

Reference List - King-Devick Saccadic Tracking Test:

- Oride, Michael & Marutani, Joel & Rouse, Michael & Deland, Paul. (1986). Reliability Study of the Pierce and King-Devick Saccade Tests. *American journal of optometry and physiological optics*. 63. 419-24. 10.1097/00006324-198606000-00005.
- Gubanich, Paul & Gupta, Resmi & Slattery, Eric & Logan, Kelsey. (2019). Performance Times for the King-Devick Test in Children and Adolescents. *Clinical journal of sport medicine : official journal of the Canadian Academy of Sport Medicine*. 29. 374-378. 10.1097/JSM.0000000000000670.
- Eddy, Ryan & Goetschius, John & Hertel, Jay & Resch, Jacob. (2018). Test–Retest Reliability and the Effects of Exercise on the King-Devick Test. *Clinical Journal of Sport Medicine*. Publish Ahead of Print. 1. 10.1097/JSM.0000000000000586.
- Nguyen, Minh & King, Doug & Pearce, Alan. (2020). A Reliability and Comparative Analysis of the New Randomized King-Devick Test. *Journal of Neuro-Ophthalmology*. 40. 207-212. 10.1097/WNO.0000000000000829.
- Subotic, Arsenije & Ting, Windsor & Cusimano, Michael. (2017). Characteristics of the King-Devick test in the assessment of concussed patients in the subacute and later stages after injury. *PLOS ONE*. 12. e0183092. 10.1371/journal.pone.0183092.
- Hecimovich, Mark & King, Doug & Dempsey, Alasdair & Gittins, Mason & Murphy, Myles. (2018). In situ use of the King-Devick Eye Tracking and changes seen with youth Sport Related Concussion: Saccadic and blinks counts. *Journal of Science and Medicine in Sport*. 21. S39. 10.1016/j.jsams.2018.09.090.
- Paul St Onge, Leonard A Temme, Aaron McAtee, Kevin J O'Brien, Brigid K Byrd, Evaluation of the Commercial, Off-the-Shelf (COTS) King-Devick Eye Tracking System, *Military Medicine*, Volume 184, Issue Supplement_1, March-April 2019, Pages 571–578, <https://doi.org/10.1093/milmed/usy380>
- Lin TP, Adler CH, Hentz JG, Balcer LJ, Galetta SL, Devick S. Slowing of number naming speed by King-Devick test in Parkinson's disease. *Parkinsonism Relat Disord*. 2014 Feb;20(2):226-9. doi: 10.1016/j.parkreldis.2013.10.009. Epub 2013 Oct 18. PMID: 24269283; PMCID: PMC3946616.
- Gold, Doria & Rizzo, John-Ross & Lee, Yuen Shan & Childs, Amanda & Hudson, Todd & Martone, John & Matsuzawa, Yuka & Fraser, Felicia & Ricker, Joseph & Dai, Weiwei & Selesnick, Ivan & Balcer, Laura & Galetta, Steven & Rucker, Janet. (2021). King-Devick Test Performance and Cognitive Dysfunction after Concussion: A Pilot Eye Movement Study. *Brain Sciences*. 11. 1571. 10.3390/brainsci11121571.
- Lawrence JB, Haider MN, Leddy JJ, Hinds A, Miecznikowski JC, Willer BS. The King-Devick test in an outpatient concussion clinic: Assessing the diagnostic and prognostic value of a vision test in conjunction with exercise testing among acutely concussed adolescents. *J Neurol Sci*. 2019 Mar 15;398:91-97. doi: 10.1016/j.jns.2018.12.020. Epub 2018 Dec 22. PMID: 30690413; PMCID: PMC7038786.
- Chiang CC, Starling AJ, Buras MR, Golafshar MA, VanderPluym JH. A pilot exploratory study comparing the King-Devick test (KDT) during and between migraine attacks. *Cephalalgia*. 2020 Mar;40(3):307-312. doi: 10.1177/0333102419885381. Epub 2019 Oct 29. PMID: 31660762.
- Bretzin, Abigail & Anderson, Morgan & Moran, Ryan & Covassin, Tracey. (2020). Long-term test-retest evaluation of the King-Devick test in youth soccer athletes. *Journal of the Neurological Sciences*. 416. 116951. 10.1016/j.jns.2020.116951.

T: 236-420-4448 F: 778-699-2067

E: reception@okanaganvisiontherapy.ca W: www.okanaganvisiontherapy.ca

Unit 201 - 880 Clement Ave, Kelowna BC, V1Y-0H8

Reliability Study of the Pierce and King-Devick Saccade Tests

MICHAEL K. H. ORIDE,* JOEL K. MARUTANI,* MICHAEL W. ROUSE†

Southern California College of Optometry, Fullerton, California

PAUL N. DeLAND‡

California State University, Fullerton, California

ABSTRACT

Two simple clinical tests which are currently being used to assess saccadic eye movements are the Pierce and King-Devick saccade tests. Although normative data have been established for both tests, little has been reported in regards to the tests' reliability. Thirty elementary schoolchildren (group I) were screened for visual abnormalities and then presented the two saccade tests in random order. These children were retested 2 weeks later by a different group of testers. A second group of 33 children (group II) who also passed the screening for visual abnormalities was administered the two saccade tests. This group was tested by the original testers who also tested the children in group I initially. Comparing initial and retest scores of group I showed poor reliability for both Pierce and King-Devick saccade tests. Thus there is some question as to the tests' clinical usefulness in both evaluating eye movements and monitoring progress of patients undergoing oculomotor therapy. Scores obtained from the initial testing of group I were compared with scores from group II. The same testers administered the tests for both groups. Results showed that tester behavior influenced the retest results of the Pierce saccade test, but results were inconclusive for the King-Devick saccade test.

Key Words: saccade testing, vision testing, pediatric vision care

Received August 19, 1985; revision received November 12, 1985.

* Optometrist, Member of Faculty.

† Optometrist, M.S.Ed., Member of Faculty, F.A.A.O.

‡ Ph.D., Member of Faculty.

The act of reading requires a person to move the eyes quickly in a series of saccades and fixations across the page of reading material. A number of studies have found that poor readers, as a group, demonstrate poorer eye movement skills, tending to make more regressions and fewer forward fixations,¹⁻⁴ even in situations where decoding and the need for comprehension have been eliminated.²⁻⁵

Based on these correlations between saccades and reading ability, it became desirable for clinicians to be able to assess eye movement skills in children experiencing reading difficulties. Because the fundamental eye movement that a reader makes is a saccade, the diagnostic task was to develop a good standardized test that would evaluate saccadic performance in a simulated reading environment. A number of clinical techniques have been developed, ranging from sophisticated infrared eye monitoring systems⁶ to simple subjective rating scales.⁷

Out of the need for a simple inexpensive quantitative method for evaluating oculomotor skill came the Pierce saccade test⁸ and later the King-Devick saccade test.⁹

The Pierce saccade test (Fig. 1) is composed of four test cards, with the first being a demonstration plate and the following three making up the actual test. Each card has 15 rows of 2 numbers. The lateral separation of the numbers on each card is 8 1/4 in. The vertical separation is 1/2 in on cards 1 and 2 and is 5/16 in on card 3. The demonstration card has arrows from left to right to simulate the type of eye movements the subject is to make. The first test card has lines between the numbers and the second and third cards have open space between them. The subject's task is to call off each of the numbers on the cards quickly and accurately. The time to complete each test and the number of errors and omissions are recorded. The subject's performance is then evaluated by comparing the total

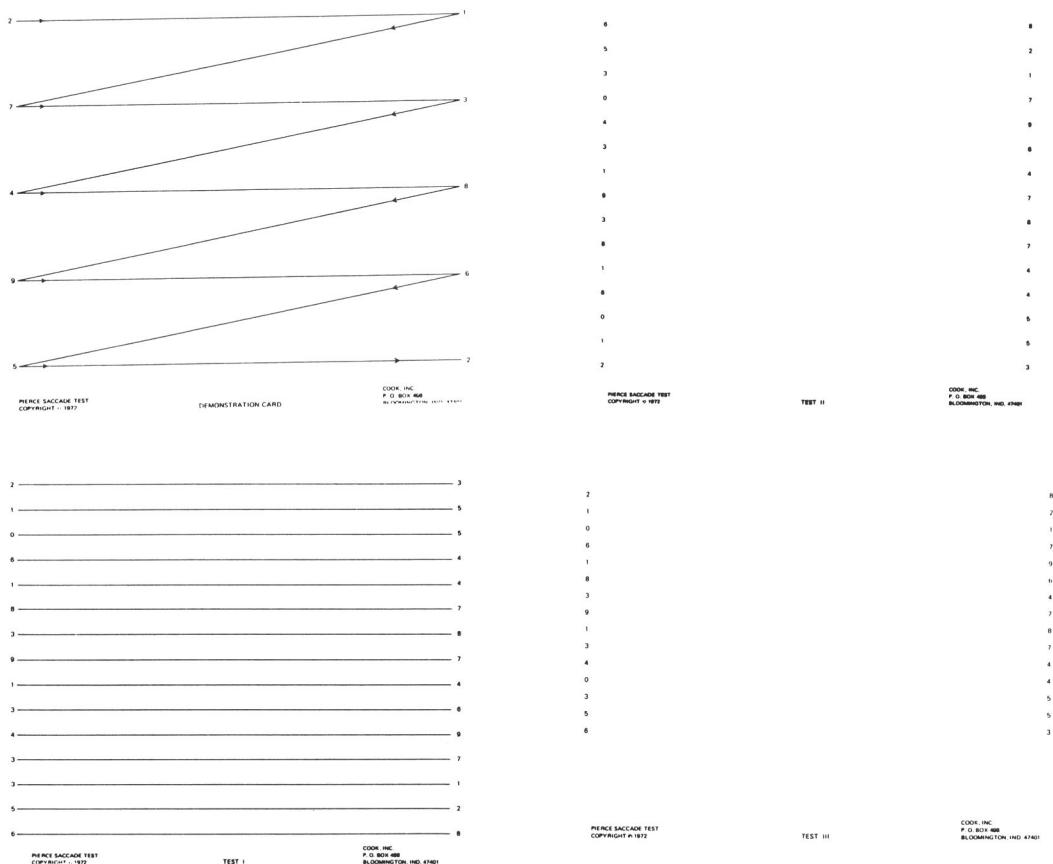


FIG. 1. Four card Pierce saccade test. First card is a demonstration card; the remaining cards make up the actual test.

corrected time (time corrected for errors) to age norms supplied with the test.

The King-Devick saccade test (Fig. 2) also has four cards, with the first being for demonstration. The three test cards have eight rows of five single digit numbers that are placed at random positions in each row. The vertical separation is $\frac{3}{4}$ in on test cards 1 and 2 and $\frac{1}{4}$ in on test card 3. The subject's task again is to call off the numbers quickly. Time and errors are recorded and the subject's performance is evaluated by separate comparisons to age-related norms of the total time for completing all tests and to the number of errors and omissions committed.

Both tests have normative data that can be used to judge whether a subject's performance is normal or abnormal. An abnormal performance is indicated by scores 1 SD below the mean for the child's age.

Most of the research to date has dealt with test design,¹⁰ developing adequate normative data,¹¹ and relating the performance on these two tests to reading ability.^{9,12} The purpose of

this study was to investigate the reliability of both the Pierce and King-Devick saccade tests.

METHODS

For this reliability study, 63 children between the ages of 7 and 12 years were tested. These subjects had all previously passed a MCT screening.¹³

Thirty children (group I) were tested initially with the Pierce and King-Devick saccade tests. The tests, in each case, were administered in random order. Two weeks later group I was retested. In addition, 33 new children (group II), same age range, were tested.

To avoid experimenter bias, the testers were trained 4th-year optometry students. Each tester read the same instructions that were printed on cards. The first test of group I was administered by Testers A. This same group of testers administered the test to group II 2 weeks later. A second group of Testers B administered the retest to group I (see Table 1 for summary).

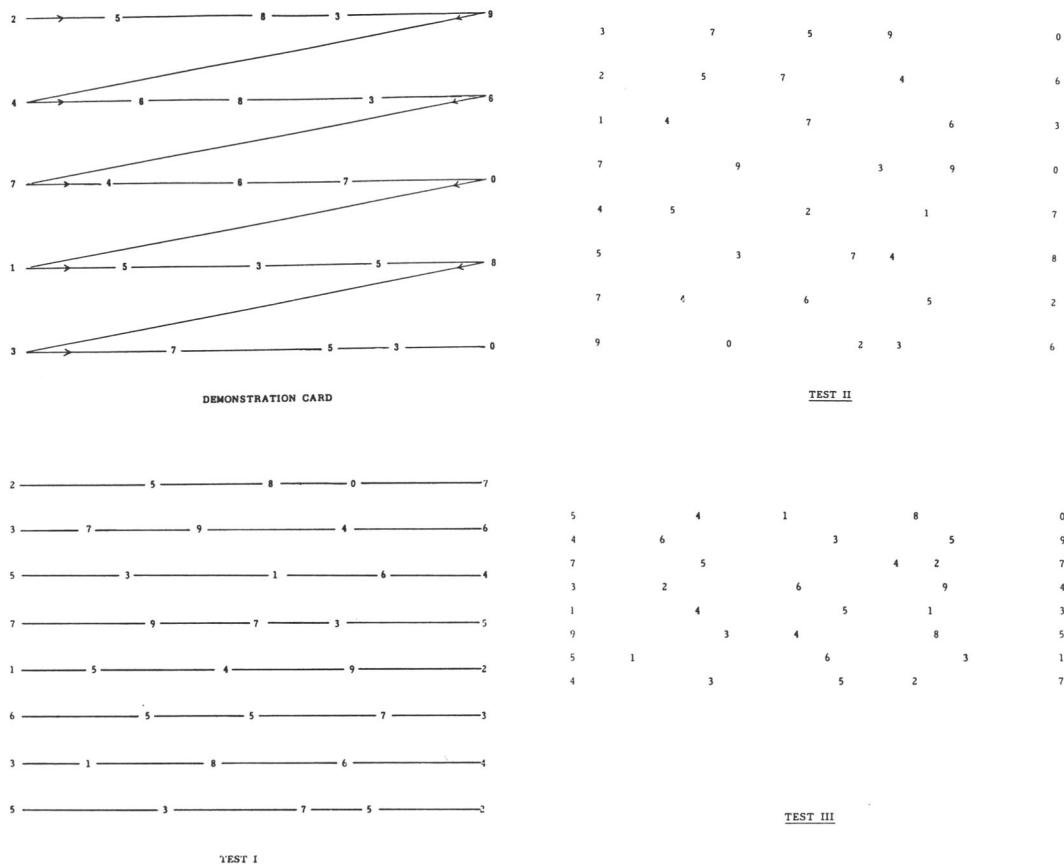


FIG. 2. Four card King-Devick saccade test. First card is a demonstration card; the remaining cards make up the actual test.

TABLE 1. Summary of testing schedule for experimental groups I and II.

	Test (wk 1)	Retest (wk 3)
Testers A	Group I	Group II
Testers B		Group I

In this way any bias generated by subject-tester familiarity in group I was eliminated. Any improvement in scores of group I would be due to a learning effect. A higher score by group II over the initial scores of group I would indicate that some aspect of the A group of testers' behavior influenced the test scores.

Instructions (Pierce saccade test) to the patient were to hold the demonstration card at 40 cm. Illumination was between 215 and 646 l/m² (20 and 60 ft. cd.) and there was no glare or shadows. The tester then instructed the subject as follows: "I want you to call out all of the numbers on this card as fast and as accurately

as possible in the following manner." The tester then pointed to the upper left-hand number, then the upper right-hand number, then the second left-hand number, etc. Instructions were repeated until the subject understood the task. The tester then proceeded with "ready, start," and began timing. The timer was stopped when all the numbers were called out. A record was made of any errors of omission or addition. The instructions were repeated and time and errors were recorded for test cards 2 and 3. The King-Devick saccade test was administered in a fashion similar to that of the Pierce saccade test.

Scoring for the Pierce saccade test was done by calculating a corrected time score, for each test card, according to the following formula:

$$\frac{30}{30 - \text{no. of errors}} \times \text{time} = \text{corrected score.}$$

The three corrected time scores were then added for a total corrected time score that was then compared against mean age norm scores pro-

vided with the test to arrive at an age equivalent score. The score on the King-Devick saccade test was evaluated slightly differently. The errors committed and the time for completing each of the three test cards was totaled. Then total time and total error scores were compared separately against mean age norm scores to arrive at independent age equivalent scores for time and errors. Although errors were recorded they were not used in our analysis of the King-Devick saccade test.

RESULTS

The individual subject test and retest scores on the Pierce and King-Devick saccade tests are displayed in Table 2. In addition, the age equivalent score was calculated for both the test and retest scores.

A comparison of the test retest scores shows a mean increase of 8.9 s (19.6 s SD) on the retest of the Pierce saccade test. A paired data one-sided t-test ($t = 2.47$, $p = 0.0098$) suggests there is a significant improvement in the retest scores compared to the initial test scores. Of the 30 subjects tested on the Pierce saccade test, 22 showed an improvement in age equivalent score. Of the 22 who improved, 11 subjects showed an increase of one year or more, and 3 an increase of more than 6 months but less than 1 year.

A comparison of the test retest scores for the King-Devick saccade test shows a mean increase of 9.7 s (11.9 s SD) on the retest. A paired data one-sided t-test ($t = 4.46$, $p = 0.0001$) suggests that there is also a significant improvement in the retest scores as compared to the initial scores. Of the 30 subjects, 23 showed an improvement in age equivalent scores. Of the 23

TABLE 2. Initial test and retest scores and age equivalents for group I subjects on the Pierce and King-Devick saccade tests.

Subject	Pierce Saccade				King-Devick Saccade			
	Test		Retest		Test		Retest	
	(Score)	Age (yr)	(Score)	Age (yr)	(Score)	Age (yr)	(Score)	Age (yr)
1	80	9.2	70	10.1	74	8.9	66	10.2
2	122	7.1	130	6.8	131	6.0	121	6.0
3	102	7.9	100	8.0	94	7.3	73	9.0
4	118	7.3	119	7.2	153	6.0	120	6.0
5	119	7.2	117	7.3	82	7.9	92	7.4
6	121	7.2	92	8.4	115	6.3	117	6.2
7	153	6.0	148	6.0	149	6.0	129	6.0
8	80	9.1	87	8.6	79	8.0	80	8.0
9	89	8.7	101	8.0	76	8.6	80	8.0
10	84	8.9	69	10.2	96	7.2	84	7.8
11	75	9.4	64	11.3	77	8.4	64	10.4
12	81	9.0	63	11.5	70	9.7	62	10.5
13	94	8.3	88	8.7	72	9.2	80	8.0
14	122	7.1	108	7.5	94	7.3	90	7.5
15	71	10.0	73	9.6	69	9.9	68	10.0
16	106	7.8	90	8.5	92	7.4	90	7.4
17	108	7.6	110	7.5	73	9.0	70	9.7
18	68	10.4	64	11.3	58	10.9	56	11.0
19	80	9.1	70	10.1	66	10.2	70	9.7
20	90	8.5	66	10.8	85	7.7	64	10.4
21	86	8.8	90	8.5	77	8.4	72	9.2
22	122	7.1	175	6.0	121	8.0	120	8.0
23	96	8.2	79	9.2	69	9.9	66	10.2
24	158	6.0	87	8.6	101	7.0	72	9.3
25	101	8.0	67	10.5	68	10.0	53	12.0
26	103	7.9	78	9.2	100	7.0	74	8.9
27	140	6.4	130	6.9	149	8.0	130	8.0
28	126	7.0	120	7.2	96	7.2	81	7.9
29	98	8.1	90	8.5	106	6.7	71	9.5
30	125	7.0	107	7.7	87	7.6	73	9.0

that improved, 9 showed an increase of one year or more, and 6 an increase of more than 6 months but less than 1 year.

At the time of retesting group I (by Testers B), 33 new elementary schoolchildren (group II) were tested (by Testers A) on both Pierce and King-Devick saccade tests. The Pierce and King-Devick saccade test scores of group II (Table 3) were compared to the initial scores obtained in group I. For the Pierce saccade test, group I had a mean time of 14.7 s (23.5 s SD) slower than group II. The pooled data one-sided t-test ($t = 2.62$, $p = 0.0056$) suggests that there is a significant improvement in the group II scores over group I. For the King-Devick saccade test, group II had a mean time of 12.2 s (37.5 s SD) slower than group I. A separate variance one-sided t-test ($t = 1.51$, NS) failed to show any statistically significant difference between the two groups.

DISCUSSION

This study can be divided into two parts. The first part was to determine the reliability of the Pierce and King-Devick saccade tests by test and retest (after a 2 week interval) of a group of 30 elementary schoolchildren (group I). A different group of testers were used for the first and second administration of the tests to prevent any biasing factors contributed by the tester such as the children becoming more comfortable or familiar with a certain tester.

The statistical analysis showed both saccadic tests to have poor repeatability with subjects showing a significant improvement in test scores on the retest. From the results, it appears that the subjects became more familiar or learned to take the test, thus improving their scores on retest. These results are in agreement with a similar study conducted by Fleming and Barney.¹⁴ Results from these two studies lead to serious questions in regards to the usefulness of either test as a tool for evaluating whether improvements, as measured by these tests, are due solely to oculomotor therapy.

The second part of the study looked at whether there was any learning or any effect contributed by the tester. Contributions might include changes in the manner the test is administered, such as showing more confidence, a change in tone, or clarity of the instructions given. In order to see if Testers A have such an influence, 33 subjects (group II) were administered the two saccadic tests on the same day group I was being retested. The same Testers A were used for testing group II as were used to initially test group I. Scores from group II were

TABLE 3. Scores and age equivalents of group II on Pierce and King-Devick saccade tests.

Subject	Pierce Saccade		King-Devick Saccade	
	Score (s)	Age (yr)	Score (s)	Age (yr)
1	60	12.0	62	10.5
2	120	7.2	130	6.0
3	91	8.5	200	6.0
4	98	8.1	200	6.0
5	75	9.4	88	7.6
6	89	8.7	73	9.0
7	73	9.6	80	8.0
8	107	7.7	109	6.5
9	56	12.8	60	10.7
10	60	12.0	64	10.5
11	60	12.0	68	10.0
12	61	11.8	71	9.4
13	66	10.8	72	9.2
14	103	7.9	107	6.6
15	81	9.0	66	10.2
16	79	9.2	111	6.4
17	74	9.5	100	7.0
18	89	8.7	123	6.0
19	108	7.6	73	9.0
20	104	7.8	113	6.3
21	96	8.2	113	6.3
22	90	8.5	105	6.7
23	88	8.7	110	6.5
24	61	11.8	86	7.7
25	80	9.1	77	8.3
26	113	7.4	170	6.0
27	95	8.3	125	6.0
28	107	7.7	147	6.0
29	127	6.9	165	6.0
30	144	6.0	164	6.8
31	98	8.1	96	7.1
32	98	8.1	108	6.6
33	92	8.4	98	7.1

compared to the initial scores of group I. A significant improvement in group II scores as compared to group I scores would indicate that some aspect of the behavior of the A group of testers, as suggested above, may have influenced the resultant scores.

For the Pierce saccade test, group II had a mean time that was 14.7 s faster than group I, which was found to be statistically significant. This would suggest that the A testers' behavior in some way contributed to improvements seen on a retest of the Pierce saccade test. For the King-Devick saccade test, group II had a mean time of 12.2 s slower than group I. It appears that in this case the testers' behavior did not contribute to any improvements on retest. How-

ever, there were four unusually large scores in group II that not only forced the group II mean to be higher, but also caused its SD to be large. This large variation in the data led to statistically insignificant results, and therefore no concrete conclusions can be drawn as to the testers' effect on the results of the King-Devick saccade test. A reevaluation of this question using the King-Devick saccade test is suggested.

SUMMARY

The results of this study confirm an earlier study reporting poor repeatability of the Pierce and King-Devick saccade tests. Subjects as a group showed significant improvements on a retest administration, showing that there is a learning effect present that influences retest results. For researchers using either test to evaluate the efficacy of vision therapy for improving oculomotor ability, a control group should be included to account for the learning effect demonstrated in this study. Clinicians should view retest results with caution, as improvements may be the result of a learning effect, as observed in this study.

In addition, at least for the Pierce saccade test, the behavior of the A group of testers from test to retest resulted in improved retest performance. This indicates the potential for a confounding effect of tester behavior in the clinical setting. In this study the experience level of the testers may be a possible factor. It is conceivable that any tester influence, as observed in this study, would diminish as a tester repeats administration of the test and brings a consistency to his/her testing technique. However, there would still remain the potential for a tester, who is also the clinician and would like to observe improved performance, to enhance a patient's retest performance. As a precaution against the possible influence of tester behavior the clinician may want to have a different tester administer the retest. An expanded study focusing on a sample of testers is suggested in order to provide illumination on this tester influence issue on the Pierce saccade test. Also, further investigation needs to be conducted before any concrete statements on the King-Devick saccade test can be made.

In conclusion, we feel these two saccade tests may be useful as a screener of oculomotor ability but we question their usefulness in the diagnos-

tic evaluation and in monitoring the progress of patients undergoing oculomotor therapy.

REFERENCES

1. Lefton LA, Lahey BB, Stagg DL. Eye movements in reading disabled and normal children: a study of systems and strategies. *J Learn Disabil* 1978;11:549-58.
2. Pavlidis G. The dyslexics erratic eye movements: case studies. *Dyslex Rev* 1978;1:22-8.
3. Ciuffreda KJ, Bahill AT, Kenyon RV, Stark L. Eye movements during reading: case reports. *Am J Optom Physiol Opt* 1976;53:389-95.
4. Elterman RD, Abel LA, Daroff RB, Dell'Osso LF, Bornstein JL. Eye movement patterns in dyslexic children. *J Learn Disabil* 1980;13:16-21.
5. Griffin DC, Walton HN, Ives V. Saccades as related to reading disorders. *J Learn Disabil* 1974;7:310-6.
6. Winter J. Clinical oculography. *J Am Optom Assoc* 1974;45:1308-13.
7. Griffin JR. Saccadic eye movements-recommended testing and training procedures. *Optom Mon* 1981;72:27-8.
8. Pierce J. *Pierce Saccade Test*. Bloomington, IN: Cook Inc, 1972.
9. King AT, Devick S. The proposed King-Devick test and its relation to the Pierce saccade test and reading levels. Senior Research Project 1976. Available from the Carl F. Shepard Memorial Library, Illinois College of Optometry, Chicago, IL.
10. Richman JE, Walker AJ, Garzia RP. The impact of automatic digit naming ability on a clinical test of eye movement functioning. *J Am Optom Assoc* 1983;54:617-22.
11. Lieberman S, Cohen AH, Rubin J. *NYSOA K-D test*. *J Am Optom Assoc* 1983;54:631-7.
12. Bond R, Painter B. A comparison of the reading performance predictabilities of the King-Devick saccade test and the Pierce saccade test. Senior Research Project 1981. Available from the MB Ketchum Library, Southern California College of Optometry, Fullerton, CA.
13. Blum HL, Peters HB, Bettman JW. *Vision Screening for Elementary Schools: The Orinda Study*. Berkeley: University of California Press, 1959:105.
14. Fleming JC, Barney P. A reliability study of tests measuring saccadic ability. Senior Research Project 1984. Available from the MB Ketchum Library, Southern California College of Optometry, Fullerton, CA.

AUTHOR'S ADDRESS:

Michael W. Rouse

Southern California College of Optometry

Optometric Center of Fullerton

2575 Yorba Linda Blvd.

Fullerton, California 92631-1699

Performance Times for the King-Devick Test in Children and Adolescents

Paul J. Gubanich, MD, MPH,*† Resmi Gupta, MS, MA,‡ Eric Slattery, MS,§ and Kelsey Logan, MD, MPH¶

Abstract

Objective: To establish the King-Devick test (KD) performance values for children and adolescents. **Design:** Prospective cohort. **Setting:** Pediatric sports medicine clinics. **Participants:** Five hundred seven athletes presenting to a pediatric sports medicine clinic for non-concussion-related evaluations. **Independent Variables:** Age, sex, and risk factors for abnormal concussion recovery. **Main Outcome Measures:** The King-Devick test time. **Results:** Four hundred eighty-three participants were included in the final analysis, which included 60.5% girls ($n = 292$) and 39.5% boys ($n = 191$). The KD test performance varied by age with a steady decrease in test time yearly from 8 years through 14 years of age, where some plateauing occurred. Baseline scores changed approximately 31 seconds over ages 8 to 18. Analysis of variance results revealed a strong effect of age on overall mean KD time ($P < 0.001$). The KD test performance was not associated with sex or other studied concussion risk factors or comorbid conditions. **Conclusions:** The KD test performance was reported in a cohort of youth aged 8 to 18 years, allowing for comparison of performance in individuals who may not have a baseline assessment. If baseline testing is desired, at least yearly intervals seems to be appropriate through childhood and early adolescence.

Key Words: concussion, pediatric, adolescent, King-Devick

(*Clin J Sport Med* 2019;29:374–378)

INTRODUCTION

Concussion awareness in the medical field and in the lay public has increased, with concussion diagnoses responsible for a higher percentage of sports injuries than seen in the past.¹ Coaches, officials, parents, and athletes have better concussion knowledge, and the recognition of the injury has improved.² All 50 states in the United States have currently implemented youth sports concussion legislation, most of which educate coaches and officials about the signs and symptoms of concussion and requires that athletes exhibiting these be removed from play and complete a medical evaluation before sport reentry.³ Not surprisingly, concussion-related visits to emergency departments (EDs) and outpatient facilities have dramatically grown.^{4,5}

Even with the large increase in ED visits, the majority of youth within the 5 to 17 age range, the ages in which sports are typically played, present to their primary care provider for their initial concussion evaluation and management.⁶ This may be partially explained by the observation that although visits to the ED have risen, severity of injury has not.^{4,5} Concussion management has thus become an important part

of many general pediatric and pediatric sports medicine clinics.

Concussion symptom reporting is considered essential to the evaluation of concussion, with the most recent consensus statement on concussion in sport encouraging use of the Concussion Recognition Tool version 5, the Sports Concussion Assessment Tool version 5 (SCAT5), and the Child SCAT5.⁷ All these tools incorporate symptom reporting. Presence of concussion symptoms and their severity remains a large part of concussion management.⁸ This is especially true in the young athlete, where objective assessments have not been studied as intensely as they have been in high school and older athletes. However, because of the subjectivity of symptom reporting in concussion, as well as the large variety of symptoms that concussed patients report, objective assessment measures are needed to help providers understand the degree of impairment beyond that of reported symptoms.

Oculomotor abnormalities have been increasingly recognized as an important part of concussion assessment and management. One study of acutely concussed adolescent athletes showed that over 75% of participants exhibited vestibulo-ocular dysfunction, and that those patients had significantly higher odds of developing postconcussion syndrome.⁹ The King-Devick test (KD), originally designed and used as a reading evaluation tool, has been suggested for use in screening for oculomotor dysfunction as part of concussion assessment and diagnosis. The KD measures rapid number naming speed, assesses function of the brain in multiple areas (language, attention, and eye movement), and requires little time to administer (<1-2 minutes).¹⁰ Several studies have reported its use as a sideline concussion diagnostic tool, with mixed results on its ability to diagnose concussion, based on an athlete's own baseline.¹¹⁻¹³ Current protocol for KD concussion evaluation requires a baseline

Submitted for publication February 9, 2018; accepted August 8, 2018.

From the *Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, Ohio; †Division of Sports Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio; ‡Division of Biostatistics and Epidemiology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio; and §Department of Kinesiology and Health, Miami University, Oxford, Ohio. ¶Division of Sports Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio.

The authors report no conflicts of interest.

Corresponding Author: Kelsey Logan, MD, MPH, Division of Sports Medicine, Cincinnati Children's Hospital Medical Center, 3333 Burnet Ave, MLC 10001, Cincinnati, OH 45229 (Kelsey.logan@cchmc.org).

Copyright © 2019 Wolters Kluwer Health, Inc. All rights reserved.

<http://dx.doi.org/10.1097/JSM.0000000000000670>

Clinical Relevance

This study provides reference times for children and adolescents completing the King-Devick test in the context of concussion management. It will allow providers to compare their patients' times against a large cohort of nonconcussed children and adolescents.

KD, to compare a possibly concussed athlete against his previously established baseline. In addition, a KD time longer (even tenths of a second) than the previously established baseline is currently considered abnormal.

Despite the KD apparently becoming more widely used in concussion evaluation, the lack of previously established baseline performance limits interpretability of the test. Baseline testing with the KD has only been performed in a limited manner and may not be practical for most youth sports programs. Furthermore, most pediatric practices that treat these patients would not have easy access to those baseline values. Therefore, the primary goal of this study was to establish performance data for the KD test for children and adolescents. The availability of these data would enable clinicians to more easily interpret KD performance in instances where baseline testing was not performed or available. Additional aims of this study were to examine the effect of age and sex on KD performance. Finally, factors that have been associated with concussive injury and/or slow recovery were also examined for their effect on KD performance at baseline.

METHODS

Subjects

Subjects who presented to 1 of 2 sports medicine clinic locations of a large Midwestern children's hospital system for evaluation and treatment from January 2015 through February 2016 were potentially eligible for the study. Participants were required to be English-speaking, between the ages of 6 to 18 years at the time of presentation, have a parent or guardian present, and have a chief complaint that did not preclude them from unrestricted completion of the testing protocol. Subjects with resolving musculoskeletal or medical conditions, or their healthy accompanying relative or sibling, were also eligible for enrollment in this baseline study, as long as they met the study criteria. Additional exclusion criteria included individuals with a history of concussion within the past 6 months or subjects currently receiving ongoing concussive care or remained symptomatic, or comorbid condition(s) that could compromise their ability to complete the baseline assessment as determined by the investigators, such as eye, facial, neck, or upper-extremity symptoms or complaints. Informed consent and assent were obtained in accordance with the subject's age and institutional review board (IRB) protocol. Cincinnati Children's Hospital Medical Center IRB approved this study.

Instrument

Originally developed in 1976, the KD test is a timed evaluation wherein participants are asked to read aloud a series of single digit numbers positioned in various

configurations on printed test cards. The 3 test cards become progressively more challenging as the spacing between numbers becomes more variable.

Testing

Subjects underwent a baseline assessment, which included performance of the KD and clinical reaction time test. The clinical reaction time testing procedures and results will be reported in a separate article. For the KD, subjects were seated, instructed on test procedure, and reviewed a demonstration card, after which they completed the 3-card test. The subjects were instructed to complete testing as quickly and accurately as possible. The total time for completion, total number of errors, and single card completion time and errors were recorded. Two trials were conducted, with the faster error-free (or best) attempt counted as the baseline score.

For the subset of participants who had clinical visits on the day of their baseline assessment, an investigation as to factors that may alter baseline performance test was undertaken. Demographic factors of age and sex were recorded. Pain level at the time of their visit was recorded as part of the Numerical Rating Scale (NRS-11).¹⁴ Finally, concussion-specific risk factors such as a diagnosis of attention deficit disorder and attention deficit hyperactivity disorder (ADD/ADHD), the use of ADHD medications, previous history of concussion, previous or current neurologic condition, and history of headache were identified. Subjects were defined as having ADD/ADHD if it was recorded on their intake medical questionnaire on the day of their visit, previously documented in their electronic medical record and reviewed on the day of their assessment, or if the subject reported taking medications to treat ADD/ADHD on their intake questionnaire or on review of their electronic medical record. All other factors were either measured on the day of their visit, obtained from their medical intake questionnaire, or from review of their electronic medical record on the day of their visit.

Statistical Analysis

Descriptive statistics were computed to describe the participant characteristics. Mean (SD) was reported for continuous measures. Categorical variables were calculated as frequencies and percentages. A quantile-quantile plot was used to determine whether continuous response variable (KD score) was normally distributed. A regression-based analysis of variance was used to assess the association between age and KD scores. Sex was included as a covariate in the initial model, but dropped later due to nonstatistical significance. The Tukey HSD adjustment for multiple testing was used for all post hoc comparisons. A paired *t* test by each age category was conducted to assess the mean difference between the 2 trials to explore the effect of repeat test administration. Finally, we performed the same analysis for mean difference without the 29 subjects who only completed one KD trial. Statistical significance level was set at $\alpha = 0.05$. All primary analyses were conducted with SAS statistical software version 9.4 (SAS Institute, Inc, Cary, NC).

Subgroup analysis was conducted on the subset of patients with clinical visits (94% of participants). Linear regression was conducted with all variables, with age and sex serving as covariates. In addition, subjects were also age- and sex-matched 1:1 to a random sample of controls for each

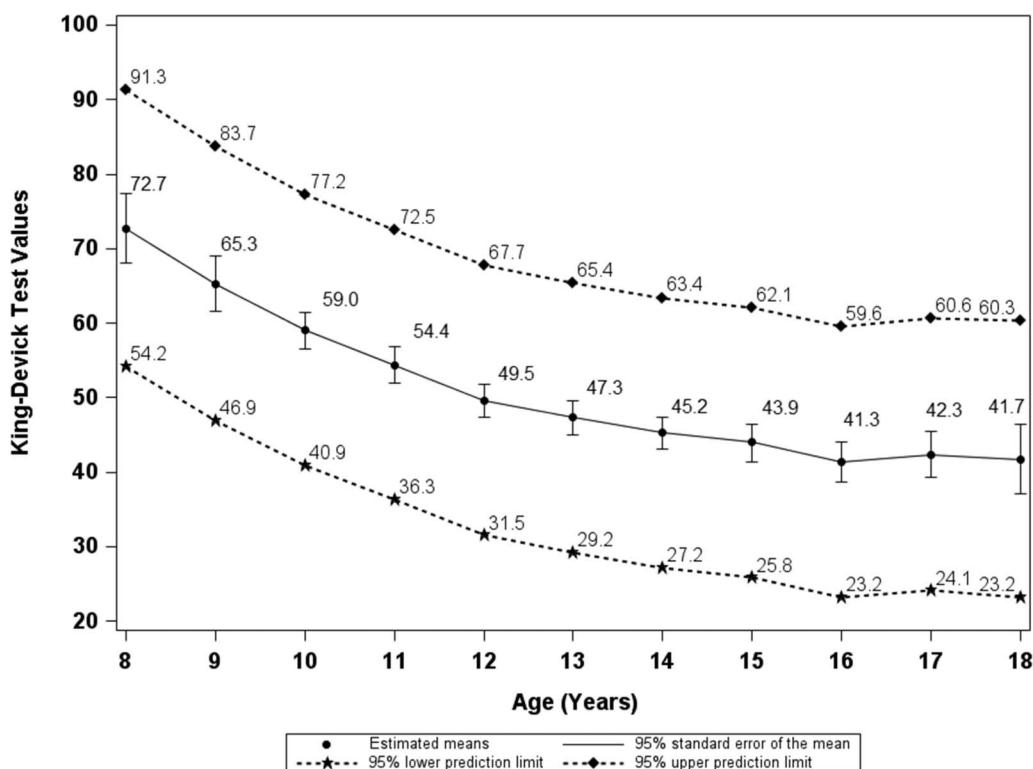


Figure. Mean KD test time (with 95% prediction interval) versus age.

categorical risk factor. T-tests were used to examine differences in the mean KD score between groups at a significance level of $\alpha = 0.05$. Subgroup analyses were conducted with MATLAB (The MathWorks, Inc, Natick, MA).

RESULTS

Five hundred and seven participants aged 6 to 18 years completed the study. The number of participants in the study was based on convenience, by presentation over the 13 months included in the study. Participants in the 6- and 7-year-old age groups were excluded due to low recruitment ($n = 7$ and $n = 9$, respectively) and difficulty in completing the baseline assessment as defined in the study protocol. More than 80% of the 6- and 7-year-old subjects had difficulty completing all 3 cards as defined in the protocol. Eight others were excluded due to concerns of valid performance, with multiple errors or line omissions on the test, despite multiple attempts (1.6%). Ultimately, 483 participants were included in the final analysis, which included 60.5% girls ($n = 292$) and 39.5% boys ($n = 191$). Subjects were then categorized by age. Analysis of variance results revealed a strong effect of age on overall mean KD time ($P < 0.001$) (Figure). The KD performance was slowest in the 8-year-old age group {mean KD time 72.7 [95% confidence interval (CI), 44.2-91.3] seconds} and improved with age group until plateauing around 14 years of age (Table 1). The mean KD time for 18 year olds was 41.7 (95% CI, 23.2-60.3) seconds, a difference of almost 31 seconds (75%) compared with the mean of 8 year olds. The prediction interval (mean + 95% CI) was found to be 37 seconds for each age group.

Pairwise comparisons showed significance across many age comparisons. The mean KD test was significantly different in 8 year olds from all age groups (10-18) except for 9 year olds. The mean KD score was significantly different in 9 year olds than in 11 to 18 years olds, but not in 8 and 10 year olds. Similarly, the mean 10-year-old KD time was significantly different from 8 and 12 to 18 year olds, but not from 9 and 11 year olds. Eleven year olds were different from 8, 9, and 13 to 18 year olds, but not from 10 and 12 year olds. After age 13, the number of significant comparisons diminished in that 13

TABLE 1. Estimated Mean Values With Corresponding 95% Prediction Intervals

Age (n = 483), yr	Estimated Mean of KD Test (s)	95% Prediction Interval, s
8 (n = 15)	72.7	54.2-91.3
9 (n = 23)	65.3	46.9-83.7
10 (n = 55)	59.0	40.9-77.2
11 (n = 53)	54.4	36.3-72.5
12 (n = 65)	49.5	31.5-67.7
13 (n = 61)	47.3	29.2-65.4
14 (n = 68)	45.2	27.2-63.4
15 (n = 49)	43.9	25.8-62.1
16 (n = 45)	41.3	23.2-59.6
17 (n = 34)	42.3	24.1-60.6
18 (n = 15)	41.7	23.2-60.3

TABLE 2. Mean Difference in KD Test Scores Between Trials 1 and 2, by Age

Age, yr	Mean Difference (95% CI), s	P
8	2.82 (-2 to 3.64)	0.22
9	0.18 (-2.7 to 3.13)	0.90
10	1.40 (1 to 1.90)	0.06
11	1.31 (-0.27 to 2.90)	0.10
12	0.39 (-0.43 to 1.22)	0.34
13	2.17 (1.11 to 3.22)	<0.01*
14	1.51 (0.58 to 2.44)	<0.01*
15	2.13 (0.84 to 3.41)	<0.01*
16	1.71 (0.72 to 2.70)	<0.01*
17	2.57 (1.33 to 3.81)	<0.01*
18	0.72 (-0.77 to 2.21)	0.31

*Denotes significant difference in average KD test scores between trials 1 and 2.

year olds showed no difference in mean KD time when compared with 12 to 15, and 17 to 18 year olds, but was significant in comparisons between 8 to 11 and 16 year olds. Similar patterns of significant comparisons were observed in 14 to 18 year old age group.

Paired evaluation of subject performance between trial 1 and trial 2 was also examined (Table 2). A small but significant learning effect was observed in the 13 to 18 age group with a range of 1.5 to 2.57 (SE 0.41-0.64) seconds. We performed the same analysis for mean difference without the 29 subjects who only completed one KD trial and found no significant differences in parameter estimates in comparison with the full cohort.

Four hundred and thirty-nine participants were included in the risk factor subgroup analysis for subjects who had completed a clinical visit (94% of sample). None of the examined factors (pain score, history of ADD/ADHD, history of concussion, history of neurologic condition, history of headache, or history of ADD/ADHD medication use) were significant in the linear regression model, when controlled for age and sex ($P > 0.10$). The univariate analysis of individuals' categorical risk factors versus age- and sex-matched controls was also unremarkable. No statistical association was observed between mean KD time and history of ADD/ADHD, history of concussion, history of neurologic condition, history of headache, or history of ADD/ADHD medication use (Table 3).

DISCUSSION

The results of this study revealed that KD times vary with age with a steady decrease in test time yearly from 8 years through 14 years of age, where some plateauing occurs. The improvement in KD test performance was much smaller in the 14 to 18 age interval in comparison to the 9 to 14 age range. This is the first study to report KD performance for children and adolescents in the large cohort of children and adolescents to date. The magnitude of this effect is rather large, changing approximately 31 seconds over this age range at baseline. These data would also suggest that if baseline tests are performed, yearly intervals may be appropriate to account for developmental-related changes in performance. Future studies are needed to examine the effect of KD performance as it related to acute injury and recovery.

TABLE 3. Comparison of Age- and Sex-Matched Risk Factors and Controls

Risk Factor (n)	Mean KD (SE), s	P
ADD/ADHD (34)	47.58 (2.02)	0.44
Control	45.34 (1.93)	
AD meds (17)	51.78 (2.87)	0.97
Control	51.92 (3.24)	
Concussion history (10)	49.17 (3.91)	0.70
Control	51.39 (4.10)	
Neurologic history (43)	48.58 (1.80)	0.52
Control	46.95 (1.76)	
Headache history (20)	44.64 (1.72)	0.54
Control	46.33 (2.11)	

Our results in part support what Alsalaheen et al found when attempting to establish normative values for high school football players. That study consisted of 157 subjects who were categorized into 13 to 15 and 16 to 18 age groups. The KD times were significantly faster for the older age group.¹⁵ The median times for those groups were slightly faster than what we found for similar ages. Similarly, the 16 and 17 year olds tested in this study showed slightly slower times, when compared with collegiate athletes¹⁶ and professional ice hockey players¹⁷ in other studies.

At ages 13 to 17, subjects were significantly faster on their second KD test, confirming that at least 2 tests are needed to establish a valid baseline. This likely represents a small learning effect and is in agreement with previous studies.^{12,16} We recommend 2 trials for baseline testing in clinical practice. For the other ages, no significant differences between trials were seen. Studies with larger cohorts of younger subjects are needed to confirm that no learning effect is seen on this test because sample sizes were smaller in the youngest age groups. It remains our practice to perform 2 tests for baselines at all ages.

Below age 8, the KD test was difficult for subjects to complete. Most children were able to complete 2 cards; however, a modified version of testing, if developed, could be beneficial. Normative values on a large group of children would need to be studied; in the younger ages of our study cohort (10 and below), the SD for testing time was above 10 seconds. We would expect an even larger SD at younger ages.

The KD provides an objective oculomotor assessment that may be used along with other objective tools during a physical examination to help providers more fully understand how the patient's brain is functioning, independent of symptom reporting. It is believed that this comparative data can help in the interpretation of performance on the KD in children and adolescence. We also examined several risk factors such as sex, history of depression, and history of neurologic diagnosis that did not seem to affect KD performance. Future studies may need to focus more on additional risk factors for vestibulo-ocular dysfunction such as dizziness, blurred vision, or difficulty focusing at the time of injury.

Several potential confounding factors such as sex, pain at the time of presentation, history of ADD/ADHD, history of concussion, history of neurologic condition, history of headache, or history of ADD/ADHD medication use were also examined, and did not seem to alter KD performance. The lack of test time

effect seen with these variables potentially makes this a good test to use with those patients presenting with comorbidities or injuries. Other objective concussion assessments are often affected by injury or neuropsychological issues; balance testing is difficult to do in the presence of arm or leg injuries, and computerized neurocognitive testing is known to be affected by the presence of ADHD,^{18,19} learning disability,¹⁹ and history of concussion.²⁰ In our study, these did not affect KD time.

There are limitations to this study. The subjects in this study presented to a large pediatric sports medicine program and did not receive this testing after injury, or in either a primary care or ED environment, nor as part of a sideline evaluation, making generalizability to any of those areas difficult. In addition, subjects from the 2 clinics used for recruitment for this project are likely more homogeneous in socioeconomic, racial, and ethnic status than would be needed for population generalization. Further testing with a large number of diverse athletes would be helpful to confirm our findings. Most of the testing was done by one provider; however, small variation is possible in how the KD was presented and timed by various providers.

Because the study was based on a convenience sample during 13 months, we did not define low recruitment. Twenty-three of the 483 subjects completed only one trial instead of the recommended 2; however, we have established that no significant differences in parameter scores were present, when those 23 subjects were removed. Although the authors attempted to examine many comorbid conditions, the strong association of age and limited frequency of several risk factors in this population precluded our ability to examine each risk factor independently. The frequency of these factors in this population was low, and the study was not powered to fully evaluate their effect. For example, only 7.7% of the population were documented to have ADD/ADHD, 2.2% reported a previous history of concussion, 9.8% reported a neurologic condition, and 4.6% reported a history of headaches. Studies with larger numbers of patients with these and other risk factors would be helpful in future investigations.

CONCLUSIONS

As a result of this study, KD values in children and adolescents were examined. Test performance was shown to vary significantly with age but not by presence of comorbid conditions relevant to concussion assessment. Because baseline KD testing is not routinely completed in youth sports, and because treating providers would not have routine access to baselines done, these data can help with KD interpretation for concussion management in youth aged 8 to 18 years. If baseline testing is desired, at least yearly intervals in childhood and adolescence would seem appropriate. Finally, a 2-trial KD test to establish a baseline time is recommended.

ACKNOWLEDGMENTS

The authors thank the certified athletic trainers and nursing staff at Cincinnati Children's Hospital Medical Center for

their assistance on this project. The authors also thank Blake Simpson and Christy Reed for help with data collection and entry. Finally, the authors thank Chris DiCesare and Scott Bonnette for contributing to data analysis and interpretation.

References

- Currie DW, Kraeutler MJ, Schrock JB, et al. Time trends in concussion symptom presentation and assessment methods in high school athletes. *Am J Sports Med.* 2017;45:3368–3373.
- Chrisman SP, Schiff MA, Chung SK, et al. Implementation of concussion legislation and extent of concussion education for athletes, parents, and coaches in Washington state. *Am J Sports Med.* 2014;42:1190–1196.
- Simon LM, Mitchell CN. Youth concussion laws across the nation: implications for the traveling team physician. *Curr Sports Med Rep.* 2016; 15:161–167.
- Hanson HR, Pomerantz WJ, Gittelman M. ED utilization trends in sports-related traumatic brain injury. *Pediatrics.* 2013;132:e859–e864.
- Haring RS, Canner JK, Asemota AO, et al. Trends in incidence and severity of sports-related traumatic brain injury (TBI) in the emergency department 2006–2011. *Brain Inj.* 2015;29:989–992.
- Arbogast KB, Curry AE, Pfeiffer MR, et al. Point of health care entry for youth with concussion within a large pediatric care network. *JAMA Pediatr.* 2016;170:e160294.
- McCrory P, Meeuwisse W, Dvorak J, et al. Consensus statement on concussion in sport—the 5th international conference on concussion in sport held in Berlin, October 2016. *Br J Sports Med.* 2017;51:838–847.
- Halstead ME, Walter KD; Council on Sports Medicine and Fitness, American Academy of Pediatrics. Clinical report—sport-related concussion in children and adolescents. *Pediatrics.* 2010;126:597–615.
- Ellis MJ, Cordingley D, Vis S, et al. Vestibulo-ocular dysfunction in pediatric sports-related concussion. *J Neurosurg Pediatr.* 2015;16: 248–255.
- Galetta KM, Barrett J, Allen M, et al. The King-Devick test as a determinant of head trauma and concussion in boxers and MMA fighters. *Neurology.* 2011;76:1456–1462.
- King D, Hume P, Gissane C, et al. Use of the King-Devick test for sideline concussion screening in junior rugby league. *J Neurol Sci.* 2015;357: 75–79.
- Leong DF, Balcer LJ, Galetta SL, et al. The King-Devick test for sideline concussion screening in collegiate football. *J Optom.* 2015;8:131–139.
- Molloy JH, Murphy I, Gissane C. The King-Devick (K-D) test and concussion diagnosis in semi-professional rugby union players. *J Sci Med Sport.* 2017;20:708–711.
- Castarlenas E, Jensen MP, von Baeyer CL, et al. Psychometric properties of the numerical rating scale to assess self-reported pain intensity in children and adolescents: a systematic review. *Clin J Pain.* 2017;33: 376–383.
- Alsalaheen B, Haines J, Yorke A, et al. King-Devick Test reference values and associations with balance measures in high school American football players. *Scand J Med Sci Sports.* 2016;26:235–239.
- Galetta KM, Brandes LE, Maki K, et al. The King-Devick test and sports-related concussion: study of a rapid visual screening tool in a collegiate cohort. *J Neurol Sci.* 2011;309:34–39.
- Vartiainen MV, Holm A, Peltonen K, et al. King-Devick test normative reference values for professional male ice hockey players. *Scand J Med Sci Sports.* 2015;25:e327–e330.
- Elbin RJ, Kontos AP, Kegel N, et al. Individual and combined effects of LD and ADHD on computerized neurocognitive concussion test performance: evidence for separate norms. *Arch Clin Neuropsychol.* 2013;28:476–484.
- Zuckerman SL, Lee YM, Odom MJ, et al. Baseline neurocognitive scores in athletes with attention deficit-spectrum disorders and/or learning disability. *J Neurosurg Pediatr.* 2013;12:103–109.
- Collins MW, Grindel SH, Lovell MR, et al. Relationship between concussion and neuropsychological performance in college football players. *JAMA.* 1999;282:964–970.

Test-Retest Reliability and the Effects of Exercise on the King-Devick Test

Ryan Eddy, MEd, ATC,* John Goetschius, PhD, ATC,† Jay Hertel, PhD, ATC,* and Jacob Resch, PhD, ATC*

Abstract

Objective: To determine the test-retest reliability and the influence of exercise on King-Devick (K-D) test performance. **Design:** Crossover study design. **Setting:** Controlled laboratory. **Participants:** Participants consisted of 63 (39 women and 24 men) healthy, recreationally active college students who were 21.0 ± 1.5 years of age. **Independent Variables:** Participants completed the K-D test using a 2-week, test-retest interval. The K-D test was administered before and after a counterbalanced exercise or rest intervention. Reliability was assessed using testing visits (visit 1 and visit 2) as the independent variables. Exercise or rest and time (baseline, postintervention) were used as independent variables to examine the influence of exercise. **Main Outcome Measures:** Intraclass correlation (ICC) coefficients with 95% confidence intervals were calculated between visits to assess reliability of K-D test completion time. A repeated-measure 2×2 analysis of variance (intervention \times time) with post hoc paired *t* tests was used to assess the influence of exercise on K-D test performance. **Results:** The K-D test was observed to have strong test-retest reliability [$\text{ICC}_{2,1} = 0.90$ (0.71, 0.96)] over time. No significant intervention-by-time interaction ($P = 0.55$) or intervention main effects ($P = 0.68$) on K-D time were observed. Mean differences of -1.5 and -1.7 seconds ($P < 0.001$) were observed between baseline and rest and exercise interventions for K-D test performance, respectively. Up to 32% (20/63) of participants were observed to have a false-positive K-D test performance before and after each intervention. **Conclusions:** Although strong test-retest reliability coefficients were observed using clinically relevant time points, a high false-positive rate warrants caution when interpreting the K-D test.

Key Words: concussion, saccades, sideline measure, exercise

(*Clin J Sport Med* 2018;0:1-6)

INTRODUCTION

Early recognition and management of a concussed athlete is critical to achieve optimal outcomes after injury.¹ The clinical examination is an essential aspect of concussion management and is often performed on the sidelines of athletic competitions and practices.² Failure to properly recognize and manage a concussed student athlete exposes the patient to risk for poor outcomes,³ as well as risk for rare, but catastrophic injuries (eg, diffuse cerebral swelling).⁴ Clinical measures of sport concussion (SC) have been developed for sideline administration to supplement clinicians' clinical examination with objective information to aid in injury recognition and guide clinical care.⁵

Most sideline measures of SC have been targeted toward identifying posttraumatic cognitive and balance deficiencies; however, recent efforts have focused on identifying SC-related impairments in visual and oculomotor function. Abnormal eye movement and saccades, a rapid movement of the eye as it changes between fixation points,⁶ are purported to be an indication of suboptimal subcortical brain function.⁷ Approximately 23% to 32% of high school and collegiate student athletes self-report visual symptomatology,

respectively.⁸ In addition, visual tracking disturbances, quantified by slower saccadic reaction time and prolonged antisaccade latencies, have been observed in patients with a history of concussion.⁹ Clinical measures that can identify abnormal saccadic eye activity may assist clinicians in identifying objective changes in brain function after SC.

The King-Devick (K-D) test is a rapid number-naming task that has recently been repurposed as a clinical measure of visual tracking and saccadic eye movements in athletes before and after a suspected SC.¹⁰ The K-D test was originally developed in the 1970s to test attention, processing, and speech¹¹ and to measure reading comprehension and learning disabilities in elementary school children.¹² Administration of the K-D test involves the patient visually following and reading aloud sequences of numbers from 3 test cards that increase in difficulty, whereas the sum time to complete all 3 cards is recorded.¹³ Postinjury increases in time to complete the K-D test, usually 5 or more seconds relative to a patient's baseline time, have been described as a sign of SC.¹¹ The K-D was first described as a measure of SC in a sample of mixed martial art fighters and boxers who had suffered head trauma. The authors of this study reported that athletes who sustained head trauma compared to those participants that did not, took an average of 18 seconds longer to complete the K-D test. The authors also reported high test-retest reliability values using a 15-minute test-retest interval.¹¹ A more recent study conducted by Molloy et al¹⁴ examined the sensitivity and specificity of the K-D test in semiprofessional rugby athletes who were 18 to 35 years

Submitted for publication September 6, 2017; accepted January 31, 2018.

