FUNDAMENTALS OF SIGHTING SIGNALS

Peter Burns MBA BAppSci MIEAust CPEng MIRSE

PYB Consulting Pty Ltd

1 Summary

The paper adopts a “bottom up” approach to determining what factors are important in positioning signals for the best sighting.

The ability of a person to detect and correctly respond to a signal depends critically on first the physical characteristics of the eye, then the way in which the brain processes these basic inputs, recognising the colour lights as a symbol, then interpretting and responding to it as required in the practical circumstances when driving the train.

The physiology of the eye means that colour can be seen only by the fovea region, within a 2° cone in the central vision. Detail of any kind is discernible only within the 8-10° cone of central vision. The importance of positioning the signal at eye level close to the track is discussed.

Once seen, the recognition and interpretting of the signal relies on various mental processes within the brain. The time that these processes take, and the degree of error involved, depend on the manner of presentation of the signal and its complexity, both in form and as a symbol for the driver to respond to.

Sources of complexity of the the signal are discussed and studies investigating the effects are reviewed.

The contribution that the light intensity of the signal lamps makes to the ease or difficulty of detection and recognition are then discussed in some detail. The actual light outputs of practical signals are placed in the context of the physiology of the eye, and its ability to respond to signals within the practical background light levels of the environment in various circumstances.

.

2 Introduction

The occurrences of SPADS (Signals Passed at Danger) is a major source of operational risk for any railway. Australian research suggests that an average of fewer than 1 stop signal in every 10,000 encountered will be passed at danger [10].[1]

Of these, the vast majority will involve simple misjudgement of stopping distance. The driver wrongly assesses the track condition, the train braking characteristics, or simply misjudges distance. Typically the train then overruns the signal be a distance less than 100m.

More serious are the 1 – 6% of these SPADS where the driver does not respond to the signal. These are termed “signal disregards” and can be further subcategorised; the most frequent being “Start on Stop” incidents, the less frequent but perhaps more hazardous being “signal disregards in running” which are often associated with secondary factors such as fatigue.

In the early 1970s, M. Mashour [8] carried out a number of studies evaluating the Human Factors elements associated with presentation and location of signals. He took the position that there was really no such thing as a driver disregarding a signal. His view was that these events could be explained by reference to the limitations of the human eye and subsequent perceptual processes at play in the particular circumstances – the driver did not see the signal, interpreted it wrongly, or forgot he had passed it (sometimes not recognising its importance).

Improving SPAD performance then becomes a case of understanding these human limitations and respecting them in the design and positioning of signals.

The purpose of this paper is to investigate those factors involved in sighting of and responding to signals.

In saying this, it has been recognised that the quality of drivers is generally quite high compared with what human factors studies lead us to expect from the general population. In a section entitled “why are drivers so reliable?”, the IRSE working group on human factors[4][2] noted that in Britain, the mean time between SPAD for a driver was 17 years. In other words it is generally a once in a lifetime event for most drivers.

3 The process for reading a signal

The Railtrack Guidance note on signal sighting[13][3] states that the process of reading a signal comprises a number of separate processes as shown in table 1:

---

[1] Nicandros (2007), Measuring railway signals passed at danger, para 4.2.

[2] IRSE Signalling Philosophy Review – 2001; report of Working Group 2, p 52.

[3] GE/GN 8537 (Appendix 3) pp 59-69.

Vigilance	Signal sighting is a vigilance task that requires sustained levels of attention and alertness
Detection	Visible features of a signal are detected in the environment
Recognition	Signal perception – the signal's form is identified and discriminated from surrounding objects Association with line – signal is recognised as appropriate for the driver's route.
Interpretation	Signal aspect is read and appropriate response is chosen
Action	Driver responds to the signal

Table 1 – Processes involved in reading a signal.

The guidance note goes on to say[13]:[4]

"The signal reading model makes the distinction between vision (detection) and perception (recognition). At the signal detection stage the raw data is of a signal are visible to the driver. At the signal recognition stage, these data are perceived, that is, the signal's form is identified and discriminated from its background and the signal is recognised as appropriate to a particular line.

"Consistent with the different meanings of vision and perception, objects can vary in their visibility and their perceptibility. In terms of signal sighting, factors that affect the a signal's visibility are those that influence the detection of the signal; that is, getting an image onto the observer's retina. Whether or not that image is then recognised as a signal and read accurately is determined by the perceptibility of the signal. Both visibility and perceptibility are properties of the signal, its design and its context".

An understanding of the physical characteristics of the human eye is important to evaluating issues of detection and recognition, particularly recognition of objects where accurately discerning colour relationships is required.

In addition, it is important to understand how the brain processes this information to give the final potential action – a brake application.

3.1 The extent of detailed vision

Handbook of Transportation Engineering states [3]:[5]

"The region of the central retina where a fixated image falls is called the fovea. The fovea has only cones for visual receptors and is 1.5 – 2 degrees in diameter. Beyond 2 degrees, cone density rapidly declines reaching a

---

[4] Ibid, pp 59-60.

