low accuracy; > 0.7 to ≤ 0.9 = moderate accuracy; > 0.9 to < 1.0 = high accuracy; and 1.0 = perfect test. Data analysis was performed using the Statistical Package for the Social Sciences (SPSS, Chicago, IL), version 22.

Table 1 - Convergence Insufficiency Symptom Survey (CISS)

		Never	Rarely	Sometimes	Often	Always
1.	Do your eyes feel tired while reading or doing close work?					
2.	Do your eyes feel uncomfortable while reading or doing close work?					
3.	Do you have headaches while reading or doing close work?					
4.	Do you feel sleepy while reading or doing close work?					
5.	Do you lose concentration while reading or doing close work?					
6.	Do you have trouble remembering what you have read?					
7.	Do you see double while reading or doing close work?					
8.	Do the words move, jump, swim, or appear to float on the page while reading?					
9.	Do you feel you read slowly?					
10.	Do your eyes hurt while reading or doing close work?					
11.	Are your eyes swollen while reading or doing close work?					
12.	Do you have eye strain while reading or doing close work?					
13.	Do the words focus and unfocus while reading or doing close work?					
14.	Do you lose your place while reading or doing close work?					
15.	Do you have to reread the same line of a text?					
		x0	x1	x2	x3	x4

Results

A total of 123 volunteers were analyzed, with a mean age of 28 ± 7 years and balanced sex distribution (50.4% male and 49.6% female). Table 2 shows moderate interrater agreement in the ocular vergence test among the three evaluators in both the first and second evaluation, with Kappa values ranging from 0.43 to 0.59 in both assessments.

Table 3 shows the intrarater reliability of the ocular vergence test. Fair agreement was detected in two of the evaluators (Kappa < 0.40), and substantial agreement in one (Kappa = 0.75).

Finally, Table 4 and Figure 1 show the results regarding the accuracy of the ocular vergence test. Low accuracy was detected in the test (AUC = 0.53), with the best cutoff point having a value of 0.50.

Table 2 - Inter-rater analysis of the ocular vergence test

Kappa assessment (95% CI)	EV1 vs EV2	EV1 vs EV3	EV2 vs EV3
First	0.54 (0.24 to 0.77)	0.49 (0.20 to 0.73)	0.43 (0.10 to 0.69)
Second	0.59 (0.11 to 0.88)	0.46 (-0.03 to 0.78)	0.58 (0.15 to 0.85)

Note: EV1 = evaluator 1; EV2 = evaluator 2; EV3 = evaluator 3.

Table 3 - Intrarater analysis of the ocular vergence test

Evaluator	Kappa
1	0.37 (0.05 to 0.63)
2	0.75 (0.46 to 0.94)
3	0.38 (0.08 to 0.66)

Table 4 - Accuracy analysis of the ocular vergence test with the Convergence Insufficiency Symptom Survey

AUC (95%CI)	Best cutoff point	Sensitivity	Specificity
0.53 (0.40 to 0.65)	0.50	0.83	0.76

Note: AUC = area under the curve.

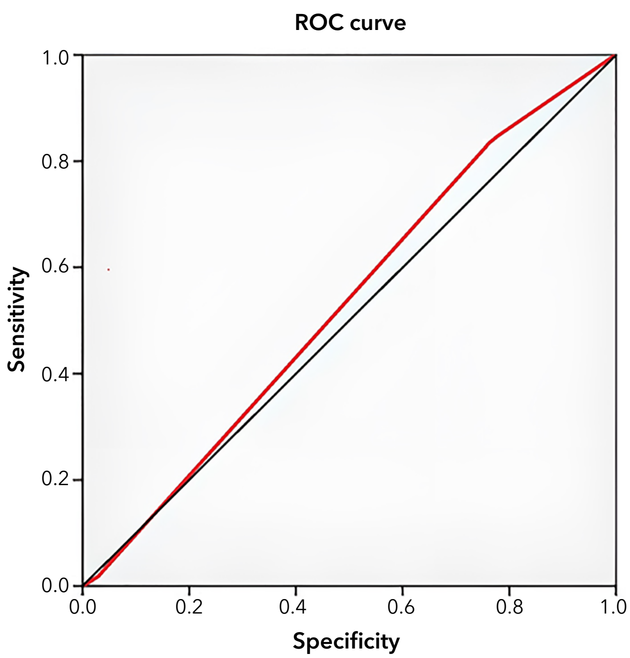

Figure 1 - ROC curve analysis of the accuracy of the ocular vergence test with the Convergence Insufficiency Symptom Survey.