From the *Department of Kinesiology, University of Virginia, Charlottesville, Virginia; and †Department of Exercise Science and Athletic Training, Adrian College, Adrian, Michigan.

The authors report no conflicts of interest.

Corresponding Author: Jacob Resch, PhD, ATC, 210 Emmet St. South, Charlottesville, VA 22904 (jer6x@virginia.edu).

Copyright © 2018 Wolters Kluwer Health, Inc. All rights reserved.

<http://dx.doi.org/10.1097/JSM.0000000000000586>

of age. The authors administered the K-D test to athletes suspected of having an SC acutely after injury and reported the K-D test correctly identified 53% and 69% of concussed and nonconcussed athletes, respectively, when compared with preinjury (baseline) values.¹⁴

Effective clinical measures of SC need to have clinically acceptable test-retest reliability because of requisite repeated administrations at baseline, postinjury, and when determining the resolution of a concussive injury. The test-retest reliability of the K-D test has been reported within specific sport populations, including mixed martial art fighters,¹¹ amateur boxers,¹⁵ junior rugby,¹⁶ collegiate football,¹⁷ and high school football;¹⁸ however, reliability has not been examined in the general, recreationally active population. In addition, previously reported test-retest reliability values have been established using variable test-retest intervals including 15 minutes,^{11,15} 1 hour,¹⁸ and full athletic seasons.^{16,17} The timing between baseline and postinjury concussion assessment is highly variable and may extend between weeks, months, and years, which limits the ecological validity of the findings of previous studies with shorter test-retest intervals. In addition, limited independent studies have been performed to assess the measurement properties of the K-D test.^{11,15,16,19}

An additional source of error, which may influence K-D test performance, is exercise because it has been demonstrated to alter saccadic eye activity²⁰ and cognitive function.²¹ Previous studies have examined the influence of exercise on visual search tasks and discovered slower saccadic reaction time after exercise compared with quicker saccadic reaction times after a rest intervention.²⁰ Opposing literature has observed an 8% decrease in saccadic velocity in 11 cyclists after 180 minutes of continuous cycling at 60% of their maximal heart rate.²² If exercise can alter performance on sideline concussion screening tests, such as the K-D test, this could limit its clinical utility as a sideline measure of SC.

The purpose of our study was to evaluate the test-retest reliability of the K-D test and the influence of exercise on K-D test performance (eg, composite time) in healthy college students. We hypothesized that the K-D test would have acceptable test-retest reliability (≥ 0.75), and that exercise performed at a high intensity would impair K-D performance.

METHODS

Study Design

This was a controlled laboratory study with a crossover design. Participants completed 2 study visits separated by a 2-week time period. At each visit, participants completed a baseline administration of the K-D test, followed by an intervention of either rest or exercise, followed by a post-intervention K-D test. Interventions were randomly assigned across visit 1 and visit 2 using a random number generator. The primary outcome was the sum time to complete the K-D test. Test-retest reliability was calculated between the baseline K-D times at visits 1 and 2. The effects of exercise on baseline and postintervention K-D times were compared with the effects of rest. This study was approved by the university's institutional review board, and all participants provided informed consent.

Participants

A convenience sample of 63 healthy, recreationally active college students between 18 and 24 years of age participated in current study (Table 1). Subjects were recruited for a large, urban university and were excluded if they had a history of concussion within 6 months of participation or if they sustained any injury anytime throughout the study protocol. Additional exclusion criteria consisted of alcohol or caffeine consumption within 12 hours of a visit, a history of a cardiovascular condition, learning disability/dyslexia, history of attention-deficit disorder (ADD) / attention-deficit hyperactivity disorder (ADHD), or noncorrected vision impairments. This study was part of a larger design, which had additional exclusion criteria, which included a vestibular condition resulting in an inability to balance, or a history of lower extremity injury or surgery in the past 6 months or 12 months, respectively.

Procedures

After enrollment, participants completed the following questionnaires for descriptive purposes; the SCAT3 Symptom Checklist²³ and the Godin Leisure Time Activity Scale.²⁴ The K-D test was administered according to manufacturer's directions and using the K-D test, Version 1, (Mayo Foundation, Oakbrook Terrace, IL) and was administered by a sole examiner. For each K-D test, participants completed a demonstration card followed by the 3 testing cards in numerical order. For each card, participants were asked to stand and hold the K-D test cards approximately 16 inches from their eyes. Participants then read a series of single digit numbers on the card from left-to-right and from top-to-bottom as quickly as possible without committing an error (eg, omitting or incorrectly reading a number). The time to complete each card was measured using an electronic stopwatch, and the sum of the 3 cards was used to calculate the primary outcome, K-D time. For the initial baseline test at visit 1, 2 complete K-D tests were performed; however, only the second baseline measure was used for analysis, per manufacturers' recommendations.

After baseline K-D testing, participants completed either the exercise or rest interventions, which were randomly assigned across visit 1 and 2. The 30-minute exercise intervention consisted of 5 continuous cycles of physical activity.^{25,26} Each cycle consisted of five minutes of inclined treadmill walking at 3.0 mph followed by 1 minute of continuous hopping exercises, alternating between 10 squat jumps and 10 lateral hops. For the first 3 cycles, the treadmill incline was increased 1% grade for each minute of walking until reaching 15% grade, at which point it remained at 15% grade for the final 2 cycles. For the rest intervention, participants were instructed to sit in a chair for 30 minutes. Borg's rating of perceived exertion (RPE)²⁷ was used to evaluate participants' perceived exertion at baseline and postintervention K-D testing for both interventions, as well as after each of the 5 cycles of the exercise protocol.

Postintervention K-D testing was performed 2 minutes after the completion of the interventions using the same procedures as baseline testing. A 2-minute delay was chosen to replicate the potential time between the removal from a practice or competition and the clinical evaluation of an athlete suspected of having a concussion.

TABLE 1. Participant Demographics (Mean \pm SD)

	Sample (n = 63)
Age, yr	21.0 \pm 1.5
Sex	39 female, 24 male
Mass, kg	66.5 \pm 13.1
Height, m	1.71 \pm 0.11
Godin leisure time activity scale	67.3 \pm 22.4
SCAT3 symptom score	1.2 \pm 1.8

Statistical Analysis

Two-way random intraclass correlation coefficients ($ICC_{2,1}$) using average measures and 95% confidence intervals (CIs) and SEM with 95% CIs [95%CI = $2(\text{SD} \times \sqrt{1 - \text{ICC}})$] were calculated²⁸ between the baseline K-D time at visits 1 and 2 to assess the test-retest reliability and precision of the K-D test, respectively. In addition, ICCs with 95% CIs and 95% SEM were calculated to assess the within-session reliability for each intervention. The effects of exercise and rest interventions on K-D time were evaluated using a 2×2 (intervention \times time) repeated-measures analysis of variance and planned post hoc paired *t* tests. Cohen's *d* effect sizes with 95% CIs ([95%]) were also calculated.²⁹ Effect size CIs were interpreted as <0.1 = minimal, <0.5 = small, <0.8 = moderate, and ≥ 0.8 = large.²⁹ Statistical significance was set with $\alpha \leq 0.05$. All analyses were performed using SPSS Version 23.0 (Armonk, NY). In addition to our statistical analyses, the number of participants who achieved a clinically meaningful change was calculated. A clinically meaningful change was defined as a total time of completion ± 3 seconds compared with participant's baseline assessment.³⁰

RESULTS

A total of 69 potential participants were recruited for our study of which 63 were included in our data analysis. Of 69 participants, 2 were excluded for inability to complete exercise protocol, 3 were lost to follow-up between visit 1 and 2, and one was determined an outlier based on the time to complete the K-D test compared with the sample. The average test-retest time interval between visits 1 and 2 was 14.0 ± 0.5 days. Strong test-retest reliability coefficients [$ICC_{2,1} = 0.90$, (0.71-0.96)] were observed between baseline K-D measurements at the 2 time points, with a combined mean K-D time of 38.4 ± 5.7 seconds and a 95% SEM of 2.6 seconds (Figure 1). For the rest intervention, strong test-retest reliability was observed [$ICC_{2,1} = 0.96$, (0.93-0.97)] between baseline and postintervention measures, with a combined mean K-D time of 37.6 ± 5.9 seconds and a 95% SEM of 2.4 seconds. For the exercise intervention, again, strong test-retest reliability was observed [$ICC_{2,1} = 0.96$, (0.94-0.98)] between baseline and postintervention measures, with a combined mean K-D time of 37.7 ± 5.5 seconds and a 95% SEM of 2.2 seconds. Depicts change scores (seconds) between baseline and postintervention assessments.

The mean values and 95% CIs for the RPE measures during exercise and rest interventions are shown in Figure 3. After each intervention, the mean RPE was 6.2 ("very, very light") after rest. The average RPE 2 minutes after exercise was 12.4 ("somewhat hard"). However, immediately after the last cycle

of the exercise intervention, an average RPE value of 16.1 ("hard") was reported. Mean values and SDs of K-D test performance at baseline and postintervention are shown in Table 2. We observed no significant intervention \times time interaction ($F_{1,62} = 0.4$, $P = 0.55$) or intervention main effect ($F_{1,62} = 0.2$, $P = 0.68$) on K-D time. No significant differences between interventions at baseline ($t_{62} = -0.1$, $P = 0.91$) or postintervention ($t_{62} = -0.6$, $P = 0.53$) were observed. A significant time main effect ($F_{1,62} = 60.2$, $P < 0.001$) for K-D time was observed. More specifically, a significant decrease in time to completion occurred from baseline to postintervention administration of the K-D test after both the rest ($t_{62} = 5.5$, $P < 0.001$) and exercise ($t_{62} = 5.9$, $P < 0.001$) interventions. However, effect sizes from baseline to postintervention were small and 95% CI crossed zero for the rest [$d = 0.27$ (-0.62 to 0.08)] and exercise [$d = 0.29$ (0.64-0.06)] interventions.

DISCUSSION

The purpose of our study was to examine the test-retest reliability and the influence of exercise of the K-D test in healthy college participants. We hypothesized that the K-D test would have acceptable test-retest reliability (≥ 0.75),³¹ and the time to complete the K-D test would increase (eg,

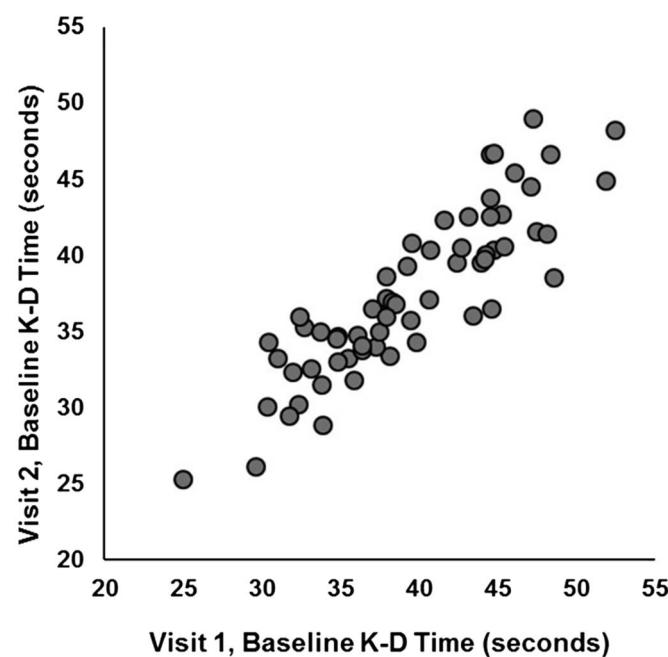


Figure 1. Scatter plot of visit 1 versus visit 2 baseline K-D test time.

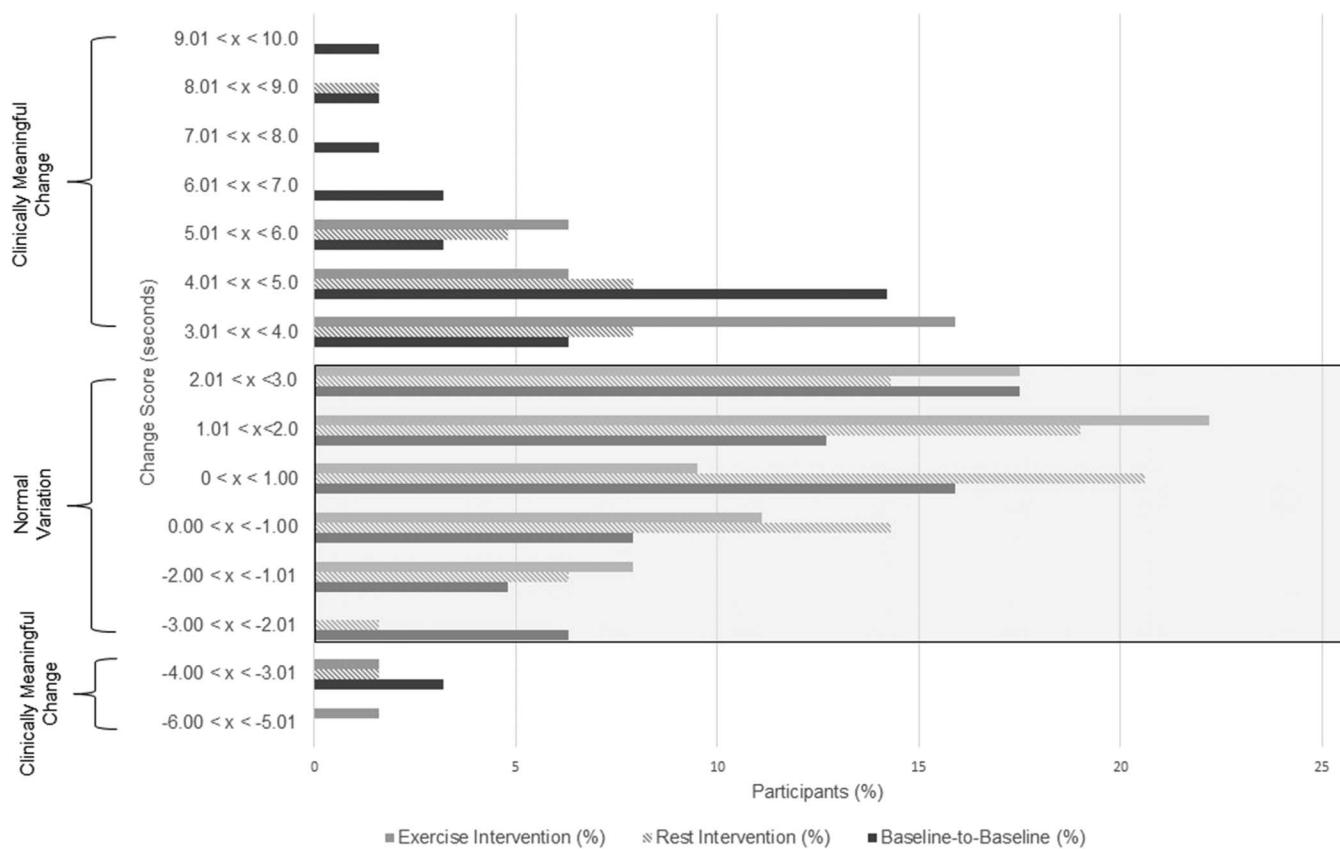


Figure 2. The frequency of race changes scores between visit 1 and 2 and before and after each intervention. The shaded area represents acceptable variation associated with the K-D test.³²

worse score) after an exercise intervention compared with a rest intervention. Overall, the K-D test was observed to have strong reliability ($ICC = 0.90$) using a two-week test-retest interval. In terms of the influence of exercise, no differences were observed for either the exercise or rest interventions on K-D test performance.

Our test-retest reliability findings are comparable with previous literature that reported stability of the K-D test to range from 0.89 to 0.97.^{11,16-18,32} We chose a 2-week test-retest interval to reflect a more clinically relevant time frame to reflect what clinicians may typically observe in the sport setting between baseline and postinjury assessments. Previous studies that have examined the reliability of the K-D have used intervals ranging from 15 minutes¹¹ to several months.^{16,17,32} Our study was also the first to examine the K-D test-retest reliability in a cohort of recreationally active young adults. Previous K-D reliability studies have focused on test-retest reliability within homogeneous sport-specific populations, including mixed martial art fighters and boxers,¹¹ amateur³² and junior rugby players,¹⁶ and collegiate athletes.¹⁷ Our findings broaden the potential clinical applicability by showing the K-D test may be a reliable tool in the recreationally active young adult population and other settings (eg, high school and collegiate athletics).

Our methodology specifically addressed the potential influence of exercise on K-D test performance. We hypothesized that exercise would have a deleterious effect on K-D time (eg, increased time to completion) contrasted with a resting state; however, our results indicated that exercise did not affect K-D

time when compared with the rest intervention. These findings suggest that the K-D test may be robust to the potential effects of exercise (ie, practice or competition). We did observe a statistically significant decrease in K-D time from baseline to postintervention for both exercise and rest; however, the effect size point estimates were small and 95% CIs crossed zero. At a group level, K-D time decreases were small (1.5-1.7 seconds), compared with what is considered a clinically meaningful change for the K-D test.³⁰ Subsequent analyses revealed high test-retest reliability between the baseline and postintervention K-D times for rest and exercise interventions and the 95% SEMs range from 2.2 to 2.5 seconds on K-D measurements that were an average of 37 seconds long. Because the mean differences were smaller than the 95% SEMs, this suggests the decrease in K-D time over 30 minutes falls within the range of the 95% SEM. This small decrease in K-D time may be the product of a practice effect with repeated measures over a short interval (~30 minutes) from baseline to postintervention.

When examining K-D test scores at the individual level, a clinically meaningful change was observed in approximately 3% to 35% of participants. Interestingly, 23.8% to 35% participants were observed to have a clinically meaningful change of 3 seconds³⁰ after a 30-minute rest period or after a two-week test-retest interval, respectively. For the postrest analysis, the 96% (23/24) of participants who demonstrated a clinically significant change were observed to meet the criteria for an impaired or “slowed” performance. Our assessment of change scores between the first and second baseline assessments

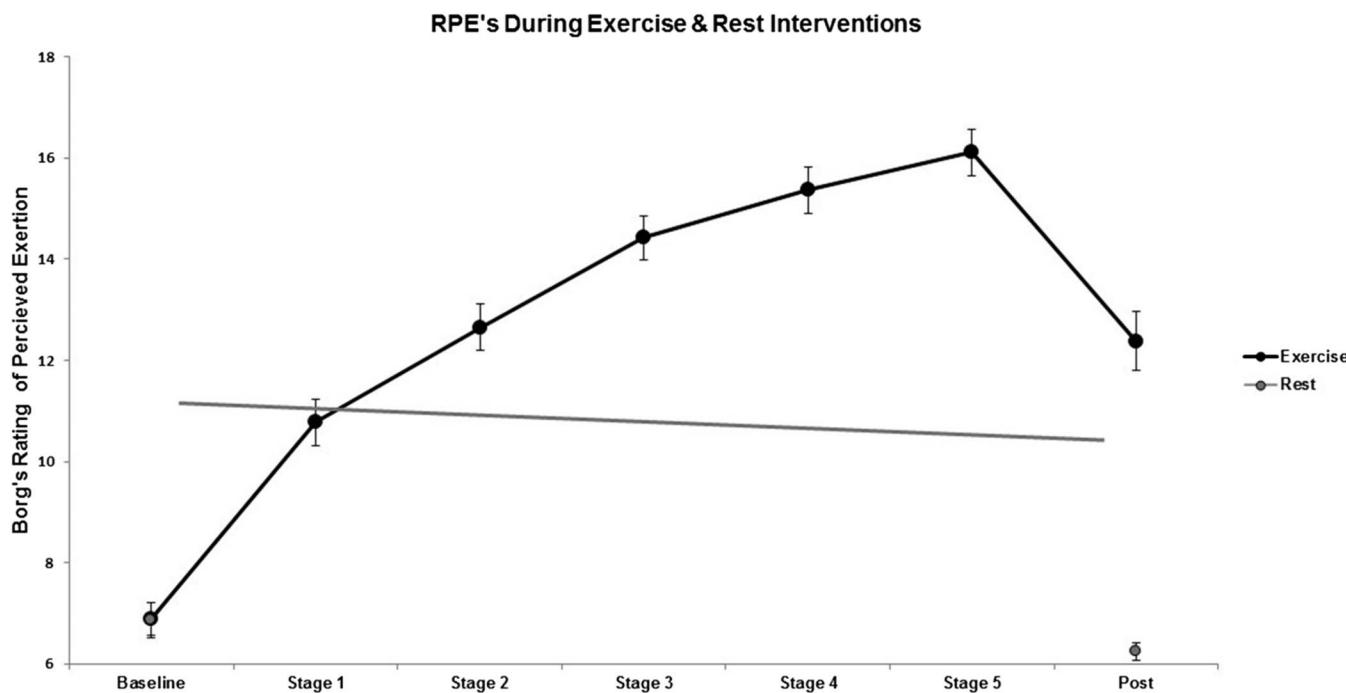


Figure 3. Mean and 95% CIs for exercise (stages) and rest intervention RPE at baseline, during, and postinterventions.

revealed 32% (20/63) would have been classified as having an impaired performance while 3.2% (2/63) were observed to have a clinically meaningful improved or “decreased” time to completion. For the exercise intervention, 90% (18/20) of participants who met the criteria for a clinically meaningful change were observed to have significantly slower K-D test completion time. Overall, these data suggest up to one-third of our healthy collegiate participants would have been classified as impaired when healthy (false-positive), which brings the clinical measure’s specificity into question. Our findings are similar to those of Molloy et al¹⁴ who reported the K-D test to have a specificity of 69%.

Our exercise intervention effectively brought participants to an average maximal exertion of “hard” on the RPE scale and, at postintervention, the RPEs were at “somewhat hard,” whereas, during the rest intervention, participants remained at a “very, very light” exertion level from baseline to post-intervention. We chose to wait 2 minutes between exercise and the postintervention K-D test in an effort to better replicate what would be observed in clinical practice. Rarely would a concussion screening test be performed immediately after the cessation of exercise. First, the patient to be removed from the field of play and a proper medical history and evaluation would be performed. Not testing the participants immediately after exercise may have led to alternative findings; however,

we feel that those findings would have been less clinically applicable.

Our findings support those of animal models’ research, which suggests that an increased levels of catecholamines such as norepinephrine, lead to an increased ability to respond to visual stimuli.³³ Our participants reported an RPE of “hard” or “somewhat hard” immediately after and after 2 minutes of rest after the prescribed 30-minute exercise protocol. Moderate-intensity exercise has been routinely shown to lead to increased cognitive function, whereas intense exercise has been demonstrated to have deleterious consequences on a variety of cognitive tasks.³⁴ This relationship is described by the inverted-U hypothesis. These deleterious effects have been hypothesized to be a result of excessive catecholamine release (norepinephrine and dopamine), hypopituitary axis stimulation resulting in release of adrenocorticotrophic hormone and subsequently cortisol, and elevated levels of serotonin.²¹ Rationale for not observing impaired K-D test performance may be due to the length of exercise protocol because most literature suggests prolonged periods of exercise (≥ 45 minutes) result in impaired cognitive performance.³⁵ That said, rarely in sports, high risk for concussion (eg, football, rugby, men’s and women’s ice hockey) do athletes have continuous periods moderate- to high-intensity exercise to evoke these physiological changes that would lead to noted

TABLE 2. Mean \pm SD (95% CIs) for K-D Test Performance at Baseline and Postrest and Exercise Interventions

	Intervention	
	Rest	Exercise
Baseline K-D time (s)	38.5 ± 5.7 (37.1 to 39.9)	38.4 ± 5.8 (37.0 to 39.8)
Post K-D time (s)	37.0 ± 5.3 (35.7 to 38.3)	36.7 ± 6.0 (35.2 to 38.2)
Average change (s)	-1.5 ± 2.2 (-2.04 to -0.96)	-1.7 ± 2.3 (-2.27 to -1.13)

cognitive deficits. An additional factor that may have influenced performance is that our study was conducted in a controlled laboratory setting. Although there are several benefits to the laboratory setting to conduct studies such as ours, it may inherently lack external validity because of the absence of intrinsic and extrinsic factors that would be present during practice or a competition.

This study is not without limitations. We chose a 2-week test-retest interval to simulate a more applicable clinical testing timeline for baseline and postconcussion testing; however, a 2-week test-retest interval may not be applicable in all clinical scenarios. In addition, the repeated exposure to the K-D test may have influenced our findings by lessening the impact of the exercise intervention as a result of practice effects. Our exercise intervention was a standardized submaximal protocol performed in a controlled laboratory setting that was non-sport specific. The results of this particular study may not be applicable to all types of exercise or sporting events. Future studies should be completed to determine the effects of uncontrolled exercise (ie, sport) on the K-D test performance. Overall, our results suggest that the K-D test has strong test-retest reliability in a large young, healthy, recreationally active population. However, despite having strong reliability values, up to 32% of participants were deemed to have a false-positive test. Future research should specifically assess the sensitivity and specificity of the K-D test in varying sports to examine its clinical utility.

CONCLUSIONS

Based on the findings of our study, the K-D test met the necessary stability reliability criteria for clinical utility in a healthy, recreationally active college-aged cohort. However, despite having strong reliability coefficients before and after rest and exercise, up to one-third of our sample met the clinical criteria for impaired performance before and after each intervention. Future research is necessary to further examine the sensitivity and specificity of the K-D test at varying levels and types of sport.

ACKNOWLEDGMENT

The authors acknowledge Catherine Donohue for her contributions to data collection.

References

- Resch JE, Kutcher JS. The acute management of sport concussion in pediatric athletes. *J Child Neurol*. 2015;30:1686–1694.
- McCrory P, Meeuwisse W, Dvorak J, et al. Consensus statement on concussion in sport—the 5th international conference on concussion in sport held in Berlin, October 2016. *Br J Sports Med*. 2017;51:838–847.
- Elbin RJ, Sufrinko A, Schatz P, et al. Removal from play after concussion and recovery time. *Pediatrics*. 2016;138:e1–e8.
- Thomas M, Haas TS, Doerer JJ, et al. Epidemiology of sudden death in young, competitive athletes due to blunt trauma. *Pediatrics*. 2011;128:e1–e8.
- Broglio SP, Cantu RC, Gioia GA, et al. National Athletic Trainers' Association position statement: management of sport concussion. *J Athl Train*. 2014;49:245–265.
- Westheimer G. Mechanism of saccadic eye movements. *AMA Arch Ophthalmol*. 1954;52:710–724.
- Heitger MH, Jones RD, Macleod AD, et al. Impaired eye movements in post-concussion syndrome indicate suboptimal brain function beyond the influence of depression, malingering or intellectual ability. *Brain*. 2009;132:2850–2870.
- Kerr ZY, Zuckerman SL, Wasserman EB, et al. Concussion symptoms and return to play time in youth, high school, and college american football athletes. *JAMA Pediatr*. 2016;170:647–653.
- Heitger MH, Jones RD, Dalrymple-Alford JC, et al. Motor deficits and recovery during the first year following mild closed head injury. *Brain Inj*. 2006;20:807–824.
- Tjarks BJ, Dorman JC, Valentine VD, et al. Comparison and utility of King-Devick and ImPACT composite scores in adolescent concussion patients. *J Neurol Sci*. 2013;334:148–153.
- Galetta KM, Barrett J, Allen M, et al. The King-Devick test as a determinant of head trauma and concussion in boxers and MMA fighters. *Neurology*. 2011;76:1456–1462.
- Bodack MI, Chung I, Krumholz I. An analysis of vision screening data from New York City public schools. *Optometry*. 2010;81:476–484.
- Sussman ES, Ho AL, Pendharkar AV, et al. Clinical evaluation of concussion: the evolving role of oculomotor assessments. *Neurosurg Focus*. 2016;40:E7.
- Molloy JH, Murphy I, Gissane C. The King-Devick (K-D) test and concussion diagnosis in semi-professional rugby union players. *J Sci Med Sport*. 2017;20:708–711.
- Leong DF, Balcer LJ, Galetta SL, et al. The King-Devick test as a concussion screening tool administered by sports parents. *J Sports Med Phys Fitness*. 2014;54:70–77.
- King D, Hume P, Gissane C, et al. Use of the King-Devick test for sideline concussion screening in junior rugby league. *J Neurol Sci*. 2015;357:75–79.
- Leong DF, Balcer LJ, Galetta SL, et al. The King-Devick test for sideline concussion screening in collegiate football. *J Optom*. 2015;8:131–139.
- Alsalameen B, Haines J, Yorke A, et al. King-Devick Test reference values and associations with balance measures in high school American football players. *Scand J Med Sci Sports*. 2016;26:235–239.
- Oberlander TJ, Olson BL, Weidauer L. Test-retest reliability of the king-devick test in an adolescent population. *J Athl Train*. 2017;52:439–445.
- Llorens F, Sanabria D, Huertas F, et al. Intense physical exercise reduces overt attentional capture. *J Sport Exerc Psychol*. 2015;37:559–564.
- McMorris T. *Exercise-Cognition Interaction: Neuroscience Perspectives*. Elsevier Science, London, UK; 2015.
- Connell CJ, Thompson B, Kuhn G, et al. Fatigue related impairments in oculomotor control are prevented by caffeine. *Sci Rep*. 2016;6:26614.
- McCrory P, Meeuwisse WH, Aubry M, et al. Consensus statement on concussion in sport: the 4th international conference on concussion in sport held in Zurich, November 2012. *Br J Sports Med*. 2013;47:250–258.
- Godin G, Jobin J, Bouillon J. Assessment of leisure time exercise behavior by self-report: a concurrent validity study. *Can J Public Health*. 1986;77:359–362.
- Goetschius J, Kuenze CM, Saliba S, et al. Reposition acuity and postural control after exercise in anterior cruciate ligament reconstructed knees. *Med Sci Sports Exerc*. 2013;45:2314–2321.
- Kuenze C, Hertel J, Hart JM. Effects of exercise on lower extremity muscle function after anterior cruciate ligament reconstruction. *J Sport Rehabil*. 2013;22:33–40.
- Borg GA. Psychophysical bases of perceived exertion. *Med Sci Sports Exerc*. 1982;14:377–381.
- Thomas JR, Nelson JK, Silverman SJ. *Research Methods in Physical Activity*. Champaign, IL: Human Kinetics; 2015.
- Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. L Erlbaum Associates, Manwah, New Jersey; 1988.
- Galetta KM, Brandes LE, Maki K, et al. The King-Devick test and sports-related concussion: study of a rapid visual screening tool in a collegiate cohort. *J Neurol Sci*. 2011;309:34–39.
- Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice*. Pearson/Prentice Hall, Upper Saddle River, New Jersey; 2009.
- King D, Gissane C, Hume PA, et al. The King-Devick test was useful in management of concussion in amateur rugby union and rugby league in New Zealand. *J Neurol Sci*. 2015;351:58–64.
- Waterhouse BD, Moises HC, Woodward DJ. Phasic activation of the locus coeruleus enhances responses of primary sensory cortical neurons to peripheral receptive field stimulation. *Brain Res*. 1998;790:33–44.
- Chmura J, Nazar K, Kaciuba-Uscilko H. Choice reaction time during graded exercise in relation to blood lactate and plasma catecholamine thresholds. *Int J Sports Med*. 1994;15:172–176.
- Hill EE, Zack E, Battaglini C, et al. Exercise and circulating cortisol levels: the intensity threshold effect. *J Endocrinol Invest*. 2008;31:587–591.

A Reliability and Comparative Analysis of the New Randomized King-Devick Test

Minh Q. Nguyen, BSc, Doug King, PhD, Alan J. Pearce, PhD

Objective: The King-Devick (K-D) test is a rapid visual screening tool that can assess underlying brain trauma such as concussion via impairments in saccadic rhythm. A new tablet version of the K-D test using randomized numbers is now available, but reliability for this new version and comparison to the traditional K-D test has not yet been reported. Known for learning effects in the test, the aim of this study was to determine test-retest reliability and to compare performance of the new “randomized” version to the “traditional” K-D test version. We hypothesized that the “traditional” K-D test would show a greater rate of improvement with repeat application, compared with the “randomized” K-D test.

Methods: Using a cross-sectional, repeated measures design in a healthy university student cohort ($n = 96$; age 21.6 ± 2.8 years; 49 women, 47 men), participants were required to complete the K-D test twice with a one-week break between testing sessions. Participants were randomly assigned into a “traditional” group, where they completed a test-retest of the established K-D protocol, using the same numbers; or the “randomized” group, where they completed test-retest protocol using 2 different sets of numbers.

Results: Reliability testing showed a strong intraclass correlation coefficient for both the “traditional” test group (control group: 0.95 [CI: 0.91–0.97]) and the “randomized test group” (0.97 [CI: 0.95–0.98]). However, contrary to our hypothesis, no differences were found between “traditional” and “randomized” groups for baseline (control: 42.5 seconds [CI: 40.2–

44.9 s] vs randomized: 41.5 [38.7–44.4], $P = 0.23$) and repeated testing between groups (control: 40.0 seconds [37.9–42.1 s] vs randomized: 39.5 [36.9–42.0], $P = 0.55$), with both groups showing improved times with repeated testing (control: 2.1 seconds [CI: 1.1–3.2 seconds] and randomized: 1.9 seconds CI: [0.9–2.9 seconds], $P < 0.001$).

Conclusions: The “randomized” version of the K-D test, using different sets of numbers, demonstrates good reliability that is comparable to the traditional K-D testing protocol that uses the same number sets. However, similar to the “traditional” K-D test, learning effects were also observed in the “randomized” test, suggesting that learning effects are not because of content memorization, but rather familiarity of the test. As a result, although either test format is suitable for sideline concussion screening or return to play decisions, comparison of data should be made to the individual’s baseline rather than to normative data sets.

Journal of Neuro-Ophthalmology 2019;00:1–6

doi: 10.1097/WNO.0000000000000829

© 2019 by North American Neuro-Ophthalmology Society

Concussion continues to be an ongoing public health concern (1). Greater awareness of concussion injury, and the effects of concussions on risks of cognitive impairment and neurodegenerative disease, has increased endeavors to develop tools to identify potentially concussed athletes quickly, efficiently, and by the field of play (2).

In addressing this concern, many sports have implemented general recognition and clinical concussion diagnostic assessments. Among the available range of sideline concussion tests, the King-Devick (K-D) test, which involves reading aloud a series of numbers as quickly as possible, has been demonstrated as a valid and reliable tool for the rapid screening of athletes suspected of concussion (2–5). This is because the visual system is important in the diagnosis of concussion (2,6,7). Visual processing uses functional connectivity of multiple areas of the cortex involving visual-spatial integration, attention, motor planning, and language function (2,5). Cortical areas involved in saccadic function include the frontal eye fields, dorsolateral prefrontal cortex (DLPFC), supplementary motor, posterior

School of Allied Health (MQN, AJP), Human Services and Sport, La Trobe University, Melbourne, Australia; and Sports Performance Research Institute New Zealand (SPRINZ) (DK), Faculty of Health and Environmental Science, Auckland University of Technology, Auckland, New Zealand.

The authors report no conflicts of interest.

A. J. Pearce is supported, in-part, by funding from *Sport Health Check* charity (Australia), and has previously received in-part funding from the Australian Football League (Melbourne, Australia), Samsung Corporation Australia (Sydney, Australia), and Impact Technologies Inc (Perth, Australia). The remaining authors report no conflicts of interest.

D. King is not affiliated with the King-Devick test or King-Devick Technologies.

Address correspondence to Alan J. Pearce, PhD, School of Allied Health, Human Services and Sport, La Trobe University, Melbourne, Australia 3083; E-mail: alan.pearce@latrobe.edu.au

parietal, and middle temporal areas, and the striate cortex (8–12). Specifically, these areas control planning and completing coordinated saccades involved in tasks such as reading. Moreover, for smooth saccades to occur during tasks such as reading, the DLPFC is also involved in antisaccades (10,12). It is also known that subcortical structures including the thalamus, superior colliculus, cerebellum, and brainstem are also involved in maintaining smooth eye movements (13). The functional connectivity of these areas toward maintaining optimal eye movements is important, and injury to brain will result in poorer performance in tasks such as reading (14). As suggested by Galetta et al (2), saccadic eye assessment is pertinent in observing the underlying neurophysiology of the brain following injury such as a concussion.

Originally described as a reading test to assess the relationship between oculomotor function and reading ability (15,16), the K-D test is now used to assess brain functionality in both experimental and clinical settings. For example, slower performances in the K-D reflecting suboptimal brain function has been observed following experimentally induced hypoxia with mean times for the K-D test 9.2 seconds slower than nonhypoxic controls (17). Similar findings of poorer K-D performance have been reported in sleep deprivation (18). Studies in clinical populations, including Parkinson disease, multiple sclerosis (MS), and amyotrophic lateral sclerosis, have shown similar outcomes of poorer performance with K-D times than age-matched healthy controls (19–21). Furthermore, the study in MS showed that poorer performance in K-D times was associated with MS patients' worse quality of life regarding vision (assessed by the 25-Item National Eye Institute Visual Functioning Questionnaire) (21). Conversely, improved K-D performance has been observed following moderate and high exercise bouts inferring increased mental arousal (22,23).

The K-D test has previously reported good construct validity to eye movement tracking metrics in experimental conditions and clinical screening tools for concussion including the sports concussion assessment tool and military acute concussion evaluation (2,23–25). However, despite reports of high reliability of the K-D test (intraclass correlation coefficient of >0.92) (2), a continuing concern revolves around improvement in test times with subsequent attempts, which has been attributed to test familiarization and learning effects. In the meta-analysis by Galetta et al (2) the weighted mean improvement in nonconcussed individuals across 15 studies was 1.8 seconds (95% CI: −3.4 to −0.1; $I^2 = 0.0\%$, $P = 0.98$). However, it has been argued that improvement in the test outcome further highlights the deleterious effects of concussion, particularly when concussed individuals show a mean worsening of times by an average of 4.8 seconds (2).

In response to this, the developers of the K-D test have updated the tablet application (Version 4) to include 3 sets

of numbers, allowing assessors to “randomize” the test application. Currently, a test–retest reliability and comparison of learning effects between the new randomized and traditional versions has not yet been undertaken. Therefore, this is the first independent study to compare intertest reliability between the “traditional” (same numbers presented) and the new “randomized” (different numbers presented) tablet application. We calculated intraclass correlation coefficients (ICC) and compared differences in repeated data between the traditional and randomized tablet versions. We hypothesized that both tests would show good reliability with repeated application. With known learning effects associated with the K-D test (2), a second hypothesis was that a significant difference in performance improvement would be observed in the “traditional” K-D test compared with the new “randomized” K-D test.

METHODS

Study Design and Participants

A convenience total of 96 participants (21.6 ± 2.8 years; 49 women and 47 men) were recruited from the university student population, based on an a priori power analysis ($f = 0.2$, $1 - \beta$ error probably 0.95, $P > 0.05$) requiring a minimum sample of 84 participants (26). With less stable data in older population groups (2), inclusion criteria required participants to be under the age of 50 years and neurologically healthy. Using previously published criteria (27), those who had a recent concussion, within the last 6 months, were ineligible to participate in the study. As exercise has been shown to influence K-D test scores (22,23), participants refrained from any physical activity 2 hours before testing. All study procedures received approval from the University human research ethics committee (LTU-HEC 18207). Individuals recruited for the study completed pretesting screening and signed an informed consent before study enrolment.

Participants were randomly assigned, via the random number generator, into one of 2 groups (Fig. 1 for study design): 1) the “traditional” K-D test where individuals completed the standard protocol using previously published numbers (Test 1 twice); and 2) the “randomized” K-D test where individuals completed the standard test the first time, but were then provided with one of 2 new tests (Test 2 or Test 3).

All K-D testing was completed using a tablet (iPad; Apple Inc., Cupertino, CA) according to the developer's recommendations (v4.2.2; King-Devick technologies Inc). All testing protocols were completed in an indoor laboratory setting, shielding participants from any background noise or visual distractions.

Using standardized protocols (3,4,22,28,29), participants were instructed to read the numbers presented for 3 trials of increasing difficulty. Errors made during reading

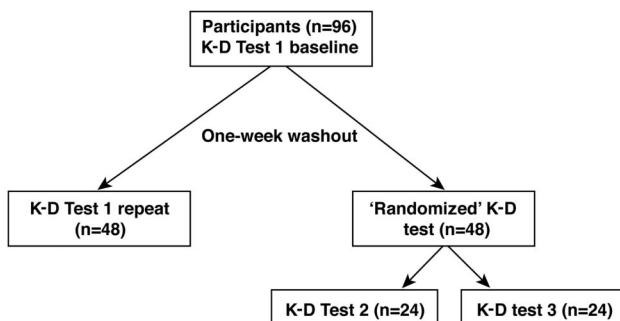


FIG. 1. Research design for the study. All participants completed “Test 1” of the King-Devick (K-D) assessment. After a 1-week washout period, half were then randomly assigned to repeat “Test 1” again, using the same number set. The other half were randomly assigned to complete “Test 2” or “Test 3” with different number sets.

and the total time for the 3 trials were included in calculating the K-D test “score.” All participants completed 2 trials of Test 1 within a short period of time (several minutes) to establish their baseline “score.” The faster time from 2 trials became the established baseline K-D test time for week 1 (3,23).

A 1-week break was provided between test–retest sessions. For the retest sessions, participants assigned to the “traditional K-D” control group repeated the protocol using the same numbers from Test 1. Participants assigned to the “randomized” group repeated the protocol from either Test 2 or Test 3. The same person conducted all baseline and repeated measurements.

Statistical Analysis

Data were screened for normal distribution using the Kologorov–Smirnov (KS) tests and found to be normally distributed (KS: 0.083, $P = 0.149$). Test–retest reliability was calculated using ICC, with 95% CI, to examine agreement between first and second baseline test scores and the repeat-testing scores.

Comparisons between groups (traditional vs randomized) were completed using a mixed-model repeated measure ANOVA. Furthermore, comparison was also made between those in the randomized group who completed Test 2 vs those who completed Test 3. Post-hoc Bonferroni tests were undertaken, where ANOVA detected differences. All data are presented as mean [\pm 95% CI], unless specified, and alpha was set at $P < 0.05$.

RESULTS

Reliability Analysis

The ICC between trials for the traditional (control) group was 0.949 (0.913–0.972). The ICC for the randomized group was 0.972 (0.951–0.984). Further reliability testing was completed within the 2 subgroups (Test 2/Test 3)

within the randomized cohort with ICCs of 0.967 (0.931–0.986) and 0.974 (0.944–0.990), respectively.

Comparison Between Traditional and Randomized Number Test Protocols

Comparison between groups (Fig. 2) showed no interaction effect ($F_{1, 95} = 1.475$; $P = 0.228$) nor differences for the main effect for group ($F_{1, 95} = 0.358$; $P = 0.551$). A main effect of time ($F_{1, 95} = 27.787$; $P < 0.001$) was observed, showing that both groups had improved in their repeated test. K-D test time for the traditional test (Test 1) improved by a mean of 2.029 seconds (1.158–2.901 seconds) whereas the combined randomized test sample (Test 2 and Test 3) showed a mean improvement of 1.269 seconds (0.358–2.180 seconds). Analyses of improvements of those who completed Test 2 and Test 3 separately were 1.014 seconds (−0.408 to 2.438 seconds) and 1.631 seconds (0.598–2.664 seconds), respectively. No interaction effect was observed ($F_{1, 47} = 0.445$; $P = 0.508$) nor a difference for the main effect for group ($F_{1, 47} = 1.736$; $P = 0.195$). A main effect for time ($F_{1, 47} = 8.195$; $P = 0.006$) was observed, also showing that both subgroups had improved in their repeated test.

DISCUSSION

The aim of this study was to assess the reliability of the newly modified K-D assessment that features different number sets to randomize the assessment protocol. A secondary question was to compare the new randomized number sets with the traditional K-D test protocol where the same set of numbers is used each time. Although the K-D test has been an important tool for the recognition of concussion and screening for other neurological conditions

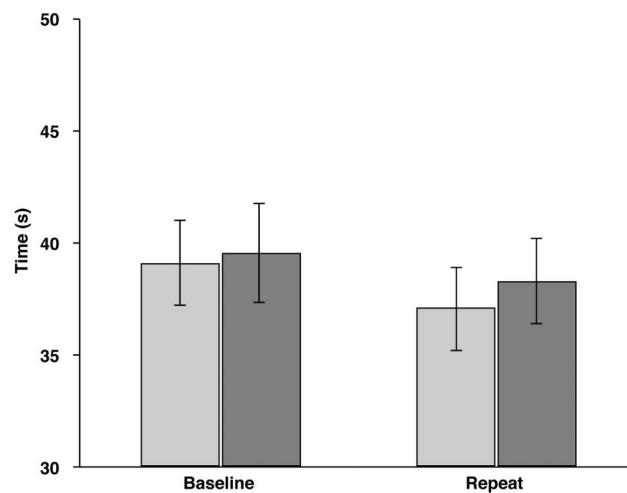


FIG. 2. Mean (\pm 95% CI) King-Devick (K-D) test times for “traditional” group (Test 1: Test 1; light-colored bars), and “randomized” group (Test 1: Test 2 or Test 3; dark-colored bars). Both groups, irrespective of the test completed in the retest session (Fig. 1), significantly improved over time.

(2), the rationale of developing 3 test options, thereby randomizing the numbers and potentially reducing learning effects, was in response to previous studies reporting improvement in times with repeat applications of the test (2,5,24,29).

We hypothesized that using new number sets (Test 2 or Test 3) would be reliable. We also hypothesized that there would be a greater learning effect using the same number sets (Test 1) compared to using a different number set (Test 2 or 3). Across all versions of the tests we found high reliability in baseline retest K-D scores. However, contrary to our second hypothesis, our data showed that there was improvement across all test options, suggesting learning effects irrespective of whether the retest used the same (Test 1: Test 1) or randomized (Test 1: Test 2 or Test 3) number sets. Furthermore, no differences were observed in improvements between the randomized subtests (Test 1: Test 2 or Test 1: Test 3).

The test-retest reliability results in this study are analogous to previous research reporting constancy of the K-D test of ICCs >0.92 (2). Similar to recent studies that used a 2-week test-retest interval (27), our interval of one week test-retest time frame was aimed to reflect general clinical practice in Australia where medical doctors may typically observe their patients between appointments, who will be rested from activities involving physical contact for 7–10 days, particularly if they are not elite-athletes (5,30). However, it is acknowledged that patients who are diagnosed with concussion may be tested more routinely after injury. Previous investigations have demonstrated reliability of the K-D test with learning effects, using test-retest intervals varying from 15 minutes (25) and 1 week (5) to several months (3,4,31).

The improvement across all 3 number sets hints that the learning effects observed are not likely to be learning or familiarization of the numbers themselves, but rather a practice effect of the testing protocol. Across cognitive testing performance literature, improvements in retesting have been well reported. Further to familiarization or learning of the content, there are a number of reasons why improvements can occur. Suggestions include reduction in anxiety with familiarity of the test (32), improved strategies in performing a test that does not reflect learning per se (33), and regression to the mean (34). However, Hausknecht et al (35) have also posited that mere repetition can affect (improve) test scores. Improvements without interventions, such as coaching or instruction, are ascribed to familiarity with the test environment, irrespective of the actual content in the assessment, or an improved understanding of the items contained in the test. We suggest that this familiarity would underlie observed improvements. Despite the K-D test protocols requiring a baseline being established from 2 error-free trials, with subsequent improvements in time establishing a new baseline for the individual (5,23), further studies are required to determine how

much repetition would be required before improvement plateaus are observed.

Although other cognitive tests may affect interpretation through improvements, improvements in K-D times, outside of interventions such as exercise (22), serve to improve recognition of suspected concussion injuries and other conditions affecting brain function. For example, in a study investigating the effects of sleep deprivation on brain function, neurology residents who averaged 2 hours of sleep showed smaller performance improvements in the K-D compared with controls, who had normal amounts of sleep (18). In concussion, recent studies have demonstrated increased (worsening) K-D times, relative to improvement in baselines, in those who were clinically diagnosed with concussion (3–5). With improvements observed in this study for Tests 2 and 3, future studies should not discourage the use of the randomization protocol for field assessments to establish baseline values. Irrespective of which test is used for retesting, it is important to note that posttest data should be compared with the individual's baseline, rather than comparison of age-normative data sets (2).

Limitations of this study include the sample tested, specifically a young, university student population. We focused on a younger population group given that most studies investigating test-retest reliability have been in those younger than 40 years of age (2), allowing for comparability to previous data. Future studies should also consider testing an older population group given their higher risk of falls, which is the largest contributor to concussions in the elderly (36).

Another limitation is that although it can be argued that the test is culturally neutral, we did not screen participants for English as their first language. Several recent studies have reported on K-D times being affected by those whose first language is not English (37,38). However, in this study we compared individuals with their own baseline, suggesting the findings would not be affected. Nevertheless, future studies that involve quantifying the effects of an intervention (such as exercise) or suspected concussion should screen participants for first language and, if feasible, individuals should complete the test using their preferred language (37).

Further studies should investigate comparing the randomized protocol using an older population, particularly those over the age of 40 years. In their systematic review, Galetta et al (2) suggest that between the ages of 18–40 years, K-D test times appear stable with slight increases in times for those older than 40 years. Future studies should also compare different intervals of application of the randomized K-D test to reflect the possibility of practice effects. Patients diagnosed with concussion may be tested routinely, such as daily or every third day for example, and thus practice effects may be an issue. Studies could compare reliability of application of the randomized test between shorter and longer time frames (2 days vs 1 week),

because the randomized version may still be prone to practice effects or issues such as test anxiety, as opposed to time.

In conclusion, this study has demonstrated reliability with the new randomized version of the K-D test, comparable with the traditional K-D testing protocol using the same number sets. Similar to the traditional protocol, we observed learning effects using the randomized number sets. This suggests that the randomized K-D test is suitable for concussion recognition using an individual comparison, rather than to any normative data sets.

STATEMENT OF AUTHORSHIP

Category 1: a. Conception and design: M. Q. Nguyen and A. J. Pearce; b. Acquisition of data: M. Q. Nguyen and A. J. Pearce; c. Analysis and interpretation of data: M. Q. Nguyen, D. King, and A. J. Pearce. Category 2: a. Drafting the manuscript: M. Q. Nguyen, D. King, and A. J. Pearce; b. Revising it for intellectual content: D. King and A. J. Pearce. Category 3: a. Final approval of the completed manuscript: M. Q. Nguyen, D. King, and A. J. Pearce.

REFERENCES

1. **Wiebe DJ**, Comstock RD, Nance ML. Concussion research: a public health priority. *Inj Prev*. 2011;17:69–70.
2. **Galetta KM**, Liu M, Leong DF, Ventura RE, Galetta SL, Balcer LJ. The King-Devick test of rapid number naming for concussion detection: meta-analysis and systematic review of the literature. *Concussion*. 2016;1:CNC8.
3. **King D**, Gissane C, Hume P, Flaws M. The King-Devick test was useful in management of concussion in amateur rugby union and Rugby League in New Zealand. *J Neurol Sci*. 2015;351:58–64.
4. **King D**, Hume P, Gissane C, Clark T. Use of the King-Devick test for sideline concussion screening in junior rugby league. *J Neurol Sci*. 2015;357:75–79.
5. **Hecimovich M**, King D, Dempsey AR, Murphy M. The King-Devick test is a valid and reliable tool for assessing sport-related concussion in Australian football: a prospective cohort study. *J Sci Med Sport*. 2018;21:1004–1007.
6. **Ventura RE**, Jancuska JM, Balcer LJ, Galetta SL. Diagnostic tests for concussion: is vision part of the puzzle? *J Neuroophthal*. 2015;35:73–81.
7. **Ventura RE**, Balcer LJ, Galetta SL. The neuro-ophthalmology of head trauma. *Lancet Neurol*. 2014;13:1006–1016.
8. **Sparks DL**, Mays LE. Signal transformations required for the generation of saccadic eye movements. *Annu Rev Neurosci*. 1990;13:309–336.
9. **Pierrot-Deseilligny C**, Rivaud S, Gaymard B, Agid Y. Cortical control of reflexive visually-guided saccades. *Brain*. 1991;114:1473–1485.
10. **Pierrot-Deseilligny C**, Rivaud S, Gaymard B, Müri R, Vermersch AI. Cortical control of saccades. *Ann Neurol Official J Am Neurol Assoc Child Neurol Soc*. 1995;37:557–567.
11. **Rivaud S**, Müri RM, Gaymard B, Vermersch AI, Pierrot-Deseilligny C. Eye movement disorders after frontal eye field lesions in humans. *Exp Brain Res*. 1994;102:110–120.
12. **Ploner CJ**, Rivaud-Péchoux S, Gaymard BM, Agid Y, Pierrot-Deseilligny C. Errors of memory-guided saccades in humans with lesions of the frontal eye field and the dorsolateral prefrontal cortex. *J Neurophysiol*. 1999;82:1086–1090.
13. **Heitger MH**, Anderson TJ, Jones RD. Saccade sequences as markers for cerebral dysfunction following mild closed head injury. *Prog Brain Res*. 2002;140:433–448.
14. **White OB**, Fielding J. Cognition and eye movements: assessment of cerebral dysfunction. *J Neuroophthalmol*. 2012;32:266–273.
15. **King A**. The Proposed King-Devick Test and its Relation to the Pierce Saccade Test and Reading Levels. Chicago, IL: Illinois College of Optometry, 1976.
16. **Oride MK**, Marutani JK, Rouse MW, DeLand PN. Reliability study of the Pierce and King-Devick saccade tests. *Am J Optom Physiol Opt*. 1986;63:419–424.
17. **Stepanek J**, Pradhan GN, Cocco D, Smith BE, Bartlett J, Studer M, Kuhn F, Cevette MJ. Acute hypoxic hypoxia and isocapnic hypoxia effects on oculometric features. *Aviat Space Environ Med*. 2014;85:700–707.
18. **Davies EC**, Henderson S, Balcer LJ, Galetta SL. Residency training: the King-Devick test and sleep deprivation: study in pre- and post-call neurology residents. *Neurology*. 2012;78:e103–e106.
19. **Lin TP**, Adler CH, Hentz JG, Balcer LJ, Galetta SL, Devick S. Slowing of number naming speed by King-Devick test in Parkinson's disease. *Parkinsonism Relat Disord*. 2014;20:226–229.
20. **Ayaz H**, Shewokis PA, Scull L, Libon DJ, Feldman S, Eppig J, Onaral B, Heiman-Patterson T. Assessment of prefrontal cortex activity in amyotrophic lateral sclerosis patients with functional near infrared spectroscopy. *J Neurosci Neuroengineering*. 2014;3:41–51.
21. **Moster S**, Wilson JA, Galetta SL, Balcer LJ. The King-Devick (K-D) test of rapid eye movements: a bedside correlate of disability and quality of life in MS. *J Neurol Sci*. 2014;343:105–109.
22. **Rist B**, Cohen A, Pearce AJ. King-Devick performance following moderate and high exercise intensity bouts. *Int J Exerc Sci*. 2017;10:619–628.
23. **Galetta KM**, Brandes LE, Maki K, Dziemianowicz MS, Laudano E, Allen M, Lawler K, Sennett B, Wiebe D, Devick S. The King-Devick test and sports-related concussion: study of a rapid visual screening tool in a collegiate cohort. *J Neurol Sci*. 2011;309:34–39.
24. **Galetta MS**, Galetta KM, McCrossin J, Wilson JA, Moster S, Galetta SL, Balcer LJ, Dorshimer GW, Master CL. Saccades and memory: baseline associations of the King-Devick and SCAT2 SAC tests in professional ice hockey players. *J Neurol Sci*. 2013;328:28–31.
25. **Galetta KM**, Barrett J, Allen M, Madda F, Delicata D, Tennant AT, Branas CC, Maguire MG, Messner LV, Devick S, Galetta SL, Balcer LJ. The King-Devick test as a determinant of head trauma and concussion in boxers and MMA fighters. *Neurology*. 2011;76:1456–1462.
26. **Faul F**, Erdfelder E, Lang AG, Buchner AG. *Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*. 2007;39:175–191.
27. **Eddy R**, Goetschius J, Hertel J, Resch J. Test-retest reliability and the effects of exercise on the king-devick test. *Clin J Sport Med*. [published ahead of print March 26, 2018] doi: 10.1097/JSM.0000000000000586.
28. **Leong DF**, Balcer LJ, Galetta SL, Evans G, Gimre M, Watt D. The King-Devick test for sideline concussion screening in collegiate football. *J Optom*. 2015;8:131–139.
29. **King D**, Clark T, Gissane C. Use of a rapid visual screening tool for the assessment of concussion in amateur rugby league: a pilot study. *J Neurol Sci*. 2012;320:16–21.
30. **Elkington L**, Manzanero S, Hughes D. Concussion in Sport Position Statement. 2019. Available at: ama.com.au/position-statement/concussion-sport-2016. Accessed March 1, 2019.
31. **Leong DF**, Balcer LJ, Galetta SL, Liu Z, Master CL. The King-Devick test as a concussion screening tool administered by sports parents. *J Sports Med Phys Fitness*. 2014;54:70–77.
32. **Messick S**, Jungeblut A. Time and method in coaching for the SAT. *Psychol Bull*. 1981;89:191.
33. **Sackett PR**, Burris LR, Ryan AM. Coaching and practice effects in personnel selection. In: Robertson CCI, eds. *International Review of Industrial and Organizational Psychology*. New York, NY: Wiley, 1989:145–183.
34. **Campbell DT**, Kenny DA. *A Primer on Regression Artifacts*. New York, NY: Guilford Publications, 1999.

