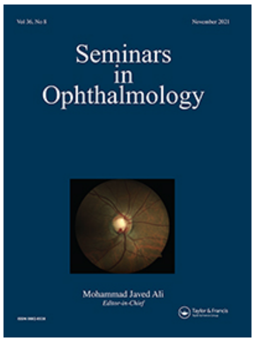

Seminars in Ophthalmology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/isio20

Effectiveness and repeatability of eye-tracking-based test in strabismus measurement of children

Leilei Zou, Tian Tian, Tamara Wygnanski-Jaffe, Oren Yehezkel, Shu Wang, Avital Moshkovitz, Xinghuai Sun, Hong Liu & Rui Liu

To cite this article: Leilei Zou, Tian Tian, Tamara Wygnanski-Jaffe, Oren Yehezkel, Shu Wang, Avital Moshkovitz, Xinghuai Sun, Hong Liu & Rui Liu (2021): Effectiveness and repeatability of eye-tracking-based test in strabismus measurement of children, Seminars in Ophthalmology, DOI: 10.1080/08820538.2021.2002918

To link to this article: https://doi.org/10.1080/08820538.2021.2002918

Published online: 23 Nov 2021.

SEMINARS IN OPHTHALMOLOGY
https://doi.org/10.1080/08820538.2021.2002918

Effectiveness and repeatability of eye-tracking-based test in strabismus measurement of children

Leilei Zou[+a,b,c,d], Tian Tian[+a,b,c], Tamara Wygnanski-Jaffe[e,f], Oren Yehezkel[g], Shu Wang[a,b,c], Avital Moshkovitz[h], Xinghuai Sun[a,b,c], Hong Liu[d], and Rui Liu[a,b,c]

[a]Department of Ophthalmology, Eye and ENT Hospital of Fudan University, Shanghai, China; [b]NHC Key Laboratory of Myopia (Fudan University), Shanghai, China; [c]Key Laboratory of Myopia, Chinese Academy of Medical Sciences (Fudan University), Shanghai, China; [d]Department of Ophthalmology, Shanghai Children's Medical Center, School of Medicine, Shanghai Jiao Tong University, Shanghai, China; [e]Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel; [f]Goldschleger Eye Institute, Tel Hashomer, Israel; [g]Novasight Ltd, Airport City, Israel; [h]School of Optometry and Vision Sciences, Faculty of Life Sciences, Bar-Ilan University, Ramat Gan, Israel

ABSTRACT

Objectives: To assess the effectiveness of an eye-tracking-based test (ETBT) to measure eye deviation angle vs. a manual prism alternating cover test (PACT) in children with strabismus.
Design: The prospective, masked, cross-sectional study included 95 children aged 1.8 years and older. Eye deviation was tested twice by each of ETBT and PACT. Each subject underwent four strabismus measurements, two by the ETBT and two by PACT. In each test, subjects were fixated on accommodative targets at 50 cm, with habitual optical correction allowed. Masked examiners compared the manual PACT results with those of the ETBT.
Results: There was a high correlation (about 90%) between the ETBT and PACT. Repeatability of ETBT was higher than that of PACT (correlation coefficients of 0.99 and 0.91 respectively, $p < .002$). Age, strabismus type, and eye deviation measurement did not affect repeatability of ETBT. However, in PACT, results could not be correlated between the two examiners when the deviation was larger than 40 prism diopters.
Conclusions: The ETBT was effective in measuring eye deviation in children as young as 1.8 years. The ETBT showed higher repeatability compared to PACT.

ARTICLE HISTORY
Received 16 August 2021
Revised 3 September 2021
Accepted 6 September 2021

KEYWORDS
Strabismus; Test; Measurement; Repeatability; Automatic

INTRODUCTION

Accurately measuring total ocular deviation in patients with strabismus is vital in clinical practice for several reasons. Monitoring the angle of deviation on consecutive visits allows clinicians to identify and quantify deterioration and can help inform parents/patients if strabismus surgery is warranted. Studies have shown that preoperative deviation is a better predictor of response to surgery than the amount of recession or resection of the muscles in the surgery.[1,2]

