RESEARCH REPOSITORY

This is the author’s final version of the work, as accepted for publication following peer review but without the publisher’s layout or pagination. The definitive version is available at:

https://doi.org/10.1016/j.jsams.2018.03.011


Copyright: © 2018 Sports Medicine Australia
It is posted here for your personal use. No further distribution is permitted.
Title: The King-Devick test is a valid and reliable tool for assessing Sport-Related Concussion in Australian Football: A Prospective Cohort Study

Authors: Mark Hecimovich\textsuperscript{a, b}, Doug King\textsuperscript{c, d}, Alasdair R Dempsey\textsuperscript{b}, Myles Murphy\textsuperscript{e, f}

\textsuperscript{a}Division of Athletic Training
University of Northern Iowa, Cedar Falls, Iowa, United States of America
mark.hecimovich@uni.edu

\textsuperscript{b}School of Psychology and Exercise Science
Murdoch University, Murdoch, Western Australia, Australia
a.dempsey@murdoch.edu.au

\textsuperscript{c}Sports Performance Research Institute New Zealand (SPRINZ) at AUT Millennium,
Faculty of Health and Environmental Science
Auckland University of Technology, Auckland, New Zealand
doug.king@clear.net.nz

\textsuperscript{d}School of Science and Technology,
University of New England
Armidale, NSW, Australia

\textsuperscript{e}School of Physiotherapy
The University of Notre Dame Australia, Fremantle, Australia
myles.physio@gmail.com

¹Medical Department, Peel Thunder Football Club, Mandurah, WA, Australia

**Corresponding author:**

Mark Hecimovich

Division of Athletic Training

003C Human Performance Center

University of Northern Iowa, Cedar Falls, Iowa, United States of America, 50614-0244

Mark.hecimovich@uni.edu

**Abstract**

**Objectives:** Sport-related concussion (SRC) research has focused on impaired oculomotor function. The King-Devick (K-D) test measures oculomotor performance and is reported to identify suboptimal brain function. The use of the K-D test in Australian football (AF), a sport involving body contact and tackling, has not been documented. Therefore, the objective of this study was to determine the test-retest reliability and diagnostic accuracy of the K-D test on a sub-elite AF team.

**Design:** Prospective cohort study

**Method:** In total, 22 male players (19.6 ± 2.3 yr) were tested and re-tested on the K-D test. Those suspected of having a SRC secondary to a significant head impact were tested. Randomly selected additional players without SRC were assessed for comparison.

**Results:** There were observable learning effects between the first and second baseline testing (48 vs. 46 s). The ICC for the first and second baseline tests was 0.91. Post-match test times were longer than the baseline times for players with SRC (n=7) (-1.9 s; z=-5.08; p<0.0001). Players tested with no signs of SRC (n=13) had an improvement in time when compared with their baseline score (3.0 s; z=-4.38; p<0.0001). The
overall sensitivity was 0.98, specificity 0.96, and a kappa of $\kappa=0.94$. The positive likelihood ratio was 11.6 and the positive predictive value was 89.0%.

**Conclusion:** This study supports the use of the K-D test due to its test-retest reliability, high sensitivity and specificity, and fast and simple use that is ideal for sports medicine professionals to make quick judgement on management and playability.

**Keywords:** Diffuse axonal injury, Neuropsychological tests, Visual motor coordination, Brain injury

**Introduction**

For healthcare providers affiliated with sport, the complexities associated with the recognition and management of sport-related concussion (SRC) is challenging. The 2017 consensus statement on concussion in sport stated that SRC is a traumatic brain injury and typically results in the rapid onset of short-lived impairment to neurological function.

One area of SRC research has focused on impaired oculomotor function following trauma to the brain. Visual processing utilizes a number of circuits throughout the brain, including visual-spatial integration, attention, motor planning, motivation and spatial organization. Increased latency and decreased accuracy of saccades (fast eye movements) is accepted to follow acute head trauma. Thus, eye function tests are a robust measure for the identification of mTBI and SRC. Heitger et al demonstrated an increased number of saccades and poor motor movement timings with slower velocities of eye movements in subjects with post-concussion syndrome. As well as reflecting changes in a number of cortical areas and pathways, saccades can be used to assess cognitive performance changes following brain trauma including attention, spatial and temporal orientation and working memory.