35. **Hausknecht JP**, Halpert JA, Di Paolo NT, Moriarty Gerrard MO. Retesting in selection: a meta-analysis of coaching and practice effects for tests of cognitive ability. *J Appl Psychol.* 2007;92:373–385.
36. **Dumire RD**. Geriatric concussions. In: Rodriguez A, Barraco RD, Ivatury RR, eds. *Geriatric Trauma and Acute Care Surgery*. Cham, Switzerland: Springer International Publishing, 2018:55–67.
37. **Dickson TJ**, Waddington G, Terwiel FA, Elkington L. The King-Devick test is not sensitive to self-reported history of concussion but is affected by English language skill. *J Sci Med Sport.* [published ahead of print August 23, 2018] doi: 10.1016/j.jsams.2018.08.013.
38. **D'Amico N**, Elbin R, Schatz P. A comparison of king-devick test baseline scores between English-speaking and Spanish-speaking high school athletes. *J Atthl Train.* 2017;52:S237.

RESEARCH ARTICLE

Characteristics of the King-Devick test in the assessment of concussed patients in the subacute and later stages after injury

Arsenije Subotic¹, Windsor Kwan-Chun Ting¹, Michael D. Cusimano^{1,2*}

1 Division of Neurosurgery, Department of Surgery, Injury Prevention Research Office, St. Michael's Hospital, Toronto, Ontario, Canada, **2** University of Toronto, Toronto, Ontario, Canada

* injuryprevention@smh.ca



OPEN ACCESS

Citation: Subotic A, Ting WK-C, Cusimano MD (2017) Characteristics of the King-Devick test in the assessment of concussed patients in the subacute and later stages after injury. PLoS ONE 12(8): e0183092. <https://doi.org/10.1371/journal.pone.0183092>

Editor: Damir Janigro, Cleveland Clinic, UNITED STATES

Received: February 17, 2017

Accepted: July 28, 2017

Published: August 31, 2017

Copyright: © 2017 Subotic et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Due to ethical restrictions imposed by the St. Michael's Hospital Research Ethics Committee, the data underlying this study are available to interested, qualified researchers upon request to: Dr. David Mazer, c/o Sharon Freitag, St. Michael's Hospital, 30 Bond Street, Toronto, Ontario, Canada, M5B 1W8.

Funding: We would like to thank the Canadian Institutes of Health Research MOP 123371 for the study funding. The funders had no role in the study

Abstract

Although the King-Devick (K-D) test has been used frequently in assessing sports related concussion early after injury, its characteristics over time after injury and in patients with prolonged persistent symptoms are unknown. The purpose of this paper was to: evaluate the ability of the K-D Test to distinguish patients seen early after concussion from those with symptoms persisting more than 3 months compared to controls, assess changes in the K-D test times over time after concussion, and determine the relationship of K-D times to the Stroop Color and Word Test scores. We performed cross-sectional comparisons of patients with recent concussive brain injury (acute group) and those with symptoms persisting more than 3 months to healthy controls on the K-D test, the Sports Concussion Assessment Tool 3 (SCAT3), and the Stroop Color and Word Test. Longitudinal comparisons of the acute group over time within the first month after injury were also made. Post-concussive syndrome (PCS) patients had significantly higher K-D times compared to controls ($p = 0.01$), while the acute group did not differ from controls ($p = 0.33$). K-D times at the second visit for the acute group were similar to those of controls (54.7 vs. 49.6, $p = 0.31$). While SCAT3 scores improved over time in the acute group, the K-D scores did not change between the first and second visit (55.2 vs. 54.7, $p = 0.94$). K-D scores correlated significantly with the Stroop scores for all three participant groups. The K-D test is likely useful very early after concussion in conjunction with baseline scores, and while scores in PCS patients remain elevated, they can be confounded by factors such as pre-morbid depression and medication use. High correlations with Stroop scores also suggest that performance on the K-D test can by proxy provide additional insight about cognitive function and predict performance on more cognitively demanding tasks.

Introduction

Rising public concern regarding the occurrence of mild traumatic brain injuries (mTBIs) including concussion in sports has led to the development of tools that can help assist in the screening, diagnosis and follow-up of these injuries. The Sport Concussion Assessment Tool 3

design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

(SCAT3), is frequently administered to concussed athletes to assess the number and degree of severity of symptoms pertaining to mTBI [1]. Symptoms on the SCAT3 can be counted and given a rating of severity, but these scores share the same limitations of all self-reported scales: they may be difficult to assess in the presence of preexisting symptoms and are open to potential bias as athletes may under or over-report their symptoms.

Due to such limitations, other tests have been investigated that can help aid clinicians in the screening and diagnosis of those suspected of brain injury. There is increasing interest in the King-Devick test (K-D) as a screening tool for concussion and mTBI. The K-D test requires that participants read a series of three test cards of numbers, which become progressively more difficult to read, as quickly as they can [2]. The total time to complete all three test cards, and the number of errors committed are included in the total score. Studies have shown the K-D test to be a sensitive marker of brain injury by detecting attentional deficits and impaired saccadic eye movements, which have been associated with higher (worse) K-D scores among athletes [3–6]. The K-D test has become a popular sideline screening tool for concussions in sports, since it is easy to administer and usually takes less than two minutes to complete. Several studies have used it in hockey league cohorts [7–9], boxing and mixed martial arts (MMA) competitions [10,11], football games [6, 12–15], as well as rugby league competitions [5, 15–18]. Despite its extensive use immediately following a suspected TBI, the efficacy of the K-D test in tracking symptom resolution longitudinally has been less widely investigated. Tjarks et al. [19] conducted a longitudinal study on concussed patients presenting at a sports clinic, but did not focus on patients affected by non-sports related injuries or those experiencing post-concussive syndrome (PCS) for more than 3 months. Most of the first visits after concussion occurred in the 6–10 day interval (40%) and after 30 days (26%), with 6% of the patients coming in the 1–5 day interval [19]. They found that K-D scores improved at each visit over the four-visit study period [19]. In addition, only Silverberg [20] and Benedict et al. [21] have examined the validity of the K-D test in assessing non-sports related TBI, but neither examined whether the K-D test could be used as an indicator of patient recovery over time after the injury. To our knowledge, previous studies have also not explored the option of correlating performance on the K-D test with tests assessing executive functions such as inhibitory control to see if the K-D test can by-proxy provide more understanding into cognitive function after head injuries.

The purpose of this study was three-fold. First, we wanted to determine if the K-D test could accurately distinguish between non-injured healthy controls, acutely injured patients, and those with persistent post-concussive syndrome with symptoms lasting more than 3 months (PCS). We hypothesized that higher (worse) K-D test scores would be seen in the acutely injured group and PCS patients compared to healthy controls. Second, we aimed to determine if K-D test scores changed over time from the original injury among the acute group patients and if it correlated to symptom resolution over time. We hypothesized that K-D scores would correlate with self-report symptom scores and show improvement over time as symptoms resolved. Our third aim was to correlate and compare K-D scores with assessments of selective attention and processing speed, namely the Stroop Color and Word Test. Our hypothesis was that scores of the Stroop assessment would negatively correlate with K-D scores.

Methods

The study was approved by the Research Ethics Board (REB) at St. Michael's Hospital (SMH). Patients were recruited from the Emergency Department, Head Injury Clinic, and the inpatient Neurosurgery Ward. Healthy controls were recruited through word of mouth to family

members or relatives of participants, as well as staff at SMH in the Emergency Department. Capacity for consent was determined by following the PATIENT Modified Aid to Capacity Evaluation (ACE) Screening Tool utilized by SMH. Participants must have been able to: communicate, understand their current medical condition, understand the purpose of the research study, understand the option of declining to participate (with no impact on medical care), understand the risks of participating, and make a decision that is not substantially based upon hallucinations, delusions, or cognitive signs of depression. If any of these criteria were not met, we did not proceed to consent. After this, written informed consent was obtained, after which participants were assessed for eligibility according to the inclusion/exclusion criteria, outlined below.

Inclusion/Exclusion criteria

Acute group patients were defined as patients who had suffered a non-penetrating head injury and exhibited a GCS score of 13–15 at the time of recruitment. PCS patients were defined as those who had sustained a mTBI three or more months prior to their first testing visit, and were still experiencing ongoing symptoms. Inclusion criteria included: being the age of 16 or over, being able to provide informed written consent, and having sufficient fluency in English. Healthy controls were matched to mTBI patients according to age (± 2.5 years), sex, and years of education (± 2.5 years). Controls participants must also not have had a history of prior head injuries. All participants were also asked to complete a screening form to ensure their eligibility. Participants were excluded if they had the following medical conditions: history of multiple sclerosis, prior hydrocephalus, prior brain irradiation, prior stroke, comorbid early dementia, comorbid Parkinson's disease, comorbid uncontrolled diabetes, comorbid eye disease causing strabismus, comorbid non-affective psychiatric illness, active substance abuse requiring treatment, comorbid alcohol related dementia, and comorbid immune-compromised (HIV/AIDS or taking immunosuppressive therapy). In addition, participants were excluded if they were unable to provide consent because of being medically unstable or intoxicated at the time of recruitment. At the time of screening we also asked for current medication use.

Participant assessments

PCS participants and controls were asked to come in for one visit, and those in the acute injury group were asked to come in for two visits. PCS and healthy controls were asked to come in at their earliest possible convenience. The first visit for acute participants was conducted within 10 days of the injury, while the second was conducted at 2–4 weeks post-injury (PI). The minimum time between the first and second visit was seven days. At each visit, participants completed the K-D test, the SCAT3 Symptom Evaluation, and the Stroop Color and Word Test. The descriptions of these tests are outlined below.

King-Devick (K-D) test

The K-D test is a saccadic eye measurement test that relies on the principle of rapid number naming [1–15]. The test contains a demonstration (practice) card and three test cards of variably spaced single-digit numbers [1–3, 5]. Participants are asked to read out aloud the numbers from left to right as quick as they can without making any errors [1–15]. The time taken for each card as well as the number of errors was recorded and summed to provide the total K-D score [1–15]. The test usually takes less than two minutes to complete [3,5].

SCAT3 symptom evaluation

The symptom evaluation is composed of 22 different symptoms, each of which are rated on a scale of 0–6, with 0 indicating absence of symptoms and 6 being most severe [17, 21]. The total score is out of 22, with symptoms being counted towards the total score if they are non-zero values [17, 21]. The symptom severity score is obtained by summing the values of the individual symptom scores, resulting in a maximum score of 132[17, 21].

Stroop Color and word test

The Stroop Color and Word Test is a neuropsychological test used to measure executive control and selective attention [22–25]. Participants are asked to name the colour of the words presented as fast as they can [22–25]. This becomes more difficult to do if the colour is incongruent with the word [23–25]. It is easier to name the word ‘red’ if it is printed in red than if it is green. In the latter situation, this conflict slows responding which leads to the ‘Stroop effect’ [23–25]. The scores on the task reflect how well participants can selectively direct attention to task relevant features while ignoring task-irrelevant features [23–25]. In our study, participants were asked to read the colour of words (blue, red, green) on a sheet of paper as fast as they can. There were 10 rows, each containing 10 words, resulting in a total of 100 words. The time limit was 45 seconds. If participants finished early, they were asked to read again from the beginning. The total number of words and number of errors are included in the final score.

Statistical analyses

Descriptive statistics were used to describe the participants. Differences in K-D Test, SCAT3, and the Stroop Color and Word Test scores between groups and time intervals were compared using one-way ANOVA. Post-hoc analysis using Fisher’s Least Significant Difference was performed after one-way ANOVA. Paired t-tests were used to do compare scores between the first and second visit for the acute group. Pearson correlation coefficients were used to calculate the correlation between the test scores. All analysis was conducted using Stata 13. Statistical significance was set at $\alpha = 0.05$.

Results

Participant characteristics

In total, 17 acute, 28 PCS and 18 controls participants completed the K-D test. The demographic data pertaining to each participant group are outlined in [Table 1](#).

Acute participants were assessed at a median of 5 days (Interquartile Range, IQR = 3–7) PI for their first visit. The second visit occurred at a median of 24 (IQR = 16–30) days PI. The single visit for the PCS patients occurred at a median of 366 days (IQR = 164–659) after injury. It was reported that 13/28 (44%) of PCS participants were using medication to relieve headache, anxiety and/or depression due to their injury, compared to 3/17 (18%) of acute participants ([Table 1](#)). Of these 12 participants in the PCS group, 3 were using benzodiazepines, 6 were using antidepressants, and 5 were using medications (ibuprofen, acetyl salicylic acid, acetaminophen) to relieve headaches.

K-D and SCAT3 score characteristics

For the first visit, K-D mean scores were 55.2 ± 12.5 s for acute group, 64.1 ± 23.4 s for PCS, and 49.6 ± 11.2 s for healthy controls ([Table 2](#)). PCS participants took significantly longer than controls($p = 0.01$), but were not significantly different from the acute group ($p = 0.13$). K-D scores were not different from the first to second visit in the acute group (55.2 ± 12.5 s versus

Table 1. Basic demographic characteristics of acute mTBI, PCS, and control participant groups.

		Acute mTBI	PCS	Controls
N				
	First Visit	17	28	18
	Second Visit	17	-	-
Age	Mean ± SD	42.7 ± 14.47	46.1 ± 12.8	36.9 ± 16.6
Sex, (%)				
	Male	47%	18%	44%
	Female	53%	82%	56%
Education	Mean ± SD	14.9 ± 3.2	16.4 ± 3.2	15.4 ± 1.8
Mechanism of injury (%)				
	Fall	47	36	-
	Sports	12	4	-
	Unintentional contact with an object	29	18	-
	MVC	6	38	-
	Intentional Assault	6	4	-
Medication use, N (% of total population)				
	Benzodiazepines	0 (0)	3 (11)	0 (0)
	Antidepressants	2 (12)	6 (21)	2 (11)
	Headache Relief	1 (6)	5 (18)	1 (6)
Initial Glasgow Coma Scale (GCS) in ED, N				
	13	1	0	0
	14	0	0	0
	15	16	28	18
History of Concussion, (%)		47	50	-

Abbreviations: SD, standard deviation; MVC, motor vehicle collision; ED, emergency department

<https://doi.org/10.1371/journal.pone.0183092.t001>

54.7 ± 15.1s, p = 0.94), and nor were they different from controls (p = 0.31). The total symptom score and symptom severity scores were higher for the acute group and PCS group compared to controls on the first visit (p<0.0001). For the acute group, symptom scores and severity

Table 2. Summary of test scores by group and by visit.

		Acute mTBI	PCS	Controls
K-D Times, Mean ± SD (s)				
	First Visit	55.2 ± 12.5	64.1 ± 23.4*	49.6 ± 11.2
	Second Visit	54.7 ± 15.1	-	-
Symptom Total Score, Mean ± SD				
	First Visit	13.2 ± 6.4	15.2 ± 6.06	1.4 ± 2.5
	Second Visit	7.5 ± 5.9	-	-
Symptom Severity Score, Mean ± SD				
	First Visit	32.6 ± 19.2	44.4 ± 31.3	2.3 ± 4.2
	Second Visit	13.4 ± 12.3	-	-
Stroop Color and Word Test, Mean number of words ± SD				
	First Visit	69.4 ± 15.9	72.3 ± 11.6	85.1 ± 12.4
	Second Visit	70.5 ± 14.1	-	-

*Post-hoc analysis. F-value = 4.66, P = 0.01.

<https://doi.org/10.1371/journal.pone.0183092.t002>

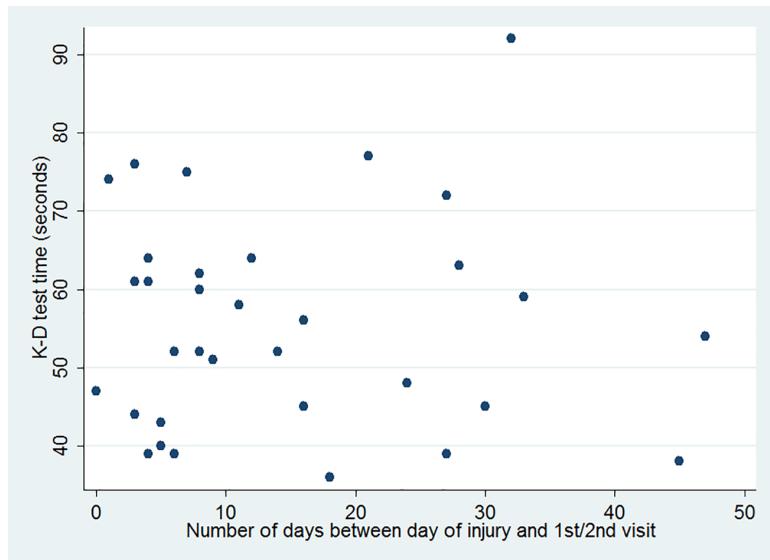


Fig 1. Relationship between K-D scores and number of days between injury and test visit in the acute group.

<https://doi.org/10.1371/journal.pone.0183092.g001>

scores improved significantly from the first visit to the second (13.2 ± 6.4 vs. 7.5 ± 5.9 , $p = 0.01$) (32.6 ± 19.2 vs. 13.4 ± 12.3 , $p = 0.002$) (Table 2).

We also performed an analysis of K-D values over time to show the pattern of scores in the acute group. A scatterplot (Fig 1) showing K-D scores and the time between date of injury and test visits showed no significant correlation or relationship. In addition, we stratified the time between the date of injury and the first/second visit into four time intervals (Fig 2) to more precisely characterize K-D performance as a function of time. Again, there were no significant differences between each time interval in terms of K-D scores.

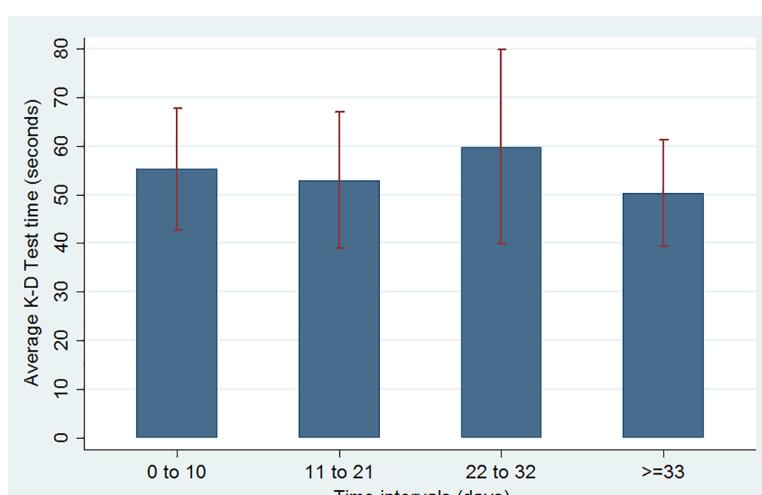


Fig 2. Average K-D score for each time interval after injury in the acute group. Abbreviations: SD, standard deviation. Symbols: \geq , greater than or equal to.

<https://doi.org/10.1371/journal.pone.0183092.g002>

Table 3. Correlations of K-D test scores with other assessment scores.

		Symptom Total Score	Symptom Severity Score	Stroop Color and Word Test
Acute: 1st Visit	r	-0.53	-0.43	-0.89
	p	0.03	0.08	<0.0001
Acute: 2nd Visit	r	0.05	0.15	-0.78
	p	0.86	0.58	0.0002
PCS	r	0.06	0.07	-0.75
	p	0.76	0.71	<0.0001
Healthy Control	r	0.30	0.22	-0.60
	p	0.23	0.39	0.001

<https://doi.org/10.1371/journal.pone.0183092.t003>

K-D and Stroop Color and word test characteristics

For the Stroop Color and Word Test, the acute group at the first visit got fewer words correct than the controls (69.4 ± 15.9 versus 85.1 ± 12.4 , $p < 0.0001$, Table 2). Similarly, at the second visit, the acute group differed significantly from controls ($p = 0.003$), but did not exhibit an improvement from the first visit ($p = 0.82$). The PCS group also scored significantly lower than controls ($p = 0.002$). A strong, inverse correlation was found for all three groups (Table 3).

Discussion

Major findings

We found that PCS patients performed significantly worse than controls and moderately worse than our acute group of patients on the K-D test. A number of potential explanations exist for this. Our PCS patients sustained their injuries more frequently in motor vehicle crashes (MVC) than the acute group and since MVC injuries are higher velocity events than those in sports, they are likely also linked with increased concussion severity and the prolongation of symptom burden [26–31]. Those suffering from PCS may have experienced greater damage to cortical and subcortical structures in the prefrontal cortex, thus resulting in ongoing cerebral impairment and slowing of saccades beyond the usual 1-3-month period in which most decrements in cognitive function resolve [26, 32]. In addition, the PCS group was more medicated for conditions like depression or anxiety which can contribute to prolongation of K-D times and post-concussive symptoms. [30, 32]. Reilly [33] found that the usage of benzodiazepines and other sedatives used to relieve anxiety have been associated with a decrease in saccadic acceleration and velocity and an increase in saccadic latency, or the time interval between two consecutive saccades, which would slow down saccadic eye movements [33]. However, it was also reported that the usage of antidepressants had no significant effect on saccadic eye movements [33]. So, it is difficult to be certain whether tests of saccadic eye movements can be used as sensitive and objective markers of brain dysfunction following injury, particularly in patients on confounding medications for depression or anxiety [33]. As we only conducted one test session for PCS participants, we were not able to determine at what point in time K-D scores worsened. A study by Heitger [26] showed saccadic eye movements among PCS patients to be significantly worse at 140 days PI compared with those with good recovery. Rizzo [32] also examined performance on the K-D test and found significant differences

between PCS and healthy controls at a median of 54 weeks PI. It is plausible that differences scores on the K-D test might already arise by 140 days PI and remain elevated past the 1 year mark. To our knowledge, no studies have so far administered the K-D test over multiple test visits in a span of multiple months to over a year PI. Future studies should focus on conducting more frequent visits over time, to better elucidate the pattern of performance on saccadic eye movement tests such as the K-D test among those suffering from chronic concussion.

In terms of symptom scores, we found higher symptom and symptom severity scores among the PCS group compared to both controls and the acute mTBI, suggesting prolongation of symptom burden in this group. A strong possibility is that symptoms related to PCS might not stem from concussion per se but are also influenced by psychological, personality, and psychosocial factors. Individuals react differently to injuries, and it is plausible that those suffering from PCS may develop a shaken sense of identity as concussion recovery takes longer than usual. This may lead to cognitive issues by suppressing attention, mental efficiency, learning and memory, therefore creating symptoms that are unrelated to those caused by the concussion itself [34]. This can in turn lead to frustration and anxiety, leading to avoidance of anxiety provoking situations leading to a buildup of depression that builds over time, resulting in heightened symptom scores [34]. There is also the possibility that symptoms were not specific to PCS, but rather a reaction to the trauma experienced by the head injury [35]. To better elucidate the precise effects of concussion in the development of symptoms in PCS, a control group consisting of patients suffering from non-head injuries would be useful to include in future analyses. In addition, as we also only conducted one study visit, we also cannot say conclusively what the pattern of symptom scores over time in those with PCS are. There has not been sufficient literature conducted so far to provide enough evidence to characterize when elevated symptom scores compared with those with acute mTBI arise and when they subside. Hou [36] did a study in which they found that there was no significant recovery in those with PCS from the 3 to 6-month period PI. McMahon et al. [37] also examined a cohort of individuals with PCS and found that the number of symptoms increased from the 3 to 6 months and from 3 to 12 months PI. Our findings of heightened and persisting symptom scores at a median of 366 days after injury are therefore in line with some of the literature examining PCS, but more studies, examining more frequent visits and over a longer period of time are needed to better characterize symptom scores.

Interestingly, we were not able to show a strong positive correlation between K-D scores and symptom scores. Even more striking is the fact that a statistically significant negative correlation was observed on the first test visit for acute participants. This suggests that the reporting of subjective symptoms does not predict performance on visual based testing scores, and that performance on the K-D test does not predict who experiences greater number of symptoms following injury. It is plausible that the pathways responsible for generating saccades in the brain are independent of those that contribute to behavioral, emotional, and cognitive states, and that both of these cognitive domains contribute to the screening for signs of concussion, albeit separately.

Although we saw total symptom and severity scores improve between the first and second assessments in the acute group, we could not show an improvement in K-D times over the two visits. Previous studies have shown that the K-D test done on the sidelines in sporting contexts could detect abnormalities among affected athletes [1, 2–15]. These abnormalities included factors such as blurred vision or attentional deficits, and the K-D scores reflected this through an increase in scores from pre-injury baseline scores. [1, 2–15]. It could thus be that our acute patients had already recovered rapid eye movements by the time we performed the K-D tests (median 5 days post injury) and so they did not change by the second visit (median 24 days post injury). Neuropsychological tests done after the acute recovery period have found that

patients with head injuries have no noticeable test differences from those of matched controls [38, 39]. Silverberg [20] found that the K-D test was not sensitive in differentiating acute mTBI patients and controls at a mean time of 31 hours after injury supporting our thesis that K-D times recover quickly after injury. Our findings therefore do not support the use of the K-D test as a screening test for those with acute mTBI without baseline scores.

Our results also showed the novel and robust finding that K-D times were significantly negatively correlated with the number of words on the Stroop Color and Word Test among all three participant groups. This meant that injured and non-injured participants that completed the K-D test faster got more words correct in 45 seconds on the Stroop Color and Word Test. To date, the K-D has not been studied for its correlation with assessments like the Stroop Color and Word Test. These findings support the idea that the K-D test can also indirectly indicate the level of interference control and selective attention measured by the Stroop Color and Word Test, providing evidence of convergent validity for the K-D test as it also assesses mental abilities such as selective attention. Future studies should consider utilizing these two tests together as the Stroop Color and Word Task can capture additional deficits in attention due to the fact that more strenuous cognitive processes, such as interference control, are often helpful in uncovering signs of concussion in what might be a seemingly non-concussed individual [22–24].

Limitations and future directions

The interpretation of our study requires a consideration of its limitations. We did not control for the time during the day or the degree of sleep the participant had prior to testing, as they came at their earliest possible convenience. These factors could have produced variations in K-D scores or subjective self-report symptom questionnaires, as previous studies by Fransson [40] and Davies [41] have shown that sleep deprivation can adversely affect eye tracking and attention that is required by the K-D test. Our study design was such that participants could not be assessed immediately following their head injury, as some were admitted into the hospital a couple of hours up until 2 days after their injury. As this was a voluntary study, many patients wanted to be assessed later as they wanted to leave the emergency room of the hospital as soon as their diagnosis was confirmed by a physician. Our mechanisms of injury were varied but when we stratified K-D times according to the mechanism of injury, we did not find any significant differences. Because of the confounding effects of medications, future studies, particularly with PCS patients should control for such medications. Future studies with larger sample sizes may wish to focus on performing an in-depth analysis on the effect of injury mechanism and time since injury, while controlling for a number of confounding factors on K-D times.

Our PCS group, which was not designed to match non-injured controls, was also biased towards females. Our study is in line with previous literature which has reported that those suffering from PCS were more likely to be female, and that there are gender disparities in seeking treatment, which may be reflected in our study [30,34–35]. However, there has not been evidence suggesting that performance on visit based tests is influenced by gender, and that females with PCS score worse than males with PCS. Benedict et al. [21] examined the influence of gender and found it to be associated with increased symptom scores, but not performance on the K-D test. Therefore, although there is bias in gender in those who develop and report PCS, it is unlikely that saccadic performance is correlated with it. Due to our small sample size and difference in number of participants between the PCS and other groups, we are not in the position to draw definitive conclusions about gender and performance on the K-D test. Future studies, with larger sample sizes and with more balanced gender distributions, should explore

the option of performing within-group and between-group analyses to better elucidate the effects that gender may play.

Although factors such as motivation can play a role in neuropsychological assessments, we did not consider it necessary to have an effort test to assess motivation, as participants who consented were concerned and eager to participate and contribute to the improvement in treatments for concussion. It is difficult to quantify effort and motivation, and it is an inherent characteristic that varies from individual to individual. To our knowledge, no studies have so far been able to elucidate the effects of effort on performance on visual based assessment tests, nor has a reliable tool been developed for this purpose. Nonetheless, future research should aim to develop an effort test to more accurately control for this potentially confounding factor. Additionally, it is important to note that vision plays only one part in the examination of concussions. The K-D test, as a vision based test, should in appropriate cases, be used as a screening tool and as a compliment to a wider number of assessments in the treatment of concussion. We did not include physical assessments in our study, and future studies should consider using assessments of balance and other tests such as the vestibulo-ocular reflex in the examination of concussions as well in order to more adequately help clinicians screen and diagnose concussion [42]. Larger sample sizes will also allow us to capture acute patients who will develop PCS, to test hypotheses of the sensitivity and specificity of the K-D test in patients with concussion and mTBI over time.

Conclusions

We found that K-D times inversely correlate highly with the number of correct Stroop words regardless of whether participants had sustained an injury or not suggesting that the K-D test can by-proxy provide insightful information about more complex cerebral functions associated with selective attention, such as response inhibition. We were not able to show any differences in K-D times between acutely injured patients at a median of 5 days post injury and neither did we find times to improve over the subsequent two weeks. In contrast, we showed that PCS patients were slower on the K-D task than controls or acutely injured patients but they were also on multiple medications that might have confounded our results. Our findings do not refute the value of the K-D as a screening test very early after injury. Its use in PCS patients is confounded by factors such as depression frequently seen in patients with PCS. Further large scale studies of the test will better delineate its characteristics acutely and longitudinally.

Acknowledgments

The authors would like to thank the staff and students at the Injury Prevention Research Office for helping recruit and test participants as well as assisting with data entry. Special thanks to Dr. Rowan Jing and Ling Chen for helping conduct the statistical data analysis. We would also like to thank Dr. Francois Mathieu for providing insightful feedback on the manuscript.

Author Contributions

Conceptualization: Arsenije Subotic, Windsor Kwan-Chun Ting, Michael D. Cusimano.

Data curation: Arsenije Subotic, Windsor Kwan-Chun Ting, Michael D. Cusimano.

Formal analysis: Arsenije Subotic, Windsor Kwan-Chun Ting.

Funding acquisition: Michael D. Cusimano.

Investigation: Arsenije Subotic, Windsor Kwan-Chun Ting, Michael D. Cusimano.

Methodology: Arsenije Subotic, Windsor Kwan-Chun Ting, Michael D. Cusimano.

Project administration: Windsor Kwan-Chun Ting, Michael D. Cusimano.

Resources: Michael D. Cusimano.

Supervision: Windsor Kwan-Chun Ting, Michael D. Cusimano.

Validation: Michael D. Cusimano.

Visualization: Arsenije Subotic, Michael D. Cusimano.

Writing – original draft: Arsenije Subotic, Windsor Kwan-Chun Ting.

Writing – review & editing: Arsenije Subotic, Michael D. Cusimano.

References

1. Galetta KM, Morganroth J, Moehringer N, Mueller B, Hasanaj L, Webb NATC, et al. Adding Vision to Concussion Testing: A Prospective Study of Sideline Testing in Youth and Collegiate Athletes. *Journal of Neuro-Ophthalmology* 2015 September; 35(3):235–241. <https://doi.org/10.1097/WNO.0000000000000226> PMID: 25742059
2. Oride M, Matrutani J, Rouse M, DeLand P. Reliability study of the Pierce and King–Devick saccade tests. *Am J Optom Physiol Opt* 1986; 63(6):419–24. PMID: 3728637
3. Galetta KM, Brandes LE, Maki K, Dziemianowicz MS, Laudano E, Allen M, et al. The King Devick test and sports-related concussion: Study of a rapid visual screening tool in a collegiate cohort. *J Neurol Sci* 2011; 309(1–2):34–39. <https://doi.org/10.1016/j.jns.2011.07.039> PMID: 21849171
4. Galetta K, Barrett J, Allen M, Madda F, Delicata D, Tenant A, et al. The King–Devick test as a determinant of head trauma and concussion in boxers and MMA fighters. *Neurology* 2011; 76(17):1456–62. <https://doi.org/10.1212/WNL.0b013e31821184c9> PMID: 21288984
5. King D, Clark T, Gissane C. Use of a rapid visual screening tool for the assessment of concussion in amateur rugby league: A pilot study. *J Neurol Sci* 2012; 320(1–2):16–21. <https://doi.org/10.1016/j.jns.2012.05.049> PMID: 22694977
6. Leong DF, Balcer LJ, Galetta SL, Evans G, Gimre M, Watt D. The King–Devick test for sideline concussion screening in collegiate football. *Journal of Optometry* 2015; 8(2):131–139. <https://doi.org/10.1016/j.joptom.2014.12.005> PMID: 25649742
7. Galetta MS, Galetta KM, McCrossin J, Wilson JA, Moster S, Galetta SL, et al. Saccades and memory: Baseline associations of the King–Devick and SCAT2 SAC tests in professional ice hockey players. *J Neurol Sci* 2013; 328(1–2):28–31.
8. Galetta KM, Morganroth J, Moehringer N, Mueller B, Hasanaj L, Webb NATC, et al. Adding Vision to Concussion Testing: A Prospective Study of Sideline Testing in Youth and Collegiate Athletes. *Journal of Neuro-Ophthalmology* 2015 September; 35(3):235–241. <https://doi.org/10.1097/WNO.0000000000000226> PMID: 25742059
9. Verna BT, Grady MF, Goodman A, Wiebe DJ, Basta L, Park Y, et al. Oculomotor and Neurocognitive Assessment of Youth Ice Hockey Players: Baseline Associations and Observations After Concussion. *Dev Neuropsychol* 2015 January 2015; 40(1):7–11.
10. Galetta KM, Barrett J, Allen M, Madda F. The King-Devick test as a determinant of head trauma and concussion in boxers and MMA fighters. *Neurology* 2011 -04-26; 76(17):1456–1462. <https://doi.org/10.1212/WNL.0b013e31821184c9> PMID: 21288984
11. Leong DF, Balcer LJ, Galetta SL, Liu Z, Master CL. The King-Devick test as a concussion screening tool administered by sports parents. *J Sports Med Phys Fitness* 2014 Feb; 54(1):70–77. PMID: 24445547
12. Duenas M, Jandial R, Whyte G. Sideline concussion testing in high school football on Guam. *Surgical Neurology International* 2014; 5(1):91.
13. Marinides Z, Galetta KM, Andrews CN, Wilson JA. Vision testing is additive to the sideline assessment of sports-related concussion. *Neurology: Clinical Practice* 2015 -02-01; 5(1):25–34.
14. Munce TA, Dorman JC, Odney TO, Thompson PA, Valentine VD, Bergeron MF. Effects of youth football on selected clinical measures of neurologic function: a pilot study. *J. Child Neurol.* 2014; 29(12): 1601–1607. <https://doi.org/10.1177/0883073813509887> PMID: 24272520

15. Seidman DH, Burlingame J, Yousif LR, Donahue XP, Krier J, Rayes LJ, et al. Evaluation of the King-Devick test as a concussion screening tool in high school football players. *J Neurol Sci* 2015; 356(1–2):97–101. <https://doi.org/10.1016/j.jns.2015.06.021> PMID: 26094155
16. King D, Brughelli M, Hume P, Gissane C. Concussions in amateur rugby union identified with the use of a rapid visual screening tool. *J Neurol Sci* 2013 20130315; 326(1–2):59–63. <https://doi.org/10.1016/j.jns.2013.01.012> PMID: 23374885
17. King D, Gissane C, Hume PA, Flaws M. The King-Devick test was useful in management of concussion in amateur rugby union and rugby league in New Zealand. *J. Neurol.Sci.* 2015; 351(1–2), 58–64. <https://doi.org/10.1016/j.jns.2015.02.035> PMID: 25748294
18. King D, Brughelli M, Hume P, Gissane C. Concussions in amateur rugby union identified with the use of a rapid visual screening tool. *J Neurol Sci* 2013 Mar 15; 326(1–2):59–63. <https://doi.org/10.1016/j.jns.2013.01.012> PMID: 23374885
19. Tjarks BJ, Dorman JC, Valentine VD, Munce TA, Thompson PA, Kindt SL, et al. Comparison and utility of King-Devick and ImPACT® composite scores in adolescent concussion patients. *J.Neurol.Sci.* 2013; 334(1–2):148–153 <https://doi.org/10.1016/j.jns.2013.08.015> PMID: 24007870
20. Silverberg ND, Luoto TM, Ohman J, Iverson GL. Assessment of mild traumatic brain injury with the King-Devick Test in an emergency department sample. *Brain Inj* 2014; 28(12):1590–1593 <https://doi.org/10.3109/02699052.2014.943287> PMID: 25093537
21. Benedict PA, Baner NV, Harrold GK, Moehringer N, Hasanaj L, Serrano LP, et al. Gender and age predict outcomes of cognitive, balance and vision testing in a multidisciplinary concussion center. *J Neurol Sci* 2015; 353(1–2):111–115.
22. Stroop JR. Studies of interference in serial verbal reactions. *J Exp Psychol: Gen* 1992 03; 121(1):15–23.
23. Olk B. Measuring the allocation of attention in the Stroop task: evidence from eye movement patterns. *Psychological Research* 2013; 77(2):106–115. <https://doi.org/10.1007/s00426-011-0405-9> PMID: 22205494
24. MacLeod CM, MacDonald PA. Interdimensional interference in the Stroop effect: uncovering the cognitive and neural anatomy of attention. *Trends Cogn.Sci.(Regul.Ed.)* 2000; 4(10):383–391.
25. Naber M, Vedder A, Brown SBRE, Nieuwenhuis S. Speed and Lateral Inhibition of Stimulus Processing Contribute. *Psychol.* 7:822.
26. Heitger MH, Jones RD, Macleod AD, Snell DL, Frampton CM, Anderson TJ. Impaired eye movements in post-concussion syndrome indicate suboptimal brain function beyond the influence of depression, malingering or intellectual ability. *Brain* 2009 2009; 132(10):2850–2870
27. Hoshizaki B, Post A, Kendall M, Karton C, Brien S. The Relationship between Head Impact Characteristics and Brain Trauma. *J Neurol Neurophysiol* 2013.
28. Post A, Hoshizaki TB, Gilchrist MD, Brien S. Analysis of the influence of independent variables used for reconstruction of a traumatic brain injury event. *Journal of Sports Engineering and Technology* 2012; 226: 290–298.
29. Wright RM, Post A, Hoshizaki B, Ramesh KT. A multiscale computational approach to estimating axonal damage under inertial loading of the head. *J Neurotrauma* 2013; 30: 102–118.
30. Ponsford J, Willmott C, Rothwell A, Cameron P, Kelly A, Nelms R, et al. Factors influencing outcome following mild traumatic brain injury in adults. *J.Inter.Neuropsych.Soc.* 2000; 6(5):568–579
31. HESSEN E, NESTVOLD K, SUNDET K. Neuropsychological function in a group of patients 25 years after sustaining minor head injuries as children and adolescents. *Scand.J.Psychol.* 2006; 47(4):245–251. <https://doi.org/10.1111/j.1467-9450.2006.00514.x> PMID: 16869857
32. Rizzo J, Hudson TE, Dai W, Birkemeier J. Rapid number naming in chronic concussion: eye movements in the King-Devick test. *Annals of Clinical and Translational Neurology* 2016 -10-01; 3(10):801–811. <https://doi.org/10.1002/acn3.345> PMID: 27752515
33. Reilly JL, Lencer R, Bishop JR, Keedy S, Sweeney JA. Pharmacological treatment effects on eye movement control. *Brain Cogn* 2008 200812; 68(3):415–435. <https://doi.org/10.1016/j.bandc.2008.08.026> PMID: 19028266
34. Broshek DK, De Marco AP, Freeman JR. A review of post-concussion syndrome and psychological factors associated with concussion. *Brain Injury* 2015 01/28; 29(2):228–237.
35. Cassidy JD, Cancelliere C, Carroll LJ, Côté P, Hincapié CA, Holm LW, et al. Systematic Review of Self-Reported Prognosis in Adults After Mild Traumatic Brain Injury: Results of the International Collaboration on Mild Traumatic Brain Injury Prognosis. *Arch Phys Med Rehabil* 2014; 95(Supplement):S132–S151.
36. Hou R, Moss-Morris R, Peveler R, Karin M, Bradley BP, Antonio B. When a minor head injury results in enduring symptoms: a prospective investigation of risk factors for postconcussional syndrome after

- mild traumatic brain injury J Neurol Neurosurg Psychiatry 2012; 83:217–223. <https://doi.org/10.1136/jnnp-2011-300767> PMID: 22028384
- 37. McMahon P, Hricik A, Yue JK, Puccio AM, Inoue T, Lingsma HF, et al. Symptomology and Functional Outcome in Mild Traumatic Brain Injury: Results from the Prospective TRACK-TBI Study. *J Neurotrauma* 2014; 31(1): 26–33.
 - 38. Iverson GL. Outcome from mild traumatic brain injury. *Curr Opin Psychiatry* 2005; 18: 301–17. <https://doi.org/10.1097/01.yco.0000165601.29047.ae> PMID: 16639155
 - 39. Schretlen DJ, Shapiro AM. A quantitative review of the effects of traumatic brain injury on cognitive functioning. *International Review of Psychiatry* 2003; 15(4):341–349. <https://doi.org/10.1080/09540260310001606728> PMID: 15276955
 - 40. Fransson PA, Patel M, Magnusson M, Berg S, Almbladh P, Gomez S. Effects of 24-hour and 36-hour sleep deprivation on smooth pursuit and saccadic eye movements. *J Vestib Res* 2008; 18(4):209–222. PMID: 19208965
 - 41. Davies EC, Henderson S, Balcer LJ, Galetta SL. Residency training: the King-Devick test and sleep deprivation: study in pre- and post-call neurology residents. *Neurology* 2012 Apr 24; 78(17):e103–6. <https://doi.org/10.1212/WNL.0b013e318251833d> PMID: 22529208
 - 42. Matuszak JM, McVige J, McPherson J, Willer B, Leddy J. A Practical Concussion Physical Examination Toolbox. *Sports Health: A Multidisciplinary Approach* 2016; 8(3):260–269.



In situ use of the King-Devick Eye Tracking test and changes seen with sport-related concussion: saccadic and blinks counts

Mark Hecimovich, Doug King, Alasdair Dempsey, Mason Gittins & Myles Murphy

To cite this article: Mark Hecimovich, Doug King, Alasdair Dempsey, Mason Gittins & Myles Murphy (2018): In situ use of the King-Devick Eye Tracking test and changes seen with sport-related concussion: saccadic and blinks counts, *The Physician and Sportsmedicine*, DOI: [10.1080/00913847.2018.1525261](https://doi.org/10.1080/00913847.2018.1525261)

To link to this article: <https://doi.org/10.1080/00913847.2018.1525261>



Accepted author version posted online: 22 Sep 2018.



Submit your article to this journal



View Crossmark data

CrossMark

Publisher: Taylor & Francis

Journal: *The Physician and Sportsmedicine*

DOI: 10.1080/00913847.2018.1525261

Original Research

In situ use of the King-Devick Eye Tracking test and changes seen with sport-related concussion: saccadic and blinks counts

Hecimovich, Mark

mark.hecimovich@uni.edu

University of Northern Iowa - Athletic Training

Murdoch University - School of Psychology and Exercise Science

King, Doug

Auckland University of Technology - Faculty of Health and Environmental Science

University of New England - Department of Science and Technology

Dempsey, Alasdair

Murdoch University - School of Psychology and Exercise Science

Gittins, Mason

Murdoch University

Murphy, Myles

The University of Notre Dame Australia - School of Physiotherapy

Abstract

Objectives: Sport-related concussion (SRC) can result in impaired oculomotor function. Oculomotor performance, measured utilizing the King-Devick/Eye Tracking test (K-D/ET), is reported to be able to identify sub-optimal brain function. The objectives of the study were to determine the diagnostic accuracy of the K-D/ET in identifying SRC occurring from game participation, and to perform a comparative analysis on saccade and blink counts for each K-D card individually and total counts between baseline and post-concussion.

Methods: Nineteen male Australian Football players were assessed on the K-D/ET test. Those suspected of having SRC secondary to a head impact were also tested.

Results: Participants recorded a slower time on the third (20.2 ± 4.6 s) screen when compared with the first ($p=0.0424$) and second ($p=0.0150$) screens. The number of blinks were higher on the third (2.9 ± 2.9) when compared with the second ($p=0.0057$) screen. There was decrease of the K-D/ET total times between pre- and post-game ($p=0.1769$). Participants who sustained a head impact recorded slower mean total K-D time ($p=0.7322$), fewer mean total saccades ($p=0.0112$) and more mean blinks ($p=0.8678$) compared with their baseline scores. The assessment of blinks was the most sensitive measure for potential SRC (0.67). The K-D/ET duration was the most specific measure for potential SRC (0.88). An increase in the number of blinks had a fair specificity of 0.69.

Conclusion: The rapid number-naming component of the K-D test is an assessment tool which quantifies impairment to oculomotor function and has been validated as a diagnostic tool for SRC. The clinical usefulness of the eye tracking component of the K-D/ET test is that it may be an effective method to assess concussions with the eye tracking component serving as a measure of progression and return to play. However, more research is required at the adult and youth level.

Keywords:

Concussion; head impact; oculomotor; saccades; eye movements; Australian football; children; youth

Accepted Manuscript

INTRODUCTION

A comprehensive approach to the evaluation and treatment of sports-related concussion (SRC) is encouraged, given the heterogeneous nature of the injury [1]. This includes a clinical examination and evaluation, symptom assessment, neurocognitive testing, and examination of vestibular (e.g. balance, vestibular ocular reflex) and oculomotor functioning [2]. The 2017 international consensus statement on SRC suggests that oculomotor screening may add clinical utility in the recognition of SRC highlighting the importance of further research in this area [3]. Eye movement abnormalities have been observed in mild traumatic brain injury [4,5] but they have not been widely, and systematically, studied in concussion.

A rapid number-naming task called the King–Devick (K–D) test has been validated as a sensitive sideline performance measure for concussion detection [6-12]. This test functions as a pseudo-reading task, broadly capturing aspects of afferent visual function, attention, language, visual fixation, and saccadic eye movements. Saccade generation involves weighing aspects of the stimulus, the goal of the eye movement, motor planning and organization [13]. Saccades must be both accurate (due to small fovea size) and fast (up to 600°/s and less than 100 ms. in duration) to efficiently acquire image information in real time. Given that speed and accuracy typically trade off in human movement control, these criteria are particularly demanding. It is therefore not surprising that saccades would be prone to malfunction from neurological trauma more readily than other eye movement types [13]. The K-D Eye Tracking System (K-D/ET) [13] allows for real-time automated eye movement data and output using measures of fixations, saccades, blinks, and pupillary dynamics. This eye tracking technology brings objective tools to the clinical management of oculomotor dysfunctions while allowing quantifiable measurement of impairment severity and status over progression and remediation

Disruption of cortical structures integral to target selection, attention, and saccade programming may result in alterations of saccades. Of particular pertinence are frontal regions that mediate saccades, such as frontal eye fields (FEF), dorsolateral prefrontal cortex (DLPFC), and cingulate eye field, as these regions are very prone to injury from head trauma [14-16]. The DLPFC plays a major role in producing anticipatory saccades when visual target location and timing can be predicted, inhibiting unwanted saccades that would disrupt vision [14-16]. The DLPFC is also involved in spatial working memory during saccades to a remembered target location after the

target has disappeared (memory- guided saccades). Dysfunction of the DLPFC has been speculated to occur in concussion, given correlations between slowed K-D reading times and a decline in general cognitive function as assessed by the Standard Assessment of Concussion (SAC) test [8]. Dysfunction of the cingulate cortex could also play a potential role in saccadic abnormalities during K-D test performance in traumatic brain injury, as the cingulate cortex has been shown by functional imaging to play a role in reading of ‘pseudo-text’ (consonant letter strings) [17] similar to reading a series of numbers on K-D test cards [13].

The standard K-D test consists of four cards (1-demonstration and 3 test cards) that can be read as hand-held cards or on a computer tablet or iPad and consists of a total 40 numbers arranged in 8 rows of five [18]. Studies of the K-D test in athletes [6-12] have reported that testing completion times are notably prolonged (worse) following concussion, but not after exercise or exertion alone [7]. Slowing of saccades that would implicate brainstem dysfunction, is speculated as a possible explanation for prolonged K-D testing times following concussive events. However, this has not yet been proven using quantitative eye movement recordings and other ocular motor deficit possibilities exist. There are several deficit possibilities and these include (1) increased duration of fixations, (2) saccadic latencies, (3) overall numbers of saccades due to backtracking or inaccurate saccades, and (4) excessive saccadic intrusions superimposed upon otherwise normal eye movements — all of which may implicate brain dysfunction [13]. The possibility also exists that ocular motor behavior on the K-D test is normal after concussion and slowed reading times may be due solely to cognitive or attentional deficits [13].

The objectives of the current study were to: (1) determine the diagnostic accuracy of the K-D Eye Tracking System (K-D/ET) test in identifying SRC that occurred from game participation, and (2) to perform a comparative analysis on saccade and blink counts for each card individually and total counts between baseline and post-injury in participants who sustained a meaningful head impact. Overall, the study seeks to examine the potential utility of a self-paced saccadic eye movement for the identification of a concussive event.

MATERIALS AND METHODS

A prospective observational cohort study was conducted on youth Australian football players competing during the 2017 competition. All members of the team were invited to participate in

the study. A total of 19 healthy male youth Australian football players (age range 13-14, mean 13.9 ± 0.3) were enrolled in the study. The participant total (19) represented a majority of players on the team and was limited to this amount due to equipment and staff (research assistant) availability. Consent was obtained from the players, parents and participating team before enrolling in the study. Exclusion criteria included a history of concussion or traumatic brain injury and a history of neurological or ophthalmological disease (other than refractive error). None of the participants had a history of concussion, neurological impairment, learning disability, or visual dysfunction. None of the participants were taking central nervous system-active medications. The researchers' University ethics committee approved all procedures (MUHREC 2016/012).

Baseline Measures

All participants (n=19) completed a baseline pre-season K-D/ET assessment. Baseline pre-season K-D/ET measures were obtained across two training session prior to the first game. Testing occurred in a darkened room out of any direct sunlight. All players were seated upright in a standard fold-up chair with their eyes level to the superior border of the computer screen from a distance of 0.6m. Players were briefed on the testing protocol using standardized instructions, followed by the completion of the familiarization screen. When ready, players signaled the tester to begin the test. The players read the numbers out aloud and all errors made were recorded. Errors in the baseline test required a second baseline test to be completed. A single assessor (MG), who was instructed on the use of the K-D/ET equipment, conducted all participant K-D/ET assessments.

The K-D test required the player to read single digit numbers aloud from left to right, from top to bottom taking approximately two minutes to complete. The reliability of the K-D test has been reported to have an inter-class correlation for test-retest reliability of 0.97 (95% CI: 0.9 to 1.0) [7] and 0.93 (95% CI: 0.68 to 0.97) and 0.92 (95% CI: 0.52 to 0.96) between concussed and non-concussed players, respectively [12].

The K-D/ET system allows for real-time automated eye movement data and output using measures of fixations, saccades, blinks and pupillary dynamics. This eye tracking technology brings objective tools to the clinical management of oculomotor dysfunctions while allowing quantifiable measurement of impairment severity and status over progression and remediation.

All participants completed one trial of a novel digitized version of the K-D/ET test under objective, video-based, infrared oculography (EyeTech VT3 Mini, EyeTech Digital Systems, Mesa, AZ, USA). The K-D card images were imported into the eye tracking software with maintenance of stimulus matching (e.g. numbers presented, spacing between numbers) [13].

Post-game Measures

For post-game measures, participants were selected to complete a follow up K-D/ET test if they were subjectively identified receiving an impact to the head during the game. The research assistant, team coaches and parents were asked to document if they witnessed a participant receiving an impact to the head during the game. Additional players who were not identified as receiving an impact to the head during the game were selected for post-season testing. The testing protocol was identical to the baseline testing utilizing standardized instructions. Over the course of the season, 12 players completed a post-game K-D/ET assessment. The same person (MG) who conducted the baseline assessments administered all post-game tests.

Qualitative data (logbooks)

Over the course of the season, participants completed a weekly logbook (Fig 1). Items included written feedback responses on recalling having a direct hit to their head with another player (their head, knee, elbow etc.) or striking their head to the ground. Further, the logbook listed nine common concussion signs and symptoms for participants to indicate if during, or after, the game they experienced. The logbooks were collected during the first training session following the weekend game. The logbook data were retrospectively reviewed to identify if the participant had received a head impact during a game, which was not witnessed, and compared to the baseline K-D/ET measurements during the statistical analysis.

Participants who sustained a head impact

Over the course of the season, 6 participants sustained an impact to the head, either witnessed during a game or respectively identified from logbook data, with two players sustaining an impact in more than one game. In all, there were 12 post-head impact K-D/ET assessments from the 6 participants who sustained a head impact. For this study, a head impact was documented if the participant sustained a player-to-player or player-to-surface head contact, or whiplash-like head movement or self-reported on the weekly logbook.

Statistical Analysis

The K-D/ET completion time (seconds), number of saccades, and number of blinks were recorded for each of the three cards as well as total measurement. This was completed offline, data were downloaded onto an Excel spreadsheet and analyzed with SPSS (IBM Corp, Released 2017. IBM SPSS Statistics for Windows, Version 25.0 Armonk, NY: IBM Corp). A repeated measures Generalized Linear Model was utilized to compare between the baseline line and post season and head impact scores of the K-D/ET test, the number of saccades and the number of blinks recorded per card. For players who were had sustained a head impact ($n=6$), only the post-game assessment was utilized for comparison with the baseline score. Diagnostic accuracy for the K-D/ET duration, saccades and blinks were determined using a 2 by 2 contingency table with 95% confidence intervals (95% CI). The level of significance was set at $p<0.05$, and all data are expressed as means and standard deviations.

RESULTS

A total of 19 participants completed the baseline K-D/ET test and 12 completed a follow-up post-game test (Table 1). Participants recorded a slower time on the third (20.2 [95% CI: 18.0 to 22.4]) screen of the K-D/ET when compared with the first ($B=0.31$, $SE_B=0.15$, 95% CI: 0.01 to 0.61, $\chi^2=4.1$, $p=0.0424$) and the second ($B=0.55$, $SE_B=0.23$, 95% CI: 0.11 to 0.99, $\chi^2=5.9$, $p=0.0150$) screens. The number of blinks recorded were higher on the third (2.9 [95% CI: 1.6 to 4.3]) screen of the K-D/ET when compared with the second ($B=0.59$, $SE_B=0.21$, 95% CI: 0.17 to 1.00, $\chi^2=7.6$, $p=0.0057$) screen. Although the time for the K-D/ET decreased between pre- and post-game ($B=0.38$, $SE_B=0.28$, 95% CI: -0.17 to 0.93, $\chi^2=1.8$; $p=0.1769$) this was not significant.

Six of the 19 participants were identified as having obtained an injury involving the head thus warranting a post-game K-D/ET test. The results (see Table 2) represent the final K-D/ET the participant underwent as two of the players were tested more than once due to sustaining an impact in more than one game. Although these participants recorded a slower (worsening) mean total K-D time (50.8 vs. 48.7 s; $B=-0.11$, $SE_B=0.34$, 95% CI: -0.79 to 0.55, $\chi^2=0.1$; $p=0.7322$) this was not significant. However, the participants who were witnessed or reported to have received an impact to the head had fewer mean total saccades (172.7 vs. 176; $B=3.12$, $SE_B=1.23$, 95% CI: 0.71 to 5.54, $\chi^2=6.4$; $p=0.0112$) and recorded more mean blinks (8.9 vs. 5.4; $B=-0.11$, $SE_B=0.67$, 95% CI: -1.43 to 1.20, $\chi^2=0.0$; $p=0.8678$) when compared with their baseline scores.

The sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (PLR) and negative likelihood ratio (NLR) were determined (see Table 3). The assessment of blinks was the most sensitive measure for potential SRC with a sensitivity of 0.67 (95% CI: 0.27 to 0.94). The K-D/ET duration was the most specific measure for potential SRC with a specificity of 0.88 (95% CI: 0.82 to 0.98). An increase in the number of blinks also had a fair specificity of 0.69 (95% CI: 0.54 to 0.79).

DISCUSSION

In this study, eye movement behavior during reading of the K-D/ET test was assessed in youth AF players over the course of a single competition season. Reviewing the summed total for the three cards of all participants, assessed at the conclusion of the season, revealed a faster K-D/ET time (decrease in time), and fewer saccade and blink counts when compared with their baseline results. Results from participants who received an impact to the head, assessed immediately following a head impact, exhibited slower read times (worsening time), fewer saccades but increased eye blink counts.

For the K-D/ET the value of time changes, via the rapid number-naming component, and the relationship to concussion has been established in adult and youth athletes [19]. That study [19] revealed the K-D rapid number-naming component detects concussion with high degrees of sensitivity and specificity, with any worsening (slowing) of the individual result when compared with their baseline, indicating a five-time greater likelihood of concussion. The findings in the current study align with those findings with the group which had received an impact to the head showing worsening time scores (slowing), albeit not significant. However, the possibility exists that ocular motor behavior on the K-D/ET test is normal after concussion and slowed reading times may be due solely to cognitive or attentional deficits [13].

