The RailCorp Lantern test

A. Casolin1, P. L. Katalinic2, G. S.-Y. Yuen2 and S. J. Dain2

1RailCorp., PO Box K349, Haymarket, NSW 1238, Australia, 2School of Optometry and Vision Science, University of New South Wales, Sydney, NSW 2052, Australia.

Correspondence to: A. Casolin, RailCorp., PO Box K349, Haymarket, NSW 1238, Australia. Tel: +61 2 8922 1956; fax: +61 2 8922 4557; e-mail: armand.casolin@railcorp.nsw.gov.au

---

Background  The RailCorp Lantern (RL) is a simulation of light emitting diode railway signals developed as a practical and consistent test for colour vision deficient workers.

Aims  To analyse errors made on the RL and correlate with diagnosis and results on other colour vision tests.

Methods  Retrospective audit of RL tests conducted between February 2006 and December 2008 and comparison between results on the RL conducted at 3 and 6 m, the Standard Farnsworth–Munsell D15 (D15) and the Farnsworth Lantern (FL).

Results  Two-hundred and seven tests were available for analysis. There were significant differences between pass rates by test with 57% passing the D15, 14% the FL and 26% for the RL at 6 m (RL6) and 47% for the RL at 3 m (RL3) (P < 0.001). Both deuts and protans had most difficulty identifying the white light of the FL followed by the yellow of the RL. Seventy-nine percent of protans made red omissions at 6 m compared with 33% of deuts (P < 0.01), and 23% of protans made red omissions at 3 m compared with 3% of deuts (P < 0.001).

Conclusions  The RL identifies individuals who can safely read railway signals and who would have been excluded from working had the FL been the sole test. It is proposed that the RL be considered for use by other rail operators.

Key words  Colour vision standards; Farnsworth Lantern; practical test; rail; RailCorp Lantern.

---

Introduction

The use of lantern tests to assess the ability of rail workers to safely distinguish colours is not a recent development. In 1904, Dr Williams’ Lantern test was adopted by the Interstate Conference of Railway Commissioners as the test for colour sense [1]. In the 1970s, it was reported that Australian Commonwealth and State railways employed a lantern described as the Edridge–Green, modified Williams or Australian Railways Lantern [2].

The Farnsworth Lantern (FL) is one of the most well-known lantern tests. It was designed in the 1940s by Commander Dean Farnsworth of the US Navy Submarine Division for the purpose of ‘identifying those with mild color vision deficiencies who are considered “color safe” for ordinary occupations’. [3] It was modelled on the British Board of Trade, Royal Canadian Navy and Air Force Lanterns and was adopted as the US military standard [4,5]. The FL has become unavailable but a commercially available substitute, the Optec 900 (Stereo Optical, Chicago, IL, USA), is accepted as an appropriate alternative [6,7]. They are both in widespread use as tests of the severity of a colour vision deficiency and to determine fitness to perform colour-critical tasks.

The Waterfall train crash in January 2003 led to the publication of the National Standard for Health Assessment of Rail Safety Workers [8] (the standard) and a nation-wide shift in the approach to the medical assessment and certification of rail workers. The critical event that led to the crash was the sudden incapacitation of the driver following a cardiac arrest [9], and while the new standard introduced the cardiac risk score to assist in the assessment and management of workers with risk factors for ischaemic heart disease, it also updated the approach to other medical conditions, including colour vision deficiency.

The evaluation process under the standard, for determining the need for a colour vision requirement, divides tasks into outdoor rail safety work involving the recognition of railway signals and work on screen-based equipment displaying colour-critical information. The process for determining the colour vision category for a particular job is outlined in Figure 1 [8]. Workers who need to distinguish reds under extreme conditions, such as high speed and poor visibility, are required to meet the...
There are two main groups of jobs regarding colour vision; external work and work involving multicolour screen based equipment.

1. External/Outdoor Rail Safety Work

<table>
<thead>
<tr>
<th>Is accurate response to &quot;reds&quot; essential for rail safety? (need to distinguish red/yellow/green)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>No colour vision requirement</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>YES</td>
</tr>
<tr>
<td>No colour vision requirement</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>NO</td>
</tr>
<tr>
<td>Requires 'Colour Vision Defective Safe'</td>
</tr>
</tbody>
</table>

2. Screen based equipment

Work on screen-based equipment with multicolours requires minimal 'Colour Vision Defective Safe B'.

*CIE Technical Report: recommendations for colour vision requirements for transport. CIE 143-2001

Figure 1. Evaluation process for colour vision [8].

'colour vision normal' standard. Workers who are 'colour vision defective safe', on the other hand, can distinguish red/green sufficiently accurately with time and may work in jobs where speed or distance is not crucial in signal recognition. Note that the terms 'colour vision defective safe' and 'colour vision defective safe A' (safe A) are used interchangeably in the standard. Safe B relates to working with multiple colour-coded information on screen-based equipment.