The prism alternating cover test (PACT) is currently considered as the gold standard for measuring strabismus; however, as a subjective test, it is affected by several factors, including patient cooperation and fatigue and examiner skill, leaving it subject to high interexaminer variability.[2,3] Recently, a non-invasive, objective, automated eye-tracking technique has been developed that can precisely locate eye gaze direction and corneal reflexes, similar to the Hirschberg test.[4] This method also automatically quantifies the sum of the dynamic and latent properties of strabismus associated with disruption of fusion.[5] However, potential limitations to the system include low levels of accuracy,[6] and a lack of effective features for precise strabismus angle evaluation.[4,7]

In this study, we evaluated an automated eye-tracking-based test (ETBT) that integrates an objective eye tracking-based device and a computerized cover test procedure EyeSwift™ system (NovaSight Ltd., Israel). We compared its effectiveness and repeatability and its correlation with PACT in a group of Chinese subjects with strabismus. Other studies have evaluated this system in Israeli children.[8] This study further analyzed the influence of age, type of strabismus and range of deviation on device repeatability.

METHOD

A total of 95 participants were recruited from the Eye and ENT Hospital of Fudan University between April 2020 and September 2020. The institutional review board of Eye and ENT Hospital of Fudan University approved the study protocol. The study adhered to the tenets of the Declaration of Helsinki. The guardian of each participant signed informed consent. Those unable to cooperate, those with low vision (≤ 20/200), those with abnormal anterior segment configuration, and those with paralytic strabismus were excluded.

CONTACT Hong Liu liuhongzef@263.net Department of Ophthalmology, Shanghai Children's Medical Center, School of Medicine, Shanghai Jiao Tong University, 1678 Dongfang Road, Shanghai 200127, China; Rui Liu lratb1@aliyun.com Department of Ophthalmology, Eye and ENT Hospital, Fudan University, 83 Fenyang Road, Shanghai, 200031, China

[+]These authors contributed equally to the work presented here and should therefore be regarded as equivalent authors.

Following a complete ocular and orthoptic assessment, each subject went through strabismus measurement four times, two by ETBT and two by PACT. The first manual test was used to determine whether the participant meets the inclusion criteria. The order of the subsequent tests, that is, the second manual test and the automated test, was randomized to eliminate any bias in the results, such as fatigue. Both tests were performed during the same visit within one hour. Subjects were fixated on accommodative targets at 50 cm; habitual optical correction was allowed. The ETBT measurements were automated; the PACT was conducted manually with two examiners who took turns to conduct the test. Each examiner was masked to the results of the other and to the results of ETBT.

Automated Test Procedure (ETBT)

The EyeSwift system is a system based on eye tracking technology, and a vector formed by the angle between the cornea and pupil reflections was calculated by the eye tracker. Each subject, with a pair of dedicated wireless glasses used to cover either eye. The automated test was similar to that of PACT, but instead of using prisms for evaluating the magnitude of the deviation, the device uses monocular targets, which are shifted to match each eye's gaze position. The target that has a short animated movie, containing accommodative fine details, was presented. One eye was covered at a time, and the algorithm searched for a change in the fixation position of the non-covered eye measuring deviation and its direction, with no binocular viewing permitted in between the alternations. The test duration depends on the participant's fixation stability, allowing the system to collect a sufficient amount of data to confirm the eye's gaze position. Once the system calculates the gaze position of the deviation, it shifts the deviating eye monocular target towards its gaze position to the point where the maximal deviation is found. After moving the target, the process is repeated, until no eye movement was detected. Next, using the combination of the obtained distance between the two monocular targets on the monitor and the sitting distance, the precise amount of deviation is calculated using the formula: Distance between target centers/sitting distance. No verbal communication was required. Once the measurement was completed, numeric and graphic results concerning the deviation type, direction and magnitude were recorded. Yehezkel et al.[1] have previously described the components of the EyeSwift system in greater detail.

Manual Examination Procedure (PACT)

A standard set of loose plastic prisms (block prisms) was used for PACT measurements, and individual prisms increased in single prism diopter (PD) increments. For horizontal deviations below 40 PD, examiners did not split prisms but chose proper prisms; for deviations greater than 40 PD, examiners split the prisms between eyes equally rather than stacked prisms. Deviations were recorded for the values of the prisms that came closest to neutralizing the misalignment.