The results of the current study revealed a reduction in overall saccades in all participants including those who sustained an impact to the head. This is not consistent to a study [20] with older participants (mean age: 31) who had a history of concussion and produced a greater number of saccades overall, when compared to a control group. However, participants who had sustained an impact to the head in the current study were younger and not assessed by a physician, or other qualified health care provider, during a game as this is not routinely done at this level of participation, and therefore it is uncertain if they had sustained a concussion. At this

level, if a player does sustain an injury which may require medical assistance, they are advised to seek care from their family physician. Although deficits in binocular vision (convergence), accommodative (focusing) and saccadic (eye movement) disorders have been reported in adults with concussion in both the civilian and military populations with a prevalence of up to 30% to 42% [21-27] the prevalence of visual diagnoses in adolescents with concussion is unknown [28]. Therefore, future studies are warranted with a focus on a younger age group.

Eye blinking is an important eye motion, and involuntary eye blinks are controlled by a central mechanism that is often associated with fatigue, lapses of attention, and stress [29-31]. Eye blinks are also useful in diagnosing the status of patients suffering from neurological disorders [32]. A reduced rate of eye blinking is associated with Parkinson's disease, and excessive blinking may indicate the onset of Tourette's syndrome, strokes, or other disorders of the nervous system [32]. Spontaneous blinks, as in those measured in the current study, are mainly dependent on cognitive processes that appear to be regulated by a central pacemaker [33,34]. This pacemaker is highly sensitive to the attentional demands and cognitive workload of the concurrent visual task [33]. Previous research [35,33], albeit noting a considerable inter-subject variability in the frequency of spontaneous blinking, have reported mean \pm SD blink rate of 7.9 ± 3.3 blinks/ minute while reading, 14.5 ± 3.3 blinks/minute in silent primary gaze and 21.5 ± 6.0 blinks/minute during a conversation. In the current study, blink rates decreased between pre- and post-season testing collectively, however, the head impact group revealed an increase in blink rate. Future research is warranted to identify if measuring blink rate independently in people with concussion can assist with concussion identification.

Reviewing the three cards individually may provide an insight into the patterns that could distinguish players who have received a meaningful impact in the absence of signs and symptoms. For all participants, notable differences on the reading times were found as the third card recorded the slowest times to complete. This result is similar to a previous study with adult participants in both control group, and those with a previous history of concussion [20], but is in conflict with another study [13]. In the current study, the head impact group recorded progressively slower means times for the third card and the total time recorded, but these were not significant. This may be due to the cards become progressively more difficult to read in a flowing manner because of increasing space between the numbers [36].

For saccades, the head impact group had a reduction in the frequency on the second card possibly indicating that these measures have the potential to identify adolescents who experience a concussive event. However, more research is needed with larger sample sizes to fully identify the clinical implication of oculomotor performance differences in these groups. In addition, an even more salient finding may be that following such an injury, young athletes incur noteworthy oculomotor deficits that could potentially put these individuals at a disadvantage in settings where quick and accurate interpretation of visual information is imperative for maintaining performance or safety.

Implications for clinicians

There is an inherent risk for injury in all contact sports with head impacts an essential component to the development of SRC. Specifically, Australian football players have been shown to receive an average of 60 head impacts per season [37]. While it is accepted that not all head impacts result in SRC [3] one issue facing clinicians is that SRC is often underreported [38,39]. Furthermore, SRC may be undiagnosed in nearly one third of athletes reporting an undiagnosed concussion in one multi-center cross-sectional study [40]. One of the reasons for the under-reporting of SRC comes from an athlete's desire to not let teammates down and the support network around an athlete (e.g. parents and coaches) [39]. In order to increase the recognition of suspected SRC employing an assessment technique such as the SCAT 5 is suggested as best management [3]. Oculomotor function can be objectively assessed for impairments in function which occur as a result of mild traumatic brain injury [4,5]. The K-D rapid number-naming test is an assessment tool which quantifies impairment to oculomotor function and has been validated as a diagnostic tool for SRC in adults [6-12]. The results of this study demonstrate that youth participants have similar oculomotor impairments following a head impact to adults and youth with SRC in reference to the rapid number-naming component with the K-D test. However, the use of the eye tracking (ET) component requires further investigation in adults and youth. If both adults and youth share oculomotor impairment following SRC implementing assessment tools such as the K-D/ET test, which quantify the extend of this deficit, may help reduce the number of undiagnosed SRC currently reported within the literature.

Limitations

Several limitations need to be noted. The sample size was limited due the study focusing on one youth AF team and, as this was an in-situ study, the K-D/ET model was limited to recording total reading time, and number of saccades and blinks. Analyses of saccade velocity, acceleration, amplitude, duration, spatial accuracy, and main sequence relationships between these parameters were not obtained. The diagnosis of a potential SRC in this study was based on self-reporting symptoms, by an adolescent athlete, to a research assistant (MG) who attended every game. This diagnosis is not ideal when compared to the gold standard diagnosis of SRC, the SCAT 5 [3], and likely introduces bias into the results. This may also explain why the sensitivity (0.67) and specificity (0.88) of the K-D/ET test in this study is worse than the previous reported sensitivity (0.98) and specificity (0.96) for adult AF players [37]. However, while this is a limitation it is also likely to be the most feasible way to assess for potential SRC in an adolescent cohort of AF players and is reflective of standard practice. Adolescent AF matches do not have medical staff in attendance who are qualified to perform a SCAT 5. Therefore, the team coaching staff and players parents will be responsible for the management of any potential SRC. This study replicates standard practice which may make it more applicable to this particular demographic. Furthermore, these results highlight that while the K-D/ET is a useful test in the diagnosis of SRC in adolescent AF players it should not be used as a stand-alone test.

CONCLUSION

Oculomotor function can be objectively assessed for impairments in function which occur as a result of mild traumatic brain injury [4,5]. The rapid number-naming component of the K-D test is an assessment tool which quantifies impairment to oculomotor function and has been validated as a diagnostic tool for SRC [6-12].

The clinical usefulness of the eye tracking component of the K-D/ET test is that it may be an effective method to assess concussions with the eye tracking component serving as a measure of progression and return to play. However, more research is required at the adult and youth level and it should not be a standalone method to evaluate for suspected SRC and the use of the eye tracking component needs to include more kinematic components.

Declaration of funding

This manuscript was not funded.

Declaration of financial/other relationships

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties. Peer reviewers on this manuscript have no relevant financial relationships to disclose.

Tables & Figures

Tables:

Table 1. Baseline and post-season King-Devick test card scores, number of saccades and blinks for all male youth Australian football participants

Table 2. Baseline, post-injury and post-season King-Devick test card scores, number of saccades and blinks for participants reporting an impact to the head with male youth Australian football participants.

Table 3. Diagnostic accuracy of the King-Devick/ Eye-tracking test for sport-related concussion in adolescent Australian Football players

Figures:

Figure 1. Administration of K-D/ET with participant

Figure 2. Example page of the player's logbook that each player was required to complete post-match.

REFERENCES

1. Alfano CA, Zakem AH, Costa NM, Taylor LK, Weems CF. Sleep problems and their relation to cognitive factors, anxiety, and depressive symptoms in children and adolescents. *Depress Anxiety*. 2009;26(6):503-512. doi: 10.1002/da.20443.
2. Collins M, Kontos A, Reynolds E, Murawski C, Fu F. A comprehensive, targeted approach to the clinical care of athletes following sport-related concussion. *Knee Surg Sports Traumatol Arthrosc*. 2014;22(2):235-246. doi: 10.1007/s00167-013-2791-6. Epub 2013 Dec 12.
3. McCrory P, Meeuwisse W, Dvořák J, et al. Consensus statement on concussion in sport—the 5th international conference on concussion in sport held in Berlin, October 2016 *Br J Sports Med*. 2017;51(11):838-847. doi: 10.1136/bjsports-2017-097699. Epub 2017 Apr 26.
4. Heitger MH, Jones RD, Anderson TJ. A new approach to predicting postconcussion syndrome after mild traumatic brain injury based upon eye movement function. In: Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE, IEEE; 2008. pp. 3570–3573. doi: 10.1109/IEMBS.2008.4649977.
5. Ciuffreda KJ, Kapoor N, Rutner D, Suchoff IB, Han M, Craig S. Occurrence of oculomotor dysfunctions in acquired brain injury: A retrospective analysis. *Optometry*. 2007;78(4):155-61. doi:10.1016/j.optm.2006.11.011.
6. Galetta KM, Barrett J, Allen M, Madda F, Delicata D, Tennant AT . Balcer LJ. The King-Devick test as a determinant of head trauma and concussion in boxers and MMA fighters. *Neurology*. 2011;26;76(17):1456-1462. doi: 10.1212/WNL.0b013e31821184c9. Epub 2011 Feb 2.
7. Galetta KM, Brandes LE, Maki K, Dziemianowicz MS, Laudano E, Allen M. The King-Devick test and sports-related concussion: study of a rapid visual screening tool in a collegiate cohort. *J. Neurol. Sci.* 2011;309(1–2):34–39. doi.org/10.1016/j.jns.2011.07.039. Epub 2011 Aug 16.
8. Galetta MS, Galetta KM, McCrossin J, Wilson JA, Moster S, Galett SL. Saccades and memory: baseline associations of the King-Devick and SCAT2 SAC tests in professional ice hockey players. *J. Neurol. Sci.* 2013;328(1–2):28–31. doi.org/10.1016/j.jns.2013.02.008. Epub 2013 March 19.

9. Leong DF, Balcer LJ, Galetta SL, Liu Z, Master CL. The King–Devick test as a concussion screening tool administered by sports parents. *J. Sports Med. Phys. Fitness.* 2014;54(1):70-77. Epub 2014 Jan 22.
10. Galetta KM, Morganroth J, Moehrnger N, Mueller B, Hasanaj L, Webb N. Adding vision to concussion testing: a prospective study of sideline testing in youth and collegiate athletes. *J. Neuroophthalmol.* 2015;35(3):235-41 doi.org/10.1097/WNO.0000000000000226. Epub 2015 Mar 6.
11. Pierrot-Deseilligny C, Muri RM, Nyffeler T, Milea D. The role of the human dorsolateral prefrontal cortex in ocular motor behaviour. *Ann. N. Y. Acad. Sci.* 2005;1039:239-251. doi.org/10.1196/annals.1325.023.
12. Hecimovich M, King D, Dempsey A, Murphy M. The King-Devick test is a valid and reliable tool for assessing Sport-Related Concussion in Australian Football: A Prospective Cohort Study. *J Sci Sport*; 2018. pii: S1440-2440(18)30095-1. doi: 10.1016/j.jsams.2018.03.011. Epub ahead of print.
13. Rizzo JR, Hudson TE, Dai W, Desai N, Yousefi A, Palsana D, Selesnick I, Balcer LJ, Galetta SL, Rucker JC. Objectifying eye movements during rapid number naming: Methodology for assessment of normative data for the King-Devick test. *J Neurol Sci.* 2016;15:362:232-9. doi: 10.1016/j.jns.2016.01.045.
14. Pierrot-Deseilligny C, Muri RM, Nyffeler T, Milea D. The role of the human dorsolateral prefrontal cortex in ocular motor behaviour. *Ann. N. Y. Acad. Sci.* 2005;1039(1):239–251. <http://dx.doi.org/10.1196/annals.1325.023> (PubMed PMID: 15826978).
15. Pierrot-Deseilligny C, Muri RM, Ploner CJ, Gaymard B, Demeret S, Rivaud-Pechoux S. Decisional role of the dorsolateral prefrontal cortex in ocular motor Behaviour. *Brain*;2003;126(Pt 6):1460-1473. Epub 2003 May 24. Cited in: PubMed; PMID12764065.
16. Pierrot-Deseilligny C, Ploner CJ, Muri RM, Gaymard B, Rivaud-Pechoux S. Effects of cortical lesions on saccadic: eye movements in humans. *Ann. N. Y. Acad. Sci.* 2002;956:216-229. Epub 2002 Apr 19.
17. Choi W, Desai RH, Henderson JM. The neural substrates of natural reading: a comparison of normal and non word text using eye tracking and fMRI. *Front Hum Neurosci.* 2014;8:1024. doi: 10.3389/fnhum.2014.01024. eCollection 2014.

18. Dziemianowicz MS, Kirschen MP, Pukenas BA, Laudano B, Balcer LJ, Galetta SL. Sports-related concussion testing. *Curr. Neurol. Neurosci. Rep.* 2012;12(5):547-559. doi: 10.1007/s11910-012-0299-y.
19. Galetta KM, Liu M, Leong DF, Ventura RE, Galetta SL, Balcer LJ. The King-Devick test of rapid number naming for concussion detection: meta-analysis and systematic review of the literature. *Concussion*. 2016;1(2). doi:10.2217/cnc.15.8.
20. Rizzo JR, Hudson TE, Dai W, Birkemeier J, Pasculli RM, Selesnick I, Balcer LJ, Galetta SL, Rucker JC. Rapid number naming in chronic concussion: eye movements in the King-Devick test. *Ann Clin Transl Neurol*. 2016;3(10):801-811.
21. Brahm KD, Wilgenburg HM, Kirby J, Ingalla S, Chang CY, Goodrich GL. Visual impairment and dysfunction in combat-injured service members with traumatic brain injury. *Optom Vis Sci*. 2009;86:817-825.
22. Doble JE, Feinberg DL, Rosner MS, Rosner AJ. Identification of binocular vision dysfunction (vertical heterophoria) in traumatic brain injury patients and effects of individualized prismatic spectacle lenses in the treatment of post-concussive symptoms: a retrospective analysis. *PM R*. 2010;2:244-253.
23. Green W, Ciuffreda KJ, Thiagarajan P, Szymanowicz D, Ludlam DP, Kapoor N. Accommodation in mild traumatic brain injury. *J Rehabil Res Dev*. 2010;47:183-199.
24. Goodrich GL, Flyg HM, Kirby JE, Chang CY, Martinsen GL. Mechanisms of mTBI and visual consequences in military and veteran populations. *Optom Vis Sci*. 2013;90:105-112.
25. Stelmack JA, Frith T, Van Koevering D, Rinne S, Stelmack TR. Visual function in patients followed at a Veterans Affairs polytrauma network site: an electronic medical record review. *Optometry*. 2009 Aug;80(8):419-424. doi: 10.1016/j.optm.2009.02.011.
26. Suchoff IB, Kapoor N, Waxman R, Ference W. The occurrence of ocular and visual dysfunctions in an acquired brain-injured patient sample. *J Am Optom Assoc*. 1999 May;70:301-308.
27. Dougherty AL, MacGregor AJ, Han PP, Heltemes KJ, Galarneau MR. Visual dysfunction following blast related traumatic brain injury from the battlefield. *Brain Inj*. 2011;25(1):8-13. doi: 10.3109/02699052.2010.536195. Epub 2010 Nov 30.

28. Master CL, Scheiman M, Gallaway M, Goodman A, Robinson RL, Master, SR, Grady MF. Vision Diagnoses Are Common After Concussion in Adolescents. *Clin Pediatr (Phila)*. 2016;55(3):260-267. doi: 10.1177/0009922815594367. Epub 2015 Jul 7.
29. Andreassi JL. Human behavior and physiological response (4th ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates; 2000.
30. Marshall SP. Identifying cognitive state from eye metrics. *Aviat Space Environ Med*. 2007 May;78(5 Suppl):B165-175.
31. Ryu K, Myung R. Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic. *International Journal of Industrial Ergonomics*. 2005;35(11):991–1009. doi:10.1016/j.ergon.2005.04.005.
32. Deuschl G, Goddemeier C. Spontaneous and reflex activity of facial muscles in dystonia, parkinson's disease, and in normal subjects. *J Neurol Neurosurg Psychiatry*. 1998;64(3):320-324.
33. Doughty MJ. Consideration of three types of spontaneous eyeblink activity in normal humans: during reading and video display terminal use, in primary gaze, and while in conversation. *Optom Vis Sci*. 2001;78:712–715.
34. Veltman JA, Gaillard AW. Physiological workload reactions to increasing levels of task difficulty. *Ergonomics*. 1998;41:656–669.
35. Ingre MA, kerstedt T, Peters B, Anund A, Kecklund G. Subjective sleepiness, simulated driving performance and blink duration: examining individual differences. *J Sleep Res*. 2006;15:47–53.
36. King-Devick test. (n.d.) Miller-Keane Encyclopedia and Dictionary of Medicine, Nursing, and Allied Health, Seventh Edition. (2003).
37. Hecimovich M, King D, Dempsey A, Murphy M. Head Impact Exposure in Junior and Adult Australian Football Players. *J Sports Med*. 2018;8376030.
<https://doi.org/10.1155/2018/8376030>.
38. Register-Mihalik JK, Guskieicz KM, Valovich McLeod TC, Linnan LA, Mueller FO, Marshall SW. Knowledge, Attitude, and Concussion-Reporting Behaviors Among High School Athletes: A Preliminary Study. *J Athl Train*. 2013;48(5):645-53. doi: 10.4085/1062-6050-48.3.20. Epub 2013 Jul 12.

39. Kroshus E, Garnett B, Hawrilenko M, Baugh CM, Calzo JP. Concussion under-reporting and pressure from coaches, teammates, fans, and parents. *Soc Sci Med*. 2015;134:66-75. doi: 10.1016/j.socscimed.2015.04.011. Epub 2015 Apr 20.
40. Meehan WP, Manniz RC, O'Brien MJ, Collins MW. The Prevalence of Undiagnosed Concussions in Athletes. *Clin J Sport Med*. 2013;23(5):339-42. doi: 10.1097/JSM.0b013e318291d3b3.

Accepted Manuscript

Accepted Manuscript

Table 1. Baseline and post-season King-Devick test card scores, number of saccades and blinks for all male youth Australian football participants

Measurement	Baseline (n=19)	Post-season (n=12)
	Mean (95% CI)	Mean ±SD
King-Devick		
Card 1	17.8 (16.3 to 19.4) s ^c	17.4 (15.9 to 18.8) s ^c
Card 2	18.4 (16.5 to 20.3) s ^c	18.2 (16.5 to 19.8) s ^c
Card 3	20.2 (18.0 to 22.4) s ^{ab}	20.0 (18.3 to 21.7) s ^{ab}
Total	56.4 (51.2 to 61.5)s	55.5 (51.5 to 59.6) s
Saccades		
Card 1	54.3 (47.9 to 60.7) ^{bc}	49.6 (45.0 to 54.1)
Card 2	57.6 (51.8 to 57.3) ^{ac}	51.6 (46.0 to 57.1)
Card 3	56.2 (50.4 to 61.9) ^{ab}	50.2 (46.2 to 54.1)
Total	168.1 (153.8 to 182.3)	151.3 (140.0 to 162.7)
Blinks		
Card 1	2.2 (1.1 to 3.4) ^b	1.3 (0.6 to 1.9)
Card 2	2.3 (0.9 to 3.7) ^{ace}	1.4 (0.6 to 2.2) ^d
Card 3	2.9 (1.6 to 4.3) ^b	1.6 (0.8 to 2.4)
Total	7.5 (3.7 to 11.3) ^e	4.3 (2.2 to 6.3) ^d

CI: = Confidence Intervals; s = seconds; Significant difference ($p<0.05$) than (a) = Card 1; (b) = Card 2; (c) = Card 3; (d) = Baseline; (e) = Post-season.

Table 2. Baseline, post injury and post-season King-Devick test card scores, number of saccades and blinks for participants reporting an impact to the head with male youth Australian football participants.

Measurement	Baseline (n=6)	Post-injury (n=6)	Post-Season (n=6)
	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)
King-Devick			
Card 1	15.5 (13.6 to 17.4) ^c	17.5 (15.0 to 20.1) ^b	15.5 (14.4 to 16.6) ^b
Card 2	16.2 (13.8 to 18.6) ^{ce}	18.8 (15.3 to 22.3) ^{acd}	16.4 (13.5 to 19.2) ^{ac}
Card 3	17.0 (14.2 to 19.7) ^{ab}	19.9 (16.1 to 23.6) ^b	17.5 (14.7 to 20.4) ^b
Total	48.7 (41.9 to 55.6) ^e	56.2 (46.7 to 65.7) ^{df}	49.3 (42.6 to 46.0) ^d
Saccades			
Card 1	58.2 (40.1 to 76.3) ^b	64.7 (42.3 to 87.0) ^b	57.5 (49.4 to 65.6) ^b
Card 2	63.3 (39.9 to 86.7) ^{ac}	59.7 (40.6 to 78.7) ^{ac}	56.1 (49.4 to 62.7) ^{ac}
Card 3	54.7 (40.2 to 69.2) ^b	67.8 (28.1 to 107.6) ^b	59.1 (45.3 to 72.9) ^b
Total	176.2 (123.3 to 229.1) ^{ef}	192.2 (113.0 to 271.3) ^d	172.7 (145.6 to 199.8) ^d
Blinks			
Card 1	2.0 (0.5 to 3.5) ^{bef}	3.0 (0.4 to 6.6) ^{cd}	2.6 (0.7 to 4.5) ^{cd}
Card 2	1.9 (0.2 to 3.6) ^{aef}	3.0 (0.2 to 5.8) ^{cd}	2.6 (0.2 to 5.0) ^c
Card 3	1.5 (0.6 to 2.4) ^{abf}	3.5 (0.6 to 7.6) ^{abd}	3.0 (0.7 to 5.3) ^{abd}
Total	5.4 (1.6 to 9.2) ^f	10.0 (0.3 to 19.7) ^d	8.2 (1.8 to 14.6)

CI = Confidence interval; s = seconds; Significant difference ($p<0.05$) than (a) = Card 1; (b) = Card 2; (c) = Card 3; (d) = Baseline; (e) = Post-injury; (f) = post-season.

Table 3. Diagnostic accuracy of the King-Devick-Eye-tracking test for sport-related concussion in youth Australian Football players.

Evaluation	King-Devick duration	Saccades	Blinks
Sensitivity (95%CI)	0.167 (0.009 to 0.432)	0.333 (0.064 to 0.728)	0.667 (0.271 to 0.935)
Specificity (95%CI)	0.875 (0.816 to 0.975)	0.438 (0.336 to 0.586)	0.688 (0.539 to 0.788)
PPV (95%CI)	0.333 (0.018 to 0.865)	0.182 (0.035 to 0.397)	0.444 (0.181 to 0.623)
NPV (95%CI)	0.737 (0.687 to 0.821)	0.636 (0.489 to 0.852)	0.846 (0.664 to 0.97)
PLR (95%CI)	1.333 (0.048 to 17.052)	0.593 (0.096 to 1.757)	2.133 (0.588 to 4.413)
NLR (95%CI)	0.952 (0.582 to 1.215)	1.524 (0.464 to 2.783)	0.485 (0.082 to 1.352)

CI = Confidence Interval

Player ID # _____

Todays' date: _____

1. Did you have a game on the weekend?

Yes No

2. What is your primary position? _____

3. Do you recall having a direct hit to your head with another player (their head, knee, elbow etc) or did you strike your head to the ground?

Yes No

If yes, please explain

4. During or after the game did you experience any of the following (please circle)

- a. Headache
- b. Feel dizzy, sluggish or foggy
- c. Bothered by light or noise
- d. Have double or blurry vision
- e. Vomit or feel sick to your stomach
- f. Have trouble focusing or problems remembering
- g. Feel more emotional or "down"
- h. Feel confused
- i. Have problems with sleep

5. If you answered yes to any of these can please provide a brief explanation

6. Did you sustain any injury to your body? If so please write down what is was

Figure 1. Example page of the player's logbook that each player was required to complete post-match.



USAARL Report No. 2018-07

Evaluation of the Commercial, Off-the-Shelf (COTS) King-Devick Eye Tracking System

By Leonard A. Temme¹, Paul St. Onge^{1,2}, Aaron
McAtee^{1,3}, Kevin J. O'Brien^{1,3}, Brigid K. Byrd^{1,3}

¹U.S. Army Aeromedical Research Laboratory

²Laulima Government Solutions, LLC

³Oak Ridge Institute for Science and Education



United States Army Aeromedical Research Laboratory

Visual Protection and Performance Division

March 2018

Approved for public release; distribution unlimited.

Notice

Qualified Requesters

Qualified requesters may obtain copies from the Defense Technical Information Center (DTIC), Fort Belvoir, Virginia 22060. Orders will be expedited if placed through the librarian or other person designated to request documents from DTIC.

Change of Address

Organizations receiving reports from the U.S. Army Aeromedical Research Laboratory on automatic mailing lists should confirm correct address when corresponding about laboratory reports.

Disposition

Destroy this document when it is no longer needed. Do not return it to the originator.

Disclaimer

The views, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other official documentation. Citation of trade names in this report does not constitute an official Department of the Army endorsement or approval of the use of such commercial items.

REPORT DOCUMENTATION PAGE
*Form Approved
OMB No. 0704-0188*

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY)				2. REPORT TYPE	3. DATES COVERED (From - To)	
4. TITLE AND SUBTITLE				5a. CONTRACT NUMBER		
				5b. GRANT NUMBER		
				5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S)				5d. PROJECT NUMBER		
				5e. TASK NUMBER		
				5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)					8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)					10. SPONSOR/MONITOR'S ACRONYM(S)	
					11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT						
13. SUPPLEMENTARY NOTES						
14. ABSTRACT						
15. SUBJECT TERMS						
16. SECURITY CLASSIFICATION OF:		17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON		
a. REPORT	b. ABSTRACT	c. THIS PAGE				
				19b. TELEPHONE NUMBER (Include area code)		

REPORT DOCUMENTATION PAGE (SF298)
(Continuation Sheet)

Foreword

The identification and characterization of sensitive biomarkers for concussion are important for military medicine to guide diagnosis and return to duty decisions. The current literature suggests that the King Devick (KD) number reading test is a sensitive screener for concussion sustained during sports activities. The KD test involves timing an individual reading aloud 120 single digit numbers, without error, printed on three standard cards; timing is performed with a stopwatch. The extensive literature commonly describes the KD test as revealing the disruption of eye movements that could be due to a number of factors including concussion. Thus, the test is sensitive if not necessarily specific. The extent to which the KD test reflects eye movements remains a matter of conjecture since it is a measure of gross behavioral reading performance. Hence, it is important to objectively measure eye movements concurrently with the performance of the KD test. Recently, a commercial, off-the-shelf (COTS), integrated KD / Eye Tracking system became available. Because of the potential importance of the contributions of this system to concussion research, two copies of the system were purchased and evaluated before they were deployed for research. The evaluation consisted of two studies; one study with 20 volunteers measured the comparability of the two systems, while the other study with 5 volunteers assessed the stability and repeatability of the measurements over 5 successive days. The results showed that several response parameters of the systems lacked face validity, calling into question the validity of the eye tracking data. Thus, we have not used these systems for further research.

This page is intentionally blank.

Table of Contents

	Page
Introduction.....	1
Methods.....	3
Volunteers	3
Equipment	4
Procedures.....	5
Design	6
Data Analysis	7
Results.....	7
Determining Limits of Agreement.....	11
Calculation of the Coefficient of Agreement.....	12
Discussion.....	14
References.....	17
Appendix A. Bland-Altman plots for the response parameters	22
Appendix B. Subject-specific agreement charts	29

List of Figures

1. The two ET/KD systems.....	4
2. The four cards comprising the KD test.	5
3. Typical eye tracking plots generated by the ET/KD.....	8
4. Difference vs mean for Total Reading Time for trial 1 of the first experiment.....	12
5. Subject-specific agreement for total response time for the first experiment.	13

List of Tables

1. The 28 eye movement output parameters generated by the ET/KD for each test card.	6
2. Selected mean measurement values by system and trial number for the first experiment.	9
3. Average values for the 1-day and 5-day experiments.....	10
4. Measures of agreement for trial 1 of the first experiment.	12
5. Measures of agreements of mean differences and COA between the two ET/KD systems for the output parameters for the first experiment.	14

This page is intentionally blank.

Introduction

The increased world-wide recognition of the importance of sports-related concussion at essentially all levels of play, from elementary school through professional leagues, is matched by the increasing interest in the identification of practicable, sensitive, sideline concussion screening methods to guide the management of concussion during sporting activities and return to play decisions (Broglio et al., 2017; Khodaei, Currie, Asif, & Comstock, 2016; Leong, Balcer, Galetta, Liu, & Master, 2014; Marshall, Guskiewicz, Shankar, McCrea, & Cantu, 2015; Ventura, Balcer, Galetta, & Rucker, 2016). The Department of Defense (DoD) has a corresponding interest in methods to guide medical fitness and return to duty decisions. The DoD reports a total of 361,092 cases of traumatic brain injury between 2000 and 2016 with concussion accounting for 297,478 of these cases (Defense and Veterans Brain Injury Center (DVBIC), DoD worldwide numbers for TBI, http://dvbic.dcoe.mil/files/tbi-numbers/DoD-TBI-Worldwide-Totals_2000-2016_Feb-17-2017_v1.0_2017-04-06.pdf). These numbers underscore the importance of concussion detection for the military. The assessment of concussion in a military context is particularly challenging because the trauma often involves a combat situation, which may include blast, the “chaos of combat,” and the vital concerns of securing oneself, others, and equipment, while ensuring mission success. These factors add to the difficulty of military medical and psychological evaluations, which may occur months and even years after the event (Davenport, 2016). The well-known comorbidity of concussion with post-traumatic stress syndrome, depression, and altered mental states further complicates the evaluation of concussion in the military context (Hoge et al., 2008; Kennedy et al., 2007; Maguen, Lau, Madden, & Seal, 2012; Schmid & Tortella, 2012; Seal et al., 2016). While blast exposure is a combat-related cause of concussion, military personnel can sustain concussion from a wide variety of activities, including those that affect the civilian sector (Regasa, Thomas, Gill, Marion, & Ivins, 2016). In combat, as well as in training scenarios, concussion can compromise performance and reduce situational awareness, which can quickly translate into catastrophic mission failure and loss of life.

Return to duty decisions should be guided by accurate and practical metrics, particularly since concussion can exacerbate the deleterious effects of subsequent concussions (Barker et al., 2017; Iverson, Echemendia, Lamarre, Brooks, & Gaetz, 2012; Ventura et al., 2016). Additionally, repeated concussions may be associated with the development of long-term neurodegeneration (see: <http://dvbic.dcoe.mil/article/cumulative-concussions>) (Ventura et al., 2016). Because of these and other complications, the military has recognized the need for clear, objective, sensitive, and specific tests and biomarkers for concussion (Marion, Curley, Schwab, & Hicks, 2011; Schmid & Tortella, 2012).

Recent reports of retrospective analyses of the medical records show that as many as 70 to 85% of concussed patients report visual dysfunctions, complaints, or impairments even when their eyes are healthy and normal upon medical examination (Brahm et al., 2009; Capo-Aponte et al., 2016; Padula, Capo-Aponte, Padula, Singman, & Jenness, 2017). In other words, following concussion, vision complaints are common in the absence of apparent or manifest ocular trauma. The list of reported visual complaints is extensive and includes visual field defects, visual neglect, visual midline shift, pupillary abnormalities, color vision deficits, photophobia, eye strain, as well as difficulties with vergence, accommodation, and reading. It should be no surprise that concussion and other traumatic brain injuries affect vision since more than 30

regions of the brain are involved with vision, seven of the twelve cranial nerves subserve vision, and more than 70% of the brain's sensory processing involves vision. Reading is among the most frequent visual difficulties following concussion (Capo-Aponte et al., 2016), and it has long been known that reading is exquisitely dependent on eye movements, which are themselves sensitive to concussion (Leigh & Zee, 2006; Liversedge, Gilchrist, & Everling, 2011).

Eye movements can be conveniently categorized as saccades, smooth pursuits, and fixation-related behaviors, which include micro-saccades to correct drift and maintain fixation (Barnes, 2011; Gilchrist, 2011; Martinez-Conde & Macknik, 2011). Saccades are characteristically rapid, ballistic, and relatively accurate sudden shifts of gaze from one area of the visual field to another. In contrast, during smooth pursuit eye movements (SPEM), gaze typically follows the continuous path of a smoothly moving target so that the image of the target remains relatively constant on the same retinal location, which is most naturally the fovea. Unlike saccades, SPEM are almost never generated in the absence of an appropriately moving stimulus (Boyer, Portron, Bevilacqua, & Lorenceau, 2017).

A recent report compared eye movements recorded in 60 blast-concussed military volunteers with eye movements recorded in 20 civilian controls (Cifu et al., 2015). Neurologic and ophthalmologic exams documented no vision problems, yet there were robust, significant, and characteristic differences in eye movements between the two groups. Specifically, the saccades of the concussed group showed larger positional errors, smaller amplitudes, slower peak velocities, slower peak accelerations, and longer durations. In general, concussed individuals tracked stepwise moving targets less accurately than the controls. These findings confirm a previous study by Heitger et al. (2009) and are consistent with the more recent report by Balaban et al. (2016) that achieved a high degree of classification (concussed vs. controls) sensitivity (89%) and specificity (97%) using measures of saccadic eye movements combined with vestibulo-ocular eye movements.

The King-Devick number reading (KDNR) test is a widely used, rapid sideline screening tool for sport-related concussion (e.g., (K. M. Galetta, Barrett, et al., 2011; K. M. Galetta, Brandes, et al., 2011; K. M. Galetta et al., 2015; M. S. Galetta et al., 2013; Leong et al., 2014; Molloy, Murphy, & Gissane, 2017; Munce et al., 2014; Rizzo, Hudson, Dai, Desai, et al., 2016). Although the literature frequently discusses the KDNR test as a direct measure of eye movements and saccades, it is more a measure of the rapid reading and verbalization of numbers (Handmaker & Waldorf, 2013). This makes the KDNR test an assessment of a chain of interdependent behaviors, some part of which is arguably dependent on saccadic eye movements and fixations crucial to reading text, even though the KDNR test does not directly reflect eye movements in isolation from other aspects of reading behavior. The KDNR test is purportedly resistant to false positives resulting from simply exercise or physical exertion associated with sports activities (Dziemianowicz et al., 2012; Leong et al., 2014). It is also sensitive to such neurological diseases such as Alzheimer's, Parkinsonism, and multiple sclerosis (Lin, Adler, et al., 2014; Lin, Rigby, et al., 2014; Moster, Wilson, Galetta, & Balcer, 2014; Rizzo, Hudson, Dai, Desai, et al., 2016), as well as to the effects of sleep deprivation (Davies, Henderson, Balcer, & Galetta, 2012) and hypoxia (Stepanek et al., 2013).

Recently, the KDNR test has been integrated with eye tracking technology to enable administration of the KDNR test simultaneously with the measurement of eye movements

(Rizzo, Hudson, Dai, Birkemeier, et al., 2016; Rizzo, Hudson, Dai, Desai, et al., 2016; Stepanek et al., 2013). Rizzo et al. (2016a,b) used such eye tracking instrumentation to record eye movements in 42 controls and 25 concussed civilians as these volunteers performed the KDNR test. While there were no differences between the two groups in saccadic velocity, duration, or amplitude, the concussed group produced a significantly larger number of saccades with significantly less precisely controlled endpoints. Furthermore, the time between successive saccades was longer for the concussed than the control group; differences that are completely consistent with the longer time the concussed group needed to perform the KDNR test. Thus, the KDNR test provides a gross measure of reading behaviors that can reflect alterations in eye movement behaviors associated with reading, and these alterations in eye movements can be associated with concussion.

Because of the need for objective concussion biomarkers, and the potential of reading-related eye movements as a concussion biomarker, the present study evaluated a recently available, commercial, off the shelf (COTS) eye tracker / KDNR test (ET/KD) system. This system is purportedly a turnkey system ready for use upon delivery. Two of these ET/KD systems were purchased to support multiple research programs assessing concussion, hypoxia, and sleep deprivation. Before deploying these new COTS ET/KD systems for extended data collection, prudence dictated the evaluation of the reliability, validity, and comparability of the two presumably interchangeable systems, particularly since they were to be deployed in parallel at different remote testing sites.

Methods

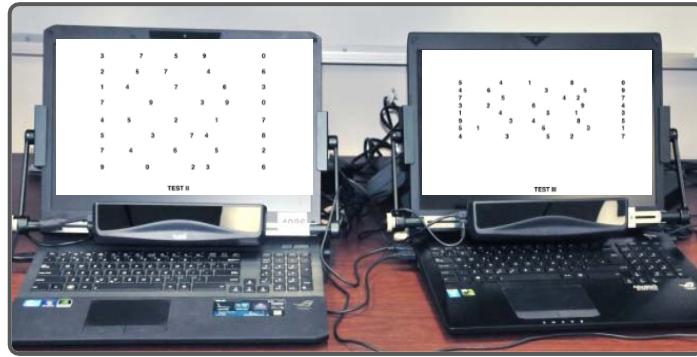
This study's protocol (No. M-10467 USAARL2015-019) was reviewed and approved by the U. S. Army Medical Research and Materiel Command Institutional Review Board (USAMRMC IRB).

Volunteers

The volunteers were a self-selected sample from the population of adults within USAARL's immediate geographic locale. All volunteers were at least 19 years of age; and as determined by screening with the Armed Forces Vision Tester, they all had normal binocular vision, and at least 20/30 near visual acuity either with or without single power corrective lenses (contacts or glasses). Volunteer exclusion criteria included the presence of an intraocular lens implant or use of bifocal, progressive, or other multi-focal corrective lenses at the time of testing. There were no exclusion criteria for this study based on gender, ethnicity, race, or other demographic characteristics. Volunteers were not financially reimbursed for their participation.

Equipment

Two copies of the same ET/KD system were ordered at the same time from the manufacturer (Figure 1). Each system consists of a laptop computer and proprietary software integrated with a third party eye tracker. The systems arrived with the software installed. Resident software functions included calibration of the system's eye tracker and volunteer's eye position and gaze, test administration, and data management. The two systems were understood



to be identical and completely functional on arrival.

Figure 1. The two ET/KD systems. The eye tracker is located below the laptop computer's flip-up screen. The screen on the left shows Test Card (TC) 2 while the screen on the right shows TC 3.

The ET/KD system presented the four standard KD cards on the computer screen. These cards are reproduced in Figure 2. The first card is the demonstration card (upper left); it contains five rows of five unevenly spaced single-digit numbers. These numbers are connected with line segments and arrow heads to indicate the path the eye's line of sight is intended to follow as the eye scans the set of 25 numbers on the card. Reading performance on the demonstration card is not scored; the card is solely for demonstration purposes. The first test card (TC 1) also contains eight rows of five single-digit numbers. All the first digits of the eight rows are vertically aligned as are all the last digits, thus defining a left and right margin, respectively. The three inner digits of each of the eight rows are spaced irregularly. Every row includes horizontal line segments between the digits to make the intended scan path explicit. The second test card (TC 2) contains 40 single-digit numbers in essentially the same layout as TC 1: 8 rows with the left and right edges justified and 3 irregularly spaced inner digits in each row. The difference between TC 1 and TC 2 is that TC 2 does not contain any line segments to aid visual scanning along the horizontal. The third test card (TC 3) again contains 40 single-digit numbers presented in 8 rows of 5 digits each. All the first digits of the 8 rows are vertically aligned as are the last digits, establishing, respectively, a left and right edge with the 3 remaining digits irregularly spaced on each row. The difference between TC 2 and TC 3 is that the vertical spacing has been compressed on TC 3 by about 50%. The expectation is that the difficulty reading the digits increases from TC 1 to TC 2 to TC 3, and that this increased reading difficulty is reflected in reading performance. The conventional KDNR test metric is the length of time required to read the three test cards, typically reported in seconds measured with a stopwatch level of precision. Additionally, the number of errors made while reading aloud the individual numbers on each of the three TCs is recorded.

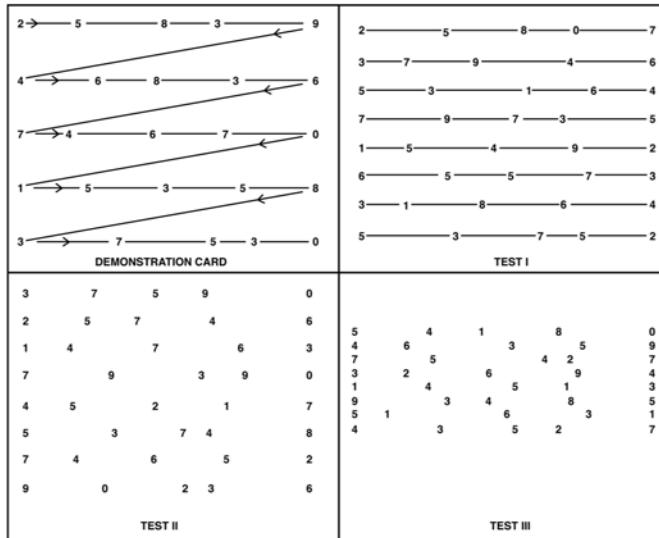


Figure 2. The four cards comprising the KD test. These same four cards are presented in the ET/KD systems. The Demonstration Card in the upper left quadrant is used only to demonstrate the test to the subject. The three remaining cards identified as TEST I, TEST II, and TEST III comprise the test itself. For all four cards, subjects read aloud each number, starting in the upper left of the card, continuing from left to right, and at the end of each line, moving down to the leftmost number on the next lower line.

Procedures

Each prospective volunteer was consented individually and in the privacy of the laboratory by one of the authors. Following the informed consent process, the volunteer's vision was screened using the Armed Forces Vision Tester to ensure that the volunteer's vision satisfied the selection criteria. Following the screening, the volunteer was seated in front of the appropriate ET/KD system, and the technician sitting beside the volunteer demonstrated the test using the second ET/KD system beside the volunteer's test system. Any additional questions the volunteer may have had as a result of the demonstration were answered.

When the volunteer was ready to begin data collection, the volunteer initiated testing with the ET/KD system, which administers the test under computer control. During testing, the test program steps through 11 screens: (1) General Instructions, (2) Calibration, (3) Calibration Results & Recalibration Option, (4) Test Instructions, (5) Demonstration Card, (6) Test Card I, (7) Test Card I Completed Screen, (8) Test Card II, (9) Test Card II Completed Screen, (10) Test Card III, and (11) Save Data Screen. The volunteer controls the progress and pacing from one screen to the next.

After calibration is complete, the Test Instructions screen presents the following text:
"The test will begin automatically after the Demonstration Screen. Please read the numbers aloud on Test Card I, starting with the number on the top left, immediately. As soon as you read the final number on each Test Card, click Next. Read aloud the numbers on each test card as quickly as you can. The Test will end when Test Card III is completed by clicking Next."

Please Note: Do not use your finger or the mouse to assist in reading the numbers on the Test Cards."

All testing was done under a photopic ambient lighting condition with fluorescent ceiling lighting of approximately 1000 lux, which is typical for office lighting. Since the manufacturer asserted that head stabilization was not necessary with the ET/KD system, the present studies did not use a head and chin rest.

For the first experiment, 20 volunteers alternated between the two ET/KD systems on successive trials. The two ET/KD systems were set up next to each other on the table. A volunteer's participation in the first experiment, from informed consent to the completion of data collection, invariably required one test session lasting less than two hours, including the volunteer's self-paced rest breaks.

The second experiment required five volunteers to return to the laboratory on five successive days at approximately the same time of day. On each of these days, a volunteer made a pair of readings on ET/KD systems, alternating between the two systems to produce four readings a day. The first day's test session included the informed consent and lasted about an hour; the 4 subsequent sessions lasted about 30 minutes per volunteer.

In both experiments, the ET/KD test reported the volunteer's eye movements along the 28 response parameters listed in Table 1. Thus, each test administration for each volunteer resulted in a spreadsheet containing values for these 28 variables. Of these parameters, 10 are averages, 9 are standard deviations, 6 are summed totals, 2 are calculated ratios, and 1 is a rate. It should be noted that the documentation for the ET/KD system did not describe methods, procedures, or algorithms used to define, derive, or calculate these response parameters. Furthermore, the ET/KD did not record the traditional measures of the KDNR test; i.e., the time required to read all 120 numbers or any errors made while reading the numbers. The technician recorded reading time and errors with paper, pencil, and a stopwatch.

Design

Two experiments were conducted. The first experiment compared the two ET/KD systems with 20 volunteers measured twice on each system, alternating between the two systems with half the volunteers starting with one system and the other half starting on the other system in an interleaved fashion. For this experiment, each volunteer's data were collected in a single session. The database of the first experiment consisted of 40 trials on each ET/KD system for a complete dataset of 80 trials. The second experiment evaluated the two ET/KD systems over 5 successive days, with 5 volunteers tested each day twice on both systems, alternating between ET/KD systems to produce a complete dataset of 100 trials.

Table 1. The 28 eye movement output parameters generated by the ET/KD for each test card.

ET/KD Reported Output Variables

Total Fixation Time (seconds)	Standard Deviation Saccade Length (mm)
Average Fixation Time Length (ms)	Total Only Saccade Length In Degrees (deg)
Standard Deviation Fixation Time Length (ms)	Average Saccade Length In Degrees (deg)
Average Fixation Size (mm)	Standard Deviation Saccade Length In Degrees (deg)
Standard Deviation Fixation Size (mm)	Average Saccade Velocity In Degrees (deg/sec)
Average Fixation Polyarea (mm ²)	Standard Deviation Saccade Velocity In Degrees
Standard Deviation Fixation Polyarea (mm ²)	Total Blinks
Average Inter Fixation Time (ms)	Blink Rate (blinks per minute)
Standard Deviation Inter Fixation Time (ms)	Average Blink Duration (ms)
Saccade Fixation Ratio	Standard Deviation Blink Duration (ms)
Total Saccades (#)	Average Inter Blink Duration (seconds)
Total Saccade Length (mm)	Left Eye Average Pupil Size (mm)
Total Only Saccade Length (mm)	Left Eye Standard Deviation Pupil Size (mm)
Average Saccade Length (mm)	Left Eye Pupillary Stress Ratio

Data Analysis

Descriptive statistics, coefficients of agreements, and differences between individuals were calculated for the 28 ET/KD eye movement output parameters as well as for the pair of conventional KDNR metrics of reading time and number of errors.

Results

Figure 3 displays eye tracking data generated by the ET/KD system for a typical volunteer successively viewing the three test cards. The left half of the figure shows the sequential scan patterns while the right half of the figure shows a heat-type plot in which color identifies the relative length of time the eye's line of sight (LOS) remained in the region of interest. The upper pair of plots displays data obtained with the TC 1. The middle pair displays data obtained with the TC 2. The bottom pair displays data obtained with the TC 3. These data are typical of the results reported for all volunteers. The data presented in the bottom left-right pair of tracings are obviously different from the data presented in the top and middle pair. The eye movements elicited by TC 3 were obviously different from the eye movements associated with either TC 1 or TC 2. In fact, not only are the eye movements different for TC 3, but the characteristic difference between TC 3 and the other two test cards is evident in the TC 3 data tracing: that is, the vertical distance between the rows of fixations or regions of interest is compressed, reflecting the compression of the TC 3. Specifically, the data obtained with TC 3 cannot be confused with the data obtained with either TC 1 or TC 2. On the other hand, there are no obvious differences in the data obtained with TC 1 and TC 2.

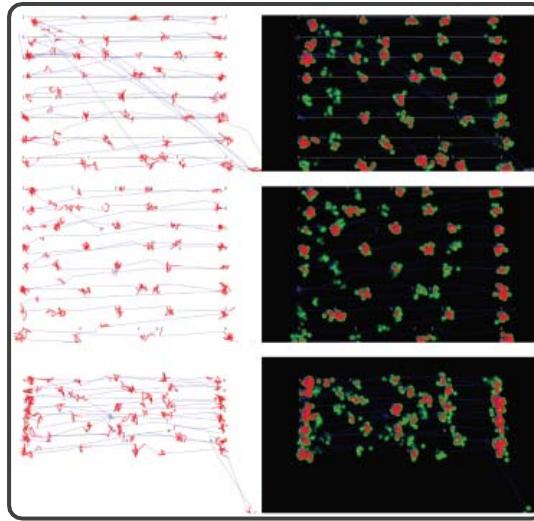


Figure 3. Typical eye tracking plots generated by the ET/KD.

The top left tracing showing the sequential scan patterns for TC 1 displays an unambiguous diagonal slewing of the LOS from the lower right to the upper left. These diagonal eye movements are a result of the test administration procedures. The volunteer starts the test by using the mouse to click a start button that is displayed on the computer screen. This start button is on the lower right corner of the display, so the experiment usually begins with the eye oriented toward the lower right corner, requiring the LOS to slew to the upper left corner to acquire the first digit. The same behavior is evident in the tracking for TC 3. However, the behavior is not as obvious in the data for TC 2, since there is some evidence the diagonal slew of the LOS begins about two thirds of the way to the upper left corner.

For the first experiment, all 20 subjects completed 2 test trials on each ET/KD system. The means for each of the 2 trials for each ET/KD system were calculated and are reported for 21 of the 28 response parameters listed in Table 1. The seven response parameters not reported in Table 2 are means of standard deviations, the interpretation of which some consider problematic.

Table 2. Selected mean measurement values by system and trial number for the first experiment.

20 subjects each completed 2 trials on the two ET/KD systems.

Measure	System 1			System 2		
	Trial 1	Trial 2	Overall	Trial 1	Trial 2	Overall
Total Time	53.19	50.16	51.68	51.87	49.55	50.71
Total Fixation Time (sec)	13.45	12.45	12.95	13.12	12.42	12.77
Average Fixation Time Length (ms)	226	216.67	221.33	227.25	229.18	228.21
Average Fixation Size (mm)	180.06	172.16	176.11	173.79	176.23	175.01
Average Fixation Polyarea (mm ²)	1156.32	1114.26	1135.29	1090.07	1102.77	1096.42
Average Inter Fixation Time (ms)	286.7	287.75	287.22	289.36	293.22	291.29
Saccade Fixation Ratio	1.5	1.54	1.52	1.54	163.28	1.52
Total Saccades	90.5	88.43	89.47	90.73	84.42	87.58
Total Saccade Length (mm)	32274	31237	31755	31675	29777	30726
Total Only Saccade Length (mm)	21066	20861	20964	21202	19805	20504
Average Saccade Length (mm)	237.48	240.2	238.84	239.3	240.88	240.09
Total Only Saccade Length (deg)	179.12	175.73	177.43	200.1	185.27	192.69
Average Saccade Length (deg)	2.02	2.03	2.02	2.27	2.27	2.27
Total Blinks	6.33	6.05	6.19	7.6	6.77	7.18
Blink Rate (blinks per minute)	21.48	22	21.74	26.44	24.38	25.41
Average Blink Duration (ms)	91.88	70.53	81.21	72.94	85.05	78.99
Average Inter Blink Duration (sec)	2.27	2.12	2.2	1.91	1.78	1.84
Left Eye Average Pupil Size (mm)	2.92	2.93	2.92	2.9	2.85	2.87
Left Eye Pupillary Stress Ratio	22.04	22.46	22.25	23.29	22.03	22.66
Errors	0.3	.035	0.32	0.2	0.3	0.25

Table 3 provides summary statistics, means, and standard deviations (SD) for the 1-day and 5-day experiments. Table 3 has 2 variables more than Table 1 because it includes Total Time and Number of Errors, which were entered by hand into the database. Thus the database contained 30 dependent variables; these were grouped according to 5 general oculometric categories addressing eye blinks, pupillary dynamics, fixations, saccades, and others. Table 3 also includes the dimensions of the response parameters. Blinks are described by 5 response parameters; pupillary dynamics are described by 3 response parameters that address only the left eye; fixations are described by 9 response parameters; saccades are described by 10 response parameters; and lastly, the “other” category contains 3 response parameters, which include the 2 hand-entered variables as well as an undefined variable called the saccade fixation ratio. These summary statistics of the 30 dependent variables were calculated across the 20 or 5 subjects for the 1-day and 5-day experiments, respectively, over TC 1, TC 2, and TC 3, as well as over the two devices and across each administration of the test.

Table 3. Average values for the 1-day and 5-day experiments. The three test cards were averaged for all subjects across all trials over both ET/KD devices.

Measure		1 day test (Mean \pm SD)	5 day test (Mean \pm SD)
Blinks	Total Blinks	6.7 \pm 6.2	7.8 \pm 6.2
	Blink Rate (blinks per minute)	23.8 \pm 21.2	31.8 \pm 25.2
	Avg Blink Duration (ms)	76.7 \pm 97.0	54.9 \pm 41.5
	SD Blink Duration (ms)	65.1 \pm 92.6	40.0 \pm 44.8
	Avg Inter Blink Duration (seconds)	2.0 \pm 2.3	1.3 \pm 1.1
Pupillary Dynamics	Left Eye Average Pupil Size (mm)	2.9 \pm 0.3	2.6 \pm 0.3
	Left Eye SD Pupil Size (mm)	0.14 \pm 0.05	0.10 \pm 0.03
	Left Eye Pupillary Stress Ratio	22.7 \pm 7.4	29.9 \pm 10.7
Fixations	Total Fixation Time (seconds)	12.4 \pm 2.7	9.5 \pm 1.7
	Avg Fixation Time Length (ms)	219.1 \pm 44.9	163.1 \pm 44.1
	SD Fixation Time Length (ms)	143.9 \pm 38.8	107.8 \pm 22.4
	Avg Fixation Size (inches)	6.8 \pm 1.6	5.1 \pm 1.3
	SD Fixation Size (mm)	122.9 \pm 35.1	99.1 \pm 23.9
	Avg Fixation Polyarea (mm ²)	1114 \pm 315	859 \pm 264
	SD Fixation Polyarea (mm ²)	1089 \pm 357	995 \pm 297
	Avg Inter Fixation Time (ms)	283.0 \pm 55.2	229.9 \pm 47.9
	SD Inter Fixation Time (ms)	169.4 \pm 115	141.2 \pm 144
Saccades	Total Saccades (#)	87.7 \pm 20.5	104.6 \pm 39.4
	Total Saccade Length (inches)	1212 \pm 212	1177 \pm 256
	Total Only Saccade Length (inches)	810.6 \pm 124	853.7 \pm 192
	Avg Saccade Length (inches)	9.4 \pm 1.2	8.5 \pm 1.3
	SD Saccade Length (mm)	199.7 \pm 42.6	178.1 \pm 45.1
	Total Only Saccade Length (deg)	183.2 \pm 28.5	183.9 \pm 41.9
	Average Saccade Length (deg)	2.1 \pm 0.3	1.8 \pm 0.3
	SD Saccade Length (deg)	1.8 \pm 0.4	1.5 \pm 0.4
	Avg Saccade Velocity (deg/sec)	70.4 \pm 9.1	61.2 \pm 7.6
	SD Saccade Velocity (deg/sec)	55.6 \pm 12.5	46.6 \pm 9.5
Others	Saccade Fixation Ratio	1.5 \pm 0.2	1.7 \pm 0.2
	Total Reading Time (seconds)	49.6 \pm 8.8	42.7 \pm 5.2
	Errors	0.3 \pm 0.5	0.0 \pm 0.0

Specifically, for the 1-day test, each volunteer was tested twice on each ET/KD system to generate 4 repeated measures. For a volunteer, the TC 1, TC 2, and TC 3 data were averaged for each test administration, and these averages were averaged to generate a single score for each subject. The individual summary data were averaged over the 20 volunteers to generate the means (\pm SD) displayed in the 1-day testing column of Table 3. A similar strategy was used to collapse the data spanning the 5 days for all 5 volunteers, taking into account the data across three test cards from both systems for each volunteer. These output summary statistics are based on the output of the eye tracking software resident on the ET/KD system.

While the means (\pm SD) of many of the output response parameters reported in Table 3 seem quite reasonable and have face validity, the values of several response parameters clearly

are not. For example, the 1-day and 5-day testing average number of Total Blinks for all three test cards is reported as 6.7 and 7.8, respectively. This is simply not possible with an average reported Blink Rate of 23.8 and 31.8 blinks per minute. Furthermore, the Average Blink Durations (76.7 and 54.9 ms) are extraordinarily brief.

The average pupil size of 2.9 and 2.6 mm is extraordinarily small with almost no evidence of variability (0.14 and 0.10), which may indicate an insensitive rather than a consistent measure. The ‘pupillary stress ratio’ was not defined. Measures of fixations are similarly suspicious. For example, Total Fixation Time (12.4 and 9.5 sec) and Average Fixation Time (219.1 and 163.1 ms) may be reasonable, but the validity of the data is undermined by the fact that the reported average size of the fixations approximate the full size of the display screen (6.8 and 5.1 inches). These are not meaningful measures. Similarly, the reported Average Saccadic Length of between 8 and 9 inches is not meaningful.

Appendices A and B present graphs of the 1-day and 5-day data, respectively, in several different formats to show that the lack of measurement precision is characteristic of these data and not due to a few isolated, idiosyncratic outliers.

Determining Limits of Agreement

An assessment of the agreement between the two ET/KD systems was made for each of the response parameters with the method described by Bland and Altman (Bland & Altman, 1996). In this method, the difference in measurements between the two ET/KD systems (System 2 – System 1) is plotted against the average of the two measurements for every subject. This assessment used the first run on each ET/KD system for each subject. Figure 4 illustrates the procedure with the Total Reading Time response parameter. The middle dotted line is the mean difference in total time for all 20 subjects (-1.3 sec), and the upper and lower dotted lines are the upper and lower limits of agreement (LOAs), respectively. The upper LOA is 6.84 seconds and the lower LOA is -9.49 seconds for total time. The LOAs are calculated by adding and subtracting 1.96 * standard deviation of the differences in measurements, s , from the mean difference in measurements, d . That is, $LOAs = d \pm 1.96 s$. Based on this sample, on 95% of the tests, the difference between the two ET/KD systems in Total Reading Time to complete the KDNR test falls between -9.49 and 6.84 seconds. The mean differences and LOAs for each response parameter are given in Table 4. Bland-Altman plots for the other response parameters are given in Appendix A.