The King–Devick (K–D) test was originally developed as a reading tool to assess the relationship between poor oculomotor functions and learning disabilities. The K–D test utilizes a series of charts of
numbers that progressively become more difficult to read in a flowing manner.\textsuperscript{7} The K-D test is reported to identify suboptimal brain function in a range of conditions and diseases including hypoxia,\textsuperscript{9} and sleep deprivation.\textsuperscript{10} Furthermore, the K-D test can be used as a rapid assessment tool for the identification of SRC with the benefit of not requiring a medically trained person for administration.\textsuperscript{11} King et al\textsuperscript{12} utilized the K-D test to assess club-level rugby union players who had not shown, or reported, any signs or symptoms of SRC but had a suspected head injury. In their 2013 study, King et al\textsuperscript{12} diagnosed 22 concussive incidents and interestingly the causative head impacts were not directly witnessed in 17/22 of these cases. However, SRC was successfully detected using the K-D test. A recent study by King et al\textsuperscript{13} investigated the ability of the K-D test to aid in identifying witnessed and unwitnessed episodes of SRC over subsequent seasons in a junior rugby league.

The use of the K-D test in rugby\textsuperscript{12, 13} is warranted as the sport is non-helmeted and associated with high levels of head impacts and head impact acceleration.\textsuperscript{14} However, the use of the K-D test in Australian football (AF), which reports fewer head impacts and less impact acceleration,\textsuperscript{15} has not been documented. This is important as AF players may be in danger of underreporting of SRC. Therefore, the primary aims of this study were to: 1) identify the test–retest reliability of the K-D test in semi-elite AF players and, 2) determine the diagnostic accuracy of the K-D test in identifying SRC that occurred from game participation. These are vital as oculomotor and saccadic disruption may indicate suspected concussion, therefore support the use of the K-D test as one of the tools in a multi-faceted approach to sideline evaluation.

Method

A prospective observational cohort study was conducted on a sub-elite West Australian Football League (WAFL) reserves team competing in the 2016 WAFL season. All members of the team were invited to participate in the study. Twenty-two male players (19.6 ± 2.3 yr; 184.2 ± 6.9 cm and 79.4 ± 7.2 kg) agreed to participate and were enrolled in the study. Consent was obtained from the players before enrolling in
the study. The researchers’ University ethics committee approved all procedures (MUHREC 2016/012). Consent was provided by the participating team and players prior to commencing the study.

All participants completed a baseline pre-game K-D assessment. 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 reported to have an inter-class correlation for test-retest reliability of 0.97 (95% CI: 0.9 to 1.0).\textsuperscript{16} A recent systematic review of the K-D test reports high sensitivity in identifying SRC (96/112 concussed athletes showed worsening; 95% CI: 78-92%), and high specificity (181/202 controls no worsening; 95% CI: 85-93%).\textsuperscript{17}

Employing the version designed for tablet device (v2.2.0), players were instructed to complete two trials within a few minutes of each other to establish their initial baseline as per K-D test recommendations. Both errors in reading and speed of reading were included in deriving a K-D test time. The faster time from two error-free trials became the established baseline K-D test time and any subsequent faster time established a new baseline.\textsuperscript{16}

Players who were suspected of having SRC secondary to a significant head impact (n=7) (as observed from the sidelines or reported to the head physiotherapist) were assessed by the club’s head physiotherapist (MM) who has post-graduate qualifications in sports physiotherapy and over 5 years’ experience in AF. Players were assessed in a quiet medical room approximately 10 minutes following ceasing participation and the diagnosis of SRC was made using a comprehensive history and detailed neurological examination using the Sports Concussion Assessment Tool V.3 (SCAT3) as per the recommendations of the 2012 Consensus Statement on Concussion in Sport.\textsuperscript{18} Those players diagnosed with SRC and randomly selected additional players without SRC were assessed with the K-D in a quiet medical room approximately 15 minutes post game (10-20 minute window).
The same person who conducted the baseline assessments and utilising the same instructions given at baseline testing, administered all of the post-game tests. The post game tests were performed at random intervals following the baseline test with all post game tests being completed within a range of 1-12 weeks post baseline test. The post-game test time was compared with the participant’s baseline to calculate reliability. Although there were 22 players, the 110 observations came from players having more than one post-game test during the season, for example a player may have completed two follow up tests at 4 weeks and 11 weeks from baseline. All collected K-D data were entered into a Microsoft Excel spreadsheet and analyzed using SPSS V24.0.0 (SPSS Inc, USA).