Statistical Analysis

Statistical analysis was performed using SAS software 9.4. A paired t-test was conducted for group comparisons. Agreement between ETBT and PACT, as well as repeatability of each method, was represented by Bland-Altman plots and by Spearman's rank concordance correlation coefficients, and when necessary, partial correlation analysis was employed with age and/or type of strabismus as covariates. Differences between the two correlation coefficients were explored by u tests.[9] A two-tailed p value less than 0.05 was considered of statistical significance.

RESULTS

A total of 95 children aged 1.8–15 years (6 ± 2.6 years) were enrolled, of whom 47 were male and 48 were female. There were 21 cases of esodeviation, 50 cases of exodeviation, 24 cases of orthophoria and eight cases of small-angle horizontal deviation combined with small-angle vertical deviation.

In the eight cases of small-angle deviation, there were four cases where both ETBT and PACT succeeded in recording horizontal deviation in every measurement. In the remaining four cases, however, ETBT successfully recorded the horizon deviation in every measurement, whereas only one examiner using PACT succeeded. Similarly, the ETBT detected vertical deviations in all cases, whereas PACT detected vertical deviations in only three cases.

There was a high positive correlation between PACT and ETBT ($R = 0.9$, $p < .002$, Figure 1), and Figure 2 illustrates the Bland-Altman plots comparing these two methods. The half-width of the 95% limit of agreement was ±13.85 PD and the mean value was 0.3 PD.

There were no significant differences in accuracy between the two tests. The average result of the repeated ETBT was 16.28 ± 16.00 PD and the average PACT deviation measured by two examiners was 16.58 ± 16.00 PD ($p = .673$).

Figure 3 shows the Bland-Altman plots comparing two tests in the ETBT. The half-width of the 95% limit of agreement was ± 4.78 PD and the mean value was 0.12 PD. The correlation coefficient between two tests in the ETBT was 0.99. Figure 4 shows the interexaminer variability of results from two examiners who conducted the PACT test. The half-width of the 95% limit of agreement was ± 14.1 PD and the mean value was −0.57 PD. The correlation coefficient between examiners in PACT test was 0.91.

Effects of three factors (age, strabismus type, and range of deviation) on repeatability were analyzed. Table 1 shows the mean ± SD and adjusted correlation coefficient for each factor, and found results were independent of the factor. When the adjusted correlation coefficient was analyzed by u tests (Table 2), a significantly higher repeatability of ETBT than that of PACT was achieved when each of the factors were assessed individually. For small-angle deviations < 20 PD and for deviations > 40 PD, ETBT showed high correlation, while PACT failed to do so.

DISCUSSION

In this study, we quantified ocular alignment in an automated method using the EyeSwift system. Our results showed excellent agreement of the ETBT with those of