[5] Handbook of Transportation Engineering, Myer Kutz, ed Sect 11.2

stable low point at about 10 degrees. Conversely, rod density rapidly increases beyond 2 degrees and reaches a maximum at about 18 degrees before dropping off …

"Functional detail vision extends to about 10 degrees, worsening in he near periphery from about 10 degrees to 18 degrees and significantly deteriorating in the far periphery from about 18 degrees to 100 degrees. …

The Traffic Control Handbook (FHWA 1983)[3] section on driver's legibility needs states:

"When the eye is in a fixed position it is acutely sensitive within a 5 or 6 degree cone, but is satisfactorily sensitive up to a maximum cone of 20 degrees. It is generally accepted that all of the letters, words, and symbols on a sign should fall within a visual cone of 10 degrees for proper viewing and comprehension."

The interest of the Handbook of Transport engineering is primarily signs, where detail colour discrimination is not important. This cone of 10 degrees (5 degrees either side of centre) sets the practical boundary for initial detection without colour.

GE/GN8537[13] recognises a slightly wider viewing cone. It states[6] that from a signal detection point of view, a signal needs to be within 8 degrees of centre in horizontal plane to be seen within the central field, and within 5 degrees of centre in the vertical plane. Once detected, the eye can shift focus quickly and unconsciously to pick up the colour information in the central field. To see the "signal at once" in colour, all the colour must be within a 2 degree cone.

Objects are seen more quickly and identified more accurately if they are positioned towards the centre of the observer's field of vision, as this is where our sensitivity to contrast is highest. Peripheral vision is particularly sensitive to movement and light.

3.2 The extent of colour and detail vision

For an object to be seen in colour, or in high detail, it must be viewed by the fovea region of the retina. This requires viewing to be in a 2 degree cone (within 1 degree either side of centre) in central vision.

Within that region, the ability to resolve small or distant objects is limited by the density of detecting cells.

On this topic, GE/GN8537[13] notes:[7]

Normal visual acuity (20/20 vision) is usually defined as the ability to resolve a spatial pattern separated by a visual angle of one minute of arc. Since one degree contains sixty minutes, a visual angle of one minute of arc is 1/60 of a degree. The spatial resolution limit is derived from the fact that one degree of a scene is projected across 288 micrometers of the retina by the eye's lens. In these 288 micrometers, there are 120 colour sensing cone cells. Therefore, if more than 120 alternating white and black lines are crowded side-by-side in a single degree of viewing space, they appear as a single grey blob to the human eye.

---

[6] GE/GN 8537 (Appendix 3) p 60

[7] GE/GN8537, p 65 (section 3.2.8)

“A person with normal visual acuity (that is, 20/20 vision) is just able to decipher a letter that subtends a visual angle of 5 minutes of arc (written 5') at the eye (5' of arc is 5/60 of a degree). It does not matter how far away something is from the eye; if it subtends an angle of 5' of arc at the eye, then a person with 20/20 visual acuity should just be able to determine what it is. However, this is only a threshold value. Reliable letter discrimination requires a much bigger angle, for example 20’, and certainly no smaller than 16’.

“Visual acuity is superior for objects that are presented in the central field of vision, and for objects that are highly illuminated. Visual acuity decreases as velocity between object and observer increases. However, the ability to resolve fine detail of moving targets improves rapidly with practice.”

There are variations in visual acuity between people, particularly of differing ages. Separate objects of differing colour will be merged into a single object due the physics at 2 minutes arc separation. People with good eyesight can reliably distinguish objects separated by 5 minutes of arc. Between those two figures lies a grey zone.

For recognition of characters and symbols, larger arc sizes are required. These requirements are set out in ISO 9241 parts 303 and 306[5][8].

ISO 9241 – 306 – table 1 states that the minimum practical character height is 10 minutes of arc. This is not suitable for all users (requires good eyesight). Recommended character is 20 – 22 minutes of arc.

ISO 9241 – 303 (sect 5.5.4) states that minimum display letter height for Roman characters is 16 minutes of arc. The minimum display height for an other symbol is 20 minutes of arc.

These numbers are consistent with those presented in GE/GN8537 [13] above. How these numbers translate to actual sizes and positions at various viewing distances is explored in section 6 below.

The importance of respecting these limitations is illustrated by the Human Factors studies discussed in the next section.

4 Human Factors Study – Mashour (1974) [8][9]

Mashour carried out some quite extensive simulation studies at a preliminary activity to recommending on a common European Signalling System. Such a system has since emerged in the rather different form of ETCS. The studies contain results which are still of value to us today.

Mashour looked at a number of dimensions of the signal sighting issue. Of particular interest, his simulation studies considered:

- Recognition times for various aspects
- Errors in recognition of various aspects

---

[8] ISO 9241: Ergonomics of human-system interaction; Part 303: Requirements for electronic visual displays; Part 306: Field assessment methods for electronic visual displays.

[9] Mashour (1974): Human Factors in Signalling Systems – Specific Applications to Railway Signalling

- Detection times for various aspects in central / peripheral vision
- Error rates (again from Engine drivers)
- Determining factors for ease/ difficulty of signal recognition

Mashour's work is simulation based. Although the simulations were quite extensive and of good quality, they were at all times faced with the high reliability with which engine drivers carry out these tasks in their normal working lives. The simulations sought to identify sources of errors by magnifying these rates somewhat. This was done by ensuring that the subjects were under rather more stress than average for each situation involved or that the task was more challenging than average (eg the task was combined with another task, the presentations were in random rather than predictable order, or the task was less familiar than usual).

It should be noted that the aspect sequences studied were actual aspect sequences then in use in a variety of countries in Europe. Neither the BR system, the NSW system or the Victorian system was studied specifically. There were more aspects to consider for these systems since there were 4 defined speeds plus stop, rather than our 3 defined speeds plus stop (including the "80" indicator as a speed).