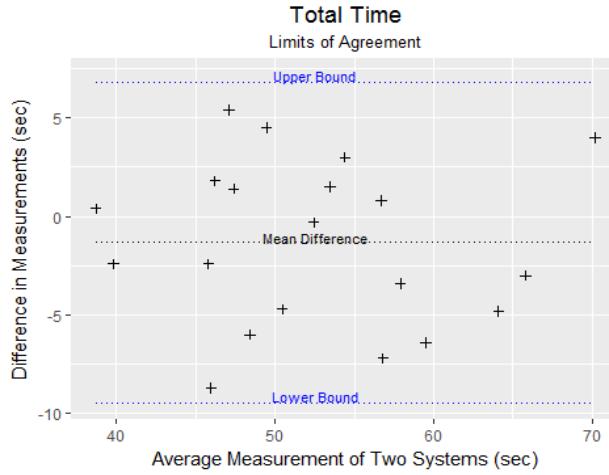


Figure 4. Difference vs mean for Total Reading Time for trial 1 of the first experiment.

Table 4. Measures of agreement for trial 1 of the first experiment.

Measure	Difference (System 2 – System 1)	
	Mean	LOAs
Total Time (sec)	-1.33	(-9.49, 6.84)
Total Fixation Time (sec)	-0.33	(-3.68, 3.01)
Average Fixation Time Length (ms)	1.25	(-93.70, 96.20)
Average Fixation Size (mm)	-6.27	(-65.43, 52.89)
Average Fixation Polyarea (mm ²)	-66.25	(-541.46, 408.96)
Average Inter Fixation Time (ms)	2.66	(-100.43, 105.75)
Saccade Fixation Ratio	0.04	(-0.35, 0.42)
Total Saccades	0.23	(-45.93, 46.40)
Total Saccade Length (mm)	-598.42	(-9957, 8761)
Total Only Saccade Length (mm)	135.98	(-7213, 7485)
Average Saccade Length (mm)	1.82	(-63.00, 66.63)
Total Only Saccade Length (deg)	20.98	(-38.90, 80.86)
Average Saccade Length (deg)	0.25	(-0.47, 0.96)
Total Blinks	1.27	(-8.15, 10.68)
Blink Rate (blinks/min)	4.95	(-28.28, 38.19)
Average Blink Duration (ms)	-18.94	(-199.60, 161.72)
Average Inter Blink Duration (sec)	-0.37	(-7.54, 6.81)
Left Eye Average Pupil Size (mm)	-0.02	(-0.23, 0.19)
Left Eye Pupillary Stress Ratio	1.25	(-9.60, 12.09)

Calculation of the Coefficient of Agreement

Another method to compare measurements of each response parameter made by the two ET/KD systems is to define a coefficient of agreement (COA), such that a coefficient close to 1.0 signifies good agreement and a coefficient near 0.0 signifies poor agreement. For this purpose, we used the appropriate COA from the options proposed by (Haber & Barnhart, 2008). Since the ET/KD systems were new and untested, we wanted to compare the two systems without considering either system as a reference. Therefore, the COA used is defined as follows:

$$COA = \frac{[MAD(X, X') + MAD(Y, Y')]/2}{MAD(X, Y)}$$

where $MAD(X, X')$ is the mean absolute difference of repeated measurements taken with ET/KD System 1, $MAD(Y, Y')$ is the mean absolute difference of repeated measurements taken with ET/KD System 2, and $MAD(X, Y)$ is the mean absolute difference of the first measurements taken on the two systems. This method makes the numerator the average within-system disagreement and the denominator the between-systems disagreement. The four individual trials for each subject were used to determine this COA.

Using the Total Reading Time measure, $MAD(X, X') = 3.565$ seconds, $MAD(Y, Y') = 3.81$ seconds, and $MAD(X, Y) = 3.605$ seconds. This resulted in a COA of 1.023, which indicates good agreement between the two ET/KD systems. Further evaluation of the agreement between the two systems was made by plotting subject-specific $MAD(X, Y)$ versus $[MAD(X, X') + MAD(Y, Y')]/2$, which is illustrated in Figure 5. Points near the 45° line indicate good agreement for that subject, while points far away from the line indicate poor subject-specific agreement.

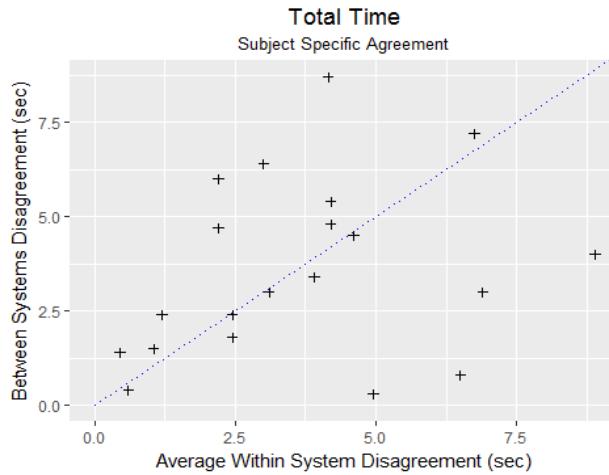


Figure 5. Subject-specific agreement for total response time for the first experiment.

The COAs, as well as the mean, standard deviation, and maximum absolute difference for the response parameters, are shown in Table 5. Subject-specific agreement charts are provided in Appendix B.

Table 5. Measures of agreements of mean differences and COA between the two ET/KD systems for the output parameters for the first experiment.

Measure	Absolute Difference (X, Y)			COA
	mean	sd	max	
Total Time (sec)	3.61	1.53	8.7	1.023
Total Fixation Time (sec)	1.36	1.01	4	1.227
Average Fixation Time Length (ms)	37.51	5.36	128.5	1.018
Average Fixation Size (mm)	24.03	4.26	88.9	1.126
Average Fixation Polyarea (mm ²)	198.86	12.00	622.6	1.078
Average Inter Fixation Time (ms)	38.87	5.79	147.4	1.142
Saccade Fixation Ratio	0.12	0.39	1	1.021
Total Saccades	16.53	4.00	103	1.060
Total Saccade Length (mm)	3483.9	56.12	16488	1.095
Total Only Saccade Length (mm)	2559.9	51.17	13045	1.018
Average Saccade Length (mm)	25.93	4.39	78.6	0.925
Total Only Saccade Length (deg)	27.98	4.84	119.7	0.788
Average Saccade Length (deg)	0.37	0.48	0.9	0.686
Total Blinks	3.57	1.82	14	1.016
Blink Rate (blinks per minute)	12.57	3.44	47.1	1.004
Average Blink Duration (ms)	52.87	8.81	559.9	1.089
Average Inter Blink Duration (sec)	1.81	1.76	17.4	0.990
Left Eye Average Pupil Size (mm)	0.08	0.26	0.3	0.844
Left Eye Pupillary Stress Ratio	4.15	1.91	16.2	1.145

Discussion

The present evaluations of the ET/KD systems were motivated largely because the recent literature widely reports the KDNR test to be a highly sensitive, easy, and reliable sideline sports concussion screener. For example, as of August 2017, a search of Pub Med on the key words “concussion” and “King Devick” produced 52 citations; of which 17 were published during the 8 months of 2017 alone, 11 during 2016, and 10 during 2015. Clearly, during this period there has been a great deal of interest and enthusiasm for the KDNR test as a possible tool for assessing concussion. This literature frequently describes the KDNR test as a measure of eye movements, saccades, or oculomotor behavior, and the KDNR test’s sensitivity to concussion is completely consistent with the well-known sensitivity of eye movements to the disruptive effects of a range of neurological conditions, including traumatic brain injury. The dependence of the KDNR test on eye movements may be logically obvious since a person has to look at the number in order to read it, so the KDNR test seems to offer a simple, intuitive eye movement assessment method using reading behavior measured with paper, pencil, and a stopwatch. The KDNR test is eminently practical and easy to use as a possible screener in a range of sports settings from football and rugby to mixed martial arts (MMA) and boxing, just to name a few contexts in which the KDNR test has been used with extraordinarily good results. A recent report describes its use in a telemedicine context (Vargas et al., 2017). The KDNR test is widely reported to have high sensitivity and specificity for sports-related concussion when the test is used properly; that is, the metric is a difference score from an individually determined baseline. Therefore, the KDNR test ideally needs to be administered before and after the suspected concussive event.

This stipulation may be a complicating factor limiting KDNR use with unpredictable concussive events.

Despite the fact that the KDNR test is widely described as a measure of eye movements or oculomotor behavior, the literature reporting objectively measured eye movements during the KDNR test is extremely thin, consisting at the time of this writing of only four published papers to our knowledge. In two of these papers the measurements were made in volunteers with some neurological condition such as concussion (Rizzo, Hudson, Dai, Birkemeier, et al., 2016) and multiple sclerosis (Hainline et al., 2017), and a third paper reports data from normal control volunteers (Rizzo, Hudson, Dai, Desai, et al., 2016). These three papers are from the same laboratory using the same objective infra-red eye tracking technology. Another research group recorded eye movements during KDNR test performance in normal volunteers while they were exposed to normobaric hypoxia (Stepanek et al., 2013). This study was intended to assess the neurocognitive effects of hypoxia in order to identify possible leading indicators of hypoxia-induced neurocognitive deficits. Thus, it seems that there is very little actual empirical information describing the relationship between eye movements and KDNR test performance. On one hand, this lack of information may be relatively moot for the use of the KDNR test as a practical concussion screener in the field where sensitivity may be more important than specificity. On the other hand, this situation means that a change in the KDNR test performance is widely attributed to changes in oculomotor behavior without actually quantifying or describing the oculomotor behavior and associated changes caused by concussion.

For example, the details of eye movements made while reading can be quantified in a great number of ways. The ET/KD system assessed here listed more than 28; disruption of almost any of these could contribute to a change in KDNR test performance. This level of detail about eye movements is invisible to the KDNR test; unfortunately, this level of detail about eye movements seems to be equally invisible to much of the discussions about the KDNR test, since it is so widely described simply as a test of eye movements. The KDNR test depends on a complex chain of behavior that culminates in an explicit verbalization. Performance is as dependent on oculomotor behavior and retinal and central visual sensory and perceptual processes as it is on motivation, cognition, and attention, as well as speech control and cardiopulmonary physiology.

A recent incidental observation made while integrating the KDNR test with our cardiopulmonary and vascular measures made us sensitive to the assumption that the KDNR test is a measure of eye movements (Temme et al., 2017). During normobaric hypoxia, we found obvious and dramatic alterations in breathing behavior that had nothing to do with eye movements as such and all to do with breath control. These were dramatic and obvious effects, and in retrospect should not have been surprising. The availability of oxygen directly affected the individual's speaking behavior. This incidental observation suggested the possibility that factors which are typically ignored completely can significantly alter KDNR test performance. This possibility further implies that the KDNR test as a measure of the effects of concussion on eye movements rests on the usually tacit assumption that all other components contributing to reading other than oculomotor behavior are constant. In other words, concussion affects nothing but the oculomotor system and its behavior. It is possible, for example, that the normal patterns of speech production are altered by concussion. Hence, it is important to measure eye movements objectively as well as other visual and behavioral components of reading.

The initial intention of these studies was to evaluate the reliability, repeatability, and comparability of the data obtained with the two ET/KD systems. The statistical approach was to use the Bland-Altman graphical limits of agreement method as well as the coefficient of agreement statistics developed by Haber and Barnhart (2008). However, because of the obvious lack of face validity of several of the reported measures (Tables 2 and 3), it seemed pointless to further pursue the questions of reliability, repeatability, and comparability, particularly since the evaluation of reliability, etc., depends on a precision of measurement that these two ET/KD systems do not support.

Eye tracking remains a highly complicated, technologically demanding area of research, which is one of the reasons the KDNR test is so popular; it is easy to use in a chaotic field environment, quick to administer, and simple to score. For these reasons, we were interested in implementing the COTS ET/KD system when it became available. We expected the ET/KD system to be a valid and reliable turnkey system which, unfortunately, turned out not to be the case, as the results of the present evaluations demonstrate. The two ET/KD systems are not usable in that they did not provide meaningful information. These findings are extraordinarily important because they prevented these ET/KD systems from being deployed for research purposes. Their use in research could have produced misleading results, which could have been an immense waste of time and money. When we discussed these results with the manufacturer, the manufacturer did not demur but readily agreed, blaming the system's shortcoming on the third-party commercial eye tracker used in fabricating the system. That is, the manufacturer of the ET/KD systems we purchased acknowledged that the ET/KD systems were inadequate, while also asserting that the third-party eye tracker used in assembling the ET/KD system was the source of the shortcoming. The manufacturer insouciantly added that a better eye tracker is now an upgrade we can purchase, and the new proprietary software version would be provided to us for free if we wished; an offer we declined. We are currently using a different eye tracking technology to support our research.

The most important conclusion of this work is that systematic test and evaluation of new equipment is essential before the new equipment can be trusted for research purposes. This is particularly important when there is limited information, or little to no publication record associated with the equipment, or available information predominantly comes from sources with a conflicting financial interest. Unfortunately, the present study, as necessary as it was, consumed substantial resources and did not appreciably advance science; but at least the study identified a potential impediment and misinformation. We recommend that the manufacturer should report formal evaluation before the device is marketed as a research tool, and failures to do this should also be documented.

References

- Balaban, C., Hoffer, M. E., Szczupak, M., Snapp, H., Crawford, J., Murphy, S., . . . Kiderman, A. (2016). Oculomotor, Vestibular, and Reaction Time Tests in Mild Traumatic Brain Injury. *PLoS One*, 11(9), e0162168. doi:10.1371/journal.pone.0162168
- Barker, T., Russo, S. A., Barker, G., Rice, M. A., Jr., Jeffrey, M. G., Broderick, G., & Craddock, T. J. A. (2017). A case matched study examining the reliability of using ImPACT to assess effects of multiple concussions. *BMC Psychol*, 5(1), 14. doi:10.1186/s40359-017-0184-1
- Barnes, G. R. (2011). Ocular pursuit movements. In S. P. Liversedge, I. D. Gilchrest, & S. Everling (Eds.), *The Oxford Handbook of Eye Movements*. Oxford: Oxford University Press.
- Bland, J. M., & Altman, D. G. (1996). Statistics notes: measurement error proportional to the mean. *Bmj*, 313(7049), 106.
- Boyer, E. O., Portron, A., Bevilacqua, F., & Lorenceau, J. (2017). Continuous Auditory Feedback of Eye Movements: An Exploratory Study toward Improving Oculomotor Control. *Front Neurosci*, 11, 197. doi:10.3389/fnins.2017.00197
- Brahm, K. D., Wilgenburg, H. M., Kirby, J., Ingalla, S., Chang, C. Y., & Goodrich, G. L. (2009). Visual impairment and dysfunction in combat-injured servicemembers with traumatic brain injury. *Optom Vis Sci*, 86(7), 817-825. doi:10.1097/OPX.0b013e3181adff2d
- Broglio, S. P., McCrea, M., McAllister, T., Harezlak, J., Katz, B., Hack, D., . . . Investigators, C. C. (2017). A National Study on the Effects of Concussion in Collegiate Athletes and US Military Service Academy Members: The NCAA-DoD Concussion Assessment, Research and Education (CARE) Consortium Structure and Methods. *Sports Med*. doi:10.1007/s40279-017-0707-1
- Capo-Aponte, J. E., Jorgensen-Wagers, K. L., Sosa, J. A., Walsh, D. V., Goodrich, G. L., Temme, L. A., & Riggs, D. W. (2016). Visual Dysfunctions at Different Stages After Blast and Non-blast Mild Traumatic Brain Injury. *Optom Vis Sci*. doi:10.1097/OPX.0000000000000825
- Cifu, D. X., Wares, J. R., Hoke, K. W., Wetzel, P. A., Gitchel, G., & Carne, W. (2015). Differential eye movements in mild traumatic brain injury versus normal controls. *J Head Trauma Rehabil*, 30(1), 21-28. doi:10.1097/HTR.0000000000000036
- Davenport, N. D. (2016). The Chaos of Combat: An Overview of Challenges in Military Mild Traumatic Brain Injury Research. *Front Psychiatry*, 7, 85. doi:10.3389/fpsyg.2016.00085
- Davies, E. C., Henderson, S., Balcer, L. J., & Galetta, S. L. (2012). Residency training: the King-Devick test and sleep deprivation: study in pre- and post-call neurology residents. *Neurology*, 78(17), e103-106. doi:10.1212/WNL.0b013e318251833d

- Dziemianowicz, M. S., Kirschen, M. P., Pukenas, B. A., Laudano, E., Balcer, L. J., & Galetta, S. L. (2012). Sports-related concussion testing. *Curr Neurol Neurosci Rep*, 12(5), 547-559. doi:10.1007/s11910-012-0299-y
- Galetta, K. M., Barrett, J., Allen, M., Madda, F., Delicata, D., Tennant, A. T., . . . Balcer, L. J. (2011). The King-Devick test as a determinant of head trauma and concussion in boxers and MMA fighters. *Neurology*, 76(17), 1456-1462. doi:10.1212/WNL.0b013e31821184c9
- Galetta, K. M., Brandes, L. E., Maki, K., Dziemianowicz, M. S., Laudano, E., Allen, M., . . . Balcer, L. J. (2011). The King-Devick test and sports-related concussion: study of a rapid visual screening tool in a collegiate cohort. *J Neurol Sci*, 309(1-2), 34-39. doi:10.1016/j.jns.2011.07.039
- Galetta, K. M., Morganroth, J., Moehringer, N., Mueller, B., Hasanaj, L., Webb, N., . . . Balcer, L. J. (2015). Adding Vision to Concussion Testing: A Prospective Study of Sideline Testing in Youth and Collegiate Athletes. *J Neuroophthalmol*. doi:10.1097/WNO.0000000000000226
- Galetta, M. S., Galetta, K. M., McCrossin, J., Wilson, J. A., Moster, S., Galetta, S. L., . . . Master, C. L. (2013). Saccades and memory: baseline associations of the King-Devick and SCAT2 SAC tests in professional ice hockey players. *J Neurol Sci*, 328(1-2), 28-31. doi:10.1016/j.jns.2013.02.008
- Gilchrist, I. (2011). Saccades. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford Handboook of Eye Movements*. Oxford: Oxford University Press.
- Haber, M., & Barnhart, H. X. (2008). A general approach to evaluating agreement between two observers or methods of measurement from quantitative data with replicated measurements. *Statistical Methods in Medical Research*, 17(2), 151-169.
- Hainline, C., Rizzo, J. R., Hudson, T. E., Dai, W., Birkemeier, J., Raynowska, J., . . . Rucker, J. C. (2017). Capturing saccades in multiple sclerosis with a digitized test of rapid number naming. *J Neurol*, 264(5), 989-998. doi:10.1007/s00415-017-8484-1
- Handmaker, H., & Waldorf, R. A. (2013). Comment: the King-Devick test and sports-related concussion: study of a rapid visual screening tool in a collegiate cohort. *J Neurol Sci*, 327(1-2), 80. doi:10.1016/j.jns.2013.01.023
- Heitger, M. H., Jones, R. D., Macleod, A. D., Snell, D. L., Frampton, C. M., & Anderson, T. J. (2009). Impaired eye movements in post-concussion syndrome indicate suboptimal brain function beyond the influence of depression, malingering or intellectual ability. *Brain*, 132(Pt 10), 2850-2870. doi:10.1093/brain/awp181
- Hoge, C. W., McGurk, D., Thomas, J. L., Cox, A. L., Engel, C. C., & Castro, C. A. (2008). Mild traumatic brain injury in U.S. Soldiers returning from Iraq. *N Engl J Med*, 358(5), 453-463. doi:10.1056/NEJMoa072972

- Iverson, G. L., Echemendia, R. J., Lamarre, A. K., Brooks, B. L., & Gaetz, M. B. (2012). Possible lingering effects of multiple past concussions. *Rehabil Res Pract*, 2012, 316575. doi:10.1155/2012/316575
- Kennedy, J. E., Jaffee, M. S., Leskin, G. A., Stokes, J. W., Leal, F. O., & Fitzpatrick, P. J. (2007). Posttraumatic stress disorder and posttraumatic stress disorder-like symptoms and mild traumatic brain injury. *J Rehabil Res Dev*, 44(7), 895-920.
- Khodaee, M., Currie, D. W., Asif, I. M., & Comstock, R. D. (2016). Nine-year study of US high school soccer injuries: data from a national sports injury surveillance programme. *Br J Sports Med*. doi:10.1136/bjsports-2015-095946
- Leigh, R. J., & Zee, D. S. (2006). *The Neurology of Eye Movements*. Oxford: Oxford University Press.
- Leong, D. F., Balcer, L. J., Galetta, S. L., Liu, Z., & Master, C. L. (2014). The King-Devick test as a concussion screening tool administered by sports parents. *J Sports Med Phys Fitness*, 54(1), 70-77.
- Lin, T. P., Adler, C. H., Hentz, J. G., Balcer, L. J., Galetta, S. L., & Devick, S. (2014). Slowing of number naming speed by King-Devick test in Parkinson's disease. *Parkinsonism Relat Disord*, 20(2), 226-229. doi:10.1016/j.parkreldis.2013.10.009
- Lin, T. P., Rigby, H., Adler, J. S., Hentz, J. G., Balcer, L. J., Galetta, S. L., . . . Adler, C. H. (2014). Abnormal Visual Contrast Acuity in Parkinson's Disease. *J Parkinsons Dis*. doi:10.3233/JPD-140470
- Liversedge, S., Gilchrist, I., & Everling, S. (2011). *The Oxford handbook of eye movements*: Oxford University Press.
- Maguen, S., Lau, K. M., Madden, E., & Seal, K. (2012). Relationship of screen-based symptoms for mild traumatic brain injury and mental health problems in Iraq and Afghanistan veterans: Distinct or overlapping symptoms? *J Rehabil Res Dev*, 49(7), 1115-1126.
- Marion, D. W., Curley, K. C., Schwab, K., & Hicks, R. R. (2011). Proceedings of the military mTBI Diagnostics Workshop, St. Pete Beach, August 2010. *J Neurotrauma*, 28(4), 517-526. doi:10.1089/neu.2010.1638
- Marshall, S. W., Guskiewicz, K. M., Shankar, V., McCrea, M., & Cantu, R. C. (2015). Epidemiology of sports-related concussion in seven US high school and collegiate sports. *Inj Epidemiol*, 2(1), 13. doi:10.1186/s40621-015-0045-4
- Martinez-Conde, S., & Macknik, S. L. (2011). Microsaccades. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye Movements*. Oxford: Oxford University Press.

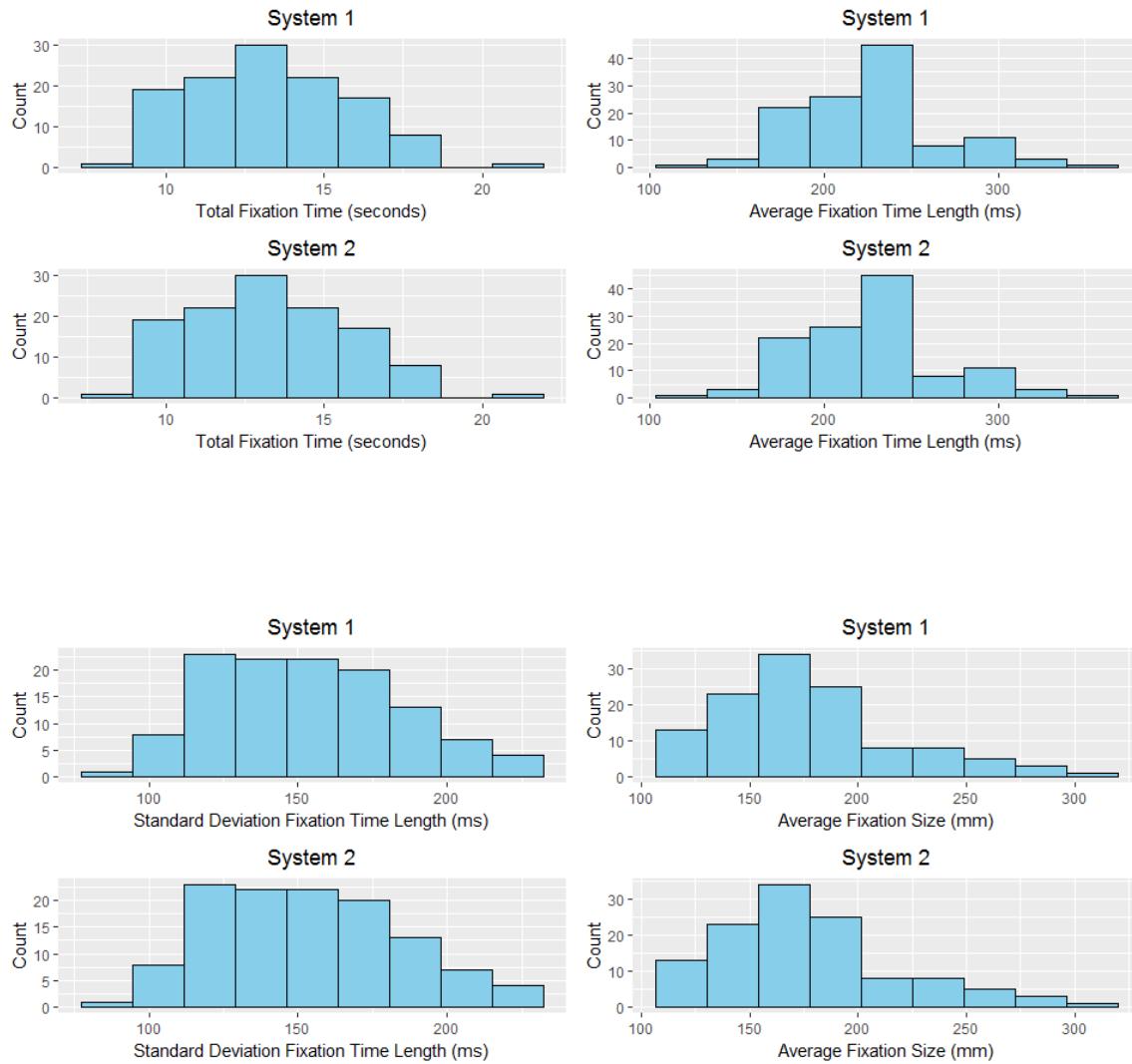
- Molloy, J. H., Murphy, I., & Gissane, C. (2017). The King-Devick (K-D) test and concussion diagnosis in semi-professional rugby union players. *J Sci Med Sport*. doi:10.1016/j.jsams.2017.02.002
- Moster, S., Wilson, J. A., Galetta, S. L., & Balcer, L. J. (2014). The King-Devick (K-D) test of rapid eye movements: a bedside correlate of disability and quality of life in MS. *J Neurol Sci*, 343(1-2), 105-109. doi:10.1016/j.jns.2014.05.047
- Munce, T. A., Dorman, J. C., Odney, T. O., Thompson, P. A., Valentine, V. D., & Bergeron, M. F. (2014). Effects of youth football on selected clinical measures of neurologic function: a pilot study. *J Child Neurol*, 29(12), 1601-1607. doi:10.1177/0883073813509887
- Padula, W. V., Capo-Aponte, J. E., Padula, W. V., Singman, E. L., & Jenness, J. (2017). The consequence of spatial visual processing dysfunction caused by traumatic brain injury (TBI). *Brain Inj*, 1-12. doi:10.1080/02699052.2017.1291991
- Regasa, L. E., Thomas, D. M., Gill, R. S., Marion, D. W., & Ivins, B. J. (2016). Military Deployment May Increase the Risk for Traumatic Brain Injury Following Deployment. *J Head Trauma Rehabil*, 31(1), E28-35. doi:10.1097/HTR.0000000000000155
- Rizzo, J. R., Hudson, T. E., Dai, W., Birkemeier, J., Pasculle, R. M., Selesnick, I., . . . Rucker, J. C. (2016). Rapid number naming in chronic concussion: eye movements in the King-Devick test. *Ann Clin Transl Neurol*, 3(10), 801-811. doi:10.1002/acn3.345
- Rizzo, J. R., Hudson, T. E., Dai, W., Desai, N., Yousefi, A., Palsana, D., . . . Rucker, J. C. (2016). Objectifying eye movements during rapid number naming: Methodology for assessment of normative data for the King-Devick test. *J Neurol Sci*, 362, 232-239. doi:10.1016/j.jns.2016.01.045
- Schmid, K. E., & Tortella, F. C. (2012). The diagnosis of traumatic brain injury on the battlefield. *Front Neurol*, 3, 90. doi:10.3389/fneur.2012.00090
- Seal, K. H., Bertenthal, D., Samuelson, K., Maguen, S., Kumar, S., & Vasterling, J. J. (2016). Association between mild traumatic brain injury and mental health problems and self-reported cognitive dysfunction in Iraq and Afghanistan Veterans. *J Rehabil Res Dev*, 53(2), 185-198. doi:10.1682/JRRD.2014.12.0301
- Stepanek, J., Cocco, D., Pradhan, G. N., Smith, B. E., Bartlett, J., Studer, M., . . . Cevette, M. J. (2013). Early detection of hypoxia-induced cognitive impairment using the King-Devick test. *Aviat Space Environ Med*, 84(10), 1017-1022.
- Temme, L. A., St Onge, P., Adams, M., Still, D. L., Statz, J. K., & Williams, S. T. (2017). A Novel, Inexpensive Method to Monitor, Record, and Analyze Breathing Behavior During Normobaric Hypoxia Generated by the Reduced Oxygen Breathing Device. *Mil Med*, 182(S1), 210-215. doi:10.7205/MILMED-D-16-00053

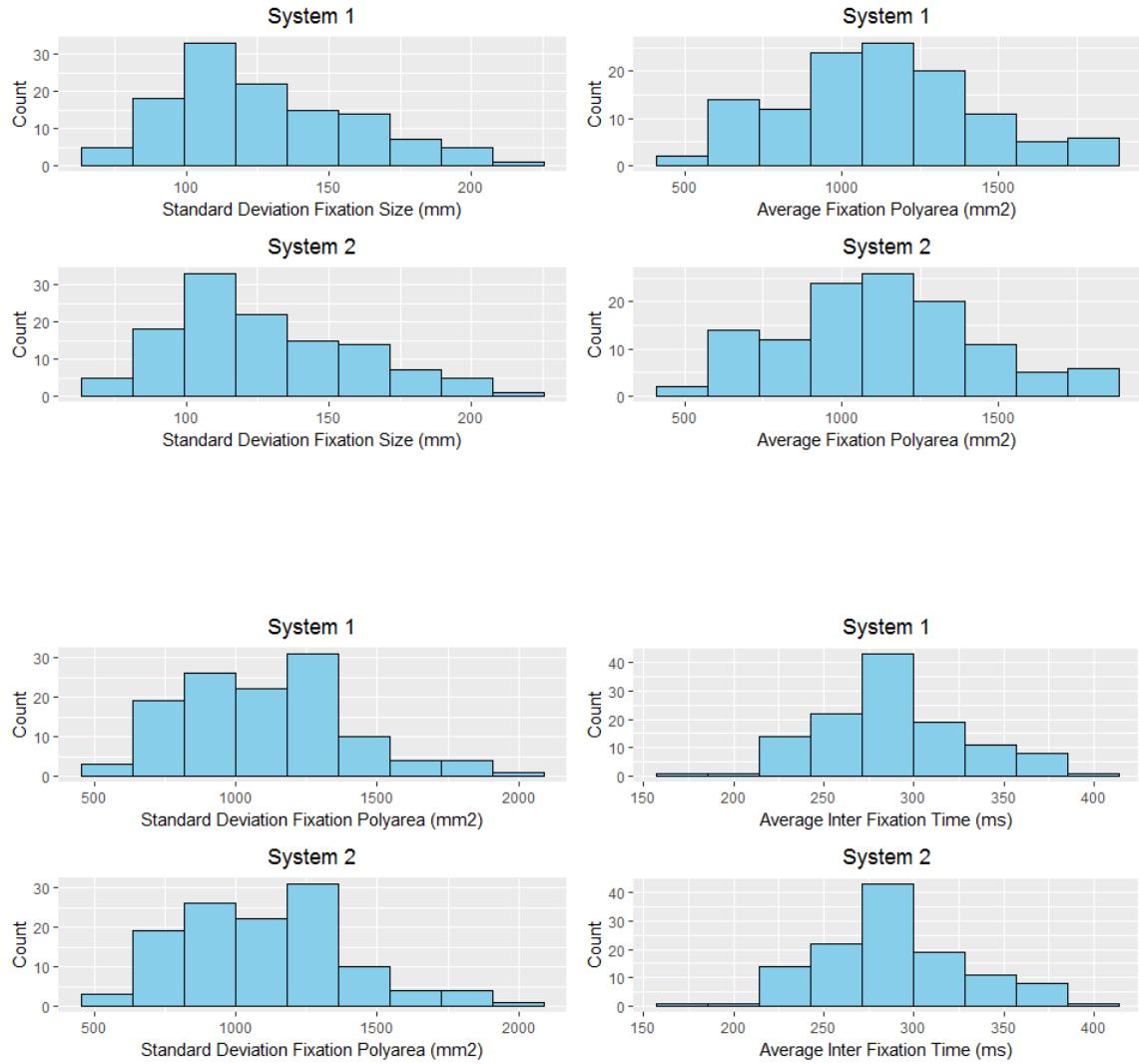
Vargas, B. B., Shepard, M., Hentz, J. G., Kutyreff, C., Hershey, L. G., & Starling, A. J. (2017). Feasibility and accuracy of teleconcussion for acute evaluation of suspected concussion. *Neurology*, 88(16), 1580-1583.

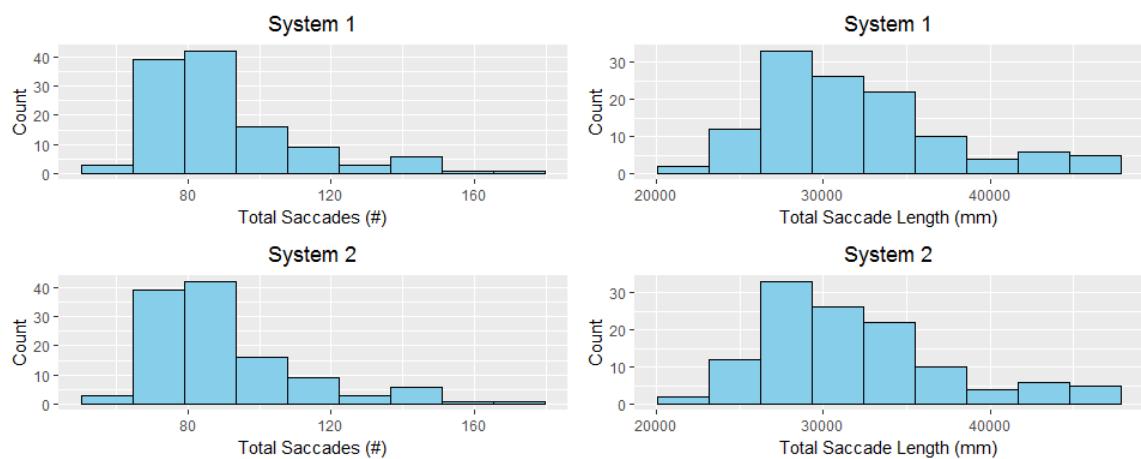
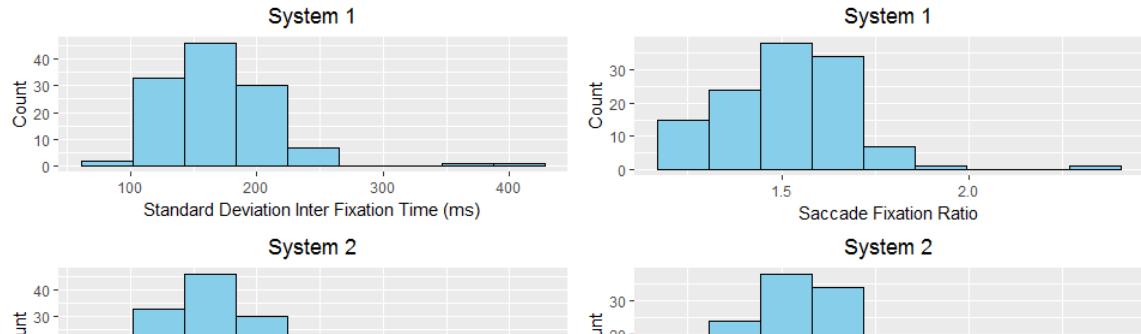
Ventura, R. E., Balcer, L. J., Galetta, S. L., & Rucker, J. C. (2016). Ocular motor assessment in concussion: Current status and future directions. *J Neurol Sci*, 361, 79-86.
doi:10.1016/j.jns.2015.12.010

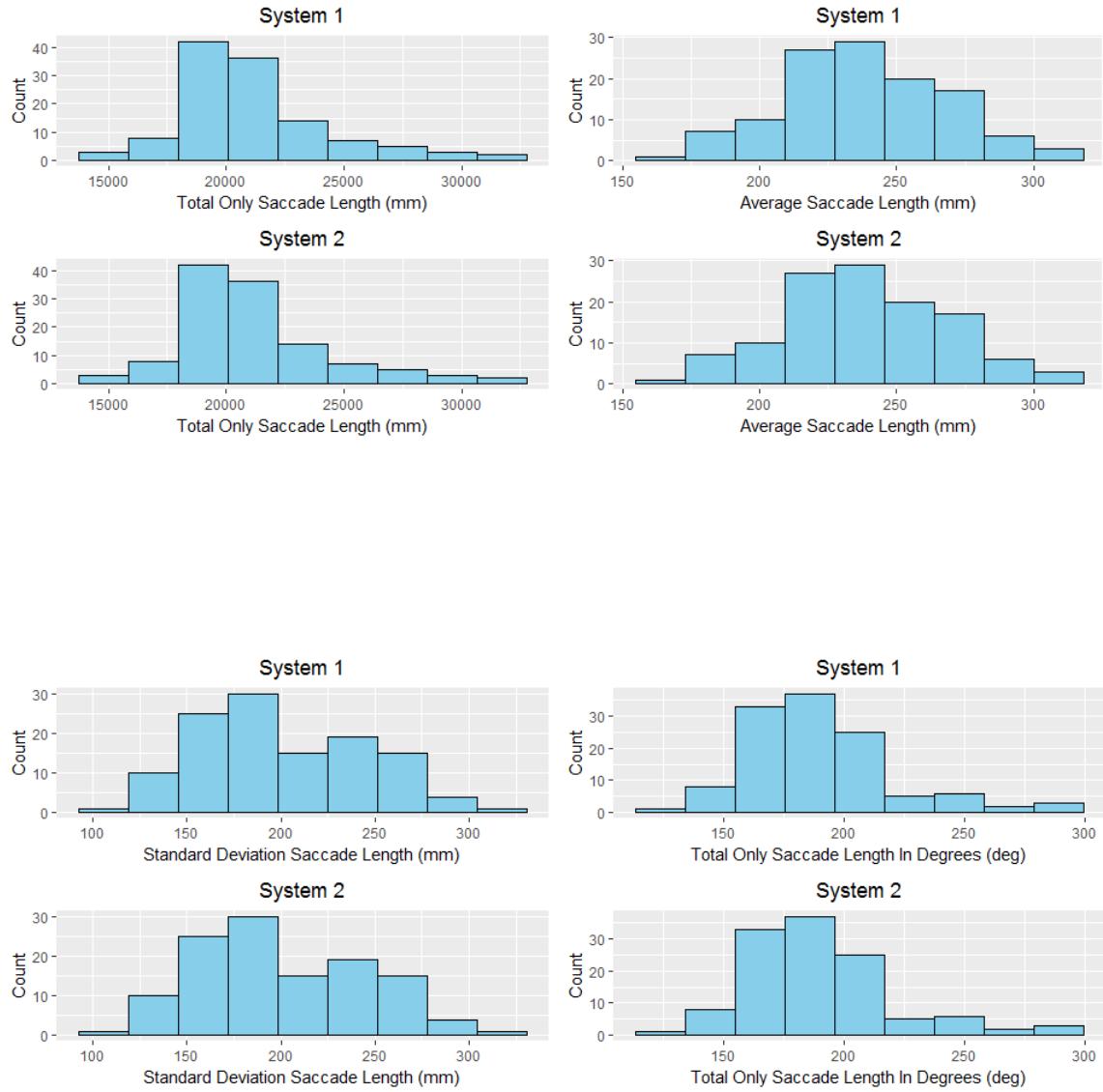
This page is intentionally blank.

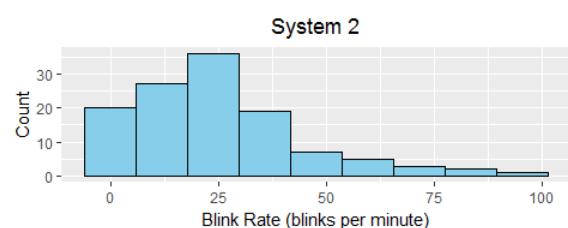
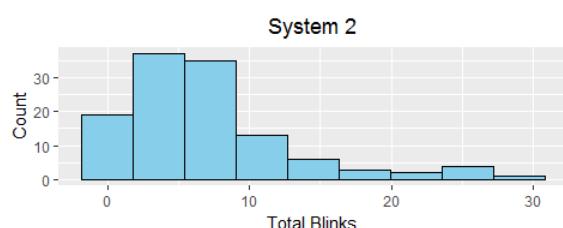
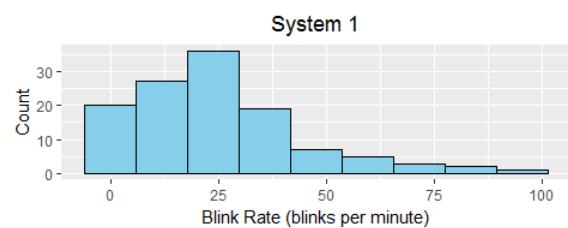
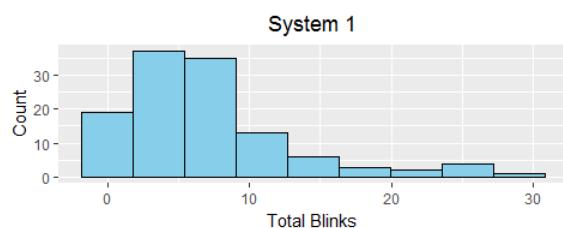
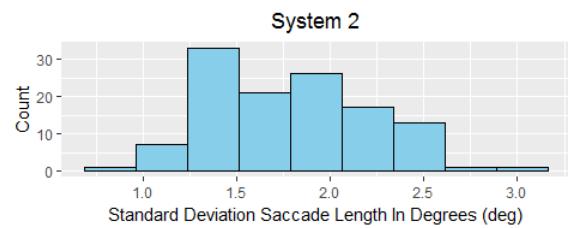
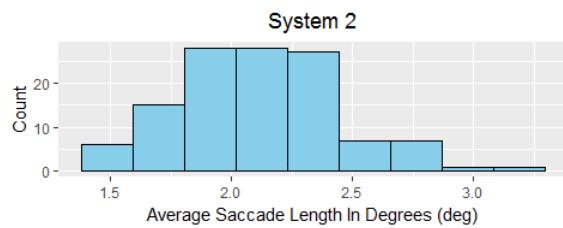
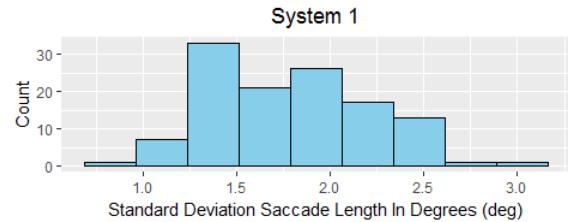
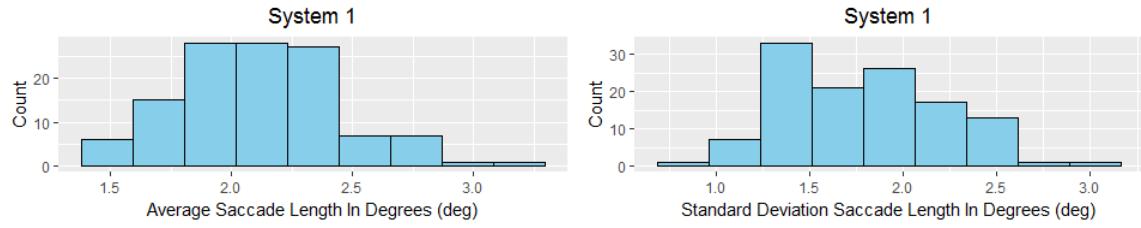
Appendix A. Bland-Altman plots for the response parameters

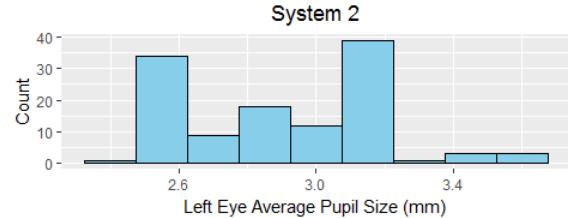
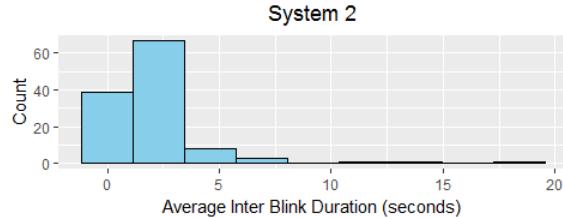
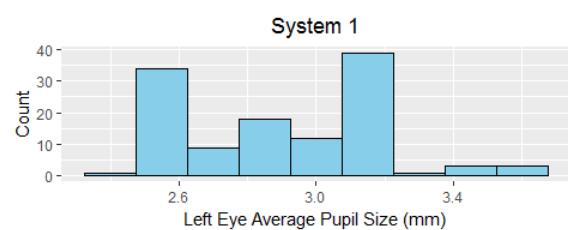
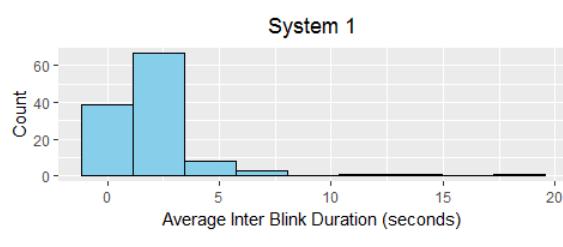
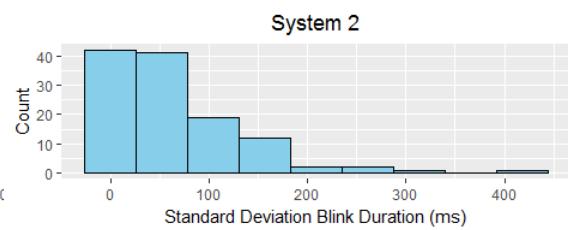
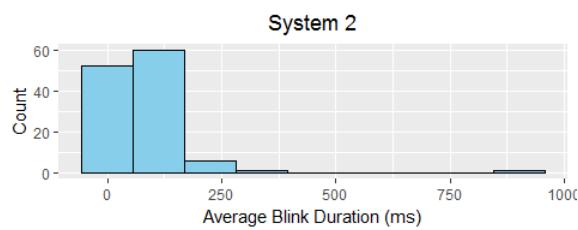
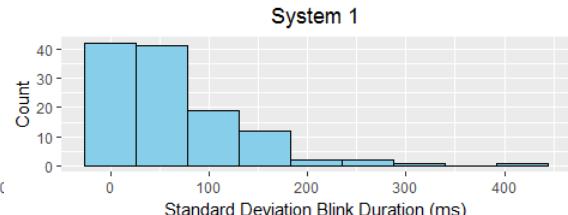
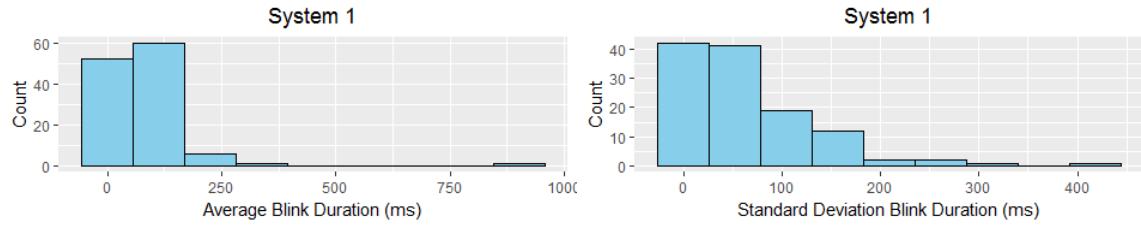


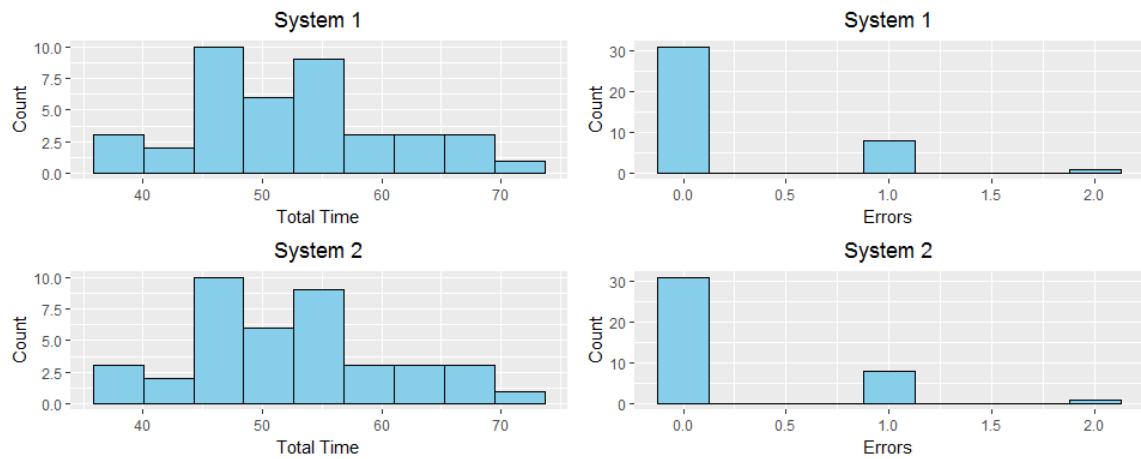
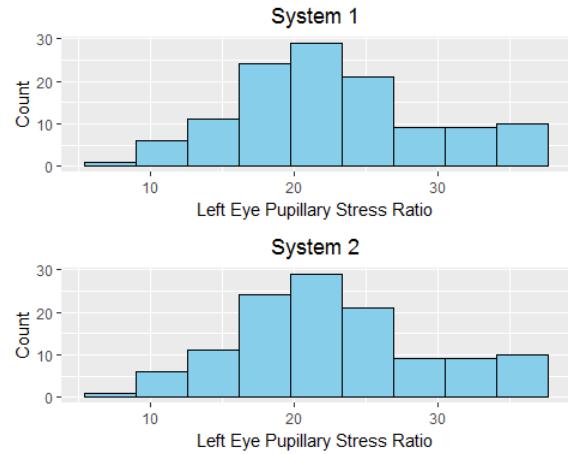
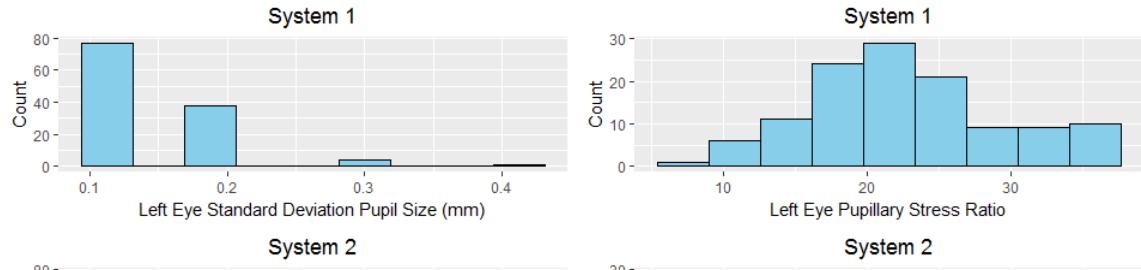