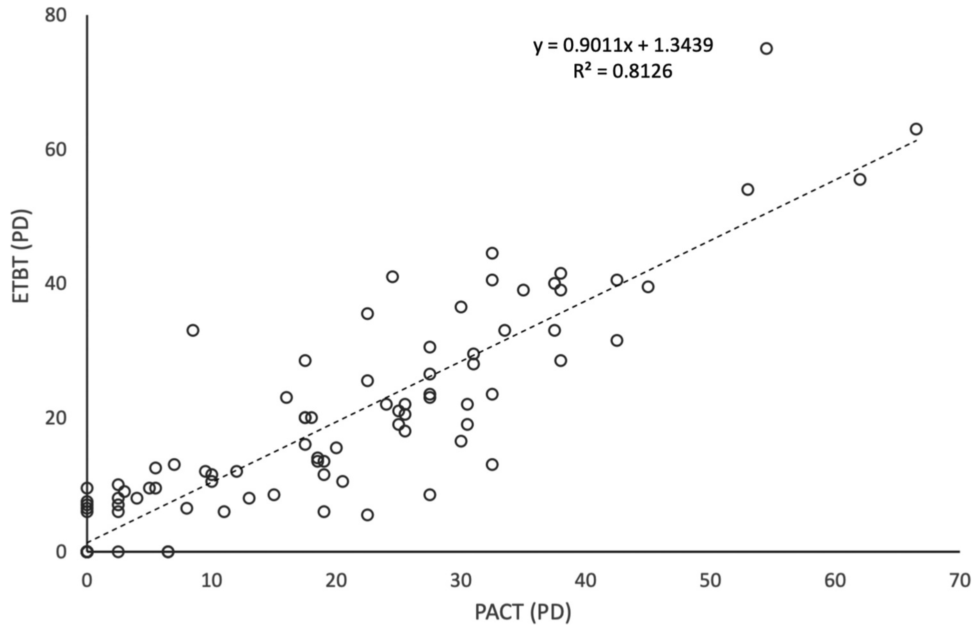

Figure 1. Correlation between PACT and ETBT. A strong correlation exists between the two measurements (correlation coefficient R = 0.90; $p < .002$). PACT: prism alternating cover test. ETBT: eye-tracking-based test.

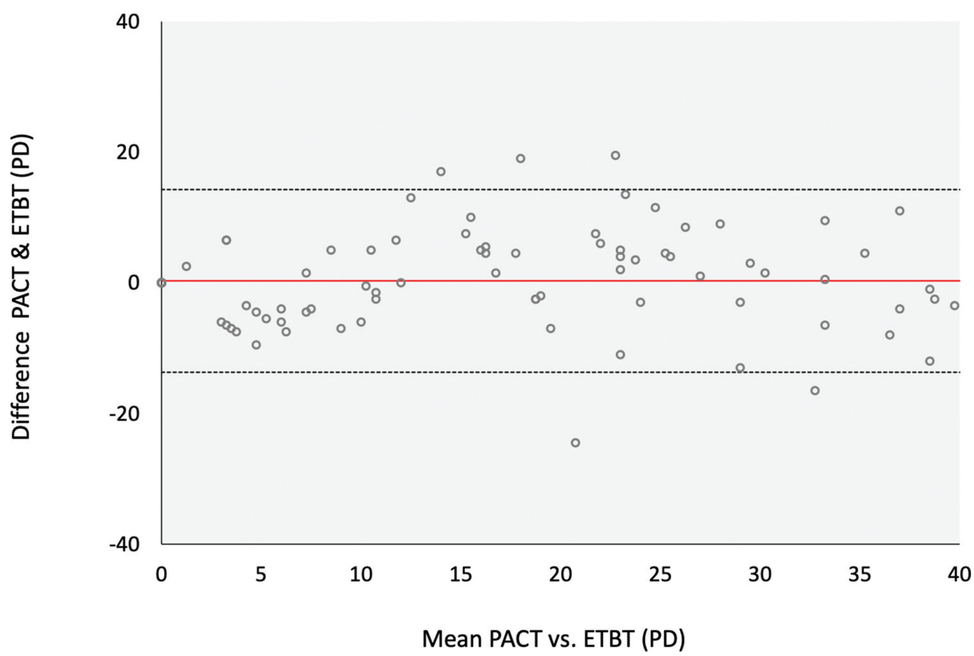

Figure 2. Bland-Altman plots showing variability between PACT and ETBT. Upper and lower dotted lines show the 95% limit of agreement. The half-width of the 95% limit of agreement was ±13.85 PD and the mean value was 0.3 PD. PACT: prism alternating cover test. ETBT: eye-tracking-based test. PD: prism diopter.

PACT (R = 0.9), and no significant difference was found between outcomes of ETBT and PACT (p = .673). This suggests that the ETBT can be successfully used to measure ocular deviation without consideration for examiner expertise.

In this analysis, the ETBT showed profoundly higher measurement repeatability compared to PACT. Our results found a substantially lower half-width of the 95% limit of agreement with ETBT than PACT, and a much lower correlation coefficient between examiners in the PACT test than with the ETBT. Because PACT measurements are performed manually, they are highly subjective and depend largely on examiner experience and proficiency and can be prone to measurement errors.[10,11] Even the most experienced eye care professionals have shown inconsistencies ranging from 6 to 12 PD when applying PACT.[12] Other studies have found that the interexaminer reliability was about 12 PD for deviations >20 PD and about 6 PD for deviations <20 PD.[3]

In this analysis, none of the factors analyzed (age, strabismus type, and range of deviation) affected repeatability of the ETBT, but each of them lowered repeatability of the PACT.