4.1 Recognition times

A number of differing systems of differing complexity were studied. Signals were displayed at simulated distance of 300m in central visual field. Results for basic 3 aspects are shown in table 2 [8][10].

The reason for variation between the signals is the difference in system complexity between systems. "R" and "G", when displayed as single colour light indications had low response times (in the order of 1.5s). This corresponds to our 3-aspect system.

There was no 3-position version of the "Y" signal included in the study. Based on other results, recognition time for this could be expected to be of the order of 1.5s.

When multiple lights were displayed, or 2-step decoding was needed, response time went up, typically by about 1s. This affect is seen in the "Y" signal which was presented as a complex signal (ie more complex than 3-aspect) in all cases. In some systems where signal meanings were largely arbitrary or involved complex understanding of inter-signal relationships, recognition times were higher still.

These more complex signal equivalents (4-aspect in our signalling) are shown in table 3.

These aspects are presented as complex signals in Victoria, so would be expected to be close to average times. The longer recognition times were associated with more arbitrary aspects[11] and would not be applicable here.

---

[10] Ibid. Summarised from section 8.2, pp 140-145. Tables are found at the end of this paper.

[11] The European signal systems evaluated as "arbitrary" were those lacking intuitive logic between successive signals in a sequence. These were the signal with the highest error rates and longest reading times. While Victorian speed signalling (4 aspect) would fall into this family in the terms of the study, it

4.2 Errors in Recognition [8][12]

It should be noted that the results, shown in table 4, are for subjects who have just learned the system, so are a measure of conceptual complexity rather than expected error rates in the field.

Same comments apply as for above. Recognition for the simplest systems were close to 100%. Accuracy reduced with complexity.

For the more complex aspects, results were as shown in table 5.

Recognition errors were found to be 37% higher for combined (or complex) signals compared with simple ones.

Common sources of error were where lights were transposed in order (eg in our system R/G and G/R confused, or R/Y and Y/R). or combinations of flashing confused with each other (not relevant here).

A reduced full simulation test was carried out using experienced engine drivers in place of students.

For these, recognition rates were much closer to 100% for aspects. The most reliably recognised aspects were R, G and "Stop at next signal" (regardless of initial speed). It was noted that error rates were in the order of 5 times higher for signals requiring change of speed from one level to another than for the simpler aspects.

4.3 Signal Detection [8][13]

Detection response times and reliability were investigated for various signal types in a variety of locations in the visual field as shown in table 6. Green, yellow or red signals were presented at simulated 300m distance at one of 3 locations in the subject's visual field.

During the experiment, the subject was required to undertake a continuous tracking task unrelated to the signal. The signal was then presented at random times in the subject's visual field. This simulated a relatively high level of stress/ distraction.

Results for detection times are shown in table 7.

The difference between Yellow detection in the central field is put down to difference in contrast. For incandescent lamps with same power, yellow lights have 4 times the luminosity of either red or green (this same relationship does not apply for LED lamps – these were not used in the study). It is noted that the luminosity (or contrast – the active variable) makes a difference in the central field, but not when in the fovea region (within 1 degree of central vision point).

Note that to incorporate these results into the previous section's, approx 1.5s needs to be added to recognition time where signal is initially presented outside fovea region.

Results for % successful detection are shown in table 8.

---

contains a lower number of actual aspects than typical in the most complex European equivalents.

[12] Ibid. Summarised from sections 8.3-4, pp 145 - 155

[13] Ibid. Summarised from section 10.3, pp 188 - 195

These figures are quite revealing. It can be seen that increasing the contrast can assist in detection rate where the signal is in the central visual field (within 5 degrees of centre), but provides limited benefit outside that region. The reduction in detection rate with position of signal in visual field is quite noticeable in this experiment.

What is also seen is that the single most effective action to improve detection rate is to ensure the signal is presented within the fovea region (within 1 degree of central vision point). At 300m, this corresponds to about within 5.2m of the central visual point.

Of note also is the relatively poor detection of the red aspect when presented at position "L", outside the fovea region but within the central region. As discussed in section 7.1 below, red light is not detectable to the scotopic vision available using this portion of the eye.

5 Sufficient Reading Time

5.1 Results from study

In section 4.1 above, it was seen that for a simple signal (eg single light) presented in central view (within 1 degree of centre vision), recognition time is less than 2 seconds. For more complex signals (eg four aspect double light), this rises to about 3.5 seconds.

Where the signal first appears off-centre to central vision (5 degrees from centre vision), 1.5 seconds need to be added to account for additional detection time. Signals presented even further from centre vision are detected with significantly lower reliability and take longer to detect. These will not be considered here as signals should always be positioned to allow adequate sighting within the "10 - 16 degree cone" (not more than 5 – 8 degrees from centre, depending on reference).

Thus the recognition task with a complex signal should take approximately 5 seconds under adverse sighting conditions. This can be reduced to 4 seconds where simple aspects are utilised.

It should be noted that these figures are averages and variations can occur.

5.2 Findings from Ladbroke

In the inquiry into the Ladbroke accident, which was a signal disregard incident, issues of signal sighting were considered in some detail.

The Railtrack Signal Sighting standard applicable at that time (since superseded) was GK/RT 0037, Issue 3, December 1997[14]. It stated:[14]

" Signals shall normally be positioned to give drivers an approach view for a minimum of seven seconds and an uninterrupted view for at least four seconds".