The colour vision assessment process begins by screening with Ishihara's test (24-plate edition). Two or more errors (any incorrect response) out of the 12 screening plates constitute a fail under the Australian Standard, as opposed to the UK rail industry where 17 plates are commonly used and a fail is any error (giving the intended abnormal response) or more than three misreadings (giving a response that is neither the normal response nor the intended abnormal response). Australian rail workers who fail Ishihara's test are referred to a colour vision clinic to confirm whether they have a colour vision deficiency, and if so, the diagnosis and severity. The FL and the Standard Farnsworth-Munsell D15 (D15) are administered and if these are passed the standard allows the worker to be certified fit for 'safe A' and 'safe B', respectively.

Where a worker does not pass these tests, the standard states that a practical test 'may be offered'. As 'colour vision normal' and 'safe A' are categories that relate to the identification of multiple aspect railway signals, the practical test must assess this task in such a way as to ensure that the worker can identify railway signals under worst case conditions such as maximum distance, poor visibility or glare.

The standard states that the practical colour vision test should be based on the ability to read a variety of multi-aspect signals typical of those encountered on the system in question. The test routine should be designed by experienced railway staff with safe working expertise in conjunction with colour vision experts. It should involve a series of random showings of at least 10 signals at a frequency and duration designed to simulate the requirements of the system. Subjects must obtain a similar score to normal controls in order to pass and, if subjects pass, they may be permitted to only work on the rail system for which they have passed the test.

In 2005, RailCorp commissioned the University of New South Wales School of Optometry and Vision Science and a simulation was constructed using light emitting diode (LED) technology, as LEDs are replacing the vast majority of signals, the consistent colour avoids the need to incorporate variations of each colour and LEDs are robust and stable. The intensities were based on a survey of 76 incandescent and 68 LED signals. The simulation that was created, now known as the RailCorp Lantern (RL) test, comprises 2 sets of 12 combinations of red, yellow and green LEDs used in the construction of railway signals (see Figure 2). In some cases, only one light is presented. These LEDs were provided by Aldridge Railway Signals Pty Ltd from their stock of LEDs used in signal construction since no chromaticity specification was in use at the time. The RL is conducted at two distances: 6 m (RL6) in the case of colour vision normal occupations (that includes train drivers) and 3 m (RL3) in the case of 'defective safe A' positions.

Test validation and setting of the pass/fail criteria were conducted using control subjects aged 18-60 years, with corrected visual acuity no worse than 6/9 binocularly and no ocular pathology. The subjects were either classified as 'experienced' or 'naïve' depending on whether they had experience observing rail signals.

Subjects were further classified as either 'normal' or 'colour vision deficient'. Normal was defined as passing Ishihara's test and having a normal matching range on the anomaloscope. Colour vision defective subjects had been previously diagnosed by the University of New South Wales colour vision clinic as either 'anomalous
trichromat' or 'dichromat'. The anomalous trichromatic subjects were further divided into three categories: mild (pass the D15 and the FL), moderate (pass the D15 but fail the FL) and strong (fail the D15).

Three types of errors were made by subjects being tested on the RL: 'misnaming', where the subject identifies a colour incorrectly; 'omissions', where the subject does not see a light and 'additions', where the subject identifies a colour when none has been presented.

The test was offered to current employees who had been assessed under the standard following its introduction in 2004 and who had not met the colour vision criteria, and also incorporated into the assessment of new recruits where 1.4% of train driver recruits, for example, have been reported to fail initial screening with Ishihara’s test [10].

The aim of this study was to analyse the types of errors that have been detected on the RL test and to determine the extent to which these errors correlate with diagnosis and results on other tests of colour vision.

Methods

Ethics approval for this study was obtained from the University of New South Wales Human Research Ethics Advisory Panel. The study population was all instances where the RL test was conducted between 28 February 2006 and 31 December 2008.

The results were de-identified and collated under the following headings: gender; diagnosis—proton, deutan, tritan; severity—1 = mild anomalous trichromat, 2 = moderate anomalous trichromat, 3 = strong anomalous trichromat, 4 = dichromat, 5 = low discrimination normal; Ishihara—number of errors; Neitz anomaloscope—midpoint and matching range; D15—pass/fail and number of crossings; FL—pass/fail, total number of errors on runs 1, 2 and 3 and number of each possible misnamings, i.e., red-green, red-white, green-red, green-white, white-green and white-red; RL3 and RL6—pass/fail, number mismatched, number omitted and number of 'additions'. Red-green means that a red light was reported as green and so forth.

The chi-squared test was used to compare errors across all colour vision tests. The chi-squared and unpaired one-tail t-tests were used to compare additions and omissions.