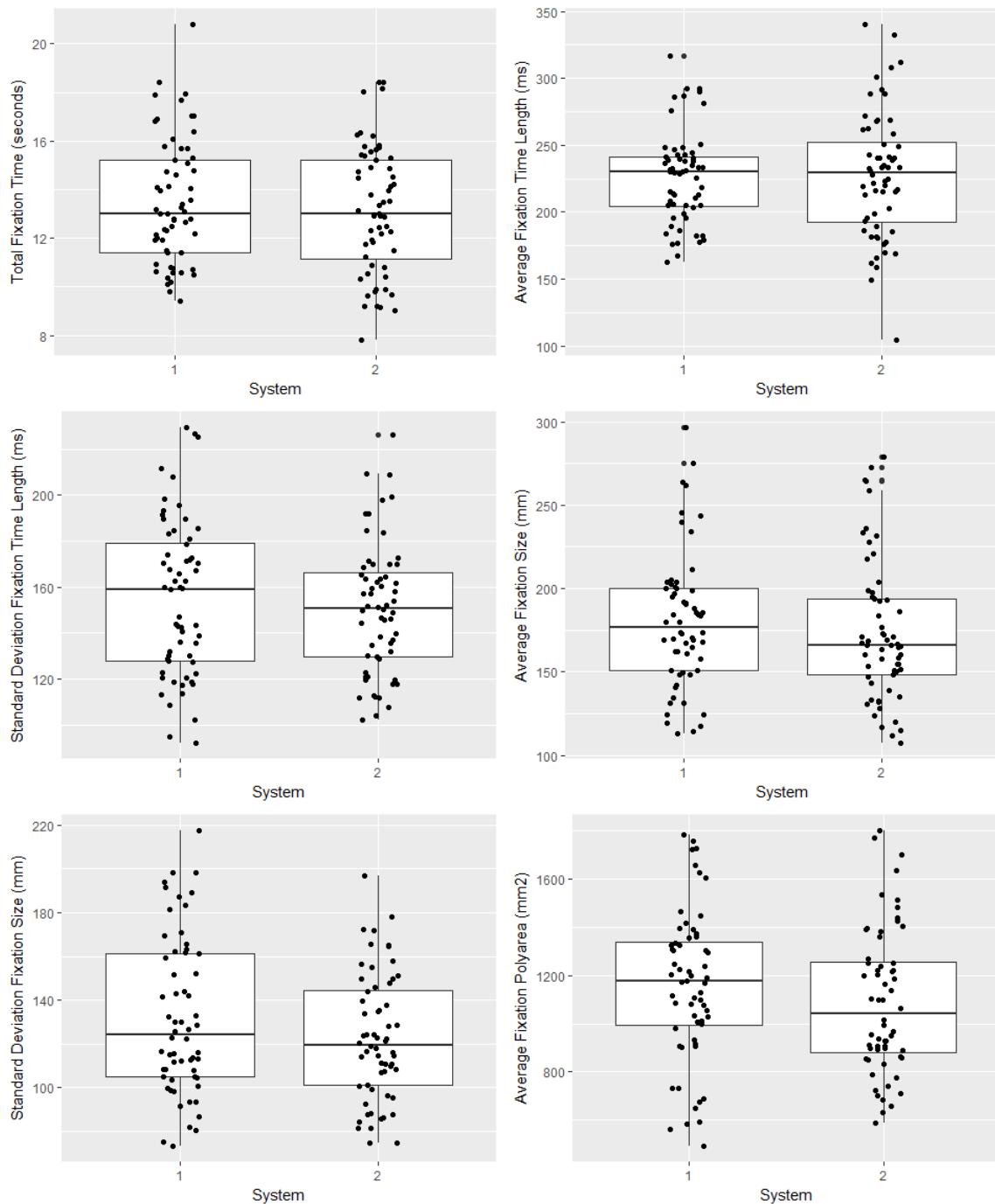


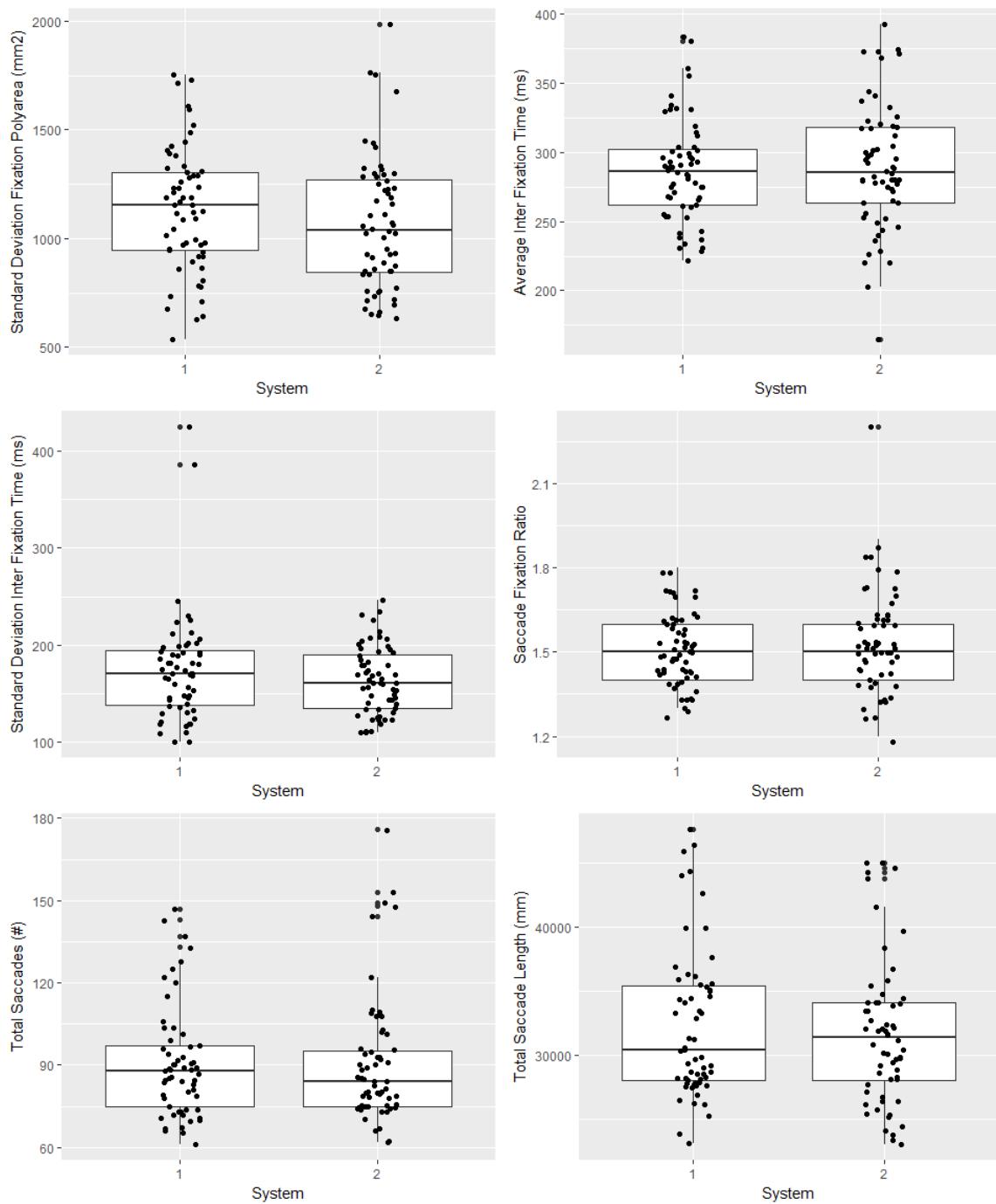


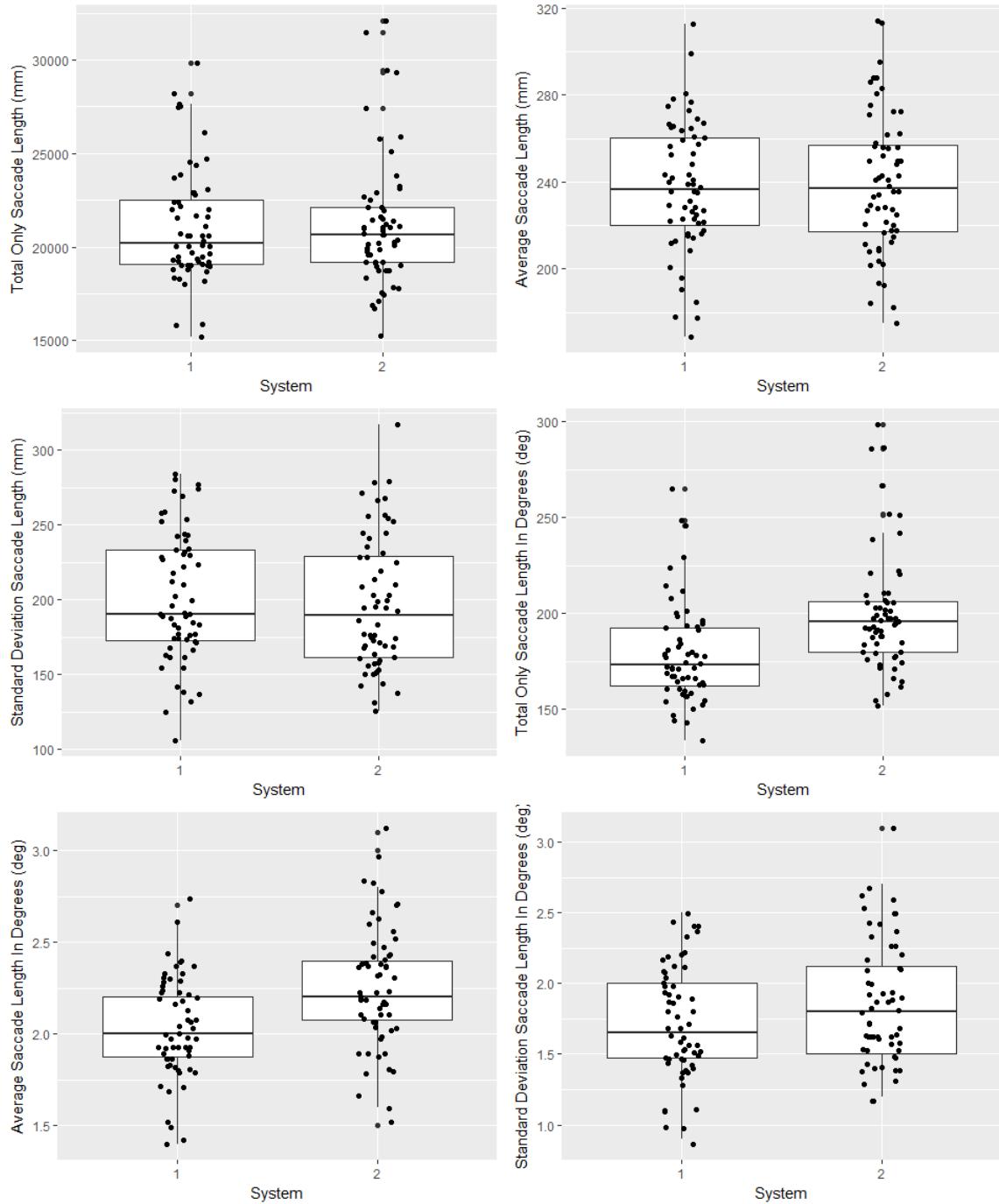


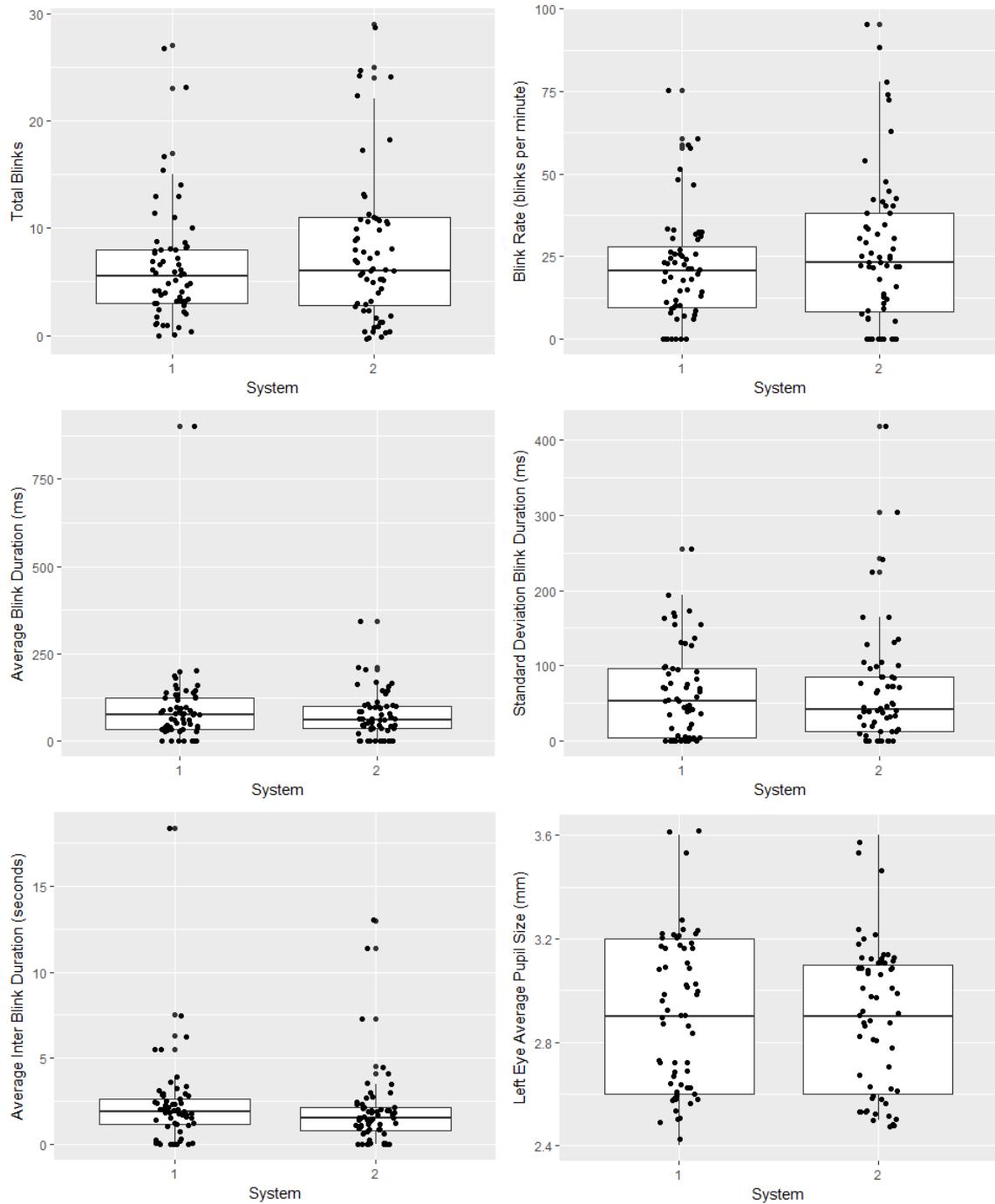


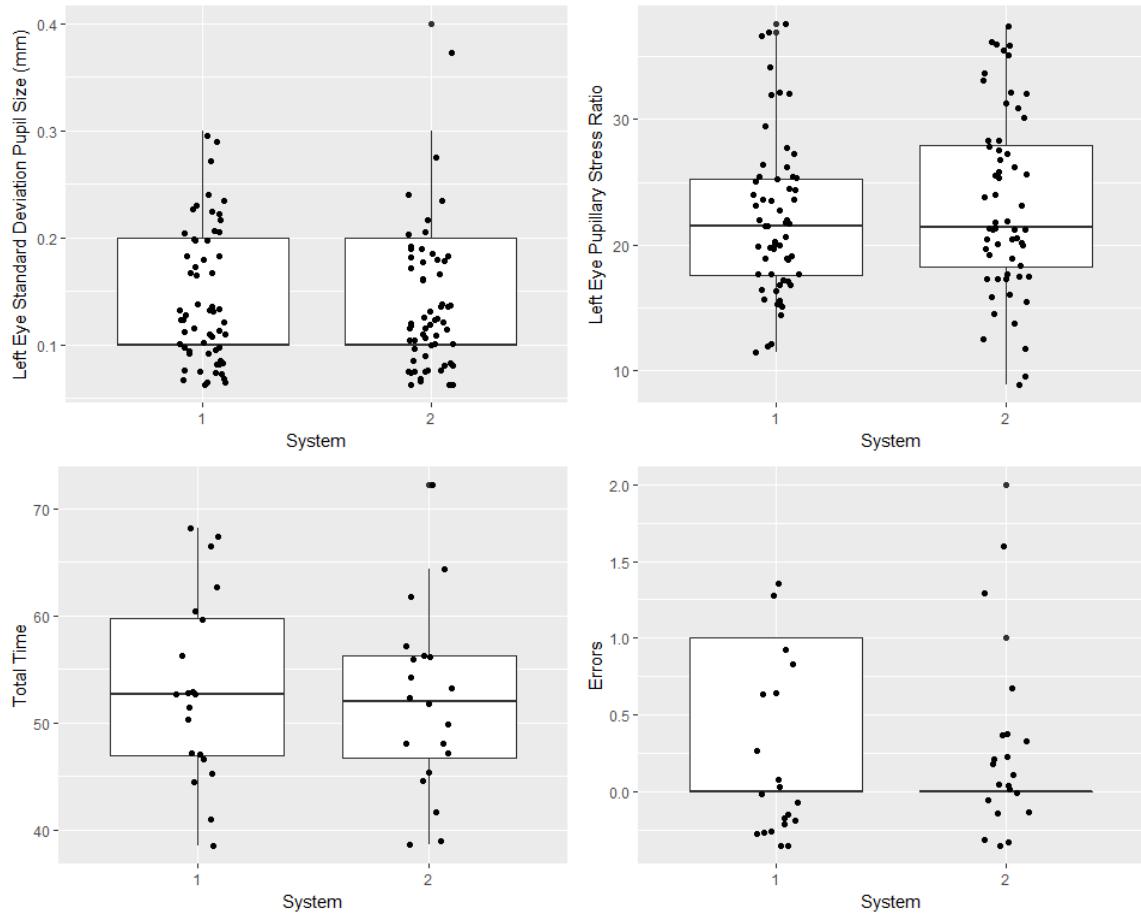
Appendix B. Subject-specific agreement charts











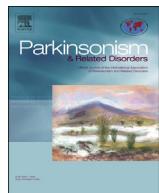


**Department of the Army
U.S. Army Aeromedical Research Laboratory
Fort Rucker, Alabama 36362-0577**

www.usaarl.army.mil



U.S. Army Medical Research and Materiel Command



Short communication

Slowing of number naming speed by King–Devick Test in Parkinson's disease



Tanya P. Lin ^a, Charles H. Adler ^{a,*}, Joseph G. Hentz ^b, Laura J. Balcer ^c, Steven L. Galetta ^c, Steve Devick ^d

^a Department of Neurology, Mayo Clinic College of Medicine, Scottsdale, AZ, USA

^b Department of Biostatistics, Mayo Clinic College of Medicine, Scottsdale, AZ, USA

^c Department of Neurology, NYU School of Medicine, New York, NY, USA

^d King–Devick Test LLC, Oakbrook Terrace, IL, USA

ARTICLE INFO

Article history:

Received 8 July 2013

Received in revised form

27 September 2013

Accepted 10 October 2013

Keywords:

Parkinson's disease

Eye movements

Cognitive function

ABSTRACT

Background: The King–Devick (KD) test measures the speed of rapid number naming, and is postulated to require fast eye movements, attention, language, and possibly other aspects of cognitive functions. While used in multiple sports concussion studies, it has not been applied to the field of movement disorders.

Methods: Forty-five Parkinson's disease (PD), 23 essential tremor (ET), and 65 control subjects were studied. Subjects performed two trials of reading out loud single-digit numbers separated by varying spacing on three test cards that were of different formats. The sum time of the faster trial was designated the KD score and compared across the three groups.

Results: PD patients had higher (worse) KD scores, with longer reading times compared to ET and control subjects (66 s vs. 49 s vs. 52 s, $p < 0.001$, adjusting for age and gender). No significant difference was found between ET and control ($\Delta = -3$ s, 95% CI: -10 to 4).

Conclusions: This is the first study of the King–Devick Test in Parkinson's disease. PD patients were found to have a slower rapid number naming speed compared to controls. This test may be a simple and rapid bedside tool for quantifying correlates of visual and cognitive function in Parkinson's disease.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Non-motor symptoms are well recognized in Parkinson's disease (PD) patients, even early in their disease process. Although clinicians routinely assess many non-motor symptoms such as those involving cognition, mood, and sleep, visual complaints are rarely evaluated quantitatively. PD patients frequently complain of blurred vision, double vision, and difficulty with reading. The underlying cause of these visual symptoms is not always well understood as many of these patients have normal or near normal visual acuity.

In addition to limb and axial motor symptoms, PD patients have ocular motor abnormalities. Studies have reported abnormal visual scanning [1], saccadic eye movement impairment [1,2], and deficiency in eye movement planning and target anticipation [2,3] in PD. Existing literature mainly focuses on laboratory recordings using electro-oculography or video-based eye tracking systems to examine saccades, antisaccades, ocular pursuit, and fixation tasks as quantitative parameters for ocular motor evaluation. The study

of eye movements is important because it provides powerful insights into neural control of volitional and reflexive processes [2]. However, since specialized equipment is required, the current eye movement studies are often done in research setting instead of clinical practice, and patient access may be limited.

The objective of our study was to find an easy-to-use quantitative bedside tool to evaluate visual function of PD patients. The King–Devick (KD) test is a rapid number naming test that requires saccadic eye movements to perform, and is postulated to also capture attention, language, and possibly other aspects of cognitive function according to recent sports-related concussion research [4,5]. This test takes about 2 min to perform and can be done in a routine office visit. To our knowledge, this is the first study using the KD test to evaluate ocular motor function of Parkinson's disease patients.

2. Patients and methods

2.1. Subjects

Forty-five PD, 23 essential tremor (ET), and 65 control subjects were studied. Subjects were tested in the movement disorders

* Corresponding author.

E-mail address: cadler@mayo.edu (C.H. Adler).

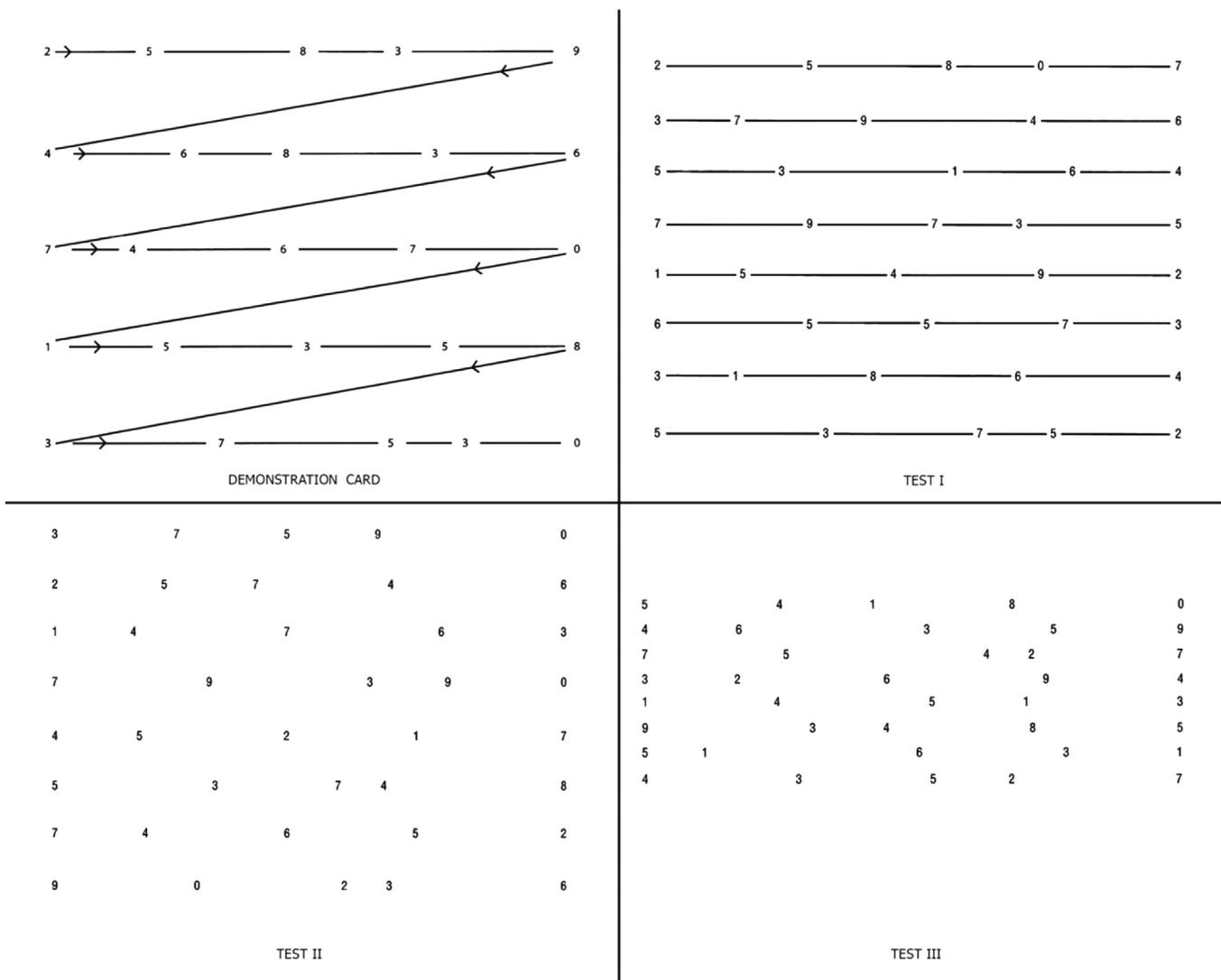


Fig. 1. Demonstration and test cards for the King–Devick (KD) Test for rapid number naming. To perform the KD test, participants were asked to read the numbers on each card from left to right as quickly as possible without making any errors. Following completion of the demonstration card (upper left), subjects are then asked to read each of the three test cards in the same manner. The times required to complete each card were recorded in seconds using a stopwatch. The sum of the three test card time scores constitutes the summary score for the entire test, the KD time score.

clinic at Mayo Clinic Arizona, or in the Arizona Study of Aging and Neurodegenerative Disorders (AZSAND) by the Arizona PD Consortium/Banner Sun Health Research Institute Brain and Body Donation Program. All participants signed written informed consents approved by the institutional IRBs. PD was clinically diagnosed according to the UK Brain Bank criteria, i.e., the presence of two of three cardinal features (resting tremor, bradykinesia, and rigidity) without atypical features (including early falls, early dementia, gaze palsy, early marked autonomic disturbance, fluctuating confusional states) or obvious secondary cause (such as stroke, drugs, toxins, arthritis). Subjects with dementia and those with a history of macular degeneration, glaucoma, untreated cataracts, or blindness were excluded from the study. Subjects were permitted to wear corrective lenses. The Unified Parkinson's Disease Rating Scale (UPDRS) was performed for all subjects.

2.2. King–Devick test

The King–Devick test consists of a demonstration card and three test cards with a series of single-digit numbers separated by varying spacing, either with or without a connecting line between

numbers (Fig. 1). Participants started with a demonstration card and read the numbers out loud from left to right and top to bottom, as quickly as possible and without making errors. The three test cards were then read in order in two consecutive trials. The sum time of the three test cards from the faster trial was designated the final test score. Accuracy was important; if errors were not immediately corrected, the score was not valid. The mean KD scores from the three groups were compared by single factor analysis of variance. Adjusted means were compared by using a generalized linear model with terms for age and gender.

3. Results

PD subjects were younger (mean \pm SD, 73.1 ± 8.4 years) compared to ET (80.8 ± 4.8 years) and control subjects (80.0 ± 6.3 years). There were more men in the PD group (67% PD vs. 43% ET vs. 32% Control). Disease duration for the PD group was 7.2 ± 5.9 years (range 1–25 years). The mean UPDRS part III score was 20.5 ± 11.8 , and Hoehn and Yahr staging was 2.2 ± 0.7 .

The mean KD score for PD (63 ± 18 s) was higher (worse) compared to scores for ET (51 ± 13 s) and control subjects

(53 ± 13 s), $p = 0.001$ (Fig. 2). After adjusting for age and gender, the mean KD score remained higher in PD (66 s) than ET (49 s) and control (52 s), $p < 0.001$. No significant difference was found between ET and control ($\Delta = -3$ s, 95% CI: -10 to 4). For the PD group, the correlation between KD scores and disease duration showed a linear correlation coefficient of $r = 0.25$ (95% CI -0.06 to 0.51, $p = 0.11$). Similarly for UPDRS part III scores, the correlation coefficient was 0.23 (95% CI -0.07 to 0.49, $p = 0.13$). Because of the small sample size of the PD group, the significance of these correlations was difficult to assess.

4. Discussion

In addition to motor symptoms, Parkinson's disease patients often develop a wide range of visual problems during their disease process. A change in their vision may be due to alterations in various factors including but not limited to visual acuity, contrast sensitivity, color discrimination, eye movements, visual perception, visuospatial processing, and visual hallucination. Some of these parameters such as visual acuity, visual perception, and visuospatial processing can be assessed clinically by routine ophthalmological exam or neuropsychological testing, while others require specialized laboratory equipments. The evaluation of eye movements is usually done by formal laboratory testing using electro-oculography or video-based eye tracking systems. The purpose of our study was to determine whether the King–Devick test, a short test based on rapid number naming, could be used as a bed-side tool to quantify ocular motor function and its associated cognitive correlate.

Abnormal eye movements in PD have been well described by formal laboratory studies. A meta-analysis of 47 studies revealed prolongation of saccadic latency in Parkinson's disease patients compared to controls [6]. Both visual guided and memory guided saccades were impaired in PD patients [7,8]. A reduction in saccadic amplitude could be observed early in the disease course [8], while increased saccade latency did not occur until later stages [7] and was associated with cognitive decline [8].

According to the current model of neurophysiology, accurate rapid eye movement control is generated by integrating cortical input from the frontal, supplemental, and parietal eye fields through circuits of thalamus, basal ganglia, and the superior colliculus [9]. Some of these neuronal pathways are involved in the neurodegeneration of Parkinson's disease. In addition, PD targets the fronto-parietal networks of attention and executive function. The close anatomical relationship between ocular motor, visual,

and higher cognitive functions may explain the visuospatial, visuoperceptual, and executive function impairment in PD patients [10]. Two recent reports found that impairment in several saccadic eye movement parameters including error rates, visual exploration strategies, prolonged fixation time, and saccadic latency correlated with cognitive dysfunction ranging from mild cognitive impairment to dementia in PD patients [9,11].

In this proof of concept study using the King–Devick test as a bedside tool, PD patients performed worse on rapid number naming compared to control subjects and patients with essential tremor. The mean KD time in PD patients was 10 s longer than controls and 12 s longer than ET patients. The advantage of the King–Devick test is that it is quantitative, brief, and easy to use. It takes about 2 min to perform and can be done in the clinic setting. However, in addition to measuring saccadic eye movements, the KD test may also be capturing attention, language, and possibly other suboptimal brain function [5]. In order to identify the contribution to the KD score by each individual component, data is needed to cross reference the KD test with laboratory eye movement tests and neuropsychological evaluation. Galetta and colleagues studied a group of professional hockey players and found cross-sectional associations of worse KD scores with worse scores for immediate memory in the Standardized Assessment of Concussion (SAC) [4]. These findings potentially implicate that the dorsolateral prefrontal cortex is involved in the generation of saccades and working memory, thus linking cognition to eye movements [4].

Another potential confounding factor is that bradykinesia may influence the speed of rapid number naming. Mixed results have been reported on correlation of ocular motor function and limb/axial motor function. Some of these studies found that eye movements correlated well with UPDRS motor scores [9,11,12], while others did not [3,10]. In our study, KD scores did not show significant correlation with PD duration or the UPDRS motor scores, possibly due to the small sample size and heterogeneous factors such as attention, medications, mood, and cognitive function that may play a role in KD testing outcome.

In conclusion, the King–Devick test may provide a simple, brief, and quantitative way to assess visual function and its cognitive correlate that routine clinical evaluation and eye exam do not capture. It does not require specialized equipment and can be done as part of a routine office visit. Further study is needed to better define the relationship between the King–Devick test and disease progression as well as cognition.

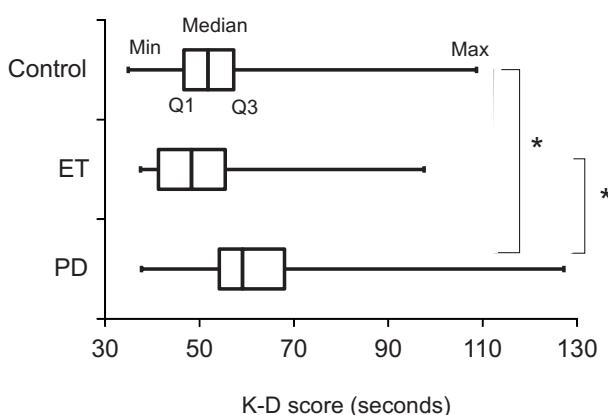
Acknowledgment

The Arizona Parkinson's Disease Consortium and the Brain and Body Donation Program are supported by the National Institute of Neurological Disorders and Stroke (U24 NS072026 National Brain and Tissue Resource for Parkinson's Disease and Related Disorders), the National Institute on Aging (P30 AG19610 Arizona Alzheimer's Disease Core Center), the Arizona Department of Health Services (contract 211002, Arizona Alzheimer's Research Center), the Arizona Biomedical Research Commission (contracts 4001, 0011, 05-901 and 1001 to the Arizona Parkinson's Disease Consortium), and Mayo Clinic Foundation.

References

- [1] Matsumoto H, Terao Y, Furabayashi T, Yugeta A, Fukuda H, Emoto M, et al. Small saccades restrict visual scanning area in Parkinson's disease. *Mov Disord* 2011 Aug;1:26:1619–26.
- [2] Chan F, Armstrong IT, Pari G, Riopelle RJ, Munoz DP. Deficits in saccadic eye-movement control in Parkinson's disease. *Neuropsychologia* 2005;43:784–96.
- [3] Helmchen C, Pohlmann J, Trillenberg P, Lencer R, Graf J, Sprenger A. Role of anticipation and prediction in smooth pursuit eye movement control in Parkinson's disease. *Mov Disord* 2012 Jul;27:1012–8.

Fig. 2. Box–Whisker plot of KD scores. The lines in the box represent the medians, and boxes delineate the interquartile range (25th–75th percentiles). Whiskers represent the range of observations. * $p = 0.001$, single factor analysis of variance.



- [4] Galetta MS, Galetta KM, McCrossin J, Wilson JA, Moster S, Galetta SL, et al. Saccades and memory: baseline associations of the King-Devick and SCAT2 SAC tests in professional ice hockey players. *J Neurol Sci* 2013 May 15;328:28–31.
- [5] Galetta KM, Barrett J, Allen M, Madda F, Delicata D, Tennant AT, et al. The King-Devick test as a determinant of head trauma and concussion in boxers and MMA fighters. *Neurology* 2011 Apr 26;76:1456–62.
- [6] Chambers JM, Prescott TJ. Response times for visually guided saccades in persons with Parkinson's disease: a meta-analytic review. *Neuropsychologia* 2010 Mar;48:887–99.
- [7] Terao Y, Fukuda H, Yugeta A, Hikosaka O, Nomura Y, Segawa M, et al. Initiation and inhibitory control of saccades with the progression of Parkinson's disease – changes in three major drives converging on the superior colliculus. *Neuropsychologia* 2011 Jun;49:1794–806.
- [8] Anderson TJ, MacAskill MR. Eye movements in patients with neurodegenerative disorders. *Nat Rev Neurol* 2013 Feb;9:74–85.
- [9] Archibald NK, Hutton SB, Clarke MP, Mosimann UP, Burn DJ. Visual exploration in Parkinson's disease and Parkinson's disease dementia. *Brain* 2013 Mar;136:739–50.
- [10] Perneczky R, Ghosh BCP, Hughes L, Carpenter RHS, Barker RA, Rowe JB. Saccadic latency in Parkinson's disease correlates with executive function and brain atrophy, but not motor severity. *Neurobiol Dis* 2011 Jul;43:79–85.
- [11] Macaskill MR, Graham CF, Pitcher TL, Myall DJ, Livingston L, van Stockum S, et al. The influence of motor and cognitive impairment upon visually-guided saccades in Parkinson's disease. *Neuropsychologia* 2012 Dec;50:3338–47.
- [12] Marino S, Lanzafame P, Sessa E, Bramanti A, Bramanti P. The effect of L-Dopa administration on pursuit ocular movements in suspected Parkinson's disease. *Neurol Sci* 2010 Jun;31:381–5.

Article

King-Devick Test Performance and Cognitive Dysfunction after Concussion: A Pilot Eye Movement Study

Doria M. Gold ¹, John-Ross Rizzo ^{1,2,3,4}, Yuen Shan Christine Lee ², Amanda Childs ², Todd E. Hudson ^{1,2}, John Martone ¹, Yuka K. Matsuzawa ², Felicia Fraser ⁵, Joseph H. Ricker ², Weiwei Dai ^{1,6}, Ivan Selesnick ⁶, Laura J. Balcer ^{1,7,8}, Steven L. Galetta ^{1,8} and Janet C. Rucker ^{1,8,*}

- ¹ Department of Neurology, New York University Grossman School of Medicine, New York, NY 10016, USA; doria.gold@nyulangone.org (D.M.G.); johnrossrizzo@gmail.com (J.-R.R.); hudson.te@gmail.com (T.E.H.); jnm2154@columbia.edu (J.M.); wd471@nyu.edu (W.D.); laura.balcer@nyulangone.org (L.J.B.); steven.galetta@nyulangone.org (S.L.G.)
- ² Department of Physical Medicine & Rehabilitation, New York University Grossman School of Medicine, New York, NY 10016, USA; yuenshan.lee@nyulangone.org (Y.S.C.L.); amanda.childs@nyulangone.org (A.C.); yuka.matsuzawa@nyulangone.org (Y.K.M.); joseph.ricker@nyulangone.org (J.H.R.)
- ³ Department of Mechanical & Aerospace Engineering, New York University Tandon School of Engineering, New York, NY 11201, USA
- ⁴ Department of Biomedical Engineering, New York University Tandon School of Engineering, New York, NY 11201, USA
- ⁵ Department of Physical Medicine & Rehabilitation, MetroHealth System, Cleveland, OH 44109, USA; ffraser@metrohealth.org
- ⁶ Department of Electrical & Computer Engineering, New York University Tandon School of Engineering, New York, NY 11201, USA; selesni@nyu.edu
- ⁷ Department of Population Health, New York University Grossman School of Medicine, New York, NY 10016, USA
- ⁸ Department of Ophthalmology, New York University Grossman School of Medicine, New York, NY 10016, USA
- * Correspondence: janet.rucker@nyulangone.org; Tel.: +1-212-263-7744



Citation: Gold, D.M.; Rizzo, J.-R.; Lee, Y.S.C.; Childs, A.; Hudson, T.E.; Martone, J.; Matsuzawa, Y.K.; Fraser, F.; Ricker, J.H.; Dai, W.; et al. King-Devick Test Performance and Cognitive Dysfunction after Concussion: A Pilot Eye Movement Study. *Brain Sci.* **2021**, *11*, 1571. <https://doi.org/10.3390/brainsci11121571>

Academic Editor: Zoi Kapoula

Received: 26 October 2021

Accepted: 24 November 2021

Published: 27 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: (1) Background: The King-Devick (KD) rapid number naming test is sensitive for concussion diagnosis, with increased test time from baseline as the outcome measure. Eye tracking during KD performance in concussed individuals shows an association between inter-saccadic interval (ISI) (the time between saccades) prolongation and prolonged testing time. This pilot study retrospectively assesses the relation between ISI prolongation during KD testing and cognitive performance in persistently-symptomatic individuals post-concussion. (2) Results: Fourteen participants (median age 34 years; 6 women) with prior neuropsychological assessment and KD testing with eye tracking were included. KD test times (72.6 ± 20.7 s) and median ISI (379.1 ± 199.1 msec) were prolonged compared to published normative values. Greater ISI prolongation was associated with lower scores for processing speed (WAIS-IV Coding, $r = 0.72$, $p = 0.0017$), attention/working memory (Trails Making A, $r = -0.65$, $p = 0.006$) (Digit Span Forward, $r = 0.57$, $p = -0.017$) (Digit Span Backward, $r = -0.55$, $p = 0.021$) (Digit Span Total, $r = -0.74$, $p = 0.001$), and executive function (Stroop Color Word Interference, $r = -0.8$, $p = 0.0003$). (3) Conclusions: This pilot study provides preliminary evidence suggesting that cognitive dysfunction may be associated with prolonged ISI and KD test times in concussion.

Keywords: concussion; King-Devick; rapid automatized naming tasks; saccades; inter-saccadic interval; cognitive dysfunction

1. Introduction

A concussion is a form of mild traumatic brain injury in which biomechanical forces to the head or body result in neurological symptoms such as headaches, dizziness, blurred vision, emotional lability, difficulty concentrating, or slowed information processing [1].

In most individuals, symptoms resolve spontaneously within days to weeks following injury. However, prolonged recovery with persistent symptoms occurs in 10–25% of individuals [2–4]. Given the lack of symptom specificity and clarity with regard to the timing of biological concussion resolution, protracted recovery with persistent symptoms may overlap with the development of other disorders such as depression and psychosocial maladjustment. Although helpful, self-reporting of subjective symptoms of concussion cannot be completely relied upon for diagnosis, as athletes have been shown to under-report or even deny symptoms in order to return to play [5–9]. As a result, sensitive sideline and outpatient diagnostic tests are needed.

Sideline diagnostic tests [10] include the Sports Concussion Assessment Tool (SCAT) Symptom Checklist [11,12], Standardized Assessment of Concussion (SAC) [13], Balance Error Scoring System (BESS) [14], Vestibular–Ocular Motor Screening (VOMS) test [15], and rapid automatized naming tests such as the King–Devick (KD) test of number-naming [16] and the Mobile Universal Lexicon Evaluation System (MULES) test of picture-naming [17]. These diagnostic tests are also increasingly utilized in the outpatient setting where concussions often arise from non-sports-related injuries and where long-lasting symptoms are common [18,19]. It is important to emphasize that the diagnosis of concussion remains a clinical diagnosis and cannot be entirely confirmed or refuted with any single diagnostic test.

Completion of the KD test requires reading numbers with variable spacing on three test cards as rapidly as possible. Scores are generated based on the total time taken to complete the test. The KD test is a performance measure that involves attention, number recognition, language retrieval, and saccadic eye movements. Worsening of the time to complete the test relative to a baseline time is consistent with the diagnosis of concussion, since test scores normally improve with practice and are not affected by physical activity [20–25]. In the outpatient setting, the KD score obtained during the initial visit for concussion evaluation has been shown to predict the total number of visits and the total number of referrals (e.g., to neuropsychology, vestibular, or vision therapy). A greater number of visits and referrals are indicators of more severe or prolonged symptoms [19].

In an effort to understand behaviors associated with slowed KD test times in concussion, quantitative analysis of eye movements (i.e., eye tracking) previously revealed that prolonged KD times in a concussion cohort with protracted post-concussive symptoms were associated with greater numbers of saccades, larger deviations of saccadic endpoints (dysmetria), and a prolongation of the inter-saccadic interval (ISI), as compared to healthy controls [26]. ISI, defined as the time between saccades, was strongly correlated with KD times. This complex interval of time captures several important aspects of test performance, including duration of fixation time on the current number, name retrieval and verbalization of the current number, attentional disengagement from the current number, planning of the saccade to the next number, and saccadic latency, as well as concentration and neurobehavioral contributions. Although these components of the ISI are intermingled and unable to be realistically compartmentalized, this pilot study aimed to retrospectively examine the relation between ISI prolongation on the KD test and cognitive performance on standard neuropsychological assessments in individuals with a history of concussion and persistent symptoms. The hypothesis of the study was that KD performance and ISI duration would relate to attention and processing speed.

2. Materials and Methods

2.1. Participants

Data from participants with a history of concussion and self-reported persistent symptoms who had previously completed a standardized clinical neuropsychological assessment and KD testing with eye tracking as part of a research protocol were retrospectively reviewed. All participants provided written informed consent to participate in eye tracking research and in a concussion registry database that included neuropsychological assessment data. The study was approved by the New York University Grossman School of Medicine

Institutional Review Board (S13-01229 and S14-02097). Exclusion criteria included incomplete neuropsychological assessment, visual impairment precluding KD performance, and moderate-severe traumatic brain injury. Failure on one or more of the freestanding and embedded measures of neuropsychological performance validity, reflecting inadequate motivation/effort, was an additional exclusion criterion. Performance validity was assessed using the Test of Memory Malingering (TOMM), the Reliable Digit Span test, and the California Verbal Learning Test—second edition Forced Choice Measure. Participants were excluded if any of these measures indicated suboptimal effort [27]. Participants were also excluded if glasses were necessary to clearly visualize calibration and visual stimulus targets during eye tracking.

Data were available for sixteen participants; two were subsequently excluded due to poor-quality eye-tracking data. For the remaining 14 (median age 34 years, range 24–61; 6 women), concussion history consisted of either a single concussion (8 participants) or multiple concussion events (6 participants). The interval of time between the most recent concussion, eye tracking, and neuropsychological assessment was variable, ranging from 2 weeks to 84 months (mean time interval between concussion and neuropsychological assessment: 10.2 months, mean time interval between concussion and eye tracking: 11.7 months) (Table 1). All participants had self-reported ongoing symptoms related to their concussion at the time of assessment.

Table 1. Patient demographics and timing of assessments relative to last concussion.

Participant	Age (at Eye Movement Visit)	Sex *	Neuropsychological Testing (Relative to Most Recent Concussion)	Eye Movement Recording (Relative to Most Recent Concussion)
1	28	M	7 months >24 months	6 months
2	28	M	(2013–11/2015)	>24 months (2013–9/2015)
3	50	F	3 months	4 months
4	61	M	29 months	30 months
5	47	M	4 months	7 months
6	32	F	18 months	17 months
7	47	F	1 month	19 months
8	57	F	10 months	13 months
9	32	M	>24 months (2009–8/2016)	>24 months (2009–6/2016)
10	38	F	2 months	5 months
11	34	M	1 month	2 weeks
12	34	F	8 months	6 months
13	30	M	8 months	8 months
14	24	M	4 months	4 months

* M = male, F = female.

2.2. Materials and Procedures

2.2.1. KD Test and Eye Tracking

All participants had previously performed a digitized version of the KD test while simultaneously undergoing binocular eye movement recordings with the EyeLink 1000 Plus, an infrared-based video-oculographic camera system (SR Research, Mississauga, ON, Canada). A forehead rest was utilized for maximum head stability while simultaneously allowing for mouth movements required for number naming. The EyeLink sampled eye position at 500 Hz with a precision of 0.1 degrees. Participants completed an Eyelink standardized 13-point spatial calibration and validation procedure prior to each testing session. The 13-point serial target presentation calibration, rather than the traditional 9-point calibration, was utilized to ensure calibration across the entire display monitor. KD numbers were presented exclusively within the calibration region. Eye position was recorded continuously during onscreen presentation of all KD cards.

The KD test consisted of three computer-generated KD test cards that maintained consistency (e.g., numbers presented, inter-number spacing) with the spiral-bound version of the KD test [21,28]. After presentation of an initial demonstration card, the three test cards of the KD test were serially presented on the computer monitor. Participants were instructed to name the numbers on each card as quickly as possible. The total test time in seconds needed to name all the numbers on the three test cards (excluding time between

cards) was recorded. The total number of errors was also recorded. The methodology of digitization of the KD test with simultaneous eye tracking and data analysis has previously been published [29].

2.2.2. Neuropsychological Assessment

Neuropsychological measures included standardized tests of performance validity, processing speed, attention and working memory, perceptual reasoning, executive functioning, and emotional functioning (Table 2). Testing was performed as part of our interdisciplinary concussion center clinical neuropsychological testing battery and followed standardized administration procedures.

Table 2. Neuropsychological assessments utilized and the corresponding cognitive domain evaluated.

Cognitive Domain	Test *
Processing Speed	SCWT Word; SCWT Color; WAIS-IV Coding
Attention/Working Memory	WAIS-IV (Digit Span Forward; Digit Span Backward; Digit Span Sequencing; Digit Span Total); TMTA
Perceptual Reasoning	WASI-II Matrix Reasoning
Executive Functioning	SCWT Interference; TMTB
Emotional Functioning	BAI; BDI-II

* Abbreviations in Table: BAI = Beck Anxiety Inventory; BDI-II = Beck Depression Inventory, Second Edition; SCWT = Stroop Color and Word Test; TMTA = Trail Making Test A; TMTB = Trail Making Test B; WAIS = Wechsler Adult Intelligence Scale, Fourth Edition; WASI-II = Wechsler Abbreviated Scale of Intelligence, Second Edition.

Processing speed was measured using the Stroop Color and Word Tests (SCWT) Word and Color Scores, as well as the Wechsler Adult Intelligence Scale 4th edition (WAIS-IV) Coding subtest [30,31]. Attention and working memory were assessed using the WAIS-IV Digit Span Forward (DSF), Backward (DSB), Sequencing (DSS), and Total (DST), in addition to the Trail Making Test Part A (TMTA). Perceptual reasoning was examined using the Wechsler Abbreviated Scale of Intelligence—2nd edition (WASI-II) Matrix Reasoning subtest. Executive functioning was evaluated using the SCWT Interference score and the Trail Making Test Part B (TMTB). Emotional functioning was assessed using the Beck Anxiety Inventory (BAI) and the Beck Depression Inventory-II (BDI-II). Standardized administration according to testing instructions was followed for all validated measures of cognitive and emotional functioning.

2.3. Data Analyses

Eye movement data were analyzed offline using custom MATLAB software (MathWorks, version 2020B, Massachusetts, MA, USA). Saccades were identified via an adaptive thresholding mechanism, and velocities and accelerations were computed from position traces using a low-pass differentiator [32]. ISIs were extracted for further analysis. Spearman correlations (non-parametric) were performed using each of the cognitive measures against the ISI values as continuous variables for the WAIS subtests. Parametric testing with Pearson correlations was used for timing data.

3. Results

3.1. KD Times and ISI Values

KD test times were substantially prolonged in this cohort of participants with a history of concussion and persistent symptoms to $72.6 (\pm 20.7)$ sec relative to previously published KD test times in healthy individuals with no history of concussion ($51.24 (\pm 9.7)$ sec [29]— $53.4 (\pm 14.04)$ sec) [26]. ISIs were measured for each participant as the median interval between all task-specific saccades across the three test cards due to the expected substantial positive skew in the distribution of these values for all participants. The median ISI for this cohort was $379.1 (\pm 199.1)$ msec, which is longer than typical ISIs in healthy individuals in our lab of $235.5 (\pm 119.1)$ msec [29] and $286.1 (\pm 49.7)$ msec [26].

3.2. ISI and Neuropsychological Assessments

Neuropsychological assessment scaled scores for assessments associated with greater ISI prolongation are shown in Table 3. Greater ISI prolongation was associated with lower scores in the cognitive domains of processing speed, attention/working memory, and executive function. The Spearman correlation coefficient was significant when comparing the ISI and TMTA ($r = -0.65, p = 0.006$) (Figure 1), DSF ($r = 0.57, p = -0.017$), DSB ($r = -0.55, p = 0.021$), and DST ($r = -0.74, p = 0.001$)—all tests of attention/working memory. Lower scores on SCWT Interference, a marker of executive function, were also significantly associated with ISI prolongation ($r = -0.8, p = 0.0003$) (Figure 2). The WAIS-IV Coding score, a marker of processing speed, was significantly associated with ISI prolongation ($r = 0.72, p = 0.0017$). The remaining measures showed a trend toward association, though they were not significant.

Table 3. Neuropsychological assessment scaled or *T*-scores *.

Subject	Trail Making Test A (T-Score)	Digit Span Forward (SS)	Digit Span Backward (SS)	Digit Span Total (SS)	SCWT Interference (T-Score)	WAIS-IV Coding (SS)
1	76	9	8	9	51	15
2	68	11	14	12	60	10
3	35	10	13	11	51	6
4	35	5	7	5	42	5
5	54	10	14	12	46	15
6	63	12	10	10	56	9
7	55	11	8	12	50	12
8	41	11	8	9	50	8
9	52	12	14	13	67	16
10	35	8	10	8	57	11
11	61	12	12	14	52	11
12	41	8	10	9	59	6
13	39	12	12	12	54	10
14	49	9	10	11	49	9

* *T*-score from 20–30 or scaled score (SS) from 1 to 4 is between -3 standard deviation (SD) and -2 SD. *T*-score from 31 to 40 or SS from 4 to 7 is between -2 SD and -1 SD. *T*-score from 41–50 or SS from 7 to 10 is between -1 SD and 0SD. *T*-score from 51 to 60 or SS from 10 to 13 is between 0SD and $+1$ SD. *T*-score from 61 to 70 or SS from 13 to 16 is between $+1$ SD and $+2$ SD. *T*-score from 71 to 80 or scaled score 16–19 is between $+2$ SD and $+3$ SD.

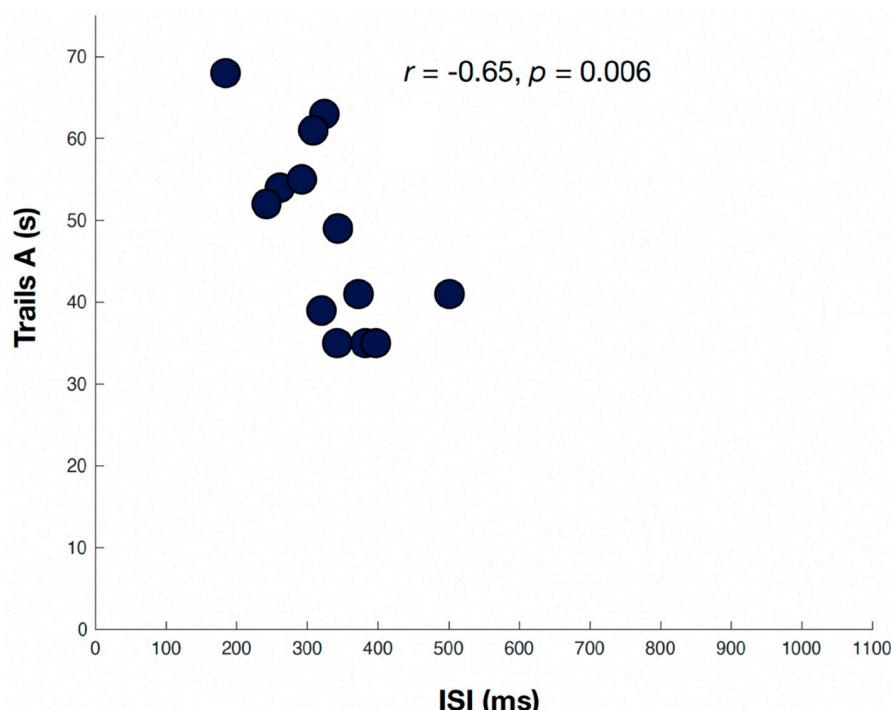


Figure 1. Relation of inter-saccadic intervals (msec) and the Trails Making Test Part A (T-score), a marker of attention/working memory. Two participants (grey circles) were excluded due to poor eye-tracking data quality.

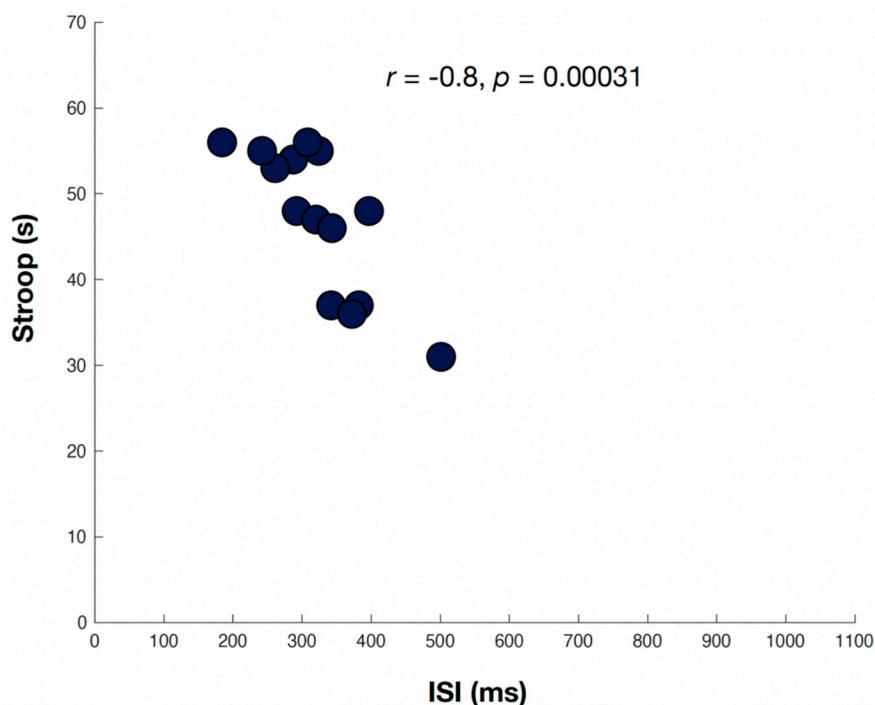


Figure 2. Relation of inter-saccadic intervals (msec) and the Stroop Color Word Interference Test (T-score), a marker of executive function. Two participants (grey circles) were excluded due to poor eye-tracking data quality.

4. Discussion

In this small pilot study, we sought to retrospectively explore the relationship between ISI prolongation during KD test performance and standardized neuropsychological assessments in a cohort of individuals with a history of concussion who had been evaluated in our concussion center. The aim was an initial exploration to advance understanding of the factors that may contribute to, and thus explain, slowed KD test performance following a concussion. Our data demonstrated that in this outpatient cohort with persistent symptoms, ISI prolongation during KD testing was associated with diminished neuropsychological function in the cognitive domains of processing speed, attention/working memory, and executive function. We will briefly review traditional applications of eye tracking and neuropsychological testing in concussion and then further consider their interactions and potential contributions to slowed KD test performance and ISI prolongation during KD testing in concussion.

4.1. Eye Tracking and Neuropsychological Assessment in Concussion

There has been growing interest in, as well as debate about, visual symptoms, eye movements, and eye tracking applications in and after concussion over the past decade [33–38]. Given that saccadic eye movements are the eye movement type predominantly employed during the KD test, we limit the discussion here to eye-tracking studies of saccades. Brain networks that govern saccades are well-delineated, widely distributed, and extend from the frontal and parietal cortices down to the brainstem premotor nuclei. These nuclei ultimately initiate the motor command for a saccade in the ocular motor nuclei that send signals to the extraocular muscles [39].

Various subtypes of saccades can be assessed with eye tracking to probe the integrity of different regions of these saccadic networks. Most studies of saccades in concussions show that simple visually guided saccades are unaffected (i.e., not slowed), thus indicating that the immediate premotor structures in the brainstem that drive saccades are typically unaffected in acute concussion and chronic symptomatic states after concussion [40–46]. The exception to simple visually guided saccades being unaffected is the finding of in-

creased saccadic latency (time between visual stimulus presentation and initiation of a saccade) for visually guided saccades in hyper-acute concussion, a finding which quickly resolves [47,48]. In keeping with studies largely showing normal visually guided saccades, saccade speeds have been shown to be normal during KD testing post-concussion, as well [26]. In contrast, abnormalities are often identified in attentionally-dependent saccade types that probe higher cortical, particularly prefrontal, structures involved in saccade generation, such as the frontal and supplementary eye fields and the dorsolateral prefrontal cortex (DLPFC). These brain regions are particularly prone to traumatic brain injury [49]. Saccade types dependent on these higher cortical structures include memory-guided saccades (e.g., saccades to the remembered location of a previously present visual target) and antisaccades (e.g., saccades in the direction opposite to a suddenly appearing visual target); these saccade types assess cognitive functions such as short-term spatial memory, response inhibition, motor-sequence programming, visuospatial processing, and visual attention [42]. Increased saccadic latencies, more directional errors, and poorer spatial accuracy in these saccade types are an established indicator of suboptimal brain function in patients with acute concussion and chronic symptomatic states after concussion [40,42,50–54].

Neuropsychological testing is also widely utilized to identify suboptimal brain function in concussion and chronic symptomatic states after concussion and can assess cognitive, behavioral, and emotional aspects of functioning. Performance on neuropsychological assessments can be impacted by a range of variables, including mood, physical symptoms (e.g., headaches, fatigue, vestibular symptoms, etc.), education level, and premorbid conditions [55,56]. Processing speed and working memory are the most sensitive measures of cognitive dysfunction in concussion, though abnormalities in executive function, attention, and cognitive flexibility may be the most persistent cognitive deficits [30,56]. These deficits can affect performance on assessments such as the Stroop Color Word tests and the Trail Making Test Part A [54,57,58].

A few studies have assessed the relationship between eye movements and neuropsychological testing by either comparing the sensitivities of higher cortical saccade types with standard neuropsychological assessments for concussion diagnosis [42] or more directly considering higher cortical saccades as measures of neurocognitive dysfunction [54]. Abnormalities in memory-guided saccades and antisaccades have been shown to remain impaired longer and to correlate better with post-concussive symptoms and impaired activities of daily living than neuropsychological assessments in individuals with persistently symptomatic post-concussive states [42]. Abnormalities of antisaccades have been shown to correlate with greater symptom burden in acute concussion [54] and with poor performance on the Stroop test of executive function, which requires response inhibition. The dorsolateral prefrontal cortex (DLPFC), in particular, is known to play a key role in working memory and in inhibition of a reflexive saccade to the suddenly appearing visual target in the antisaccade task and may play a key role in deficits in concussion [59]. Indeed, the DLPFC has been shown to have transient alterations in its metabolic profile following head acceleration events, a proxy for sports-related concussion [60].

4.2. Interactions between KD Performance and Neuropsychological Assessments

The KD test and other rapid automatized naming tasks are performance measures that harness a number of different neurological systems, including vision and saccadic eye movements, cognitive aspects of attention and processing speed, and language function. As such, they have the capacity to capture dysfunction in concussion and have been shown to be sensitive measures for diagnosis on the sidelines of sport [16,21–23] and to be predictive of recovery in the outpatient arena [19]. To date, research on the exact contributing factors that underlie prolonged test times on these vision-based performance measures in the setting of concussion has been sparse.