At the inquiry, there was the following discussion concerning this standard and its rationale [2]:[15]

"Mr Wilkins stated that the periods of seven and four seconds had stood for the past 20 or 25 years. In Annexe A to their report W S Atkins comment on this rule as being "probably the paramount measure". They also state that the

---

[14] GK/RT 0037, Issue 3, December 1997, section 4.1.2

[15] Ladbroke Report s 11.11, p179.

reputed rationale for the seven seconds was two seconds to identify the aspect, three seconds to assimilate meaning and two seconds for loss of sight as the train closed on the signal. The standard for signal sighting which was first issued in October 1994 introduced the word "normally" and added the note which appears in the current issue."

It was noted that there were some factors which might warrant ensuring a greater sighting time than the minimum [2]:[16]

"Mr Wilkins said that it might be necessary to take the view that where there were a large number of signals which were simultaneously visible, and the driver had to differentiate between them before he could read them, the sighting time should be increased, perhaps to ten seconds. As I have already narrated in Chapter 5, he strongly distinguished signals standing on their own from signals which were mounted on a gantry."

This comment needs to be placed in the context of Lord Cullen's understanding of a "gantry". At Ladbroke, each gantry presents several signals in a row. The "signal of interest" needs to be identified by the driver by "counting across" before the processes of response discussed in this current paper can commence.

Considering all these factors Lord Cullen stated [2]:[17]

"The minimum distance is nothing better than a minimum distance. Good practice is to ensure that twice the minimum distance is available where this can be reasonably achieved".

This would imply that, for a complex layout like Ladbroke, where "counting across" was needed, 14 seconds of sighting from "first sighting" should have been sought if it was available. This translates to 600 metres at 160 km/hr or 500 metres at 130 km/hr.

The current RailTrack standard [12][18] sets a "minimum reading distance" of 8 seconds sighting. A note is also made to the affect that it is permissible to determine the sighting distance to a stop signal taking into account the braking curve of the train approaching that signal.

5.3 Victorian Practice

The Signal Sighting Standard in Victoria has historically called for 6 seconds of clear sighting or 10 seconds of interrupted sighting. This is better than the prior Railtrack standard but not quite as good as Lord Cullen's recommendation on good practice.

For a train travelling at 80 km/hr, 10 seconds' sighting corresponds to 220m. For a train travelling at 130 km/hr, 6 seconds' sighing corresponds to 217m.

5.4 Optimal Sighting Range

In Railtrack practice, the AWS magnet is placed 183m ahead of the signal to which it refers. This is approximately the 4 second sighting point for trains travelling at 160 km/hr (the common speed for fast trains prior to the introduction of the 200 km/hr HSTs). According to Railtrack guidelines,

---

[16] Ladbroke Report s 11.12, p180.

[17] Ladbroke Report s 11.7, p179

[18] GE/RT 8037, section B 5.4

signals should generally be focussed to a point 183m distant from the signal, 3m above track height.

The sound from the AWS calls the driver's attention to the signal which should at that time be clearly in centre view. The driver then has time to see the signal and respond to it prior to passing it.

This view of optimal sighting aligns well with the findings of Mashour in has simulation studies.

The study finds that simplicity of sequence is key here and that the key signal which needs to be responded to is the warning signal rather than the stop signal. It is important that the warning signal requires a response and it is important that the driver is aware of it as he passes. It finds [8][19]:

> *"Both the size and apparent brightness of a signal increase as the observation distance decreases, and so, consequently does the probability of detection …*
>
> *"This significant factor – shortening of observation distance – should be employed as far as possible to promote safety. The most effective way of doing this is to install simple foresignals (SF)[20] at braking distance. Detection of SF signal at 50m or even less would not be too late or dangerous for safety, since the adjustment of speed should begin at SF signals, not before."*

Fog, background light, etc can be controlled at these distances. At long distance these factors cannot be controlled and judging distance is difficult.

Put another way: what is important for a yellow signal is not that it is visible at 1000m, but that it is "in your face" and clearly understood at 50m. This is the principle behind AWS and other similar systems which work by generating an audible warning in the driver's cab as the SF signal is approached by the train.

Optimal sighting then, should above all ensure clear sighting in the driver's central field of view in the key range between 50 and 200m ahead of the signal.

5.5 Short range viewing

In addition to the above, as an effective control measure against "start on stop" SPADs, clear sighting in central field of view should be ensured to as close as possible (within 15 – 25m) of the signal itself.

In section 6.1, the signal positioning needs which flow from this recommendation are discussed. The Railtrack guideline also considers these issues in some detail.

On this topic, GE/GN 8537[13] states:[21]

> *"For close range viewing and when stationary at a signal, the best lateral position for a signal is to the left hand side of the track and as close to the track as reasonably practicable (Railtrack standard is 900mm from left track[22]). For long range viewing, such a location may not be suitable because of obstructions such as platform buildings*

---

[19] Mashour (1974) p 176

[20] "simple foresignal" translates to a normal speed warning (Y/R) in Victoria or a distant signal (Y) in NSW.