Results

The RL was administered on 207 occasions during the study period—205 males and 2 females. One-hundred and forty-eight (71%) were deutan and 56 (27%) were proton. Three (1.4%) individuals failed Ishihara’s test but subsequently fell into the normal matching range when tested using the Neitz anomaloscope and were excluded from analysis.

All subjects completed Ishihara’s test and the RL6. Six subjects did not complete the D15 and five did not complete the FL. Forty-five individuals did not attempt the RL3 because they were part of a group of 53 individuals who passed the RL6.

The overall results for each diagnostic category and level of severity is outlined in Table 1. The overall pass rates differed significantly and were 57% for the D15, 14% for the FL and 26% for the RL6 (P < 0.001). The pass rate for the RL3 was 31%; however, if the 45 individuals who passed the RL6 and did not attempt
the RL3 are included, then the pass rate increases to 47%.
For the purpose of analysis in this study, it has been assumed that these 45 subjects would have made no errors at 3 m.

By definition, all mild deuteranomals and mild protanomals who attempted the FL passed and all moderate protanomals, strong deuteranomals, strong protanomals and dichromats who attempted the FL failed, as did all but two moderate deuteranomals who passed the FL but were assessed as moderate.

The RL was easier than the FL, particularly for moderate deuteranomals and protanomals. Among strong deuteranomals, strong protanomals and dichromats, four individuals who were unable to pass the FL passed the RL6 and 10 individuals passed the RL3. These findings suggest that the standard categorizations of mild, moderate and strong, which are based on the results of the FL and D15, are not precisely predictive of ability to identify railway LEDs.

The percentage of specific colour misnaming errors made by deutans on the FL and RL is illustrated in Figure 3 and Figure 4 depicts the colour misnaming errors made by protans.

In the case of deutans, the most difficult task was identifying the white light of the FL followed by the yellow of the RL. With regard to the critical task of identifying reds, mild deuteranomals made no red errors on either the FL or the RL. Moderate deuteranomals and deuteranopes had similar rates of red misnaming on the FL and RL6, but less on the RL3. Strong deuteranomals had similar rates of red misnaming on the FL and the RL3, but more on the RL6.

Identifying the white light of the FL was also the main source of colour misnaming errors for protans. Strong protanomals and protanopes had a higher red misnaming rate on the FL as opposed to the RL6 and the RL3.

The above error rates relate to colour misnaming. The RL test also records 'omissions' and 'additional' colours reported when no signal was present. These results are outlined in Table 2.

The majority of omissions and additional signals reported were red. At 6 m, there were 93 individuals who omitted one or more red signals compared with 1 yellow omission and 13 green omissions. At 3 m, there were 17 individuals who missed one or more red signals and two who missed one or more green. The instances of signals being reported when none were displayed were few and mainly confined to the 6-m test. One strong deuteranomal added a green and another added two green and four red. The additional signals reported by protans were all red. Due to the low number of green and yellow errors, the results outlined in Table 2 and the statistical analysis are confined to red errors.
No trend of increasing additions or omissions with increasing severity within the deutan and protan groups was observed; however, 79% [99% confidence limit (CL) 62–91%] of protons made red omissions at 6 m compared with 33% (99% CL 23–47%) of deutsans (P < 0.01), and 23% (99% CL 12–40%) of protons made red omissions at 3 m compared with 3% (99% CL <1–8%) of deutsans (P < 0.001). Protons also made more red omissions on average than deutsans at both 6 m (P < 0.001) and at 3 m.

Eight protons reported 'additional' red signals at 6 m compared with one deutan (P < 0.01).

**Discussion**

This study found that the pass rate for the FL was lower than both the RL6 and the RL3. Fourteen percent of the study population passed the FL versus 26% passing the RL6 and 47% the RL3. The figures for the FL were consistent with those found previously [6,11].

The strengths of this study are that it included all individuals who undertook the RL since its implementation, allowing a comparison of this new test with existing tests and direct comparisons of the colour misnamings for each category of colour vision deficiency. A general trend of increased misnamings with increasing severity of anomalous trichromatism was shown.

The weaknesses of this study were that the numbers in some subgroups were small and 45 individuals did not attempt the RL3 as they had passed at 6 m. The assumption was made that these individuals would have also passed at 3 m because the closer distance makes the lights brighter and easier to see.

The most common colour misnaming that was made by participants in this study was the white of the FL. As white does not form part of fixed railway signals, the most common error made on the FL is therefore not actually relevant to the task of identifying fixed railway signals. Given the context in which Farnsworth developed his lantern test, the issue of concern was likely to have been indicator lights within the vessel rather than signal lights, thus the colours of the FL do not lie within the chromaticity limits typical of signal lights.