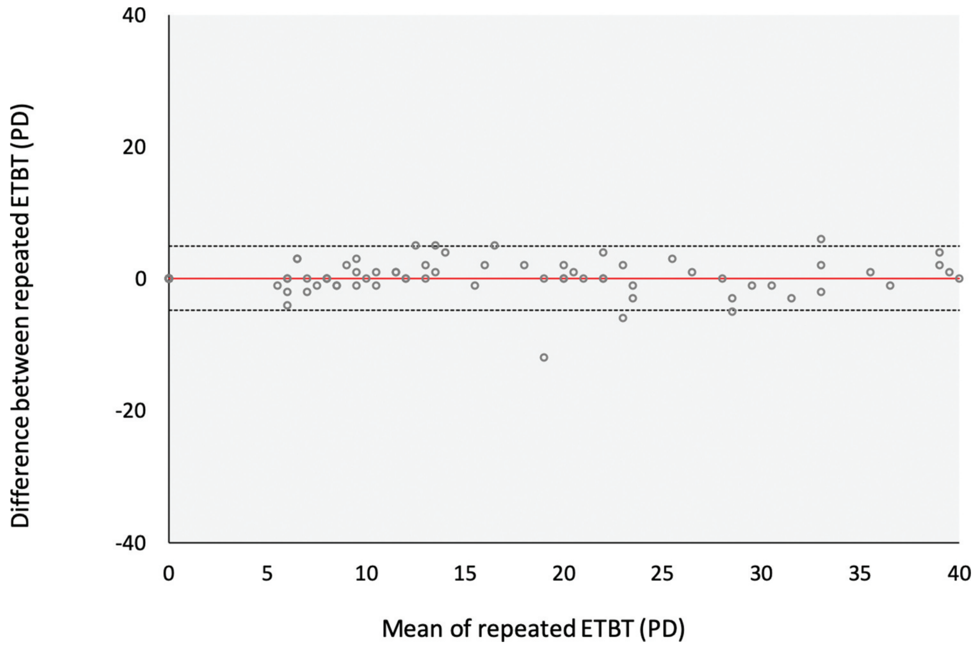

Figure 3. Bland-Altman plots showing variability between the repeated ETBT. Upper and lower dotted lines show the 95% limit of agreement. The half-width of the 95% limit of agreement was ± 4.78 PD and the mean value was 0.12 PD. ETBT: eye-tracking-based test. PD: prism diopter.