Data were presented as median [25th to 75th inter-quartile range] for K-D test times. Differences in K-D test times from pre-competition (baseline establishment) were calculated, baseline and post-match K-D test times were compared using the Wilcoxon signed-rank test. Sensitivity, specificity, positive predictive value and likelihood ratios for the K-D test were calculated using a 2-by-2 contingency table with 95% CI. Cohen kappa (κ) with 95% CI were utilized to assess for intra-rater concordance. Test-retest reliability was also estimated utilising the intra-class correlation coefficient (ICC), with 95% CI, to examine agreement between first and second baseline test scores and the post-match scores. Statistical significance was set at p=0.05.

Results
There were observable learning effects between the first and second K-D test baseline trials (48 vs. 46 s; z=-1.05; p=0.2937) (see Table 1). Over the duration of the study the K-D test demonstrated a sensitivity of 0.98 (0.87 to 1.00) a specificity of 0.96 (0.80 to 1.00), and a kappa of κ=0.94 (0.85 to 1.00). The positive likelihood ratio was 11.6 (1.7 to 77.0) and the positive predictive value was 89.0% (54.5 to 98.2%). The ICC between the baseline trials was 0.91 [0.80 to 0.96]. Post-game K-D test times were increased, or worse, for players with SRC (-1.9 s; z=-5.08; p<0.0001) when compared with their baseline time (see Table 2).
Post-game non-injured control players with no signs of SRC (n=13) had an improvement (faster) time when compared with their baseline score (3.0 s; z=-4.38; p<0.0001).

A post hoc power analysis was conducted using the formulae provided by Hajian-Tilaki\(^2\) for calculating sample size for sensitivity and specificity. The sensitivity and specificity used for these analyses were 0.98 and 0.96 respectively, as per our results section. The alpha level used for these analyses were \(\alpha < 0.05\) and the precision (maximal marginal error) used was \(d=0.10\). The prevalence used for these analyses were 0.2 which has been previously reported as the prevalence of concussion in the WAFL.\(^2\) A sample of 38 observations were required for sensitivity and 18 observations were required for specificity but as per the instructions by Hajian-Tilaki\(^2\) the larger number should be used when calculating both sensitivity and specificity showing this study is adequately powered given the trial included 110 observations.

**Discussion**

This prospective study investigated the efficacy of the K-D test with results demonstrating that it is an accurate screening tool for SRC in AF players. This is vital, as AF is a high-intensity sport involving body contact, collision and tackling\(^2\) with risk of SRC. However, unlike rugby and American football, it has lower rates of SRC and impact exposures.\(^15\) Therefore, the need for an accurate screening test is required, as players may not exhibit gross outwardly signs and symptoms of SRC.

The test–retest reliability of the K–D test in this study demonstrated interclass correlations of 0.91 (95% CI) between baseline 1 and baseline 2, 0.92 (95% CI) between baseline and post-game for all players, and 0.93 (95% CI) and 0.92 (95% CI) between concussed and non-concussed players, respectively. These are comparable to previously reported data in amateur boxers (ICC = 0.90, 95% CI 0.84, 0.97),\(^11\) and junior rugby players (\(r_s = 0.86\)).\(^13\)
The results demonstrated that players who sustained SRCs were shown to have increased (worsening) read times relative to baseline and conversely those without had shown improvement, or faster read times. The clinical significance of these findings need to be viewed with caution as SRC can exhibit a variety of symptoms. Future research could investigate the correlation between the amount of time worsening from baseline and frequency, magnitude and distribution of head impacts as measured by head-mounted sensors at various age levels.