On the RL, the most common misnaming involved yellow followed by green. Deutsans had similar rates of red misnamings on the FL and RL; however, fewer red misnamings were made by protons on the RL. This is likely to be because the RL is based on LED railway signals that, like modern road traffic signals, are a much more

![Figure 4. Percentage of misnaming errors made by protons on FL and RL. P1, mild protanomal; P2, moderate protanomal; P3, strong protanomal; P4, protanope; FL red, misnaming of a red light on FL; RL6 red, misnaming of a red light on RL6; RL3 red, misnaming of a red light on RL3; FL green, misnaming of a green light on FL; RL6 green, misnaming of a green light on RL6; RL3 green, misnaming of a green light on RL3; FL white, misnaming of a white light on FL; RL6 yellow, misnaming of a yellow light on RL6 and RL3 yellow, misnaming of a yellow light on RL3.](image)

**Table 2. Omission of red lights and reporting of additional red signals**

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>RL6</th>
<th>RL3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total number</td>
<td>Number of subjects omitting reds</td>
</tr>
<tr>
<td>-----------</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>D1</td>
<td>25</td>
<td>6</td>
</tr>
<tr>
<td>D2</td>
<td>65</td>
<td>22</td>
</tr>
<tr>
<td>D3</td>
<td>24</td>
<td>6</td>
</tr>
<tr>
<td>D4</td>
<td>34</td>
<td>15</td>
</tr>
<tr>
<td>Sub-total</td>
<td>148</td>
<td>49</td>
</tr>
<tr>
<td>P1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>P2</td>
<td>29</td>
<td>22</td>
</tr>
<tr>
<td>P3</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>P4</td>
<td>13</td>
<td>10</td>
</tr>
<tr>
<td>Sub-total</td>
<td>56</td>
<td>44</td>
</tr>
<tr>
<td>Total</td>
<td>204</td>
<td>93</td>
</tr>
</tbody>
</table>

D1, mild deuteronomal; D2, moderate deuteronomal; D3, strong deuteronomal; D4, deutanope; P1, mild protanomal; P2, moderate protanomal; P3, strong protanomal and P4, protanope.
saturated red than the lights on the FL and represent a greater colour difference with white or yellow. This result is consistent with previously reported findings for traffic signals [12].

The green LED signal lies within the region defined by the Commission Internationale de l’Éclairage [13] as being for use when colour-deficient persons are using the code. This excludes the yellow-greens, which are confusable with red and yellow by colour vision-deficient observers. This study demonstrates that it cannot be safely assumed that colour-deficient individuals will be able to recognize green LED signals, particularly under adverse conditions. This is also true of traffic signals [12]. The error rate for misnaming green on the RL was higher than the FL and ranged up to 37% for severe protanomals.

The RL test assesses ‘omission’ and ‘addition’ errors that are not considered by other tests. Omissions are of particular concern in the event known as signal passed at danger, which is an uncomfortably frequent event across the rail industry [14], even for drivers with normal colour vision. This study has shown that red omissions are a particular problem for protans, with 79% of protans omitting one or more red signals at 6 m.

In recent times, legal judgements such as MacDonald, Conrad and Cherti versus Queensland Rail [15] have found that a practical test should be conducted in situations where the colour characteristics of the FL do not replicate the task to be performed. Practical tests conducted in the field, however, can be difficult to conduct in a standardized and reproducible way and worst-case viewing conditions difficult to recreate. To overcome these problems and to meet the requirements of the National Standard [8], RailCorp commissioned the development of the RL test as a simulation of railway signals. Similar work has been conducted in Canada in relation to the Holmes-Wright Lantern [16] and the Canadian National Lantern [17,18].

The RL has been subject to legal scrutiny in the Australian Industrial Relations Commission [19], with a finding by Commissioner Raffaelli that ‘in all the circumstances, I find that the RailCorp Lantern Test is a practical test to be applied according to the National Standards’.

The reason that some individuals fail the FL but pass the RL is that the two lanterns were developed for different purposes and have different colour characteristics. The RL performs as well as the FL in terms of detecting red mismatings and has the added advantage of also measuring omissions. This demonstrates that the RL is able to detect those individuals who can safely identify railway signals and who would have been excluded from working had the FL been the sole test offered.

It is proposed that the RL test be considered for use by other rail operators and adopted as a national practical test for the Australian rail industry.

### Key points

- The RailCorp Lantern is a simulation of modern light emitting diode railway signals that permits individuals with a colour vision deficiency to be tested in a consistent manner on a simulation of the actual lights that they are required to identify in the workplace.
- The most common error on the Farnsworth Lantern involves the white light and this is not actually relevant to the task of identifying fixed railway signals.
- The RailCorp Lantern performs as well as the Farnsworth Lantern in terms of detecting red mismatings and has the added advantage of also measuring omissions.

### Conflicts of interest

None declared.

### References