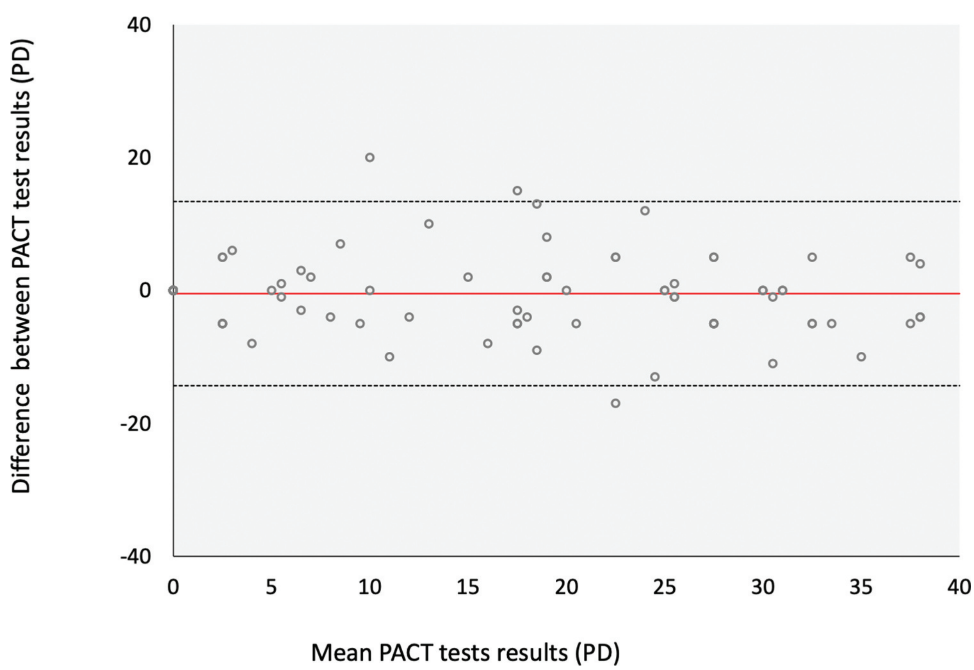

Figure 4. Bland-Altman plots showing variability of two examiners in PACT. Upper and lower dotted lines show the 95% limit of agreement. The half-width of the 95% limit of agreement was ± 14.1 PD and the mean value was −0.57 PD. PACT: prism alternating cover test. PD: prism diopter.

Our experience suggests the ETBT is more efficient when evaluating younger patients than PACT. In PACT, plastic prisms are alternately placed in the frontal plane position before one eye with a cover paddle occluding the other eye. Subjects must then be re-fixated on a static accommodative target with the just-uncovered eye, and for younger patients, this is often problematic as they have difficulty in concentrating on the target due to the moving cover paddle. Also, even in cooperative patients, a high degree of observational skill and experience is required to determine correctly when the examiner watches for elimination of refixation eye movements with gradually increasing prism power.[13] By comparison, the ETBT presents a dynamic fixation target, usually a short animated movie, to the patient, which has an advantage of enticing younger patients to fixate on the movie, thereby allowing greater accuracy.

In this study, there was no statistically significant difference between repeatability of subjects in ETBT, regardless of age. We had expected lower repeatability of PACT on subjects younger than 6 years old, and while the repeatability of PACT tended to be lower for this age group than in the

Table 1. The mean ± SD and adjusted correlation coefficient of the three factors.

		PACT				ETBT			
Factors	n	mean ± SD (PD)		correlation coefficients		mean ± SD (PD)		correlation coefficients	
		test-1	test-2	r	*p* value	test-1	test-2	r	*p* value
Age (in years)									
<6	68	17.5 ± 16.4	18.2 ± 18.0	0.87*	<0.001	17.2 ± 17.3	17.1 ± 17.40.98*		<0.001
>6	35	14.0 ± 14.7	14.0 ± 14.0	0.88*	<0.001	14.7 ± 13.1	14.5 ± 13.20.96*		<0.001
Direction									
ESO	21	21.4 ± 13.4	23.9 ± 14.6	0.79#	<0.001	19.0 ± 15.0	19.0 ± 14.90.97#		<0.001
EXO	50	24.3 ± 14.5	24.5 ± 15.3	0.87#	<0.001	24.5 ± 15.1	24.2 ± 15.40.97#		<0.001
Range (PD)									
<20△	38	9.4 ± 7.8	9.0 ± 7.3	0.54 USD	0.001	11.4 ± 7.3	10.8 ± 7.0 0.92 USD		<0.001
20–40	34	28.1 ± 5.9	29.9 ± 6.1	0.59 USD	<0.001	26.1 ± 10.4	26.5 ± 10.50.94 USD		<0.001
>40	7	51.9 ± 9.6	52.7 ± 16.5	0.01 USD	0.983	51.4 ± 15.0	51.1 ± 15.40.97 USD		0.007

* adjusted for direction; # adjusted for age; $ adjusted for age and direction; △ patients diagnosed with orthophoria by both PACT and ETBT were excluded from this group; PACT: prism alternate cover test; ETBT: Eye-tracking-based test; PD: prism diopter; ESO: esotropia; EXO: extropia.

Table 2. Difference between PACT and ETBT in adjusted correlation coefficients by u tests.

Factors	Z_{rPACT} vs. Z_{rETBT}	*p* value
Age (in years)		
<6	6.036	<0.001
> = 6	2.438	0.015
Direction		
ESO	2.861	0.004
EXO	3.443	<0.001
Range (PD)		
<20$^{\triangle}$	4.243	<0.001
20–40	4.240	<0.001
>40	2.869	0.004

$^{\triangle}$Subjects diagnosed with orthophoria by both PACT and ETBT were excluded from this group. PACT: prism alternate cover test; ETBT: Eye-tracking-based test; PD: prism diopter; ESO: esotropia; EXO: exotropia; PD: prism diopter.