[21] GE/GN 8537, P27-8

[22] This corresponds to 1.62m from track centre

> *or the side of a cutting on a left-hand curve."*
>
> *"A right hand signal may appear to provide a better option for viewing the signal throughout the required viewing distance and possibly avoid the need for a co-actor or banner repeater signal. …*
>
> *"Where platforms are on the right-hand side, consideration should also be given to the ability of platform staff observing the signal aspect. The most likely option would be a right-hand signal and a train-stop mark sufficiently far from the signal to give the driver an acceptable view."*
>
> *"On lines with only one track signalled in the direction being considered, the likelihood of the signal being associated with the wrong track is low and therefore right hand signals may be suitable to improve reading time."*
>
> *"On two track railways with bi-directional signalling, it is common practice to place the "wrong road" signals on the right-hand side; however these should always be positioned with particular regard to close-range viewing or where the approach makes the observance of both signals difficult, such as after a bend."*

For trains stopping very close to the signal, GE/GN 8537[13] states:[23]

> *"Existing driving policies encourage drivers to stop between 15m and 25m back from the signal.*
>
> *"Technically a train could legitimately stop in line with the signal, but it is not practical either for the signal to be readable from this distance or for the cab sightlines to cater for this.*
>
> *"The following strategy may be applied:*
>
> *"At the planned stopping point (eg 20m ahead) the driver should always be able to view the signal from the normal driving position.*
>
> *"If the train stops between A and B the driver should still be able to view the signal from the normal driving position.*
>
> *"If the train stops between points B and C the driver should be able to view the signal but may need to lean forward, or stand, in order to achieve this to overcome limitations imposed by cab design.*
>
> *"If the train stops closer than point C the driver may not be able to view the signal from the cab and may need to get out of the train to view the signal."*

Concerning optimal positioning of signals for close viewing – (height and lateral to track), GE/GN 8537[13] states:[24]

> *"For close-range viewing of a main colourlight signal, the ideal height is at drivers's eyelevel.*
>
> *"Unfortunately, driver's eyelevel is not consistent. Humans, cabs, and seat design all vary to make it impractical to define a specific value. The current use of hoods makes observing from above difficult. The signal should thus be as low as possible but no lower than the highest expected driver eye level. Consequently a height of 3.3m is recommended.*
>
> *"…*

---

[23] GE/GN 8537 P30-31

[24] GE/GN 8537 P36-7

> *"Any consideration to positioning the signal further from the running line than usual should consider carefully the potential negative affects on line association and close-up viewing."*

A table is provided at P48 of GE/GN 8537 [13] which shows minimum stopping distance from signal for good viewing for various signal positions. Stopping distance should be 25m from signal. GE/GN 8537 App 1 identifies that the signal should be viewable within 8 degree cone (this should actually be 0-4 degrees from centre). Table shows that 15m is closest stopping for left hand signal at 3.3m height, 900mm from rail, 18m is minimum for 5m mounted signal, while 25m closest for right hand signal (unless mitigated due to station activity directing driver's vision towards platform).

5.6 Starter and Dwarf Signals

GE/GN 8537[13] states[25] that when the train is starting from signal (eg terminal station or dwarf), "sufficient reading time" is not relevant.

What is important is that the signal is in clear view (centre field) from the stopped position. Planned stop position should be min 20m from signal to allow this.

For dwarf, light can be angled up for viewing (Guideline mentions 50m which may not be realistic for a dwarf – matches Mashour recommendation for other types of signals). A problem with "counting across" has been suggested in situations where a row of dwarf signals is visible from mainline at the same time. The strategy of providing a blue aspect for these is current in some locations. This strategy fails in cases where the dwarfs display "proceed".

A more traditional approach has been to limit the viewing range and focus of dwarf signals with the use of viewing angles and hoods.

If the signal is angled up and focus kept narrow, spillover can be controlled to some extent. More recent practice of mounting dwarf signals at 3m (eye level) tends to defeat this strategy.

6 Detailed parameter when approaching a Signal

In evaluating the effectiveness of signal positioning, the most important zone is that on the final approach to the signal. For a yellow signal, it is important that the driver sees and responds to the signal while in this zone.

For a red signal, it is important that the driver is continuously aware of the signal and its aspect while in this zone. There should be no possibility that the driver should believe that the signal shows an aspect other than red when it does not.

This zone can be taken as commencing (fairly arbitrarily) at the 200m mark (5.5 seconds sighting at 130 km/hr), extend through intermediate points at 100m and 50m up to the close sighting point at 20m.

The minimum close sighting point is taken as 15m. At this distance the driver should be able to observe the signal in his/her central viewing zone without leaving his/her seat.

This approach from 200m is the first part of the review. The characteristics of the

---

[25] GE/GN 8537 p 8

approach from longer distances up to 200m is then considered.

6.1 Approaching from 200m to shorter distances

6.1.1 Sighting at 200m from signal

At 130 km/hr, the train is 5.5s sighting away from the signal. At 115 km/hr the train is 6.2 seconds away. At 80 km/hr it is 9 seconds away.

Modern LED signals are 200mm in diameter. At 200m, this subtends 3.4 minutes of arc to the eye. Thus it is seen as a point source by the eye (minimum arc distinguishable from a dot is 5 minutes)

Where a symbolic indicator is also provided, the size of the symbol presented needs to be approximately 1.2m to be easily readable from this distance.

The recommended focussing point in Railtrack practice is 183m behind signal, 3m above rail. AWS is also set at this distance.

6.1.2 Sighting at 100m from signal

At 130 km/hr, the train is 2.7s sighting away from the signal. At 115 km/hr the train is 3.1 seconds away. At 80 km/hr it is 4.5 seconds away.