Discussion

The primary findings of this study indicate that the ocular vergence test demonstrates both inter-rater (Kappa = 0.43 to 0.59) and intra-rater (Kappa = 0.40 to 0.75) reliability, but low accuracy (AUC = 0.53) for the methodology used. The use of measurement instruments requires knowledge, attention, and skill in application, since the analysis and reliability results are influenced by the methodological rigor of the assessments and the subjectivity of the measurements.[16,17]

A literature search revealed no studies that assessed the reliability of clinical vergence tests using the same methodology as the present research, with the exception of the CISS. However, based on evaluation of the NPC, the ocular vergence test is used as a measure complementary to the CISS for diagnosing individuals with convergence insufficiency (CI),[18] as described in the optometry standards and routines manual published by the Brazilian Optics and Optometry Council.[8] These tests are also used to assess the effectiveness of visual thera-py in CI treatment. Wolf and Taglietti[19] observed that visual therapy was effective for treating CI, with a reduction in the NPC and symptoms such as eye fatigue, headaches, blurry vision, diplopia, sleepiness, burning sensation, excessive tear production, difficulty concentrating, and difficulty understanding after short reading periods or activities that require prolonged near vision. This emphasizes the relevance of the test's applicability, since it is used as an instrument to assess the effectiveness of CI treatment.

Other instruments used to administer the convergence test include the Royal Air Force, a sliding device with a centimeter scale, pencil, examiner's index finger, light, red filter, and digital caliper. Studies have found no significant differences between using a pencil or the examiner's index finger to measure the NPC. Consequently, based on cost-benefit analysis and clinical practice, examiner preference is recommended.[20-22] This study therefore used the most cost-effective material offering the greatest benefit to facilitate the assessment and enhance its utility within the intended clinical setting.

A study with a design similar to the present research assessed the reliability of the NPC through retesting by different examiners after one week, reporting an ICC of 0.95. However, the instrument used was the Royal Air Force, and only 10 individuals were assessed for reliability.[23] By contrast, our study presents the cost-benefit instrument with comparable evidence regarding the results and a larger number of participants to strengthen the quality of the evidence.

Another study found a significant and statistically moderate correlation between the symptoms reported in the CISS and the NPC.[24] Consistent with our findings, it concluded that associations with the questionnaire indicate its effectiveness in diagnosing CI. Reinforcing the current evidence, the present study also highlights the need to assess the accuracy of the questionnaire.

Studies indicate that the lack of standardization in instruments and the disagreement regarding the number of test repetitions needed to produce meaningful clinical information are potential biases in the evaluation. This is because some individuals may exhibit an NPC < 6 cm after only a few minutes of near work.[21,24]

The significance of studies aimed at diagnosing extraocular muscle (EOM) function is underscored by the increasing use of electronic devices such as computers, tablets, smartphones, and gaming consoles, which require greater effort for near vision. This increased near work demands greater accommodation, convergence and enhanced visual attention. When this effort is sustained (more than two hours per day), the adaptation mechanisms will fail, leading to exhaustion of the ocular muscles (intrinsic and extrinsic) and subsequent visual fatigue (asthenopia), thereby impairing the ability to perform intended tasks.[25]

As such, further research is recommended to improve the understanding of the validity of ocular convergence tests through the NPC. This research should use different instruments, a larger number of evaluators, and varied levels of evaluator experience, given that it has limitations in terms of the non-systematized evaluation protocol, the use of only one instrument, three evaluators, and an average of three years of experience. Based on the present literature, establishing the validity of the ocular convergence test is crucial for diagnosing changes in extraocular muscles and for assessing the outcomes of the proposed treatment, given its easy clinical applicability, as a quick and low-cost test.

Conclusion

The ocular vergence test was found to be reliable and accurate and can be recommended for clinical assessment of ocular vergence. Future studies should use more rigorous evaluation protocols and consider the impact of the number of repetitions, time averaging, materials, and reassessment rest periods. Moreover, recruiting a larger number of participants and evaluators, and conducting more assessments are necessary.

With respect to methodological quality, special attention should be given to blinding the evaluators to ensure the results are truly representative.

Authors' contributions

VCGP and FJJ conceived and approved the study design and project. VCGP, PFP, TTDA, and GLSL collected the data and, along with FJJ, analyzed and interpreted them. TTDA was responsible for statistical analysis, while VCGP, FJJ, OBBA, FMB, and RACA developed and reviewed the manuscript. All the authors approved the final version.

References

1. Figueiredo VFR. Análise da função oculomotora e movimentos cervicais na convergência visual [master´s thesis]. Brasília: Universidade de Brasília; 2020. https://repositorio.unb.br/handle/10482/38885

2. Erickson GB. Sports Vision: Vision care for the enhancement of sports. St. Louis, MO: Butterworth-Heinemann; 2007.

3. Jang JU, Jang JY, Tai-Hyung K, Moon HW. Effectiveness of vision therapy in school children with symptomatic convergence insufficiency. J Ophthalmic Vis Res. 2017;12(2):187-92. https://doi.org/10.4103/jovr.jovr_249_15

4. Dornbos B, Kokotas V, Kitchener G. Convergence insufficiency. Points de Vue Int Rev Ophthalm Opt. 2019:1-5.

5. Von Noorden GK, Campos EC. Binocular vision and ocular motility: theory and management of strabismus. 6th ed. New York: Mosby; 2002.