ETBT, the differences did not reach statistical significance. We recommend a further analysis with a larger sample size to confirm our findings.

We also found that ETBT performed well, regardless of the strabismus type. The ETBT showed significantly higher repeatability than PACT in both esotropia and exotropia measurements. Given that variability in measurements of horizontal deviation may originate from fusion control that depends largely on the patient's current general health, alertness, attention span and anxiety level,[14] employing occlusion glasses and animation targets containing fine accommodative details in ETBT may contribute to a steadier result. Five participants were determined to have small vertical deviations only when tested by ETBT, and three cases were detected in both the ETBT and PACT. To verify the validity of the ETBT, a third examiner (results not shown) repeated the PACT in these five participants and found that these cases were all dissociated vertical deviation (DVD). It has been noted in previous studies that the measurement of DVD is challenging[15] and that it is not possible to measure DVD using a prism and alternate cover test since there is no definitive end point.[16] The automated ETBT does not have those same obstacles and can measure movements of the covered eye because its eye tracker is capable of recording the covered eye through the closed LCD shutter. This allows for the different deviations between the two eyes to be quantified.

We also found the ETBT to be more accurate than the PACT, especially for small-angle deviations < 20 PD and for deviations > 40 PD. In the group of small-angle deviation (< 20 PD), four subjects were quantified when tested by ETBT, but were only quantified by one examiner of PACT and were judged as orthophoria by the other. This may reflect the difficulty in determining the magnitude of small-angle deviations and even in determining by PACT the existence of a small deviation. In our study, these four small-angle cases had both horizontal deviations coexisting with vertical deviations. Such coexistence can interfere with the PACT measurement of each type of deviation: when observing ocular movements during covering and uncovering of the eyes, the examiner can miss or be confounded by very small movements. Although most of these small-angle deviations are too small to affect surgical planning, recognition of the presence of the deviations may better aid clinicians in dispensing prismatic correction.[17] Using the ETBT to quantify tropias may prove to be useful in both clinical and research settings in cases of small-angle misalignment. When the deviation was larger than 40 PD, the repeatability of the ETBT was not affected, but with the PACT, the results from the two examiners were uncorrelated (p = .983). In PACT, loose plastic prisms for ophthalmic use were available in assorted powers up to 40 to 50 PD; however, for deviation over 40 PD, instead of using a single prism, prisms were split between eyes. Although loose plastic prisms are generally positioned with the posterior face of the prism in the coronal plane of the head to minimize measurement errors,[18] mechanical difficulties can arise when examiners try to simultaneously hold separate prisms in front of each eye in the frontal position. Furthermore, proper positioning is difficult and may induce a horizontal or vertical prismatic error.[19] Finally, loose plastic prisms would degrade acuity and impact fixation ability especially at higher angles. By comparison, the ETBT measures deviations without a prism, thereby eliminating these potential issues.

Our present study has several limitations. There were a small number of over- or under-corrections measured with the ETBT compared to the PACT because of the existence of variability in measurements of the angle of deviation in horizontal deviations. Diagnostic occlusion testing should be applied to the patients before both tests. The limited sample size of subjects with large-angle deviations may contribute to statistical error. Moreover, our sample size was equally small for measurements of vertical deviations, and the deviation was limited to a small range as well. We recommend and embrace a larger randomized prospective study that is in need to confirm these findings. Finally, the ETBT reported here was not designed to measure torsion, similar to PACT.

In summary, the ETBT shows excellent agreement with standard PACT and offers an accurate and objective method for automated measurements of ocular deviation. Its satisfactory repeatability also facilitates clinicians to assess and diagnose patients with strabismus and follow their clinical course.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Nature Science Foundation of China [grant numbers 82070997, 81900901 and 81770957] and Shanghai Municipal Health Commission Program [grant number 201940485].