Our focus has been on understanding eye-movement behaviors associated with slowed KD test times in concussion, which are predominantly correlated with longer ISIs [26]. Certainly, it is anticipated that neuropsychological abnormalities might be one of

the factors contributing to prolonged test times. Indeed, it has been previously reported that longer (worse) KD completion times are associated with lower (worse) scores on the Sports Concussion Assessment Tool 2 (SCAT2), Standardized Assessment of Concussion (SAC) Immediate Memory Score, and on the overall SAC score [23,61]. In this exploratory pilot study, ISI prolongation during KD testing was associated with impaired neuropsychological function in the cognitive domains of processing speed, attention/working memory, and executive function. It is thus possible that the processing speed for the KD test may be adversely impacted by impaired visuospatial attention. KD performance requires a constant “updating” of attention and motor planning, and concussed individuals have been shown to have difficulty with visual disengagement [62]. We only included neuropsychological data interpreted as valid based on performance validity tests. Self-reported measures of mood were not significantly associated with ISI findings, suggesting interconnectivity of ISI and injury sustained by concussion that was independent of mood. Also notable in this study was the fact that the association between ISI and neuropsychological assessments was present even in the absence of objective cognitive impairment based on scoring parameters.

4.3. Study Limitations

Participants for this exploratory, retrospective pilot study were included based on the availability of relevant data for retrospective analysis, which led to a small number of participants being included and very high variability in the timing between the eye-tracking and neuropsychological assessments. Nonetheless, associations persisted in this pilot study and would likely be even more robust if timing intervals were standardized in future studies. In keeping with the retrospective nature of the study, the study population was also heterogeneous with regard to the number of concussions and time since the most recent injury. Thus, the generalizability of the results is presently unknown, and it was not possible to control for levels of fatigue or medications at the time of testing. Future prospective studies exploring the relationship between KD and neuropsychological testing performance will allow the opportunity for the inclusion of a control participant group and evaluation of the impact of age on performance.

5. Conclusions

Quantitative assessment of the ISI during rapid automatized naming tests, likely in conjunction with other concussion-based diagnostics, is an objective, quantifiable eye-tracking metric of potentially high importance. This study provides preliminary evidence that cognitive dysfunction may be one element underlying prolonged ISI and KD test times after concussion. It is likely that other factors may also play a role. Given that all visual targets (e.g., all KD numbers) are displayed simultaneously during the KD test, we cannot directly measure true saccadic latency or assess capacity for disengagement from numbers. Exploration of these components will be our next step, as we assess the relationships between ISI and traditional measures of saccades, including the latencies of visually guided saccades and other saccade types, such as gap saccades (in which the fixation target disappears prior to the appearance of the visual target for saccade initiation, which facilitates disengagement from the prior number). A comprehensive understanding of the underlying components contributing to the prolongation of KD test times in and after concussion will help to elucidate what rapid number naming tasks are capturing and where these deficits may localize in the brain.

Author Contributions: Conceptualization and methodology: D.M.G., J.-R.R., Y.S.C.L., A.C., T.E.H., J.H.R., L.J.B., S.L.G., J.C.R.; data curation and formal analysis: D.M.G., Y.S.C.L., A.C., Y.K.M., J.M., F.F., J.M., W.D., I.S.; writing—original draft preparation: D.M.G., J.-R.R., Y.S.C.L., A.C., J.C.R.; writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded in part by 5K12HD001097 NICHHD and NCMRR, National Institutes of Health Rehabilitation Medicine Scientist Training Program (JRR). Empire Clinical Research Investigator Program (ECRIP).

Institutional Review Board Statement: All research protocols were approved by the NYU Institutional Review Board (Code: S13-01229).

Informed Consent Statement: Written informed consent for participation was obtained from each participant.

Data Availability Statement: Original data will be made available upon request.

Conflicts of Interest: No author has received any financial compensation or consultant fees from King-Devick Test, Inc. No author has other disclosures pertinent to this study.

References

1. Davis, G.A.; Ellenbogen, R.G.; Bailes, J.; Cantu, R.C.; Johnston, K.M.; Manley, G.T.; Nagahiro, S.; Sills, A.; Tator, C.H.; McCrory, P. The Berlin International Consensus Meeting on Concussion in Sport. *Neurosurgery* **2018**, *82*, 232–236. [[CrossRef](#)] [[PubMed](#)]
2. Hipple, C.; Dufort, P.A.; Davis, H.S.; Wennberg, R.A.; Tartaglia, M.C.; Mikulis, D.; Hazrati, L.N.; Tator, C.H. Longitudinal Study of Postconcussion Syndrome: Not Everyone Recovers. *J. Neurotrauma* **2017**, *34*, 1511–1523. [[CrossRef](#)]
3. Cassidy, J.D.; Cancelliere, C.; Carroll, L.J.; Cote, P.; Hincapie, C.A.; Holm, L.W.; Hartvigsen, J.; Donovan, J.; Nygren-de Boussard, C.; Kristman, V.L.; et al. Systematic review of self-reported prognosis in adults after mild traumatic brain injury: Results of the International Collaboration on Mild Traumatic Brain Injury Prognosis. *Arch. Phys. Med. Rehabil.* **2014**, *95*, S132–S151. [[CrossRef](#)] [[PubMed](#)]
4. Polinder, S.; Cnossen, M.C.; Real, R.G.L.; Covic, A.; Gorbunova, A.; Voormolen, D.C.; Master, C.L.; Haagsma, J.A.; Diaz-Arrastia, R.; von Steinbuechel, N. A Multidimensional Approach to Post-concussion Symptoms in Mild Traumatic Brain Injury. *Front. Neurol.* **2018**, *9*, 1113. [[CrossRef](#)] [[PubMed](#)]
5. Delaney, J.S.; Lamfookon, C.; Bloom, G.A.; Al-Kashmiri, A.; Correa, J.A. Why university athletes choose not to reveal their concussion symptoms during a practice or game. *Clin. J. Sport Med.* **2015**, *25*, 113–125. [[CrossRef](#)]
6. Kerr, Z.Y.; Register-Mihalik, J.K.; Kroshus, E.; Baugh, C.M.; Marshall, S.W. Motivations Associated With Nondisclosure of Self-Reported Concussions in Former Collegiate Athletes. *Am. J. Sports Med.* **2016**, *44*, 220–225. [[CrossRef](#)]
7. Kerr, Z.Y.; Register-Mihalik, J.K.; Marshall, S.W.; Evenson, K.R.; Mihalik, J.P.; Guskiewicz, K.M. Disclosure and non-disclosure of concussion and concussion symptoms in athletes: Review and application of the socio-ecological framework. *Brain Inj.* **2014**, *28*, 1009–1021. [[CrossRef](#)]
8. Conway, F.N.; Domingues, M.; Monaco, R.; Lesnewich, L.M.; Ray, A.E.; Alderman, B.L.; Todaro, S.M.; Buckman, J.F. Concussion Symptom Underreporting Among Incoming National Collegiate Athletic Association Division I College Athletes. *Clin. J. Sport Med.* **2020**, *30*, 203–209. [[CrossRef](#)] [[PubMed](#)]
9. Torres, D.M.; Galetta, K.M.; Phillips, H.W.; Dziemianowicz, E.M.; Wilson, J.A.; Dorman, E.S.; Laudano, E.; Galetta, S.L.; Balcer, L.J. Sports-related concussion: Anonymous survey of a collegiate cohort. *Neurol. Clin. Pract.* **2013**, *3*, 279–287. [[CrossRef](#)] [[PubMed](#)]
10. Yue, J.K.; Phelps, R.R.L.; Chandra, A.; Winkler, E.A.; Manley, G.T.; Berger, M.S. Sideline Concussion Assessment: The Current State of the Art. *Neurosurgery* **2020**, *87*, 466–475. [[CrossRef](#)] [[PubMed](#)]
11. McCrory, P.; Johnston, K.; Meeuwisse, W.; Aubry, M.; Cantu, R.; Dvorak, J.; Graf-Baumann, T.; Kelly, J.; Lovell, M.; Schamasch, P. Summary and agreement statement of the 2nd International Conference on Concussion in Sport, Prague 2004. *Br. J. Sports Med.* **2005**, *39*, 196–204. [[CrossRef](#)] [[PubMed](#)]
12. Echemendia, R.J.; Meeuwisse, W.; McCrory, P.; Davis, G.A.; Putukian, M.; Leddy, J.; Makdissi, M.; Sullivan, S.J.; Broglio, S.P.; Raftery, M.; et al. The Sport Concussion Assessment Tool 5th Edition (SCAT5): Background and rationale. *Br. J. Sports Med.* **2017**, *51*, 848–850. [[CrossRef](#)] [[PubMed](#)]
13. McCrea, M.; Kelly, J.P.; Kluge, J.; Ackley, B.; Randolph, C. Standardized assessment of concussion in football players. *Neurology* **1997**, *48*, 586–588. [[CrossRef](#)]
14. McCrory, P.; Meeuwisse, W.; Johnston, K.; Dvorak, J.; Aubry, M.; Molloy, M.; Cantu, R. Consensus Statement on Concussion in Sport: The 3rd International Conference on Concussion in Sport held in Zurich, November 2008. *Br. J. Sports Med.* **2009**, *43* (Suppl. 1), i76–i90. [[CrossRef](#)] [[PubMed](#)]
15. Mucha, A.; Collins, M.W.; Elbin, R.J.; Furman, J.M.; Troutman-Enseki, C.; DeWolf, R.M.; Marchetti, G.; Kontos, A.P. A Brief Vestibular/Ocular Motor Screening (VOMS) assessment to evaluate concussions: Preliminary findings. *Am. J. Sports Med.* **2014**, *42*, 2479–2486. [[CrossRef](#)] [[PubMed](#)]
16. Galetta, K.M.; Liu, M.; Leong, D.F.; Ventura, R.E.; Galetta, S.L.; Balcer, L.J. The King-Devick test of rapid number naming for concussion detection: Meta-analysis and systematic review of the literature. *Concussion* **2016**, *1*, CNC8. [[CrossRef](#)]
17. Fallon, S.; Akhand, O.; Hernandez, C.; Galetta, M.S.; Hasanaj, L.; Martone, J.; Webb, N.; Drattell, J.; Amorapanth, P.; Rizzo, J.R.; et al. MULES on the sidelines: A vision-based assessment tool for sports-related concussion. *J. Neurol. Sci.* **2019**, *402*, 52–56. [[CrossRef](#)] [[PubMed](#)]

18. Thurman, D.J.; Alverson, C.; Dunn, K.A.; Guerrero, J.; Sniezek, J.E. Traumatic brain injury in the United States: A public health perspective. *J. Head Trauma Rehabil.* **1999**, *14*, 602–615. [CrossRef] [PubMed]
19. Kyle Harrold, G.; Hasanaj, L.; Moehring, N.; Zhang, I.; Nolan, R.; Serrano, L.; Raynowska, J.; Rucker, J.C.; Flanagan, S.R.; Cardone, D.; et al. Rapid sideline performance meets outpatient clinic: Results from a multidisciplinary concussion center registry. *J. Neurol. Sci.* **2017**, *379*, 312–317. [CrossRef]
20. King, D.; Clark, T.; Gissane, C. Use of a rapid visual screening tool for the assessment of concussion in amateur rugby league: A pilot study. *J. Neurol. Sci.* **2012**, *320*, 16–21. [CrossRef] [PubMed]
21. Galetta, K.M.; Barrett, J.; Allen, M.; Madda, F.; Delicata, D.; Tenant, A.T.; Branas, C.C.; Maguire, M.G.; Messner, L.V.; Devick, S.; et al. The King-Devick test as a determinant of head trauma and concussion in boxers and MMA fighters. *Neurology* **2011**, *76*, 1456–1462. [CrossRef]
22. Galetta, K.M.; Brandes, L.E.; Maki, K.; Dziemianowicz, M.S.; Laudano, E.; Allen, M.; Lawler, K.; Sennett, B.; Wiebe, D.; Devick, S.; et al. The King-Devick test and sports-related concussion: Study of a rapid visual screening tool in a collegiate cohort. *J. Neurol. Sci.* **2011**, *309*, 34–39. [CrossRef]
23. Galetta, M.S.; Galetta, K.M.; McCrossin, J.; Wilson, J.A.; Moster, S.; Galetta, S.L.; Balcer, L.J.; Dorshimer, G.W.; Master, C.L. Saccades and memory: Baseline associations of the King-Devick and SCAT2 SAC tests in professional ice hockey players. *J. Neurol. Sci.* **2013**, *328*, 28–31. [CrossRef] [PubMed]
24. King, D.; Gissane, C.; Hume, P.A.; Flaws, M. The King-Devick test was useful in management of concussion in amateur rugby union and rugby league in New Zealand. *J. Neurol. Sci.* **2015**, *351*, 58–64. [CrossRef]
25. King, D.; Hume, P.; Gissane, C.; Clark, T. Use of the King-Devick test for sideline concussion screening in junior rugby league. *J. Neurol. Sci.* **2015**, *357*, 75–79. [CrossRef]
26. Rizzo, J.-R.; Hudson, T.E.; Dai, W.; Birkemeier, J.; Pasquilli, R.M.; Selesnick, I.; Balcer, L.J.; Galetta, S.L.; Rucker, J.C. Rapid number naming in chronic concussion: Eye movements in the King-Devick test. *Ann. Clin. Transl. Neurol.* **2016**, *3*, 801–811. [CrossRef]
27. Denning, J.H. The efficiency and accuracy of the Test of Memory Malingering trial 1, errors on the first 10 items of the test of memory malingering, and five embedded measures in predicting invalid test performance. *Arch. Clin. Neuropsychol.* **2012**, *27*, 417–432. [CrossRef]
28. Oride, M.K.; Marutani, J.K.; Rouse, M.W.; DeLand, P.N. Reliability study of the Pierce and King-Devick saccade tests. *Am. J. Optom. Physiol. Opt.* **1986**, *63*, 419–424. [CrossRef]
29. Rizzo, J.-R.; Hudson, T.E.; Dai, W.; Desai, N.; Yousefi, A.; Palsana, D.; Selesnick, I.; Balcer, L.J.; Galetta, S.L.; Rucker, J.C. Objectifying eye movements during rapid number naming: Methodology for assessment of normative data for the King-Devick test. *J. Neurol. Sci.* **2016**, *362*, 232–239. [CrossRef]
30. Donders, J.; Strong, C.A. Clinical utility of the Wechsler Adult Intelligence Scale-Fourth Edition after traumatic brain injury. *Assessment* **2015**, *22*, 17–22. [CrossRef]
31. Alvarez, J.A.; Emory, E. Executive function and the frontal lobes: A meta-analytic review. *Neuropsychol. Rev.* **2006**, *16*, 17–42. [CrossRef]
32. Dai, W.; Selesnick, I.; Rizzo, J.R.; Rucker, J.; Hudson, T. A nonlinear generalization of the Savitzky-Golay filter and the quantitative analysis of saccades. *J. Vis.* **2017**, *17*, 10. [CrossRef]
33. Ventura, R.E.; Balcer, L.J.; Galetta, S.L.; Rucker, J.C. Ocular motor assessment in concussion: Current status and future directions. *J. Neurol. Sci.* **2016**, *361*, 79–86. [CrossRef] [PubMed]
34. Akhand, O.; Balcer, L.J.; Galetta, S.L. Assessment of vision in concussion. *Curr. Opin. Neurol.* **2019**, *32*, 68–74. [CrossRef] [PubMed]
35. Akhand, O.; Rizzo, J.R.; Rucker, J.C.; Hasanaj, L.; Galetta, S.L.; Balcer, L.J. History and Future Directions of Vision Testing in Head Trauma. *J. Neuroophthalmol.* **2019**, *39*, 68–81. [CrossRef]
36. Barton, J.J.S.; Ranalli, P.J. Vision Therapy: Ocular Motor Training in Mild Traumatic Brain Injury. *Ann. Neurol.* **2020**, *88*, 453–461. [CrossRef]
37. Rucker, J.C.; Rizzo, J.R.; Hudson, T.E.; Balcer, L.J.; Galetta, S.L. Concerning Vision Therapy and Ocular Motor Training in Mild Traumatic Brain Injury. *Ann. Neurol.* **2020**, *88*, 1053–1054. [CrossRef] [PubMed]
38. Barton, J.J.S.; Ranalli, P.J. Reply to “Concerning Vision Therapy and Ocular Motor Training in Mild TBI”. *Ann. Neurol.* **2020**, *88*, 1054–1055. [CrossRef]
39. Leigh, R.J.; Zee, D.S. *The Neurology of Eye Movements*, 5th ed.; Oxford University Press: Oxford, UK; New York, NY, USA, 2015; 1109p.
40. Heitger, M.H.; Anderson, T.J.; Jones, R.D.; Dalrymple-Alford, J.C.; Frampton, C.M.; Ardagh, M.W. Eye movement and visuomotor arm movement deficits following mild closed head injury. *Brain* **2004**, *127*, 575–590. [CrossRef]
41. Kraus, M.F.; Little, D.M.; Donnell, A.J.; Reilly, J.L.; Simonian, N.; Sweeney, J.A. Oculomotor function in chronic traumatic brain injury. *Cogn. Behav. Neurol.* **2007**, *20*, 170–178. [CrossRef]
42. Heitger, M.H.; Jones, R.D.; Macleod, A.D.; Snell, D.L.; Frampton, C.M.; Anderson, T.J. Impaired eye movements in post-concussion syndrome indicate suboptimal brain function beyond the influence of depression, malingering or intellectual ability. *Brain* **2009**, *132*, 2850–2870. [CrossRef] [PubMed]
43. Clough, M.; Mutimer, S.; Wright, D.K.; Tsang, A.; Costello, D.M.; Gardner, A.J.; Stanwell, P.; Mychasiuk, R.; Sun, M.; Brady, R.D.; et al. Oculomotor Cognitive Control Abnormalities in Australian Rules Football Players with a History of Concussion. *J. Neurotrauma* **2018**, *35*, 730–738. [CrossRef] [PubMed]

44. Wetzel, P.A.; Lindblad, A.S.; Raizada, H.; James, N.; Mulatya, C.; Kannan, M.A.; Villamar, Z.; Gitchel, G.T.; Weaver, L.K. Eye Tracking Results in Postconcussive Syndrome Versus Normative Participants. *Investig. Ophthalmol. Vis. Sci.* **2018**, *59*, 4011–4019. [[CrossRef](#)]
45. Webb, B.; Humphreys, D.; Heath, M. Oculomotor Executive Dysfunction during the Early and Later Stages of Sport-Related Concussion Recovery. *J. Neurotrauma* **2018**, *35*, 1874–1881. [[CrossRef](#)] [[PubMed](#)]
46. Matuseviciene, G.; Johansson, J.; Moller, M.; Godbolt, A.K.; Pansell, T.; Deboussard, C.N. Longitudinal changes in oculomotor function in young adults with mild traumatic brain injury in Sweden: An exploratory prospective observational study. *BMJ Open* **2018**, *8*, e018734. [[CrossRef](#)]
47. Pearson, B.C.; Armitage, K.R.; Horner, C.W.; Carpenter, R.H. Saccadometry: The possible application of latency distribution measurement for monitoring concussion. *Br. J. Sports Med.* **2007**, *41*, 610–612. [[CrossRef](#)] [[PubMed](#)]
48. Cochrane, G.D.; Christy, J.B.; Almutairi, A.; Buseytini, C.; Swanson, M.W.; Weise, K.K. Visuo-oculomotor Function and Reaction Times in Athletes with and without Concussion. *Optom. Vis. Sci.* **2019**, *96*, 256–265. [[CrossRef](#)] [[PubMed](#)]
49. Levin, H.S.; Williams, D.H.; Eisenberg, H.M.; High, W.M., Jr.; Guinto, F.C., Jr. Serial MRI and neurobehavioural findings after mild to moderate closed head injury. *J. Neurol. Neurosurg. Psychiatry* **1992**, *55*, 255–262. [[CrossRef](#)] [[PubMed](#)]
50. Crevits, L.; Hanse, M.C.; Tummers, P.; Van Maele, G. Antisaccades and remembered saccades in mild traumatic brain injury. *J. Neurol.* **2000**, *247*, 179–182. [[CrossRef](#)]
51. Heitger, M.H.; Anderson, T.J.; Jones, R.D. Saccade sequences as markers for cerebral dysfunction following mild closed head injury. *Prog. Brain Res.* **2002**, *140*, 433–448. [[CrossRef](#)] [[PubMed](#)]
52. Heitger, M.H.; Jones, R.D.; Dalrymple-Alford, J.C.; Frampton, C.M.; Ardagh, M.W.; Anderson, T.J. Motor deficits and recovery during the first year following mild closed head injury. *Brain Inj.* **2006**, *20*, 807–824. [[CrossRef](#)]
53. Johnson, B.; Hallett, M.; Sloboounov, S. Follow-up evaluation of oculomotor performance with fMRI in the subacute phase of concussion. *Neurology* **2015**, *85*, 1163–1166. [[CrossRef](#)] [[PubMed](#)]
54. Ting, W.K.; Schweizer, T.A.; Topolovec-Vranic, J.; Cusimano, M.D. Antisaccadic Eye Movements Are Correlated with Corpus Callosum White Matter Mean Diffusivity, Stroop Performance, and Symptom Burden in Mild Traumatic Brain Injury and Concussion. *Front. Neurol.* **2015**, *6*, 271. [[CrossRef](#)]
55. Iverson, G.L.; Lange, R.T.; Brooks, B.L.; Rennison, V.L. “Good old days” bias following mild traumatic brain injury. *Clin. Neuropsychol.* **2010**, *24*, 17–37. [[CrossRef](#)] [[PubMed](#)]
56. Rohling, M.L.; Binder, L.M.; Demakis, G.J.; Larrabee, G.J.; Ploetz, D.M.; Langhinrichsen-Rohling, J. A meta-analysis of neuropsychological outcome after mild traumatic brain injury: Re-analyses and reconsiderations of Binder et al., (1997), Frencham et al., (2005), and Pertab et al., (2009). *Clin. Neuropsychol.* **2011**, *25*, 608–623. [[CrossRef](#)]
57. Schretlen, D.J.; Shapiro, A.M. A quantitative review of the effects of traumatic brain injury on cognitive functioning. *Int. Rev. Psychiatry* **2003**, *15*, 341–349. [[CrossRef](#)]
58. Carney, N.; Ghajar, J.; Jagoda, A.; Bedrick, S.; Davis-O'Reilly, C.; du Coudray, H.; Hack, D.; Helfand, N.; Huddleston, A.; Nettleton, T.; et al. Concussion guidelines step 1: Systematic review of prevalent indicators. *Neurosurgery* **2014**, *75* (Suppl. 1), S3–S15. [[CrossRef](#)]
59. Pierrot-Deseilligny, C.; Muri, R.M.; Ploner, C.J.; Gaymard, B.; Demeret, S.; Rivaud-Pechoux, S. Decisional role of the dorsolateral prefrontal cortex in ocular motor behaviour. *Brain* **2003**, *126*, 1460–1473. [[CrossRef](#)]
60. Bari, S.; Svaldi, D.O.; Jang, I.; Shenk, T.E.; Poole, V.N.; Lee, T.; Dydak, U.; Rispoli, J.V.; Nauman, E.A.; Talavage, T.M. Dependence on subconcussive impacts of brain metabolism in collision sport athletes: An MR spectroscopic study. *Brain Imaging Behav.* **2018**, *13*, 735–749. [[CrossRef](#)]
61. Subotic, A.; Ting, W.K.; Cusimano, M.D. Characteristics of the King-Devick test in the assessment of concussed patients in the subacute and later stages after injury. *PLoS ONE* **2017**, *12*, e0183092. [[CrossRef](#)]
62. Drew, A.S.; Langan, J.; Halterman, C.; Osternig, L.R.; Chou, L.S.; van Donkelaar, P. Attentional disengagement dysfunction following mTBI assessed with the gap saccade task. *Neurosci. Lett.* **2007**, *417*, 61–65. [[CrossRef](#)] [[PubMed](#)]



The King-Devick test in an outpatient concussion clinic: Assessing the diagnostic and prognostic value of a vision test in conjunction with exercise testing among acutely concussed adolescents

Justine B. Lawrence^{a,*}, Mohammad N. Haider^{a,b}, John J. Leddy^b, Andrea Hinds^b, Jeffery C. Miecznikowski^c, Barry S. Willer^d

^a Department of Neuroscience, University at Buffalo, SUNY, Buffalo, NY 14214, United States

^b UBMD Department of Orthopedics and Sports Medicine, University at Buffalo, SUNY, Buffalo, NY 14214, United States

^c Department of Biostatistics, University at Buffalo, SUNY, Buffalo, NY 14214, United States

^d Department of Psychiatry, University at Buffalo, SUNY, Buffalo, NY 14214, United States

ARTICLE INFO

Keywords:

King-Devick test
Concussion
Sport-related concussion
Persistent post-concussion syndrome
Sideline assessment

ABSTRACT

Objective: This study investigated the diagnostic and prognostic value of the King-Devick (K-D) test in conjunction with treadmill testing in adolescents after sport-related concussion (SRC) in an outpatient concussion management clinic without baseline measures.

Design: Prospective cohort.

Methods: The K-D test was administered pre- and post-exercise on a graded treadmill test to acutely concussed (AC, < 10 days from injury, n = 46, 15.4 ± 2.1 years) participants for 2 clinic visits (1 week apart) and to matched controls (MC, n = 30, 15.8 ± 1.4 years) for 2 visits (1 week apart). Initial K-D test times were compared between MC and AC. Changes in times from pre- to post- exercise during a treadmill test were compared for MC and AC and from Visit 1 to Visit 2. Smooth pursuits and repetitive saccades were compared with initial visit K-D test performance.

Results: Comparison of pre-exercise K-D test times at Visit 1 distinguished MC from AC (46.1 ± 9.2 s vs. 53.7 ± 13.0 s, p = .007). Comparison of pre- and post-exercise K-D test times revealed significant improvements for MC (46.1 ± 9.2 s vs. 43.1 ± 8.5 s, p < .001) and AC who recovered by Visit 2 (Fast Recovery Group [FRG], n = 23, 50.4 ± 10.0 s vs. 47.3 ± 9.8 s, p = .002). No significant difference was seen in pre- and post-exercise K-D test times on Visit 1 for AC who took longer than 2 weeks to recover (Slow Recovery Group [SRG], n = 23, 57.0 ± 15.0 s vs. 56.0 ± 16.3 s, p = .478). At Visit 1, AC had more abnormal smooth pursuits than MC (17% vs. 3%, non-significant, p = .064). AC, however, had significantly more abnormal repetitive saccades than MC (37% vs. 3%, p = .001) and AC with abnormal repetitive saccades took significantly longer to complete the Visit 1 pre-exercise K-D test than AC with normal repetitive saccades (58.6 ± 16.0 s vs 50.8 ± 10.2 s, p = .049).

Conclusion: The study supports utility of the K-D test as part of outpatient concussion assessment. Lack of improvement in K-D test performance after an exercise test may be an indicator of delayed recovery from SRC.

1. Introduction

Sport-Related Concussion (SRC), a subset of mild traumatic brain injury (mTBI), is a growing public health concern [1]. An estimated 1.6–3.8 million SRCs occur annually in the United States [2]. Concussions are caused by blunt mechanical forces to the head, resulting in a pathophysiological process that alters brain function and causes widespread neurological and cognitive deficits [3,4]. The most commonly

used sideline assessment tools are the Sports Concussion Assessment Tools (SCAT) [5]. The most recent version, the SCAT-5, includes a symptom checklist, the Standardized Assessment of Concussion (SAC), the modified Balance Error Scoring System (mBESS), the timed tandem gait (TTG), and a brief neurological exam. Recently, there has been increasing evidence for the King-Devick (K-D) test to be used as an additional sideline concussion assessment tool [6]. The Buffalo Concussion Treadmill Test (BCTT) is a graded exercise test used in clinical

* Corresponding author at: 160 Farber Hall, 3435 Main St, Buffalo, NY 14214, United States.

E-mail address: jblawren@buffalo.edu (J.B. Lawrence).

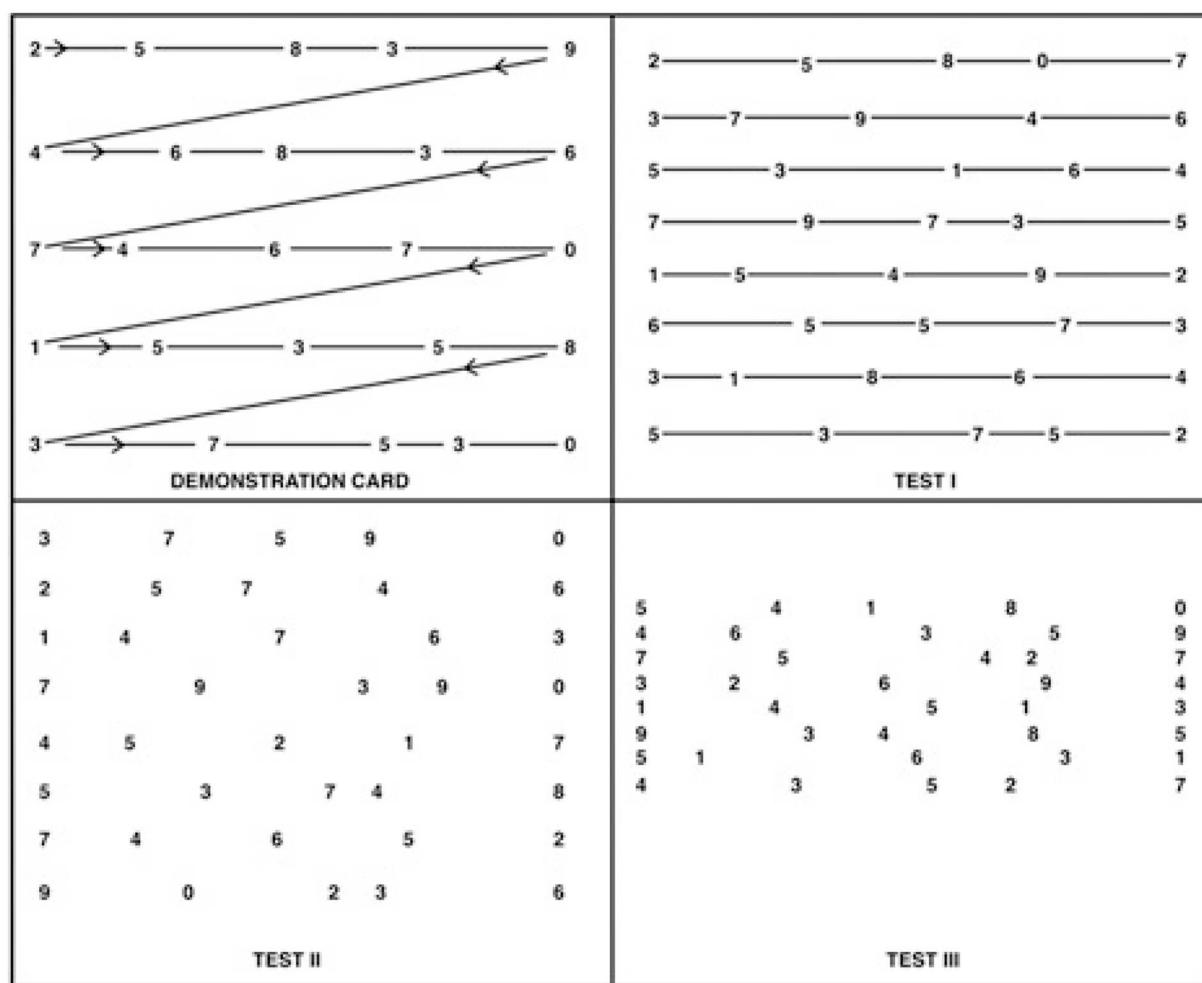


Fig. 1. Demonstration and test cards for the K-D test.

settings to diagnose physiological dysfunction after concussion [7]. The BCTT can quantify clinical severity of concussion and can be used to prescribe safe aerobic exercise treatments for recovery [7].

The K-D test is a rapid number naming task originally developed in 1976 to evaluate school-aged children with reading disabilities [8]. The K-D test evaluates saccadic eye movements, processing speed, and visual tracking [9]. Cortical structures, including the frontal eye fields, supplementary eye fields, dorsolateral prefrontal cortex, parietal eye fields, and brainstem areas are involved in saccade generation [10–13]. The incidence of oculomotor abnormalities after SRC is high and presence of these deficits is associated with prolonged recovery times [14]. Hence, the K-D test is becoming a complement to other sideline concussion assessment tools [15,16] because of ease of use and portability [17]. Furthermore, the test can be effectively administered by non-medically trained parents and coaches [18].

Designed for sideline usage, post-exercise assessment of performance is an important aspect of the K-D test. Mild to moderate exertion has been shown to have immediate improvement on cognitive speed, so a single bout of activity could significantly affect K-D test scores [19]. Leong et al. [20] showed a significant improvement in K-D test scores after a 2.5 h exercise session. Galetta et al. [21] showed a mean 2.8 s ($p = .03$) second improvement on average on the K-D test score in basketball players before and after 2 h of intense activity, Galetta et al. [17] observed a mean improvement of 1.2 s ($p = .0001$) in boxers and MMA fighters before and after a fight, and Dhawan et al. [22] showed a 1.8 s ($p < .05$) improvement on the KD Test before and after a high school hockey game.

Relatively few studies have evaluated the K-D test in an outpatient clinical setting. Subotic et al. [23] assessed the diagnostic value of the K-D test in a cross-sectional comparison among acutely concussed patients, persistent post-concussive syndrome (PPCS) patients, and in healthy controls. PPCS patients took longer than healthy controls to complete the K-D test but participants with PPCS did not differ from acutely concussed patients and no differences were found between acutely concussed patients and controls. The authors concluded that the K-D test has more assessment value immediately following a concussive hit on the sideline and was most useful when athletes had pre-injury baseline scores [23]. Unfortunately, baseline K-D test scores are not available for every athlete. The purpose of this study was to examine the diagnostic and prognostic value of the K-D test without individual baseline scores in a clinic setting and to see the effect of the BCTT on K-D test values. We hypothesize that K-D test performance would be significantly different between acutely concussed and healthy controls on the initial visit, between those who recovered and those who did not, and would be affected by exercise on BCTT.

2. Methods

2.1. Study participants

Participants for this study were obtained from a larger study investigating the role of exercise in concussion recovery (clinicaltrials.gov NCT02710123). Acutely concussed (AC) participants were recruited following a clinical visit at a University Concussion Management Clinic.

A diagnosis of acute concussion was made during the initial clinic visit (Visit 1) by a physician based on history, a concussion symptom checklist, a physical examination, and demonstrated exercise intolerance [24,25]. Inclusion criteria included individuals aged 13–18 years with physician-diagnosed SRC within 10 days of injury. Exclusion criteria were previous exposure to the K-D test, history of learning disorders including ADHD and reading disorders, history of ophthalmic disorders except corrective lens, inability to understand English, and presence of other orthopedic injuries that would affect research interventions. Age-matched controls (MC) were recruited from both in- and out-of-season sports from local adolescent sport teams. Inclusion criteria included participants aged 13–18 years with no history of concussion within the past year. Exclusion criteria were the same as for AC. For analysis, the AC group was further divided according to duration of recovery. The Institutional Review Board (IRB) at the University at Buffalo approved all study protocols. Informed consent and/or age appropriate assent was obtained from each participant.

2.2. Demographic & previous medical history questionnaire

All participants completed a questionnaire regarding their previous concussion history and general health. Demographic information, including student status and athletic involvement, was also obtained.

2.3. King-Devick (K-D) test

The K-D test is an assessment tool involving a series of rapid number naming tasks [17,21,25] that take < 2 min to administer [26]. The participant reads aloud a sequence of single digit numbers from left to right. The paper version of the K-D test was used for this study. The K-D test includes one demonstration card and three visually distinct test cards of increasing difficulty (Fig. 1). Card 3 represents the most difficult card with high levels of visual complexity. Standardized instructions were given to read the single digit numbers from left to right as quickly as possible without errors. The time to complete each individual card and number of errors were recorded. Participants started with the demonstration card (upper left quadrant) to familiarize them with the task. Participants were then asked to read each test card with the same instructions. Number of non-corrected errors, including omissions and misspeaks, was recorded as per the instructions in the K-D test manual [17,27]. The K-D test was completed twice at Visit 1 prior to exercise, with the faster of the two trials being recorded as the baseline time. This was performed pre-exercise on Visit 1 and once at all subsequent time points.

2.4. Eye movement assessment

The physician-performed physical examination included orthostatic vital signs, cranial nerve function, cervical motion, motor coordination, and oculomotor function [28]. For this study, only the eye movement data for smooth pursuits and repetitive saccades were used because they represent eye movements assumed to be tested during the K-D test. For smooth pursuits, the patient was asked to visually track an object moving slowly in the horizontal direction (20°/s) with the head stationary. Target movement was limited to 30° from neutral to avoid eliciting end-gaze nystagmus. Abnormal eye movement included saccadic (or jerking) eye motion, corrective (catch-up or back-up) saccades, or loss of visual fixation. For repetitive saccades, the physician held both index fingers three feet apart at half an arm length's distance from the patient. The patient was instructed to move the eyes from one finger to the other in rapid succession in the horizontal visual plane thirty times. Abnormal responses include delayed initiation of eye movement, slow velocity, or inaccurate movements such as over/undershooting with > 1 re-fixation saccade, or symptom provocation of increased headache or dizziness.

2.5. Buffalo concussion treadmill test (BCTT)

The BCTT is a graded exercise test that can diagnose physiological dysfunction after concussion and has demonstrated prognostic value [7]. The test is a modification of the cardiac Balke protocol. Patients begin with a speed of 3.2–3.6 mph (depending on height) at 0% incline. At the first minute and every minute thereafter, the incline is increased by 1°. Speed remains constant until a maximum incline is reached or the patient cannot continue. Rating of perceived exertion (RPE, Borg scale) [29], symptoms (10-point Visual Analog Scale [VAS]), heart rate (Polar HR monitor, Model #FIT N2965; Kempele, Finland) are recorded every minute. The test is stopped when there is symptom exacerbation (≥ 3 -point increase from the pre-treadmill VAS value) or exhaustion (RPE of 18 or more). The test is also stopped if there are any visible signs of distress. The treadmill test has a high degree of interrater reliability (95%) for distinguishing symptom exacerbation [30].

2.6. Study design

The initial research visit (Visit 1) occurred immediately after the first clinic visit where participants completed the demographic and history questionnaire. The K-D test and the Post- Concussion Symptom Scale (PCSS) from the Sport Concussion Assessment Tool (SCAT)-3 were administered once before and once after each BCTT trial [31]. SCAT-3 was the current version at the time the study was conducted. Symptoms, RPE, heart rate, and time spent on the treadmill were recorded. The post-treadmill K-D test was administered immediately following a 2-minute cool-down. AC followed up with the physician and performed the BCTT after one week. Recovery was defined a-priori as a normal level of symptoms, a normal physical examination, and normal exercise tolerance. Medical records were searched to find the date of recovery. Research visits immediately followed clinic visits. At study completion, AC was divided into two groups based on recovery time. The fast recovery group (FRG) included AC participants recovered by the second clinic visit. The slow recovery group (SRG) included AC participants not recovered by the second clinic visit. MC completed 2 trials of the BCTT one week apart. At the initial visit, participants completed the demographic and history questionnaire. Controls were given a physical exam prior to each BCTT trial. The K-D test and the PCSS from the SCAT 3 were administered before and after each BCTT following the same procedures as in the AC group.

2.7. Statistical analysis

ANOVA was used to assess group-wise differences in age, height, weight, days to initial visit, recovery time, and K-D test times (single card and cumulative). Chi-square was used to assess group-wise differences in sex, athletic status, history of concussions, and vision tests. Bonferroni correction was performed on single card K-D test times and a p-value of 0.016 (0.05/3) was considered significant. MC Visit 1 and Visit 2 pre-exercise scores were compared to assess the learning effect. Visit 1 pre-exercise scores were compared between MC and AC to assess diagnostic ability. Pre- and post-exercise score were compared for the effect of exercise on test performance and compared between groups. All data were analyzed using SPSS Version 20 (IBM Corporation USA).

3. Results

3.1. Demographics

Demographic data for MC and AC are shown in Table 1. The mean recovery time for AC was 16.6 ± 12.0 days from date of injury. The mean pre-treadmill PCSS score for MC was 0.69 ± 1.0 and for AC was 30.79 ± 18.1 ($p < .001$). Demographic data for FRG and SRG are shown in Table 2. FRG was slightly older than SRG and had a history of more concussions but this difference was not significant. Mean pre-

Table 1

Demographic and oculomotor information for matched controls (MC) and acutely concussed (AC) participants.

	MC (n = 30)	AC (n = 46)	P-value*
Age (years)	15.8 ± 1.4	15.4 ± 2.1	0.287
Height (centimeters)	172.1 ± 10.5	168.4 ± 10.3	0.131
Weight (kilograms)	65.1 ± 11.1	63.0 ± 12.4	0.466
Sex			0.138
Male	22 (73%)	26 (56.5%)	
Female	8 (27%)	20 (43.5%)	
Athlete	29 (97%)	44 (95.7%)	0.770
History of concussion	7 (23%)	20 (43.5%)	0.281
Post-Concussion Symptom Scale (max = 132)	0.69 ± 1.0	30.79 ± 18.1	< 0.001
Days from injury	–	5.0 ± 2.4	–
Recovery time (days)	–	16.6 ± 12.0	–
Abnormal Smooth Pursuits	1 (3%)	8 (17%)	0.064
Abnormal Repetitive Saccades	1 (3%)	17 (37%)	0.001

"Bold" indicates significant p-values.

* P-values ≤ .05 considered significant.

Table 2

Demographic information for fast recovery group (FRG) and slow recovery group (SRG).

	FRG (n = 23)	SRG (n = 23)	P-value*
Age (years)	16.0 ± 1.8	14.8 ± 2.1	0.051
Height (centimeters)	171.6 ± 9.9	165.2 ± 9.9	0.034
Weight (kilograms)	66.4 ± 8.7	59.5 ± 14.8	0.064
Sex			1.00
Male	13 (57%)	13 (57%)	
Female	10 (43%)	10 (43%)	
Athlete	22 (96%)	23 (100%)	0.312
History of concussion	7 (30%)	13 (57%)	0.137
Post-Concussion Symptom Scale (max = 132)	24.95 ± 13.6	37.20 ± 20.4	0.027
Days from injury	4.9 ± 2.6	5.0 ± 2.2	0.853
Recovery time (days)	9.4 ± 3.4	23.8 ± 13.1	< 0.001
Abnormal Smooth Pursuits	3 (13%)	5 (22%)	0.437
Abnormal Repetitive Saccades	7 (30%)	10 (43%)	0.359

"Bold" indicates significant p-values.

* P-values ≤ .05 considered significant.

exercise PCSS scores were significantly different between FRG and SRG (24.95 ± 13.6 vs. 37.20 ± 20.4, p = .027).

3.2. K-D test performance

The mean K-D test times for MC and AC at Visit 1 are presented in Table 3. The pre- and post-treadmill K-D test times for AC were significantly higher (worse) compared with MC at Visit 1 except for pre-exercise Card 1, which was not significant after Bonferroni correction.

Table 3

Visit 1 K-D test performance, mean time in seconds.

	MC (n = 30)	AC (n = 46)	P-value*
Visit 1 Pre-test			
Card 1	15.1 ± 3.1	17.1 ± 4.3	0.038
Card 2	14.9 ± 2.8	17.6 ± 4.3	0.002
Card 3	16.0 ± 3.7	19.0 ± 5.0	0.007
Total	46.1 ± 9.2	53.7 ± 13.0	0.007**
Visit 1 Post-test			
Card 1	14.3 ± 2.7	16.5 ± 5.0	0.013
Card 2	14.0 ± 3.2	17.1 ± 4.7	0.001
Card 3	14.7 ± 3.0	18.1 ± 4.8	< 0.001
Total	43.1 ± 8.5	51.6 ± 14.0	0.001**

"Bold" indicates significant p-values.

* P-values ≤ .016 considered significant.

** P-values ≤ .05 considered significant.

Statistically significant differences in mean total K-D test time were found between the MC and SRG for Visit 1 pretest (46.1 ± 9.2 vs 53.7 ± 15.0, p = .002) and Visit 2 pretest (42.9 ± 9.1 vs 50.0 ± 8.6, p = .006). FRG and SRG times differed significantly only on Card 3 on the Visit 1 pre-exercise test (17.5 ± 3.1 vs 20.6 ± 6.1, p = .039). No statistically significant differences were found between MC and FRG at Visit 1 pretest (46.1 ± 9.2 vs 50.4 ± 10.0, p = .111) and Visit 2 pretest (42.9 ± 9.1 vs 42.9 ± 7.5, p = .996). Comparison of pre- and post-treadmill K-D test times for MC, FRG, and SRG are presented in Table 4. Both MC and FRG showed an improvement in K-D test times after the treadmill test at Visit 1 and Visit 2. MC improved by 3.0 s (p < .001) at Visit 1 and by 2.3 s (p < .001) at Visit 2. FRG improved by 3.1 s (p = .002) at Visit 1 and 1.5 s (p = .016) at Visit 2. SRG, however, did not significantly improve post-treadmill K-D test times at Visit 1 (p = .478) or at Visit 2 (p = .772).

3.3. Vision assessment

Abnormal vision assessment findings from the Visit 1 physical examination are summarized in Tables 1 and 2. AC had significantly higher rates of abnormal repetitive saccades but not smooth pursuits during Visit 1 when compared with MC. There were no significant differences in abnormal smooth pursuits or repetitive saccades when comparing FRG and SRG. Mean K-D test times and vision assessment findings are presented in Table 5. AC participants with abnormal smooth pursuits did not differ in K-D test times on Visit 1, but AC participants with abnormal repetitive saccades had significantly longer times for Card 3 (p = .012) and for total K-D time (p = .049).

4. Discussion

This study showed that acutely concussed adolescents who did not improve their post-treadmill test exercise K-D test times took significantly longer to recover from SRC when compared with adolescents whose post-exercise times improved. The data suggest that lack of improvement in K-D test performance after an exercise test is prognostic for delayed recovery from SRC, which is a clinically useful result. The sensitivity of the K-D test for detecting SRC has been validated in multiple studies across various sports including rugby, mixed martial arts, boxing, football, and hockey, and in multiple age cohorts including youth, high school, and collegiate athletes [15–18,20–21,28,32–37]. Our study confirms that K-D test performance distinguished the concussed cohort from the non-concussed cohort, even without baseline K-D measures. Such an effect has been well demonstrated in adolescent athletes tested on the sideline but if they had a baseline measurement [28,32,33]. Interestingly, when our concussed cohort was subdivided into those who recovered fast (FRG) versus those who did not (SRG), only group SRG was distinguishable from the healthy controls. SRG was slightly younger and had a history of more concussions than FRG, which has been associated with longer recovery and worse K-D test times, but these differences were not significant.

Other researchers have found no difference in K-D performance between acutely concussed patients and healthy controls in the clinic setting [23]. Our results may differ because Subotic et al. [23] studied much older non-athlete participants (concussed and controls had a mean age of 43 and 37 years, respectively) while we studied adolescent athletes after SRC. Additionally, SRC may be a different clinical entity than non-SRC because SRC patients recover faster, have fewer total symptoms, have lower severity scores, and have fewer clinic visits [38]. Thus the diagnostic value of the K-D test in a clinical setting may vary with respect to age and mechanism of injury. Moran et al. [39] found that healthy athletes aged 8–11 had slower KD test times and more errors than individuals aged 12–14 years (p < .001). Weise et al. [40] showed that KD scores improved with age in junior high and high school athletes and that they were highly variable within these age groups. Alsalaheen et al. [41] showed that 16–18 year old athletes

Table 4

Comparison of pre- and post-treadmill K-D test times at Visit 1 and Visit 2.

	Visit 1			Visit 2		
	Pre-test	Post-test	P-value*	Pre-test	Post-test	P-value*
MC						
Card 1	15.1 ± 3.1	14.3 ± 2.7	0.006	14.2 ± 3.2	13.2 ± 2.6	< 0.001
Card 2	14.9 ± 2.8	14.0 ± 3.2	0.006	13.9 ± 2.8	13.3 ± 3.2	0.036
Card 3	16.0 ± 3.7	14.7 ± 3.0	< 0.001	14.7 ± 3.5	14.2 ± 3.6	0.057
Total	46.1 ± 9.2	43.1 ± 8.5	< 0.001**	42.9 ± 9.1	40.6 ± 9.2	< 0.001**
FRG						
Card 1	16.2 ± 3.7	15.2 ± 3.2	0.015	13.9 ± 2.5	13.3 ± 2.3	0.058
Card 2	16.7 ± 3.8	15.6 ± 3.3	0.035	14.0 ± 2.5	13.8 ± 2.4	0.569
Card 3	17.5 ± 3.1	16.6 ± 3.8	0.036	15.0 ± 2.9	14.3 ± 2.7	0.026
Total	50.4 ± 10.0	47.3 ± 9.8	0.002**	42.9 ± 7.5	41.4 ± 6.8	0.016**
SRG						
Card 1	17.9 ± 4.8	17.9 ± 6.0	0.996	16.2 ± 2.9	16.2 ± 3.7	0.951
Card 2	18.5 ± 4.6	18.5 ± 5.6	0.998	16.4 ± 3.1	16.1 ± 3.9	0.411
Card 3	20.6 ± 6.1	19.6 ± 5.2	0.117	17.5 ± 3.1	17.6 ± 4.1	0.766
Total	57.0 ± 15.0	56.0 ± 16.3	0.478**	50.0 ± 8.6	49.7 ± 11.3	0.772**

"Bold" indicates significant p-values.

* P-values ≤ .016 considered significant.

** P-values ≤ .05 considered significant.

Table 5

Vision assessment and mean K-D test times (in seconds) among AC at Visit 1.

	AC with normal repetitive saccades (n = 29)	AC with abnormal repetitive saccades (n = 17)	P-value
Card 1	16.3 ± 3.3	18.3 ± 5.6	0.147
Card 2	16.8 ± 3.8	18.9 ± 4.8	0.110
Card 3	17.6 ± 3.5	21.4 ± 6.3	0.012*
Cumulative	50.8 ± 10.2	58.6 ± 16.0	0.049**
	AC with normal smooth pursuits (n = 38)	AC with abnormal smooth pursuits (n = 8)	
Card 1	16.9 ± 4.6	17.6 ± 3.2	0.715
Card 2	17.4 ± 4.4	18.6 ± 3.9	0.459
Card 3	18.9 ± 5.0	19.4 ± 5.3	0.811
Cumulative	53.3 ± 13.5	55.6 ± 11.3	0.648

"Bold" indicates significant p-values.

* P-values ≤ .016 (0.05/3) considered significant.

** P-values ≤ .05 considered significant.

achieved faster KD times than 13–15 year old athletes whereas Anderson et al. [42] and Benedict et al. [43] showed that KD test scores become worse with age in adults (mean age 40.5 and 36 years, respectively). These studies suggest that KD test scores, eye tracking, and/or visual processing improve with age, peak in the late adolescent and young adult phase, and then decline. On the other hand, McIntyre et al [44], and Oberlander [45] found no correlation between age and KD test scores. Future studies should explore the role of age and concussion etiology in the diagnostic ability of the K-D test and clinicians should consider this information when administering the K-D test in the outpatient setting.

The lack of a learning effect on K-D test performance after a standardized exercise tolerance test (the BCTT) identified patients that took significantly longer to recover from SRC. This is an interesting and potential useful addition to concussion assessment in the outpatient clinic setting, particularly in those doing exertional testing to manage concussed adolescents. Exercise typically has a beneficial effect on K-D test performance in non-concussed athletes. King et al. [46] reported a significant improvement in KD test time (50 vs. 45 s, $p < .001$) between the first and second test after a 10 min interval in amateur rugby players, consistent with a learning effect. Oberlander et al. [45] found acceptable re-test reliability of the KD test ($ICC = 0.81$, 95% CI 0.73–0.87) because there was a 4.3 ± 0.5 s ($p < .001$) improvement between first and second test and a 6.9 ± 0.5 s ($p < .001$)

improvement between first and third test. Our data show that both MC and FRG improved K-D test performance following the BCTT at the initial visit: MC by 3.0 s ($p < .001$) and FRG by 3.1 s ($p = .002$). SRG, however, did not improve K-D test performance after the Visit 1 treadmill test. Similar results were found for Visit 2, with both MC and FRG continuing to improve post-treadmill cumulative K-D test times while SRG did not. Therefore, K-D test performance following the BCTT was able to distinguish AC participants who recovered within a week of their initial visit (mean 9.4 days post-injury) from AC who took longer to recover (mean 23.8 days post-injury). Thus, lack of improvement in K-D test performance after a standardized exercise test within a week of SRC could help clinicians identify adolescent patients at risk for prolonged recovery from SRC, which is important for school and athletic team planning purposes.

The reason for the lack of improvement in K-D test times in those slow to recover is unclear. It is speculated that the SRG group had exacerbation of symptoms following exercise that interfered with reading the K-D test cards. Human experimental data indicate that excessive activity after concussion can increase symptoms and worsen neurocognitive performance [46]. A study of neurocognitive performance following SRC showed that concussion can be associated with a lack of expected practice effect rather than a decline per se, particularly on traditional measures of speed and complex attention [47]. Another possibility for the lack of improvement in post-exercise K-D test times is that SRG experienced a more severe initial injury compared to FRG. Pre-treadmill symptom evaluation at Visit 1 revealed that SRG had significantly more symptoms than FRG, indicating a difference in initial concussion severity.

Lastly, vision involves the relay of information through a number of dispersed brain structures that include cortical and brainstem areas [10–13]. The wide distribution and location of neural networks involved make the pathways susceptible to damage during concussion. The vision assessments used in the present study test two separate pathways that may be injured together or separately. On average, participants with abnormal repetitive saccades took 8 s longer to complete the K-D test than participants with normal repetitive saccades. It was perhaps unsurprising to find an association of abnormal repetitive saccades with longer K-D test times because saccadic testing assesses quick fixation and re-fixation moves between two targets. We did not find a relation between abnormal smooth pursuits and K-D test scores, perhaps because the saccades in a smooth pursuit task are corrective, not defective [48]. These results suggest that the K-D test may also be a

useful clinical tool for assessing abnormal repetitive saccades in adolescents.

The current study has several limitations. Although we were as uniform as possible with our instructions, the K-D test may be affected by patient motivation. Researchers were not blinded to AC and MC groups so there is a possibility of bias. All participants were tested using the paper version of the K-D test because the study began before the computer version of the test was released. There are no documented limitations for using the paper version, however, the results of this study may not generalize to the computerized K-D test. This study had a higher percentage of males in the MC cohort compared with the AC cohort but this difference was not statistically significant. Further assessments should include sex-matched controls to eliminate sex as a confounder. The BCTT is a graded treadmill test and involves limited rotational head movement. Some symptoms of concussion, specifically vestibular and oculomotor symptoms, may be triggered by rotational head movement in sport activities that should be considered in the future. Lastly, the study focused on a specific cohort of adolescent athletes suffering from SRC. Therefore, results may not easily generalize to other ages, non-athletes, non-SRC patients, or patients with learning disorders. Clinicians should keep this in mind when administering the K-D test within a clinical setting.

5. Conclusion

Our study shows that a lack of improvement in K-D test performance immediately after a standardized exercise test (the BCTT) performed within a week of SRC was associated with significantly prolonged recovery in adolescent athletes when compared with those who demonstrated the typical post-exercise improvement in performance. This could help some clinicians give their athletic patients important prognostic information that has relevance for school and athletic team planning. We also confirmed that the K-D test is useful in identifying acutely concussed adolescents from healthy controls, even without baseline values. This may add objective data to the clinician's arsenal for helping to determine physiological recovery from SRC.

Funding

Research reported in this publication was supported by the National Institute of Neurological Disorders and Stroke of the National Institutes of Health under award number 1R01NS094444. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Funding was also provided by the Ralph and Mary Wilson Foundation and the Robert Rich Family Foundation.

Conflict of interest

The authors declare no conflict of interest.