A 200mm LED signal can be perceived as a distinct disc (14 minutes of arc). This has less intensity than a point source of similar light output.

Where a symbolic indicator is also provided, the size of the symbol presented needs to be approximately 600mm to be readable from this distance. This is the size of currently utilised for banner indicators and theatre route indicators. Thus 100m represents the effective viewing limit for such indicators.

Railtrack statement re Banner Indicators (630mm symbols)[13]:[26]

“The banner is classified as a category 2 device which guarantees readability from 250m only[27].

“The display limitation has the effect of making the banner repeater signal more applicable to circumstances where there is a benefit in alerting the driver to a signal possibly displaying a stop aspect.

Taking 4s as the minimum viewing time, it can be concluded that these indicators are not reliably visible for trains travelling faster than 80 km/hr.

The 2 degree cone for this distance (the highest acuity central vision) extends to 5.2 metres. The signal including colours and symbol indicator can be viewed as a unit from this distance by keeping all visual elements within this 5.2m cone.

6.1.3 Sighting at 50m from signal

Mashour [8] recommends this as the optimum viewing distance for signals.

At 130 km/hr, the train is 1.4s sighting away from the signal. At 115 km/hr the train is 1.6 seconds away. At 80 km/hr it is 2.3 seconds away.

A 200mm LED signal can be perceived as a distinct disc (28 minutes of arc). This has less intensity than a point source of similar light output.

---

[26] GE/GN 8537, P44-5

[27] This statement assumes that extended readability is available due to the simplicity of interpreting a single line compared with a more complex symbol. The 10’ arc provided at this range (200m) is at the limit of viewability.

Where a symbolic indicator is also provided, the size of the symbol presented needs to be approximately 300mm (discs 1.5 times signal light diameter) to be readable from this distance. Currently utilised for banner indicators and theatre route indicators are easily readable from this distance.

The 2 degree cone for this distance (the highest acuity central vision) extends to 1.7 metres. The signal including colours and symbol indicator can be viewed as a unit from this distance by keeping all visual elements within this 1.7m cone. Where the peripheral elements of the signal have no colour (eg route indicators), these can be located outside the cone (a 10 degree cone is available where colour is not required to be distinguished)

6.1.4 Sighting at 20m from signal

This is recommended minimum placement distance for starter signals from platforms.

At 130 km/hr, the train is 0.6s sighting away from the signal. At 115 km/hr the train is 0.5 seconds away. At 80 km/hr it is 0.9 seconds away.

A 200mm LED signal can be perceived as a distinct disc (53 minutes of arc). This can almost be seen as an array of LEDs. The intensity is less than for a filament.

The 2 degree cone for this distance (the highest acuity central vision) extends to 0.7 metres. Signals comprising two colour lights separated by 0.7m cannot easily be viewed as a unit from this distance without shifting focus. Where intense red marker lights are provided as part of the signal, the “after-image” risk raised in GE/GN8537 [13] (appendix, sect 3.2.7, p.65) may need to be considered.

For comfortable viewing for driver mainly focussed on the road ahead, the top element of the signal should be ideally within 5 degrees but no more than 8 degrees viewing of horizontal. This corresponds to 4.7m – 5.8m above rail (based on head height 3m above rail)

Note Railtrack standards set limit of 5m above rail for signal height (red lamp).

It should be noted that the colour of the signal cannot be seen while focussing on the road ahead unless the signal aspect is lower than about the 4m zone.

6.1.5 Sighting at 15m from signal

This is minimum Railtrack sighting distance “from seated position”.

GE/GN8537 [13] (p.48) states that at 15m, 8 degree lateral cone is 2.11m. This translates to a signal not more than 2.8m from track centre (left side).

Railtrack specifies 1.7m signal clearance from track centre as standard (900mm from near rail).

GE/GN8537 [13] (p.48) states that at 18m, 8 degree vertical cone is 2.53m. This translates to a signal not more than 5.1m above rail, not more than 900mm from left rail.

GE/GN8537 [13] (p.48) states that at 25m, 8 degree lateral cone is 3.51m. This translates to a signal 3.3m above rail, not more than 900mm from right rail.

6.1.6 Sighting at 8m from signal

At this distance, signal is outside optimum viewing (laterally, even if placed at eye height).

Driver cannot see colour of signal without looking with deliberation, and cannot focus on both A and B arms simultaneously.

Whether driver can see the signals at all from seated position depends on cab design.

Railtrack recommendation on viewing from this distance and closer is discussed in section 5.5 above.

6.2 Approaching to 200m from longer distances

6.2.1 Sighting at 400m from signal

At 130 km/hr, the train is 11s sighting away from the signal. At 115 km/hr the train is 13 seconds away. At 80 km/hr it is 18 seconds away.

An LED signal presents as a point light source at this distance. A 600mm background is just distinguishable as a disc (5 minutes of arc); a 900mm background presents more clearly with nearly 8 minutes of arc.

Signal light separation is clearly discernible even at closest light separations.

6.2.2 Sighting at 800m from signal

At 130 km/hr, the train is 22s sighting away from the signal. This is the furthest sighting distance considered as relevant in Railtrack system.

It is stated that the signal cannot be distinguished as the “signal of interest” beyond this distance. In addition, distances are difficult to judge and the signal cannot be utilised effectively as a target stopping point.

The LED signal presents as a point source with not even a 900mm background distinguishable as a separate object.