In this study, at baseline, measurements for all players demonstrated a noticeable learning effect between the first and second tests. Previous research has shown that in the absence of SRC, the K-D test has a learning effect associated with repeated testing.5, 7, 11 Previous investigations of adult athletes without SRC showed an improvement on average of 2.2 - 3.1 seconds between the two baseline test trials11, 12, 16 and improvements have been reported in the adolescent as well.23 Additionally, improvement rather than degradation with fatigue post-exercise has been demonstrated.24 This highlights the need for further research that involves large sample size and additional clinical testing to determine significance of variations between repeated attempts. Additionally, in athletes without SRC, pre-season and post-game testing showed an improvement of 3 seconds. This is more than reported by Galetta et al16 and Leong et al11 in collegiate level athletes, which were reported to be 0.72 and 1.9 seconds, respectively.

The results in this study showed those who did suffer from a SRC recorded increased (worsening) baseline to post-game read times by 2 seconds. Note that this may include slower read-times due to correction of reading errors therefore not likely to be clinically significant. Previous studies have reported slower read time scores ranging from 4 to 7 seconds.5, 12, 16

In this study, the results demonstrated high sensitivity (0.98) and specificity (0.96) and is consistent with previous studies in other contact sports12, 16, 23 indicating that increasing read times (diminishing scores) from pre-season or pre-game baseline is an accurate and sensitive indicator of SRC.17 In other SRC
assessment tools (Post-Concussion Symptoms Scale, Standardized Assessment of Concussion, Sport Concussion Assessment Tool, Immediate Post-Concussion Assessment and Cognitive Testing), the sensitivity, specificity, validity and reliability remain largely undefined, particularly among different age groups, cultural groups and settings\textsuperscript{25} and can be more time consuming than the K-D test. The SCAT is routinely used and accepted as the standard in AF\textsuperscript{26} and although it does not objectively assess vision or eye movement function\textsuperscript{27} and psychometric properties of all SCAT subtests have not been firmly established\textsuperscript{28}, Marindes et al\textsuperscript{29} showed the K-D is additive to components of it, thus supporting the need for multiple tools as not everyone responds to concussion in the same way.

A key benefit of the K-D test for sports medicine professionals, besides the strong test-retest reliability and high sensitivity/specificity, is a simple and efficient delivery in order to assist in greater accuracy on the sideline. The K-D test involves the functional integration of the brainstem, cerebellum, and cerebral cortex via visual processes therefore performance correlates with suboptimal brain function.\textsuperscript{30} Athletes with a worsening change in time is suggestive of a meaningful neurological event. Additionally, performance in ocular tests have been shown to improve with fatigue, rather than degrade\textsuperscript{24} which is vital in AF due to its high level aerobic nature.

The limitations in this study include a low amount of reported SRCs (n=7). Although the SCAT3 was used at baseline and follow-up on suspected SRCs, not all those without suspected SRCs underwent a follow-up so comparison was not done in the analysis. Furthermore, the individual who conducted all baseline and post-tests was not masked to the K-D test results. However, this person was the teams head physiotherapist and the diagnosis was determined prior to K-D testing using the SCAT.

**Conclusion**

Multi-sport research has supported the use of the K-D test as a fast and simple screening tool for SRC evaluation and follow-up recovery. The uniqueness of the K-D test is its measure of aspects of ocular motor
functioning that are not captured by measures imbedded in commonly used screening tests such as the SCAT. Therefore, the K-D test adds an additional value in screening for SRCs. Furthermore, aside from the test-retest reliability, high sensitivity and specificity, the benefit of the K-D test for AF is that it is a simpler and more time efficient screening test for sports medicine professionals to make quick judgement on management and playability. This is vital for sports medicine personnel involved in sports like AF, which has an intrinsic risk for head injury and SRC, and the players they attend, would benefit from the use of K-D test.

Practical Implications

- The King-Devick test is a fast and simple screening tool which may be added to the sport-related concussion evaluation process
- The King-Devick test demonstrated strong test-retest reliability and high sensitivity/specificity in a cohort of AF players.
- The K-D involves integration of functions of the brainstem, cerebellum, and cerebral cortex via visual processes and an athletes change in time is suggestive that a meaningful neurological event has occurred.

Compliance with Ethical Standards

The authors declare that there are no competing interests associated with the research contained within this manuscript. According to the definition given by the International Committee of Medical Journal Editors (ICMJE), the authors listed above qualify for authorship on the basis of making one or more of the substantial contributions to the intellectual content of the manuscript.