References

- [1] D.J. Wiebe, R.D. Comstock, M.L. Nance, Concussion research: a public health priority, *Inj. Prev.* 17 (1) (2011) 69–70.
- [2] J.A. Langlois, W. Rutland-Brown, K.E. Thomas, Traumatic Brain Injury in the United States: Emergency Department Visits, Hospitalizations, and Deaths, Department of Health and Human Services, Centers for Disease Control and Prevention, Division of Acute Care, Rehabilitation Research and Disability Prevention, National Center for Injury Prevention and Control, 2004.
- [3] N. Carney, et al., Concussion guidelines step 1: systematic review of prevalent indicators, *Neurosurgery* 75 (2014) S3–S15.
- [4] P. McCrory, et al., Consensus statement on concussion in sport—the 5th international conference on concussion in sport held in Berlin, October 2016, *Br. J. Sports Med.* 51 (2017) 838–847 (p. bjsports-2017-097699).
- [5] R.J. Echemendia, et al., The Sport Concussion Assessment Tool 5th edition (SCAT5), *Br. J. Sports Med.* 51 (2017) 851–858 (p. bjsports-2017-097506).
- [6] A. Bohm, et al., Sideline testing in youth and collegiate athletes: what does vision add to the concussion puzzle? *Neurology* 84 (2015).
- [7] J.J. Leddy, B. Willer, Use of graded exercise testing in concussion and return-to-activity management, *Curr. Sports Med. Rep.* 12 (6) (2013) 370–376.
- [8] A. King, The Proposed King-Devick Test and its Relation to the Pierce Saccade Test and Reading Levels [Senior Research Project], Illinois College of Optometry, Chicago, Ill, 1976.
- [9] S. Lieberman, A. Cohen, J. Rubin, NYSOA KD test, *J. Am. Optom. Assoc.* 54 (7) (1983) 631–637.
- [10] T. Paus, et al., Role of the human anterior cingulate cortex in the control of oculomotor, manual, and speech responses: a positron emission tomography study, *J. Neurophysiol.* 70 (2) (1993) 453–469.
- [11] C. Pierrot-Deseilligny, D. Milea, R.M. Müri, Eye movement control by the cerebral cortex, *Curr. Opin. Neurol.* 17 (1) (2004) 17–25.
- [12] C. Pierrot-Deseilligny, et al., Cortical control of ocular saccades in humans: a model for motricity, *Prog. Brain Res.* 142 (2003) 3–17.
- [13] R.J. Leigh, D.S. Zee, *The Neurology of Eye Movements*, vol. 90, Oxford University Press, USA, 2015.
- [14] A.J. Anzalone, et al., A positive Vestibular/Ocular Motor Screening (VOMS) is associated with increased recovery time after sports-related concussion in youth and adolescent athletes, *Am. J. Sports Med.* 45 (2) (2017) 474–479.
- [15] Z. Marinides, et al., Vision testing is additive to the sideline assessment of sports-related concussion, *Neurol. Clin. Pract.* 5 (1) (2015) 25–34.
- [16] M.S. Galetta, et al., Saccades and memory: baseline associations of the King-Devick and SCAT2 SAC tests in professional ice hockey players, *J. Neurol. Sci.* 328 (1–2) (2013) 28–31.
- [17] K. Galetta, et al., The King-Devick test as a determinant of head trauma and concussion in boxers and MMA fighters, *Neurology* 76 (17) (2011) 1456–1462.
- [18] D.F. Leong, et al., The King-Devick test as a concussion screening tool administered by sports parents, *J. Sports Med. Phys. Fitness* 54 (1) (2014) 70–77.
- [19] C.L. Hogan, J. Mata, L.L. Carstensen, Exercise holds immediate benefits for affect and cognition in younger and older adults, *Psychol. Aging* 28 (2) (2013) 587.
- [20] D.F. Leong, et al., The King-Devick test for sideline concussion screening in collegiate football, *J. Opt.* 8 (2) (2015) 131–139.
- [21] K.M. Galetta, et al., The King-Devick test and sports-related concussion: study of a rapid visual screening tool in a collegiate cohort, *J. Neurol. Sci.* 309 (1–2) (2011) 34–39.
- [22] P. Dhawan, et al., King-Devick test identifies symptomatic concussion in real-time and asymptomatic concussion over time. (S11. 003), *Neurology* 82 (10 Supplement) (2014) (p. S11. 003–S11. 003).
- [23] A. Subotic, W.K. Ting, M.D. Cusimano, Characteristics of the King-Devick test in the assessment of concussed patients in the subacute and later stages after injury, *PLoS One* 12 (8) (2017) e0183092.
- [24] M.N. Haider, et al., A systematic review of criteria used to define recovery from sport-related concussion in youth athletes, *Br. J. Sports Med.* 0 (2017) 1–14 (p. bjsports-2016-096551).
- [25] M.K. Oride, et al., Reliability study of the Pierce and King-Devick saccade tests, *Optom. Vis. Sci.* 63 (6) (1986) 419–424.
- [26] K.M. Galetta, et al., Adding vision to concussion testing: a prospective study of sideline testing in youth and collegiate athletes, *J. Neuroophthalmol.* 35 (3) (2015) 235–241.
- [27] King-Devick technologies, i, King-Devick Test Online System v4.0 Instructions [(cited 2018 10/10/2018]; Available from:], 2018, <https://kingdevicktest.com/wp-content/uploads/2018/06/KDTOS-Instructions-v4.0-.pdf>.
- [28] J.M. Matuszak, et al., A practical concussion physical examination toolbox: evidence-based physical examination for concussion, *Sports Health* 8 (3) (2016) 260–269.
- [29] G. Borg, Borg's perceived exertion and pain scales, Champaign (IL): Human Kinetics, 1998.
- [30] J.J. Leddy, et al., Reliability of a graded exercise test for assessing recovery from concussion, *Clin. J. Sport Med.* 21 (2) (2011) 89–94.
- [31] E.Y. Chin, et al., Reliability and validity of the Sport Concussion Assessment Tool-3 (SCAT3) in high school and collegiate athletes, *Am. J. Sports Med.* 44 (2016) 2276–2285 (p. 0363546516648141).
- [32] D.H. Seidman, et al., Evaluation of the King-Devick test as a concussion screening tool in high school football players, *J. Neurol. Sci.* 356 (1) (2015) 97–101.
- [33] M. Duenas, G. Whyte, R. Jandial, Sideline concussion testing in high school football on Guam, *Surg. Neurol. Int.* 5 (2014).
- [34] D. King, et al., Concussions in amateur rugby union identified with the use of a rapid visual screening tool, *J. Neurol. Sci.* 326 (1) (2013) 59–63.
- [35] D. King, T. Clark, C. Gissane, Use of a rapid visual screening tool for the assessment of concussion in amateur rugby league: a pilot study, *J. Neurol. Sci.* 320 (1) (2012) 16–21.
- [36] D. King, et al., The King-Devick test was useful in management of concussion in amateur rugby union and rugby league in New Zealand, *J. Neurol. Sci.* 351 (1) (2015) 58–64.
- [37] D. King, et al., Use of the King-Devick test for sideline concussion screening in junior rugby league, *J. Neurol. Sci.* 357 (1) (2015) 75–79.
- [38] G. Kyle Harrold, et al., Rapid sideline performance meets outpatient clinic: results from a multidisciplinary concussion center registry, *J. Neurol. Sci.* 379 (2017) 312–317.
- [39] R. Moran, T. Covassin, Risk factors associated with baseline King-Devick performance, *J. Neurol. Sci.* 383 (2017) 101–104.
- [40] K.K. Weise, et al., King-Devick and pre-season visual function in adolescent athletes, *Optom. Vis. Sci.* 94 (1) (2017) 89–95.
- [41] B. Alsalaheen, et al., King-Devick Test reference values and associations with balance measures in high school American football players, *Scand. J. Med. Sci. Sports* 26 (2) (2016) 235–239.

- [42] H.D. Anderson, S.A. Biely, Baseline King–Devick scores for adults are not generalizable; however, age and education influence scores, *Brain Inj.* 31 (13–14) (2017) 1813–1819.
- [43] P.A. Benedict, et al., Gender and age predict outcomes of cognitive, balance and vision testing in a multidisciplinary concussion center, *J. Neurol. Sci.* 353 (1–2) (2015) 111–115.
- [44] L. McIntyre, M. Campo, Descriptive values for dancers on baseline concussion tools, *J. Athl. Train.* 52 (11) (2017) 1035–1040.
- [45] T.J. Oberlander, B.L. Olson, L. Weidauer, Test-retest reliability of the King-Devick test in an adolescent population, *J. Athl. Train.* 52 (5) (2017) 439–445.
- [46] C.W. Majerske, et al., Concussion in sports: postconcussive activity levels, symptoms, and neurocognitive performance, *J. Athl. Train.* 43 (3) (2008) 265–274.
- [47] A. Collie, et al., Cognition in the days following concussion: comparison of symptomatic versus asymptomatic athletes, *J. Neurol. Neurosurg. Psychiatry* 77 (2) (2006) 241–245.
- [48] D. Purves, et al., Types of eye movements and their functions, *Neuroscience* (2001) 361–390.



A pilot exploratory study comparing the King-Devick test (KDT) during and between migraine attacks

Chia-Chun Chiang¹, Amaal J Starling¹, Matthew R Buras², Michael A Golafshar² and Juliana H VanderPluym¹

Abstract

Background: The King-Devick test is a timed rapid number naming task that involves complex cerebral functions. The objective of this pilot exploratory study is to determine whether there is a difference in the King-Devick test during a migraine attack compared to the interictal phase.

Methods: We evaluated 29 adult subjects with migraine with aura or migraine without aura. For each participant, we performed King-Devick tests during migraine attacks and interictal phases. Subjects served as their own controls.

Results: The King-Devick test was slower during the migraine attack compared to the interictal baseline (median 4.6 sec slower, $p < 0.001$). The slowing of the King-Devick test during migraine attack was more prominent in those with migraine with aura compared to subjects with migraine without aura (median 7.5 vs. 2.8 sec, $p = 0.028$).

Conclusions: This exploratory, observational study shows changes in the King-Devick test during migraine compared to the interictal phase. Future studies are required to determine if the King-Devick test may be used as a rapid and simple tool to objectively characterize migraine-associated disability.

Keywords

King-Devick test, rapid eye movement, migraine attack, migraine-related disability

Date received: 5 October 2018; revised: 11 July 2019; 17 August 2019; 13 September 2019; 5 October 2019; accepted: 7 October 2019

Introduction

Migraine is accompanied by a myriad of symptoms such as sensory sensitivities (photophobia, phonophobia), visual impairment and cognitive dysfunction, which reflects altered central nervous system functioning. Past studies have demonstrated alterations in central nervous system functioning with impaired neuropsychological performances (1). However, the studies were done comparing migraine patients to healthy controls; therefore, it was unclear how migraine attacks affect cerebral function compared to one's interictal phase. Additionally, the neurocognitive exams could take a long time to perform. It would be beneficial to have a simple, fast and objective testing modality that could be applied to quantify altered task-performing ability and characterize migraine-related dysfunction and disability during attacks.

The King-Devick test (KDT) is a rapid number naming test performed on an iPad or tablet that requires intact vision, saccadic eye movements,

language, concentration and attention. Participants are asked to read the numbers on the test cards from left to right as quickly as possible without making errors (Figure 1). The time required to complete each test card and the number of errors are recorded. The KDT was originally designed to assess reading difficulties in children. It has since been applied to characterize cerebral function in different neurological disorders. KDT is now a validated sideline concussion screening tool (2), and it has also been shown to be sensitive to hypoxia (3) and sleep deprivation (4), as well as the progression of neurological diseases such as multiple sclerosis (5) and Parkinson's disease (6).

¹Department of Neurology, Mayo Clinic, Scottsdale, AZ, USA

²Division of Health Sciences Research, Mayo Clinic, Scottsdale, AZ, USA

Corresponding author:

Juliana H VanderPluym, 13400 E Shea Blvd, Scottsdale, AZ 85259, USA.
Email: VanderPluym.Juliana@mayo.edu

The objective of this exploratory, observational study is to determine whether KDT: a) Demonstrates a difference during a migraine attack compared to the interictal phase, b) demonstrates between-group differences among patients with migraine with aura (MwA) and migraine without aura (MwoA).

Methods

This study was approved by the institutional review board at Mayo Clinic in August 2016. All subjects provided signed written informed consent prior to enrollment. The enrollment and data collection took place from September 2016 to November 2017.

Subjects

Forty subjects were screened. Among them, 31 subjects were enrolled and nine were excluded as they met exclusion criteria. To best accommodate the time requirements and need for the KDT to be administered by the same trained examiner, we aimed to enroll Mayo

Clinic employees. We included adult subjects that met the ICHD-3 beta diagnostic criteria for migraine with and without aura, either episodic or chronic migraine. Patients with chronic migraine had to have at least 15, but less than 25 headache days per month to be enrolled. We were able to evaluate at least one migraine attack and one interictal phase in 29 subjects, and those were the subjects included in our statistical analysis. For the other two subjects, we collected data during their interictal phase; however, we were not able to collect data during their migraine attack and therefore those two subjects were not included in the statistical analysis. For the 29 subjects, three had chronic migraine, 26 had episodic migraine; 13 had MwoA, and 16 had MwA. All 16 subjects with aura had visual aura and one of the 16 had both visual and sensory aura. The majority of the subjects had already received the diagnosis of episodic or chronic migraine with or without aura by their neurologist or family physician before being enrolled in this study. Regardless of whether they had an established diagnosis by a medical professional prior to the enrollment,

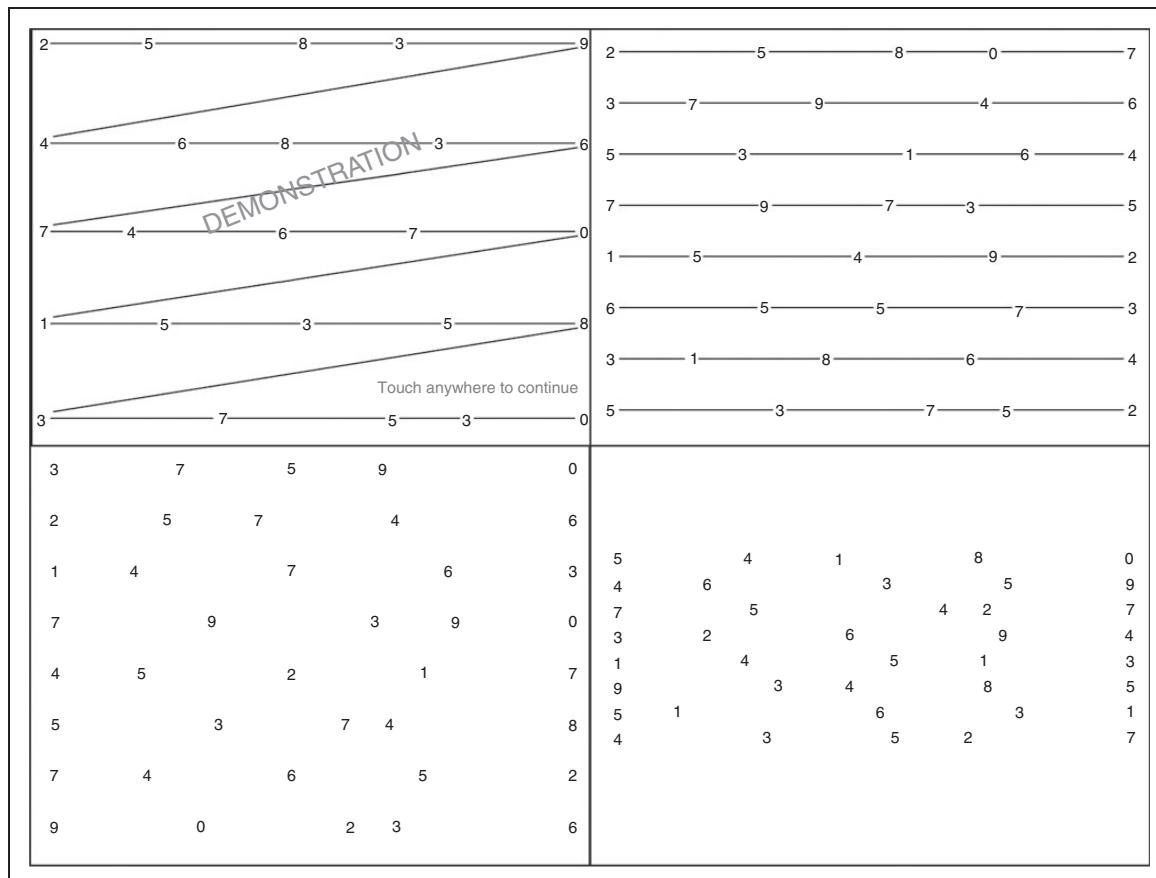


Figure 1. The King-Devick test. Subjects were asked to read the numbers on the cards from the left to the right as quickly as they could, without making any errors. This requires intact saccadic eye movement, vision, language, concentration and attention.

at the enrollment interview, the diagnosis of episodic or chronic migraine with or without aura was verified by a neurology professional by carefully going through their headache history and characteristics using the ICHD-3 beta diagnostic criteria with each individual subject. Exclusion criteria included medication overuse, history of concussion within the past 2 years, other neurological diseases, history of intracranial surgeries, and ocular conditions. Subjects served as their own controls, as we were comparing their performance between their interictal baseline phase and during migraine attacks. The average age of the group was 36 years old (range 24–60 years old). There were 25 females and four males. The mean headache frequency was 6.7 days per month (range 1–20 days per month). The age, female/male ratio, and headache frequency were similar between the two groups of MwA and MwoA (Table 1).

Methods

For each participant, we planned to obtain KDT data during a migraine attack, and between migraine attacks (interictal phase). The interictal phase was defined by a duration of at least 48 hours since the termination of the last migraine attack, and without experiencing any headache or premonitory symptoms. We asked our study participants to contact us to complete KDT tests during a migraine attack and an interictal phase. The KDT tests were conducted by the same examiner (C, C-C). To avoid the learning effect, participants underwent testing in random sequence; testing was performed either first during the interictal phase then later during the migraine attack phase, or vice versa.

During the first visit with a subject, we confirmed the diagnosis of migraine and collected their migraine characteristics before proceeding with the KDT. Migraine characteristics included headache frequency, whether they had aura and the types of aura, the types of acute headache medications, the frequency of acute medication use, and whether they were using preventive medications and the types of preventive medications. During their migraine attack we collected information on the time of headache onset, whether they had taken acute headache medications, the headache pain level on visual analog scale (from 0–10), and symptoms they were experiencing other than headache, such as aura symptoms, photophobia, phonophobia, blurry vision, nausea, allodynia, neck stiffness, and cognitive dysfunction (including mental fogginess, mental slowing, and difficulty with concentration).

For the KDT, participants were asked to read numbers on tests cards on an iPad from left to right without error as quickly as possible. The amount of time and the number of errors made were recorded. At least two

Table 1. Demographic and disease characteristics of patients with migraine with aura and migraine without aura.

	Total group	With aura	Without aura	p-value
N	29	16	13	
Mean age (SD)	35.9 (9.2)	37.3 (11.0)	34.2 (6.3)	0.66 ¹
Sex, n (%)				1.0 ²
Male	4 (13.8)	2 (12.5)	2 (15.4)	
Female	25 (86.2)	14 (87.5)	11 (84.6)	
Migraine frequency, n (%)				1.0 ²
Episodic	26 (89.6)	15 (93.8)	12 (92.3)	
Chronic	3 (10.3)	1 (6.3)	1 (7.7)	
Headache days per month (SD)	6.7 (4.9)	7.0 (4.7)	6.3 (5.3)	0.51 ¹

¹Wilcoxon Rank Sum test.

²Fisher's exact test.

trials were performed during the interictal and acute migraine phases. The shortest time for which they could read without making any errors was used as their score. The iPad was held by the participants approximately 40 cm from their eyes. The iPad display was set at maximum brightness for each test and the ambient light setting was similar among clinic rooms where tests were performed. Collected data was kept in a Redcap database for analysis.

Statistical analysis

As a pilot exploratory study, power analysis was not utilized. All paired comparisons were performed using the Wilcoxon Signed Rank test, while the Wilcoxon Rank Sum test was used to perform all non-paired comparisons between groups. Correlation between numerical variables was quantified using Spearman's correlation. All analyses were performed using SAS v9.4 (SAS Institute; Cary, NC). All hypothesis tests performed were two-sided.

Results

KDT comparing migraine attack vs. interictal

There was a difference in the performance of KDT during the migraine attack compared to the interictal baseline (median of differences 4.6 seconds, $p < 0.001$) (Table 2).

Migraine attack vs. interictal baseline in migraine without aura vs. migraine with aura

For those with MwoA, the KDT was 2.8 seconds slower during a migraine attack compared to the

Table 2. The mean differences in the KDT performances during a migraine attack vs. interictal baseline of the group (n=29).

Tests (n=29)	Interictal baseline	During migraine attack	Difference in scores	p-value ¹	n (%) of patients with increased score	n (%) of patients with no Change	n (%) of patients with Decreased score
KDT (in seconds)				<0.001	25 (86.2%)	0 (0%)	4 (13.8%)
Mean (SD)	48.7 (7.3)	55.1 (8.2)	6.5 (7.6)				
Median (Q1, Q3)	47.8 (43.5, 54.3)	55.6 (49.3, 61.1)	4.6 (2.2, 9.2)				

¹Wilcoxon Signed Rank test.

Table 3. Differences in KDT performance during a migraine attack compared to the interictal baseline for migraine without aura versus migraine with aura.

MwoA (n=13)	MwA (n=16)	p-value ¹	Effect size ²
KDT differences (in seconds)		0.028	0.74
Mean (SD)	3.7 (6.8)	8.7 (7.6)	
Median (Q1, Q3)	2.8 (1.7, 5.4)	7.5 (4.1, 10.4)	

¹Wilcoxon Rank Sum test.

²Common language effect size (10).

interictal baseline phase. For those with MwA, the KDT was 7.5 seconds slower during a migraine attack compared to the interictal baseline phase. There was a difference between the two groups ($p=0.028$) in the degree of KDT slowing during a migraine attack compared to the interictal baseline (Table 3). Of note, none of the participants were actively having an aura during the KDT.

The effect of pain level on change in KDT scores

The average pain level during migraine attacks at the time of testing was 3 on an 11-point visual analog scale (VAS) among all subjects (ranged 1 to 7). There was no correlation between pain level and the change in KDT scores.

Effect of non-headache symptoms during migraine attack on the change in KDT scores

Analyses of the effect of non-headache symptoms during migraine attack showed no difference between the KDT scores and the presence of symptoms including photophobia or cognitive deficits during migraine attack. We did not observe enough events of the following symptoms to evaluate their respective effects on KDT scores: Vision changes, language deficits, nausea, neck pain, and allodynia.

Effect of headache frequency and the time of migraine onset on the KDT scores

KDT tests were performed between 1–144 hours (mean 19 hours) after headache onset, and our subjects' headache frequency ranged from 1 day per month to 20 days (mean 6.7 days) per month. We did not find any correlation between the headache frequency, or the timing between the headache onset to test time, of the KDT slowing during migraine attack compared to their interictal baseline (KDT Spearman correlation = 0.16, $p=0.40$, and Spearman correlation = -0.33, $p=0.08$, respectively).

Preventive medications

Median KDT was 4.0 seconds slower during migraine attack compared to the interictal baseline in subjects taking preventive medications (n=9; including six subjects on topiramate, one on amitriptyline, one on magnesium, two on onabotulinumtoxinA, one on propranolol, and one on paroxetine). For subjects not on preventive medications, the KDT was 5.0 seconds slower (n=20) compared to the interictal baseline. There was no between-group difference ($p=0.81$).

Discussion

The KDT is a brief and simple test of rapid number naming. It requires a number of complex central nervous system functions including intact vision, saccadic eye movements, attention, processing, reading and speech skills. Our exploratory pilot study showed that the KDT was slower during a migraine attack compared to the interictal baseline (median of differences 4.6 seconds). Between MwA and MwoA, the effect of migraine attack on change in KDT performance is more prominent in those with MwA compared to MwoA (median 7.5 vs. 2.8 seconds slower during migraine attack compared to baseline, $p=0.028$). Of note, none of the testing was performed while subjects were experiencing visual aura. A previous functional MRI study reported that patients with MwA had an

altered resting-state visual network connectivity compared to those with MwoA and healthy controls during the interictal phase (7), which could be related to more slowing during the KDT testing.

Migraine is starting to be recognized as more than just a headache disorder and therefore having ways of assessing migraine beside pain levels is important. We did not find any correlation between the change in KDT scores and the pain level during tests. This requires further evaluation, but may suggest that changes in the KDT are capturing components of migraine-associated disability beyond pain associated disability. In regard to the effect of non-headache symptoms during testing on the change in KDT scores, there was no difference in KDT scores depending on the presence of photophobia or cognitive deficits. Ten of our subjects reported cognitive deficits during migraine attacks when testing was performed. For those subjects, their KDT score was 8.5 seconds (median) slower than their interictal baseline. For those that did not perceive cognitive deficits while testing, the KDT was 4.0 seconds (median) slower than their interictal baseline. This suggests possible cerebral function changes even when there was no subjective cognitive dysfunction reported. Many reports of migraine-associated disability rely on subjective patient report. Having an objective test that can capture dysfunction regardless of subjective patient report may be useful. One subject reported language difficulties during migraine while testing, and the KDT score was 22.1 seconds slower compared to the interictal baseline. While this observation was interesting, the number of subjects was too small to make any conclusion.

Since migraine typically presents with photophobia and pain, it might be a straightforward assumption that performance on visual tasks, such as the KDT, might be poorer during a migraine attack compared to the interictal phase. Currently, there are limited ways to objectively measure the dysfunction and disability experienced by patients during a migraine attack, especially measures that are brief and simple. To establish the practical utility of KDT in migraine, further studies are required to test whether the KDT may be used as an objective and quantitative measure of migraine disability, which could provide clinicians and researchers with another indicator of disease severity and treatment efficacy in addition to the Visual Analog Scale for headache pain during the migraine attack phase.

Furthermore, examining whether there is a difference in the KDT during other migraine phases (premonitory and postdrome) or leading up to these phases, may allow us to further understand migraine pathophysiology. Additionally, future studies are warranted to duplicate our current findings, and see whether these findings are unique to migraine by

comparing KDT scores among different headache disorders (for example tension-type headache).

We chose to use the iPad version to perform KDT due to benefits of electronic testing including ease of storing test history and consistency in timing with automatic timing. A previous study comparing the paper and electronic version of KDT showed high reliability using the same platform and incompatibility across platforms (8). Migraine is often accompanied by photophobia and difficulty using screens, which might make use of an iPad test more difficult for patients. Given that the majority of people use screens in their everyday lives, we believe using the iPad version of the KDT is a reasonable testing format.

The results of this exploratory, observational study are promising but future studies will be needed. The small sample size is a limitation of this study. It is possible that the results of this study could be from selection bias in a small group of subjects.

Although the KDT testing takes no more than a few minutes, the version of the KDT used for this study required a trained test administrator. From a clinical standpoint, the requirement of a test administrator makes the testing more difficult and might limit the clinical utility. A new version of the KDT has been developed which uses voice-recording software so the testing can be individually administered. A study evaluating the individually administered KDT is pending, and if it shows similar changes to those seen in the current study would have more clinical utility given the ease of use and portable nature. Additionally, the individually administered KDT would allow for more recording opportunities and therefore the ability to potentially capture test results in the premonitory and postdrome phase.

Previous concussion studies reported learning effects for the KDT. For athletes without a head trauma, the post-fight KDT score was 1.9 seconds faster (9). To avoid learning effects in our study, we chose to administer the tests in random sequences. For eight of the 29 subjects, data during migraine attack was collected prior to interictal testing. For 21 subjects, interictal testing was performed first. As the majority of the subjects underwent interictal testing first, the slowing during migraine attack would go against a practice effect.

Overall, this is an exploratory study to determine whether there would be differences between the KDT during a migraine and the interictal phase and differences between MwA and MwoA. These potential differences may serve as an objective assessment of the migraine attack phase versus the interictal state. Performance of the KDT requires intact saccadic eye movements, vision, language, concentration and attention. Our study does not allow us to determine whether a deficit among one, some or all of these required skills explains the changes in KDT performance. Future

research should confirm these findings in a larger sample, evaluate the KDT during the premonitory and postdrome phase, and compare KDT results among different headache disorders.

Conclusion

The KDT, an objective quantitative measure of rapid number naming that requires intact saccadic eye movements, vision, language, concentration, and attention, showed slowing during migraine attacks compared to

the interictal baseline. Furthermore, the slowing of the KDT score was more prominent for those that had MwA compared to those with MwoA. Future studies are required to see if the KDT could be an effective tool to objectively characterize migraine-associated disability and its clinical utility.

Abbreviations

KDT: King-Devick test; SD: standard deviation; MwA: migraine with aura; MwoA: migraine without aura.

Clinical implications

- The KDT score was slower during the migraine attack compared to the interictal baseline, and the slowing of KDT during migraine attack was more prominent in those with MwA compared to subjects with MwoA.
- The KDT may be a rapid and simple tool to objectively characterize migraine-associated disability.

Declaration of conflicting interests

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: C-CC has nothing to disclose; AJS has received compensation from serving on advisory boards and/or consulting within the past 5 years for Alder, Amgen, Eli Lilly & Company, eNeura, and Novartis. AJS has also received funding for travel and speaking from the American Academy of Neurology and American Headache Society. MRB and MAG have nothing to disclose; JVdP reports speaking fees for Amgen and Novartis, and consulting for Healink and Teva. She has received a research grant from Amgen.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

References

1. Le Pira F, Zappala G, Giuffrida S, et al. Memory disturbances in migraine with and without aura: A strategy problem? *Cephalgia* 2000; 20: 475–478.
2. King D, Hume P, Gissane C, et al. Use of the King-Devick test for sideline concussion screening in junior rugby league. *J Neurol Sci* 2015; 357: 75–79.
3. Stepanek J, Cocco D, Pradhan GN, et al. Early detection of hypoxia-induced cognitive impairment using the King-Devick test. *Aviat Space Environ Med* 2013; 84: 1017–1022.
4. Davies EC, Henderson S, Balcer LJ, et al. Residency training: The King-Devick test and sleep deprivation: Study in pre- and post-call neurology residents. *Neurology* 2012; 78: e103–e106.
5. Moster S, Wilson JA, Galetta SL, et al. The King-Devick (K-D) test of rapid eye movements: A bedside correlate of disability and quality of life in MS. *J Neurol Sci* 2014; 343: 105–109.
6. Lin TP, Adler CH, Hentz JG, et al. Slowing of number naming speed by King-Devick test in Parkinson's disease. *Parkinsonism Relat Disord* 2014; 20: 226–229.
7. Tedeschi G, Russo A, Conte F, et al. Increased interictal visual network connectivity in patients with migraine with aura. *Cephalgia* 2016; 36: 139–147.
8. Raynowska J, Hasanaj L, Zhang I, et al. Agreement of the spiral-bound and computerized tablet versions of the King-Devick test of rapid number naming for sports related concussion. *Neurology* 2016; 86(16 S): I13.001.
9. Galetta KM, Barrett J, Allen M, et al. The King-Devick test as a determinant of head trauma and concussion in boxers and MMA fighters. *Neurology* 2011; 76: 1456–1462.
10. McGraw KO and Wong SP. A common language effect size statistic. *Psychol Bull* 1992; 111: 361–365.



Long-term test-retest evaluation of the King-Devick test in youth soccer athletes

Abigail C. Bretzin^{a,*}, Morgan Anderson^b, Ryan N. Moran^c, Tracey Covassin^b

^a Penn Injury Science Center, University of Pennsylvania, 423 Blockley Hall, Room 937, Philadelphia, PA 19104, United States

^b Michigan State University, 308 W. Circle Drive, Room 134, East Lansing, MI 48823, United States

^c The University of Alabama, 478 Russell Hall, Box 873011, Tuscaloosa, AL 35487, United States

ARTICLE INFO

Keywords:

Concussion
Youth sport
Baseline

ABSTRACT

Despite the clinical utility of baseline comparisons during concussion assessments, little evidence exists on long-term test-retest reliability of baseline tests in youth athletes. In addition, sex differences in baseline performance are inconsistent in youth athletes, warranting further research. The purpose was to examine sex differences, prevalence of false-positive scores, and long-term test-retest reliability of the King-Devick (KD) test. Healthy youth athletes (23 males, 28 females) completed the KD test prior to the Spring 2016 and Fall 2017 seasons. Two-way random-effects intraclass correlation coefficients (ICCs) were utilized to determine test-retest reliability. A mixed between-within ANOVA with post-hoc *t*-tests were used to identify the interaction between sex and season, and frequencies were used to determine abnormal test score prevalence. The KD test demonstrated good test-retest reliability (0.77[95% CI, 0.43–0.89]), with 11.8% of youth athletes having clinically meaningful improvements between Season 1 to Season 2. There was a significant sex*season interaction ($F_{(1,49)} = 4.67$, $p = .04$), with significantly greater improvements between seasons in male youth athletes compared to female youth athletes. However, 33–35% of youth athletes displayed abnormal test scores in Season 2 relative to Season 1. The KD test demonstrated good reliability and only a small percentage had clinically meaningful changes, however a high prevalence of false-positive scores were observed in this sample.

1. Introduction

Sport-related concussion (SRC) recognition and diagnosis in young athletes should be approached using a multifaceted strategy [1] comprised of a thorough history, detailed assessments of neurological, cognitive, ocular, and vestibular function, and gait and balance [2]. Appropriately, clinicians commonly use clinical exams, symptom scales, balance assessments, and computerized neurocognitive assessments during SRC management; however, less than half include an oculomotor assessment [3]. An example of an objective assessment is the King-Devick (KD) test, which assesses cognitive visual impairment [4]. The KD test was originally designed to identify learning disability and reading fluency problems caused by saccadic rhythm dysfunction, common in individuals suffering from dyslexia [5,6]. Previous research has identified the KD test as a SRC screening tool [7,8], especially in the acute phase after SRC [9]. The KD test can be a beneficial assessment for SRC, especially in youth sports settings with minimal resources, as it is an objective assessment requiring little time and equipment to

complete. However, more evidence supporting the utility of a rapid number naming test at baseline and post-concussion is needed, especially in younger populations [1].

The recent SRC consensus statement suggests that baseline testing, although not required, may aid in the interpretation of clinical assessments during SRC evaluations [2]. Baseline assessments allow the athletic trainer or team physician to have an individual performance measure, to make post-concussion and return-to-play decisions. Efforts have also been made to include multifaceted assessments at baseline, consisting of symptom reporting and balance assessment [2]. While, including objective measures like neurocognition and vestibular and ocular motor assessments may benefit clinical decision making by further identifying deficits or impairments resulting from SRC. One influence on clinical decision making using baseline assessments is the reliability of the baseline tests, which allows clinicians to differentiate normal variation within the test from variations resulting from the concussive injury itself [10]. A test-retest approach allows clinicians to interpret findings relative to the concussive outcomes and the test

* Corresponding author.

E-mail addresses: abigail.bretzin@pennmedicine.edu (A.C. Bretzin), ande1997@msu.edu (M. Anderson), rnmoran@ches.ua.edu (R.N. Moran), covassin@msu.edu (T. Covassin).

sustainability, rather than random error. Test-retest scores of the KD test across two baseline trials produced high reliability in high school aged [11] and junior Olympic athletes [12]. Adolescent athletes demonstrated good reliability ($ICC = 0.81$) across 1, 30, and 45-day intervals, with a 6.9 s improvement over 3 test administrations [13]. Similarly, Worts and colleagues [14] examined the reliability of the KD test at pre-season and during the season in high school athletes. Approximately 40 days after initial assessment, the KD test time yielded high reliability ($ICC = 0.91$), with the prevalence of false positive scores at 36%. False positive scores were defined as any increase in time or committing errors. However, intervals for the test-retest time in the study varied from 12 to 69 days, which may reflect the high prevalence of false-positive scores.

There are no consistent or evidence-based guidelines regarding how often baseline testing should be performed or when baseline testing should be updated, especially in younger athletes. Annual baseline testing is a common practice, while some researchers suggest re-administering a baseline assessment to athletes following SRC [15,16]. Others suggest updating baseline tests every 2–3 years in youth athletes to account for cognitive development and maturation [16,17], warranting research on the long-term reliability of concussion assessment tools especially in younger athletes. To better understand the long-term stability of the KD test across multiple years, the National Collegiate Athletic Association (NCAA)-Department of Defense (DoD) Concussion Assessment, Research, and Education (CARE) Consortium [10] conducted 1-year test-retest of the King-Devick test in collegiate athletes and military service cadets. After 1 year, there was good reliability ($ICC = 0.74$) with a 3-s improvement between the two baseline assessments. In another study of college-aged athletes and military cadets, there was good test-retest reliability with 1 year between test administrations ($ICC = 0.89$); however, 27% of participants had a slower or worse performance at the second administration [18]. Similarly, Naidu et al. [19] reported good reliability ($ICC = 0.88$) in Canadian Professional Football players, and 68% were within 5 s of their baseline scores. Importantly, the aforementioned studies evaluate long-term test-retest reliability of the KD test in older and elite athletes (e.g., college, professional). To date, it is unknown whether the KD test has long-term stability, in test administrations greater than 1 year apart, in youth athletes (i.e., ≤ 13 yrs).

In youth athletes, younger individuals (i.e., 9th grade and below compared to 10th grade and above) are reported to take longer to complete the KD test [20], and demonstrate significant improvements across time [13,21]. However, in addition to differences in age, male and female athletes also demonstrate performance differences in various baseline assessments [22–24]. Yet, further research should investigate differences between males and females specific to the youth sport setting. Previously, female youth athletes demonstrated better performance on the KD test compared to males [21]. However, in another sample of youth athletes, there were no significant differences between sexes for time to complete the KD test at baseline and post-baseline administrations [13]. These inconsistencies in current literature warrant further evaluation to determine if sex differences exist in youth athletes during baseline KD assessments across seasons. Therefore, the purpose of this study was to examine the long-term test-retest reliability of the KD test. In extension of test-retest reliability, we examined the prevalence of abnormal test scores, slower performance or errors committed on the KD test at the second test administration compared to the first. Second, to examine if sex differences exist in KD test performance on two baseline assessments at test and retest intervals (Season 1 and Season 2). We hypothesized the KD test would have good and clinically acceptable reliability across seasons. Based on prior research that reflected sex differences on the KD test [21], we also hypothesized that females would have faster test times compared to male athletes.

2. Materials and methods

2.1. Participants

A total of 51 athletes completed two trials of the KD test at two separate occasions. Participant demographics, including age, sex, and concussion history were included. Athletes completed the first baseline test prior to the Spring 2016 season (10.61 ± 1.5 years old) and again completed a second baseline test (12.30 ± 1.3 years old) prior to the Fall 2017 season. Thus, the time between test sessions was approximately 1.5 years apart. Athletes completed the testing at the beginning of their season prior to the start of their practice, and therefore it is assumed athletes were equipped with any needed visual corrective lenses (e.g., contacts or glasses) to participate. All participants were deemed healthy and medically qualified for physical activity.

2.2. Study protocol

Institutional Review Board approval was received prior to the start of data collection. In addition, prior to the start of each season, approval was obtained from league administrators during a pre-season meeting to recruit participants. The primary investigator communicated with team coaches and managers to inform them of the study protocol and set testing dates and times during each season. All parents and athletes completed informed consent and assent. Demographic variables including age, sex, and previous history of concussion were recorded for each participant. Prior to the start of each season, research assistants received proper training to administer the KD test. The KD test was completed with hard copy test cards. All assessments were completed on the sideline in a designated distraction-free area away from practices and individually with trained researchers. The participants first completed a demonstration card, followed by three test cards progressing in difficulty. Research assistants recorded the total time to complete each of the three test cards, which were used to calculate the total time for each trial. Participants were asked to complete two trials (Trial 1, Trial 2) during each pre-season assessment. The fastest trial (Best Trial) was recorded as the baseline time. All procedures were repeated for the second pre-season administration. Each test session was completed in less than 5 min.

2.3. Statistical analysis

Athlete demographics were assessed using descriptive statistics. Frequencies were used to determine the prevalence of false positive scores. Abnormal scores were defined as the percentage of athletes that had slower KD test times or errors at Season 2 compared to Season 1, as previously used by Worts et al. [14]. Test-retest reliability of Trial 1, Trial 2, and the Best Trial were assessed for Season 1 and Season 2 with separate two-way random effects intraclass correlation coefficient ($ICC_{2,k}$) because each athlete completed two trials of the KD test. $ICC_{2,k}$ were interpreted as poor ($ICC_{2,k} \leq 0.5$), moderate ($ICC_{2,k} = 0.5–0.75$), good ($ICC_{2,k} = 0.75–0.9$), and excellent ($ICC_{2,k} \geq 0.9$) reliability [25]. To determine the mean differences for the reliability assessment, a Friedman's test was completed due to non-normal distribution of the data. In addition, the standard error of measurement (SEM) was used to provide a measure of variability. SEM was calculated using the formula $SEM = SD\sqrt{1 - ICC}$. The minimal detectable change (MDC) was used to ensure the change wasn't a result of measurement error and was clinically relevant, commonly used in baseline reliability assessments [26]. The MDC was calculated with the formula $MDC = 1.96 * SEM * \sqrt{2}$. A mixed between-within analysis of variance (ANOVA) was used to identify any significant interactions between sexes at the two time points with statistical significance set a priori at $p < .05$. For significant sex*season interaction, post-hoc Wilcoxon signed rank tests were used to identify significant changes between baseline assessments for each sex, using Bonferroni correction. All data

Table 1

King-Devick trial 1, trial 2, and best trial total time (absolute-agreement, 2-way random-effects model).

Average measures	Intraclass correlation	95% Confidence Interval		F test with true value 0				
		Lower bound	Upper bound	Value	df1	df2	Sig.	
Trial 1	0.723	0.404	0.859	4.496	50	50	0.000	
Trial 2	0.805	0.444	0.913	7.267	50	50	0.000	
Best Total Time	0.774	0.430	0.893	5.931	50	50	0.000	

were analyzed using SPSS version 24.

3. Results

Fifty-one youth athletes (23 male, 28 female) participated in both pre-season KD test administrations, that took place approximately 1.5 years apart (e.g., Season 1: spring 2016, Season 2: fall 2017). The average age for males increased from 10.22 ± 1.4 years to 11.91 ± 1.1 years across seasons, and females' average age increased from 10.92 ± 1.5 years to 12.61 ± 1.3 years. The two-way random effects ICC_{2,k} revealed good reliability between the two seasons. ICC_{2,k} estimates and 95% confidence intervals are presented in Table 1. This study identified a SEM of 4.94 s and an MDC of 16.34 s for KD Best Trial total time. The percentages of athletes whose Best Trial total time exceeded the SEM was 60.8% (faster: 51.1% $n = 26/51$; slower: 9.8% $n = 5/51$) and the MDC was 11.8% (faster: 9.8% $n = 5/51$; slower: 2.0% $n = 1/51$), respectively.

The prevalence of abnormal retest scores using the percentage of athletes with abnormal KD test performance based on an increase in time and/or errors from Season 1 to Season 2 are presented in Table 2. Overall, 33–35% of athletes had abnormal KD performance on the two trials at Season 2 respective to Season 1; almost 14% of athletes demonstrated a slower Best Trial total time between seasons (Table 2). Overall, The Best Trial total time significantly improved between Season 1 (53.38 ± 12.4 s) and Season 2 (47.46 ± 10.4 s), $\chi^2(1) = 26.843$, $p < .001$). The mixed between-within ANOVA revealed a significant interaction of sex*season ($F_{(1,49)} = 4.67$, $p = .04$). Post-hoc Wilcoxon signed ranks test revealed male athletes (8.7 ± 10.1 s; $Z = -3.437$, $p < .001$; Cohen's d = 0.73) had significantly greater improvements from Season 1 to Season 2 compared to female (3.6 ± 6.7 s; $Z = -2.619$, $p = .009$; Cohen's d = 0.36) athletes. The average Best Trial total time between sexes are presented in Table 3.

Table 2

Prevalence of Abnormal* King-Devick performance between season 1 and season 2 ($N = 51$).

	n	%
Increase in time from Season 1		
Trial 1 Total Time	11	21.5
Trial 2 Total Time	8	15.6
Best Trial Total Time	7	13.7
Increase in errors from Season 1		
Trial 1 Errors	11	21.5
Trial 2 Errors	14	27.4
Increase in time or errors from Season 1		
Trial 1	18	35.3
Trial 2	17	33.3
Increase in time and errors from Season 1		
Trial 1	4	7.8
Trial 2	5	9.8

* Abnormal King-Devick test performance is indicated as an increase in time and/or errors During Season 2 respective to Season 1.

Table 3

King-Devick best trial total time.

	n	Season 1		Season 2	
		M (sec)	SD	M (sec)	SD
Females	28	49.17	10.6	45.55	9.5
Males	23	58.50	12.8	49.77	11.2
Total	51	53.38	12.4	47.46	10.4

4. Discussion

This study examined the long-term test-retest reliability of the KD test in youth soccer athletes. In addition, this study assessed the relationship between sex and time on KD performance, as well as the percentage of abnormal scores on the KD test in the current sample of youth soccer athletes. The results of the current study suggest that the KD test has good long-term and clinically relevant reliability as only 11.8% exceeded the MDC in youth soccer athletes. However, there was also a 33–35% abnormal KD test prevalence in youth athletes, suggesting that despite the good reliability healthy youth athletes may perform slower or commit errors on the KD test in subsequent test administrations which should be considered when using the KD test to make SRC decisions. In addition, these findings suggest female soccer athletes had faster KD test times in Season 1 and Season 2; however, male soccer athletes improved significantly more between seasons than female soccer athletes.

The results of the current study revealed good test-retest reliability (ICC's: 0.72–0.81) of the KD test across an 18-month period in youth soccer athletes. These results are similar to previous researchers who found good test-retest reliability from day 1 to day 45 (ICC = 0.81) [13], and excellent test-retest reliability (ICC = 0.91) across an average of 41 days in adolescent athletes [14]. The current study extends this test-retest time interval to 18 months, suggesting that the KD test is a reliable tool to use in healthy youth athletes longitudinally.

In addition to the test-retest reliability of the KD test, the results of KD test completion times in the current study are similar to those of previous research, as youth athletes in this sample Best Total Time to complete the KD test was 53.38 ± 12.4 s in Season 1 and 47.46 ± 10.4 s in Season 2 [12,13]. King-Devick test performance, in the current study, significantly improved by 5 s across the testing period. This is similar to previous findings that suggest a 3–7 s improvement at the test-retest interval [10,13]. Oberlander et al. [13] reported an improvement of 6.9 s in total KD test time in adolescent athletes, while Broglio et al. [10] observed a 3 s improvement across 1 year among collegiate athletes and military service academy cadets. The differences in Best Total time could be due to the difference in ages between the two samples.

Differences in KD test total time improvement may be attributed to the younger age in the current study, as older aged athlete's performance may be more stable over time, due to both physical and cognitive maturation [21]. It was anticipated that an improvement would be observed in the current study, as previous research has reported age to be a factor in KD test performance [11,12,21]. King-Devick test times have been reported to improve by 3.7 s for every 1-year increase in age [12]. In addition, in another study, 12–14 year olds performed 10 s faster than 8–11 year olds, and 16–18 year olds performed 3 s faster than 13–15 year olds [21]. Given that the athletes in this sample were approximately 10 years old during the first KD test administration and approximately 12 years old during the KD test re-test, the significant increase in KD test performance was anticipated.

The current study also examined the effect of sex on KD test performance across time. In the current study, female youth soccer athletes demonstrated faster KD test times during both Season 1 and Season 2. This finding is supported by previous research that observed a 5.6-s faster KD test time in female athletes compared to males among a

sample of youth athletes [21]. These results could be attributed to the notable sex differences observed for processing speed and reaction time, with female athletes demonstrating faster performance on both assessments [27,28]. However, in the current study, male athletes improved by 9 s on the KD test compared to a 3 s improvement in KD test performance in female athletes. Interestingly, previous researchers [13] reported no significant sex differences on KD test performance in an older sample (age range 12–18 years; mean age 15.4 ± 1.9 years) of adolescent athletes. Therefore, it's possible that sex differences on KD test performance only exist in younger populations as there may be maturational differences between males and females.

Finally, the current study examined false positive scores of youth athletes who demonstrated slower KD test time and increased errors at the retest session. The percentage of athletes who either demonstrated slower KD test time or committed increased errors was 36%. Interestingly, Worts and colleagues [14] observed a false-positive of 36% among high school athletes. As per the KD test instructions, any increase in KD test time or increase in error from the baseline performance is indicative of a concussion [14]. Given the relatively high false positive scores observed in this study, caution should be taken when interpreting results following a suspected concussion. Therefore, athletic trainers or sports medicine professionals should consider utilizing the MDC of 16.34 s increase in KD test time as a clinically relevant change, which was found in KD test performance in the current sample. An increase of 16.34 s may be interpreted as clinically relevant change in KD test performance and may be a better indicator of concussion and may reduce the false-positives on the KD test. Regardless, the results of this study warrant the need for a multifaceted assessment approach when screening and managing SRC. In addition, clinicians should consider using changes exceeding the MDC when making SRC decisions with the KD test.

This study is not without limitations. First, the traditional plastic, hard copy cards were utilized in this study. Recently, the KD test has transitioned to test administration via an iPad [18]; however, more research is needed to examine the usefulness of administering this assessment with this method. Therefore, based on differences in digital versions of the KD test, these results should not be generalized to all KD measures [29]. Second, this study utilized previously published methodology [21,30], which recorded KD test errors to mimic post-concussion assessment. Traditionally, the KD test is recorded at baseline as the fastest time of two trials for all three-test cards [4]. In addition, at baseline, the KD test should be completed error free, regardless of the number of attempts that it requires for the athlete to complete it error free. This creates potential for a learning curve with subsequent, repeated trials [13]. However, a post-injury assessment on the KD test is completed one time, with any errors indicative of a concussion [4]. If an error is made, but quickly corrected, the athlete's time continues and the number of errors is recorded [30]. Other commonly utilized concussion assessment tools, including computerized neurocognitive, balance, and vestibular/ocular motor assessments, do not follow this error-free at baseline approach, which may affect the results of the baseline assessment. Therefore, it was determined that errors would be recorded at baseline utilizing the post-injury assessment approach. In addition, as this study was performed on the sideline due to limited access to controlled laboratory within youth the sports arena as well as purposefully mimicking a testing scenario in this population. Therefore, the background, stadium noise, time of day, and presence of light was not controlled for in this setting. However, trained members of the research team completed assessments in a separate space individually with each participant. In addition, although athletes were equipped with any visual corrective lenses (e.g., contacts or glasses) needed for sport participation, it is unknown if all athletes wore their corrective lenses. Also, given the relatively small sample size, which only included youth soccer athletes, these results may not be generalizable to all youth athletes. Finally, the retest administration of the KD test was approximately one and a half years apart rather than 1 or 2 years apart. Due to the

partnership with the soccer league and timing of the soccer seasons, researchers administered baseline testing prior to the beginning of each seasons in order to have a valid baseline assessment in the event of a concussion.

5. Conclusions

This study documents the long-term test-retest reliability of the KD test assessment in youth soccer athletes. The results of this study suggest that the KD test assessment demonstrates good and clinically relevant reliability in a sample of youth soccer athletes across 18 months. In addition, female soccer athletes demonstrated faster KD test times; however, males had significantly faster improvements at the retest period compared to female soccer athletes. Finally, 36% of the sample demonstrated increased KD test time or increased errors at the retest period, suggesting relatively high false positive scores on this concussion measure. Future research should continue to explore the usefulness of the KD test in acutely concussed youth and collegiate samples.

References

- [1] A. Lumba-Brown, K.O. Yeates, K. Sarmiento, et al., Centers for disease control and prevention guideline on the diagnosis and management of mild traumatic brain injury among children, *JAMA Pediatr.* 172 (11) (2018) e182853.
- [2] P. McCrory, W. Meeuwisse, J. Dvorak, et al., Consensus statement on concussion in sport—the 5th international conference on concussion in sport held in Berlin, October 2016, *Br. J. Sports Med.* 51 (11) (2017) 838–847.
- [3] L.B. Lempke, J.D. Schmidt, R.C. Lynall, Athletic trainers' concussion-assessment and concussion-management practices: an update, *J. Athl. Train.* 55 (1) (2020) 17–26.
- [4] K.M. Galetta, L.E. Brandes, K. Maki, et al., The King-Devick test and sports-related concussion: study of a rapid visual screening tool in a collegiate cohort, *J. Neurol. Sci.* 309 (1–2) (2011) 34–39.
- [5] M.K. Oride, J.K. Marutani, M.W. Rouse, P.N. DeLand, Reliability study of the pierce and King-Devick saccade tests, *Am. J. Optom. Physiol. Optic* 63 (6) (1986) 419–424.
- [6] M.T. Kulp, P.P. Schmidt, Reliability of the NYSOA King-Devick saccadic eye movement test in kindergartners and first graders, *J. Am. Optom. Assoc.* 68 (9) (1997) 589–594.
- [7] M. Hecimovich, D. King, A.R. Dempsey, M. Murphy, The King-Devick test is a valid and reliable tool for assessing sport-related concussion in Australian football: a prospective cohort study, *J. Sci. Med. Sport* 21 (10) (2018) 1004–1007.
- [8] A.M. Dassy, F.J. Yuk, A.Y. Maniya, et al., Review of assessment scales for diagnosing and monitoring sports-related concussion, *Cureus* 9 (12) (2017) e1922.
- [9] A. Subotic, W.K. Ting, M.D. Cusimano, Characteristics of the King-Devick test in the assessment of concussed patients in the subacute and later stages after injury, *PLoS One* 12 (8) (2017) e0183092.
- [10] S.P. Broglio, B.P. Katz, S. Zhao, M. McCrea, T. McAllister, C.C. Investigators, Test-retest reliability and interpretation of common concussion assessment tools: findings from the NCAA-DoD CARE Consortium, *Sports Med. (Auckl. NZ)* 48 (5) (2018) 1255–1268.
- [11] B. Alsalaheen, J. Haines, A. Yorke, J. Diebold, King-Devick test reference values and associations with balance measures in high school American football players, *Scand. J. Med. Sci. Sports* 26 (2) (2016) 235–239.
- [12] V.M.C. Smolyansky, S.A. Hitzman, S. Beckerman, Test-retest reliability of the King-Devick test in elite junior Olympic athletes, *Optom. Vis. Perf.* 4 (3) (2016) 147–154.
- [13] T.J. Oberlander, B.L. Olson, L. Weidauer, Test-retest reliability of the King-Devick test in an adolescent population, *J. Athl. Train.* 52 (5) (2017) 439–445.
- [14] P.R. Worts, P. Schatz, S.O. Burkhardt, Test performance and test-retest reliability of the vestibular/ocular motor screening and King-Devick test in adolescent athletes during a competitive sport season, *Am. J. Sports Med.* 46 (8) (2018) 2004–2010.
- [15] A. Collie, D. Darby, P. Maruff, Computerised cognitive assessment of athletes with sports related head injury, *Br. J. Sports Med.* 35 (5) (2001) 297–302.
- [16] T.C. Valovich, D.H. Perrin, B.M. Gansneder, Repeat administration elicits a practice effect with the balance error scoring system but not with the standardized assessment of concussion in high school athletes, *J. Athl. Train.* 38 (1) (2003) 51–56.
- [17] A. Collie, P. Maruff, D.G. Darby, M. McStephen, The effects of practice on the cognitive test performance of neurologically normal individuals assessed at brief test-retest intervals, *J. Int. Neuropsychol. Soc.* 9 (3) (2003) 419–428.
- [18] K.M. Breedlove, J.D. Ortega, T.W. Kaminski, et al., King-Devick test reliability in National Collegiate Athletic Association Athletes: a National Collegiate Athletic Association-Department of Defense Concussion Assessment, Research and Education Report, *J. Athl. Train.* 54 (12) (2019) 1241–1246.
- [19] D. Naidu, C. Borza, T. Kubitowich, M. Mrazik, Sideline concussion assessment: the King-Devick test in Canadian professional football, *J. Neurotrauma* 35 (19) (2018) 2283–2286.
- [20] K.K. Weise, M.W. Swanson, K. Penix, M.H. Hale, D. Ferguson, King-Devick and pre-season visual function in adolescent athletes, *Optom. Vis. Sci.* 94 (1) (2017) 89–95.
- [21] R. Moran, T. Covassin, Risk factors associated with baseline King-Devick performance, *J. Neurol. Sci.* 383 (2017) 101–104.

- [22] E.Y. Chin, L.D. Nelson, W.B. Barr, P. McCrory, M.A. McCrea, Reliability and validity of the sport concussion assessment Tool-3 (SCAT3) in high school and collegiate athletes, *Am. J. Sports Med.* 44 (9) (2016) 2276–2285.
- [23] R.S. Moser, L. Olek, P. Schatz, Gender differences in symptom reporting on baseline sport concussion testing across the youth age span, *Arch. Clin. Neuropsychol.* 34 (1) (2019) 50–59.
- [24] S.J. Ozinga, S.M. Linder, M.M. Koop, et al., Normative performance on the balance error scoring system by youth, high school, and collegiate athletes, *J. Athl. Train.* 53 (7) (2018) 636–645.
- [25] T.K. Koo, M.Y. Li, A guideline of selecting and reporting Intraclass correlation coefficients for reliability research, *J Chiropr Med.* 15 (2) (2016) 155–163.
- [26] R. Morrison, K. Petit, C. Kuenze, R.N. Moran, T. Covassin, Pre- to post-season changes on the BTrackS force-plate in a sample of collegiate athletes, *J. Sport Rehabil.* (2019) 1–12.
- [27] T. Covassin, R. Elbin, A. Kontos, E. Larson, Investigating baseline neurocognitive performance between male and female athletes with a history of multiple concussion, *J. Neurol. Neurosurg. Psychiatry* 81 (6) (2010) 597–601.
- [28] D.F. Halpern, Sex differences in intelligence. Implications for education, *Am. Psychol.* 52 (10) (1997) 1091–1102.
- [29] J.R. Clugston, S.P.D. Chrisman, Z.M. Houck, et al., King-Devick test time varies by testing modality, *Clin. J. Sport Med.* (2018).
- [30] R.N. Moran, T. Covassin, King-Devick test normative reference values and internal consistency in youth football and soccer athletes, *Scand. J. Med. Sci. Sports* 28 (12) (2018) 2686–2690.