At 800m, the limit of effective viewing, signal lights (A and B arm) can be clearly distinguished when separated by 1.2m (5 minutes of arc). Arms separated by 1.5m are clearly distinguishable.

6.2.3 Sighting at 2000m from signal

This is approximately GW40 breaking distance.

Signal lights cannot be distinguished from other point sources of light. “A” and “B” arm lights merge unless separated by distances in the order of 3m. Thus green + marker at similar intensity merge with standard arm separation distances (whether 1.5m or 2m) to appear a shimmering yellow.

Distances cannot be reliably judged.

Reliable viewing of the signal is not possible from this distance.

6.2.4 Sighting at 5000m from signal

The fact that a signal is ahead may be discernible from this distance, but little else of value.

7 Signal light output, detection and glare

Apart from position, the most important chacteristics of a signal are its colour and light output. A signal must be effective in providing its message on a bright sunny day, but also at night.

In this section, the issues surrounding appropriate levels of light output in the context of its colour are discussed.

7.1 Adaptation, scotopic and photopic vision

The eye is able to respond to light over a very large range of light levels, ranging from the light of a moonless night though to the light reflected from snow on a sunny day. This represents a total range of more than 10^6 (brightest: dimmest) in terms of luminance levels responded to.

The eye is not able to respond to the full range of light levels at one time. It adapts to the available environmental light level and is able to view over a range of about 10^3 (brightest: dimmest) light levels at one time.

Scotopic vision

At very low light levels, the eye relies on its rod cells for vision. Rod cells, whilst sensitive to very low light levels, cannot detect colour or resolve fine detail. This form of vision is effective for light levels between 10^{-3} cd m^{-2} and 10^0 cd m^{-2}.

The photopic and scotopic visual systems can to some extent be regarded as separate systems with their own individual characteristics.

The sensitivity curve of the eye using scotopic vision differs markedly to that using photopic vision. As can be seen in diagram 1, the peak sensitivity is to green light and the eye is all but insensitive to red light. One consequence of this is that “night vision” is not affected by red light. Another consequence is that green objects can appear bright in low light levels – green and blue light can appear white under low light conditions.

Diagram 1 also provides some context for reviewing the error rates quoted in section 4.3 above in detection of signals of various colours outside the fovea region.

Photopic vision

When light levels reach approximately 3 cd m^{-2}, colour emerges as a characteristic in the visual field. This is the light level where photopic vision, where the eye is able to use cone cells for vision, becomes operable. Dependent on adaption level, environmental light levels up to 10^4 cd m^{-2} can be seen in full detail using photopic vision.

The sensitivity curve of the eye using photopic vision peaks in the yellow range of light. It drops off at both the red and the violet ends of the spectrum as shown in diagram 1.

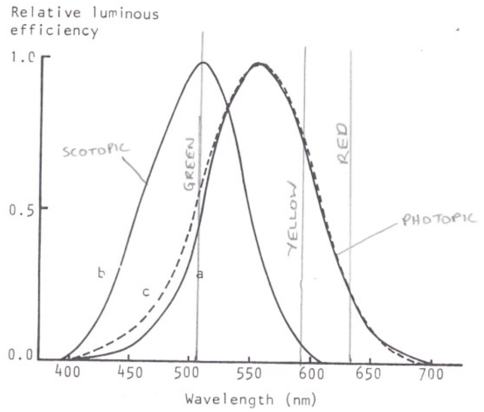

Diagram 1: relative sensitivity of eye to scotopic and photopic vision

Adaption

At any particular moment, the eye is adapted to light of a particular level. Around that adaption level, a range of light levels will be visible in detail. Ranging from light levels perceived as “dark shadow”, light up to 10^3 times that level will be visible. Higher light levels will dazzle the eye and be perceived as glare.

Diagram 2 shows the range of usable vision at various light levels (this diagram should be regarded as indicative only in nature, illustrating the principles rather than defining specific light levels applicable).

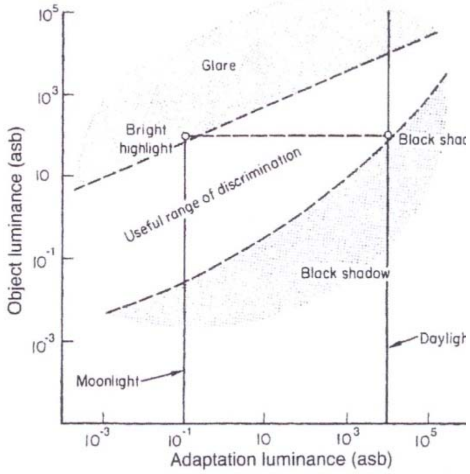

Diagram 2: range of usable vision at various light levels

When light levels change, or attention is drawn to bright or dark objects, the level of adaption changes. This process occurs on a number of dimensions over various periods of time. The time constants involved are different for cone cells compared with rod cells.

The fastest processes of adaption can occur over a few seconds. The major part takes approximately 7 minutes. Full adaption to the new light level does not occur for about 2 hours.

7.2 Contrast

For seeing an object in an environment, the quality it must possess is “contrast”. Contrast is a relatively complex concept. At its most basic, it represents the difference in light level between object and background (contrast is available for either a light object against a dark background, or a dark object against a bright background), but it is complicated as contrast can also be a consequence of differences in colour without the need for a difference in light level (eg red object against green background).

For the performance of its basic function, a signal must be able to stand out against the background on a sunny day. Contrast is the quantity which determines whether such a light will stand out against the background of the environment.