Financial and Material Support: This study was provided financial support by the Australian Football League.
Acknowledgements

We would like to thank the Australian Football League for their funding of this research project (RES-16656: Quantifying Head Trauma in ARF 2016). We would like to thank exercise science student Mason Gittins for his work with the data collection process.
References


Table 1: King-Devick test times in seconds by median [25th to 75th inter-quartile range] for all participants (n=22) at baseline assessment (test 1 and 2), baseline established time and post season with differences, intra-class correlation coefficient, specificity and sensitivity.

<table>
<thead>
<tr>
<th></th>
<th>K-D Baseline 1, s, median [IQR]</th>
<th>Errors, average (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>48.3 [44.0 to 56.9]</td>
<td>0.1 (±0.4)</td>
</tr>
<tr>
<td>K-D Baseline 2, s, median [IQR]</td>
<td>45.9 [42.2 to 56.1]</td>
<td></td>
</tr>
<tr>
<td>Errors, average (±SD)</td>
<td>0.0 -</td>
<td></td>
</tr>
<tr>
<td>Difference, s, median [IQR]</td>
<td>-3.4 [-0.8 to -7.8]&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>K-D Baseline, s, median [IQR]</td>
<td>45.7 [42.2 to 52.2]&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>ICC (95% CI) K-D Baseline1 vs. Baseline 2</td>
<td>0.91 [0.80 to 0.96]</td>
<td></td>
</tr>
<tr>
<td>K-D Post-Season, s, median [IQR]</td>
<td>42.7 [39.0 to 49.0]&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Difference, s, median [IQR]</td>
<td>-0.9 [-2.9 to 1.9]&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>ICC (95% CI) K-D Baseline vs. Post-Match total</td>
<td>0.92 [0.87 to 0.95]</td>
<td></td>
</tr>
<tr>
<td>Sensitivity K-D Test (95% CI)</td>
<td>0.98 (0.87 to 1.00)&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Specificity K-D Test (95% CI)</td>
<td>0.96 (0.80 to 1.00)</td>
<td></td>
</tr>
</tbody>
</table>

IQR = Inter-quartile Range; s = seconds; SD = Standard Deviation; CI = Confidence Interval; ICC = Intra-class Correlation Coefficient; Significant difference (p<0.05) than: (a) = Baseline; (b) = Post-Season

Table 2: King-Devick test times in seconds by median [25th to 75th inter-quartile range] for players with no signs or symptoms, and players with signs or symptoms of sports-related concussion at baseline scores, post-game times, errors made and the intra-class correlation coefficient between baseline and post-match assessments.

<table>
<thead>
<tr>
<th>Players with SRC signs or symptoms</th>
<th>K-D Baseline, s, median [IQR]</th>
<th>Errors, average (±SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-D Baseline</td>
<td>39.8 [35.9 to 53.9]&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>K-D Post-Match</td>
<td>42.8 [38.3 to 56.0]&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Difference, s, median [IQR]</td>
<td>+1.9 [-5.6 to -1.0]&lt;sup&gt;d&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Errors, average (±SD)</td>
<td>0.2 (±0.5)</td>
<td></td>
</tr>
<tr>
<td>ICC (95% CI) K-D Baseline vs. Concussive Post Match</td>
<td>0.93 [0.68 to 0.97]</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Players with no SRC signs or symptoms</th>
<th>K-D Baseline, s, median [IQR]</th>
<th>Errors, average (±SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-D Baseline</td>
<td>44.8 [42.0 to 50.5]&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>K-D Post-Match</td>
<td>42.5 [39.4 to 47.5]&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Difference, s, median [IQR]</td>
<td>-3.0 [1.4 to 4.9]&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Errors, average (±SD)</td>
<td>0.0 -</td>
<td></td>
</tr>
<tr>
<td>ICC (95%CI) K-D Baseline vs. Non-concussive Post Match</td>
<td>0.92 [0.52 to 0.96]</td>
<td></td>
</tr>
</tbody>
</table>

IQR = Inter-quartile Range; s = seconds; SD = Standard Deviation; ICC = Intra-class Correlation Coefficient; Significantly difference (p<0.05) than: (a) = Baseline; (b) = Post-Match; (c) Players with SRC signs or symptoms; (d) Players with no SRC signs or symptoms