6. Scheiman MM. Optometric management of learning-related vision problems. New York: Mosby; 2005.

7. Sonoda RT, Moura KL, Santos EM. Forias: métodos de avaliação objetivos e subjetivos. Recima21. 2023;4(4):e443019. https://doi.org/10.47820/recima21.v4i4.3019

8. Rocha F. Manual de normas e rotinas da optometria. Brasília: Conselho Brasileiro de Óptica e Optometria; 2023.

9. Gonçalves VP, Scharlach RC. Avaliação oculomotora em adultos: um estudo do efeito da idade e de alterações visuais. Audiol Commun Res. 2016;21:e1704. https://doi.org/10.1590/2317-6431-2016-1704

10. World Health Organization. World report on vision. Geneva: WHO; 2020. https://www.who.int/publications/i/item/9789241516570

11. Kottner J, Audigé L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. J Clin Epidemiol. 2011;64(1):96-106. https://doi.org/10.1016/j.jclinepi.2010.03.002

12. Tavares C, Nunes AMMF, Nunes AJS, Pato MV, Monteiro ML. Tradução e validação de Conver-gence Insufficiency Symptom Survey (CISS) para o português – resultados psicométricos. Arq Brasil Oftalmol. 2014;77(1):21-4. https://doi.org/10.5935/0004-2749.20140007

13. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther. 2005;85(3):257-68. https://doi.org/10.1093/ptj/85.3.257

14. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;327(8476):307-10. https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(86)90837-8/fulltext

15. Akobeng AK. Understanding diagnostic tests 3: Receiver operating characteristic curves. Acta Paediatr. 2007;96(5):644-7. https://doi.org/10.1111/j.1651-2227.2006.00178.x

16. Souza AC, Alexandre NMC, Guirardello EB. Propriedades psicométricas na avaliação de instrumentos: avaliação da confiabilidade e da validade. Epidemiol Serv Saude. 2017;26(3):649-59. https://doi.org/10.5123/S1679-4974201700030 0022

17. Echevarría-Guanilo ME, Gonçalves N, Romanoski PJ. Propriedades psicométricas de instrumentos de medidas: bases conceituais e métodos de avaliação - Parte I. Texto Context Enferm. 2017;26(4):e1600017. https://doi.org/10.1590/0104-07072017001600017

18. Horwood AM, Toor SS, Riddell PM. Change in convergence and accommodation after two weeks of eye exercises in typical young adults. J AAPOS. 2014;18(2):162-8. https://doi.org/10.1016/j.jaapos.2013.11.008

19. Wolf CM, Taglietti M. Exercícios oculares na insuficiência de convergência: série de casos. Rev Soc Port Oftalmol. 2019;43(1). https://doi.org/10.48560/rspo.14040

20. Siderov J, Chiu SC, Waugh SJ. Differences in the near-point of convergence with target type. Ophthalmic Physiol Opt. 2001;21(5):356-60. https://doi.org/10.1046/j.1475-1313.2001.00609.x

21. Sharma IP. RAF near point rule for near point of convergence — A short review. Ann Eye Sci. 2017;2:16. https://aes.amegroups.org/article/view/3639/html

22. Giffard P, Daly L, Treleaven J. Influence of neck torsion on near point convergence in subjects with idiopathic neck pain. Musculoskelet Sci Pract. 2017;32:51-6. https://doi.org/10.1016/j.msksp.2017.08.010

23. Sampaio MFT. Relação entre sinais e sintomas em pacientes com disfunções da visão binocular, acomodativas e oculomotoras [master´s thesis]. Braga: Universidade do Minho; 2015. https://hdl.handle.net/1822/35654

24. Hassan LI, Ibrahim SM, Abdu M, MohamedSharif A. Prevalence of convergence insufficiency among secondary school students in Khartoum, Sudan. Oman J Ophthalmol. 2018;11(2):129-33. https://doi.org/10.4103/ojo.ojo_170_2017

25. Henriques S, Lopes AS, Roque J, Gonçalves S, Condado P, Gomes MV, et al. Astenopia Digital - Estudo Konecta. In: Vaz FT, editor. Perguntas e Respostas em Ergoftalmologia. Lisboa: Sociedade Portuguesa de Oftalmologia; 2018. p. 38-41. https://ergophthalmology.com/pt-pt/book/42-astenopia-digital-estudo-konecta