References

1. Archer SM. Why strabismus surgery works: the legend of the dose-response curve. *J AAPOS*. 2018;22(1):1.e1–6. doi:10.1016/j.jaapos.2017.12.001.
2. Schutte S, Polling JR, van der Helm FC, et al. Human error in strabismus surgery: quantification with a sensitivity analysis. *Graefes Arch Clin Exp Ophthalmol*. 2009;247(3):399–409. doi:10.1007/s00417-008-0961-x.
3. Pediatric Eye Disease Investigator Group. Inter-observer reliability of the prism and alternate cover test in children with esotropia. *Arch Ophthalmol*. 2009;127:(1). 59–65. 10.1001/archophthalmol.2008.548.
4. Chen ZH, Fu H, Lo WL, et al. Eye-tracking-aided digital system for strabismus diagnosis. *Healthcare Technology Letters*. 2018;5(1):1–6. doi:10.1167/tvst.8.1.19.
5. Yang HK, Seo JM, Hwang JM, et al. Automated analysis of binocular alignment using an infrared camera and selective wavelength filter. *Invest Ophthalmol Vis Sci*. 2013;54(4):2733–2737. doi:10.1167/iovs.12-11400.
6. Mestre C, Otero C, Díaz-Doutón F, et al. An automated and objective cover test to measure heterophoria. *PLoS One*. 2018;13(11):e0206674. doi:10.1167/iovs.12-11400.
7. Valente TLA, de Almeida JDS, Silva AC, et al. Automatic diagnosis of strabismus in digital videos through cover test. *Comput Methods Programs Biomed*. 2017;140(3):295–305. doi:10.1007/s00417-008-0961-x.
8. Yehezkel O, Belkin M, Wygnanski-Jaffe T. Automated diagnosis and measurement of strabismus in children. *Am J Ophthalmol*. 2020;213:226–234. doi:10.1016/j.ajo.2019.12.018.
9. Weaver B, Wuensch KL. SPSS and SAS programs for comparing Pearson correlations and OLS regression coefficients. *Behav Res Methods*. 2013;45(3):880–895. doi:10.3758/s13428-012-0289-7.
10. Holmes JM, Leske DA, Hohberger GG. Defining real change in prism-cover test measurements. *Am J Ophthalmol*. 2008;145(2):381–385. doi:10.1016/j.ajo.2007.09.012.
11. Hrynchak PK, Herriot C, Irving EL. Comparison of alternate cover test reliability at near in non-strabismus between experienced and novice examiners. *Ophthalmic Physiol Opt*. 2010;30(3):304–309. doi:10.1111/j.1475-1313.2010.00723.x.
12. de Jongh E, Leach C, Tjon-Fo-Sang MJ, et al. Inter-examiner variability and agreement of the alternate prism cover test (APCT) measurements of strabismus performed by 4 examiners. *Strabismus*. 2014;22(4):158–166. doi:10.3109/09273972.2014.972521.
13. Pundlik S, Tomasi M, Liu R, et al. Development and preliminary evaluation of a smartphone app for measuring eye alignment. *Transl Vis Sci Technol*. 2019;8(1):19. doi:10.1167/tvst.8.1.19.
14. Yang HK, Hwang JM. The effect of target size and accommodation on the distant angle of deviation in intermittent exotropia. *Am J Ophthalmol*. 2011;151(5):907–13e901. doi:10.1016/j.ajo.2010.11.021.
15. Christoff A, Raab EL, Guyton DL, et al. DVD–a conceptual, clinical, and surgical overview. *J AAPOS*. 2014 4;18(4):378–384. doi:10.1016/j.jaapos.2014.03.009.
16. Raab EL. Dissociated vertical deviation. *Int Ophthalmol Clin*. 1985;25(4):119–131. doi:10.1007/BF02173296.
17. Bitner DP, Adesina OO, Ding K, et al. Comparison of objective and subjective techniques of strabismus measurement in adults with normal retinal correspondence. *J Pediatr Ophthalmol Strabismus*. 2017;54(4):216–220. doi:10.3928/01913913-20170322-02.
18. Thompson JT, Guyton DL. Ophthalmic prisms, measurement errors and how to minimize them. *Ophthalmology*. 1983;90(3):204–210. doi:10.1016/s0161-6420(83)34572-3.
19. Bishop JE. Magnetic prism alignment system for measuring large-angle strabismus. *J AAPOS*. 2014;18(1):101–102. doi:10.1016/j.jaapos.2013.09.016.