Contrast is defined as $C = \frac{(L_o - L_b)}{L_o}$ where L_0 = Illuminance of object, while L_b = Illuminance of background [6][28].

For good visibility of signals, positive contrast is recommended.

Some measure of “ideal” contrast can be obtained from that practiced for reflectorised road signage. The following is quoted from the Handbook of Transport Engineering [3][29]:

“Sivak and Olsen (1985)[15] derived perhaps the most well-accepted optimum contrast for sign legibility. These researchers reviewed the sign legibility literature pertaining to sign contrast and came up with a contrast

---

[28] Lighting: Basic Concepts; Julian (2006), p 58.

[29] Handbook of Transport Engineering; ed M Kutz (2004), p11.10

ratio of 12:1 for 'fully reflectorised' or positive contrast signs … . This 12:1 ratio would, for example, result in a sign with a 24 cdm^{-2} legend and a 2 cdm^{-2} background. This single, optimal ratio was expanded in a 1995 synthesis report by Staplin (1995)[16] that gave a range of acceptable internal contrast levels between 4:1 and 50:1."

The studies went on to suggest an optimum nighttime sign legend of 75 cdm^{-2}. For daytime, legibility distance continued to improve with increases in luminance up to 850 cdm^{-2} after which performance levelled off.

7.3 Glare

Glare is in its essence a quantity separate from contrast. At its most basic, glare can be seen as a consequence of viewing objects in the "glare" section of diagram 2 above. Two types of glare are recognised: "disability glare" and "discomfort glare".

Disability Glare

Disability Glare is a quantity well understood by science, although the precise mechanism causing it is not always agreed.

Glare occurs when an object above the adaption range appears in the visual field, and the affect of this object is to temporarily disable the ability of the eye to see objects in the darker background around it.

This disabling may be due to:

- In the case of light shining on the periphery of the eye, due to the affects of the light shining on internal structures of the eye and creating interference with the viewing of objects in the direct field of view (the sun is a common source of this type of glare); or
- In the case where the light is directly in the field of view, due to the light source being too bright for the current adaption level of the eye. In this case, the eye can see neither the object nor the background effectively till the eye has become adapted to the brighter light level.

This second is the mechanism of interest for the viewing of signals at night.

As can be seen in diagram 2, it is possible for a light level which appears as dull in bright daylight to dazzle an eye adapted to darkness. The balance for practical signals is discussed in section 9.2 below.

Discomfort Glare

Discomfort glare is less well understood than disability glare. Many studies have been done with results not always aligned with the theoretical framework proposed.

Discomfort glare occurs when a light source is visible in peripheral vision, distracting from a primary task. In this sense, the glare is dependent on the adaption level of the eye, the light intensity of the distracting source, and the angle of that source from central vision.

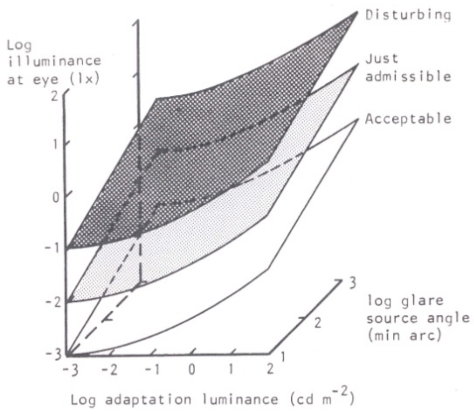

Diagram 3: factors contributing to discomfort glare

Discomfort glare is not relevant in a context where the source of the glare is the focus of the primary task, rather than a distraction from that task. In the studies investigating discomfort glare, the subject performs a task such as data entry over an extended period (generally hours) with a distracting light source set at various positions in the visual field.

Boyce (1982) [1][30] characterises discomfort glare as follows:

"People at work do not keep their eyes on their task all the time they are working. At frequent intervals they look away into the surroundings, although not necessarily fixated on any particular part of them. If the brightness of these surroundings is very much higher or very much lower than the brightness of the task area, the eye will begin automatically to adapt to the change in the brightness. This process, continually repeated during the working day, can produce fatigue and discomfort and possibly a lowering of visual performance."

In the case of viewing signals, the light provided by signals is intermittent to the driver and intended to be the focus of attention rather than a distraction from a primary task.

Discomfort glare from signals is not considered to be a mechanism of concern for a driver of a train. By satisfying the need to avoid disability glare as discussed above, any needs with respect to discomfort glare should also be satisfied.

8 Character of signals in the environment

Signals need to provide positive contrast to the environment in daytime, but not dazzle the eye at night. The relative light levels associated with LED signals compared with the environment are discussed in the following sections.

8.1 Background light levels

Table 9 gives background illuminations under various circumstances [1][31].

This provides the standard "typical" value for sunny day illumination. In practice during the daytime, background luminance varies from 1,000 Cd/m^2 to 10,000 Cd/m^2 depending on factors such as amount of sunshine. The highest luminance quoted is 10,000 cdm^{-2} for fresh snow in sunshine (not so common here).

During the day, a signal needs to stand out against grass or trees in sunlight with luminance of typical value 2,900 cdm^{-2}.

During the night, the background will be brighter than moonlight since it will be lit

---

[30] Boyce (1982); Human Factors in Lighting, p.66

[31] Human Factors in Lighting; Boyce (1982), p 8.

by the train's headlight. As a typical value, background in this situation may be taken as around 50 cd m^{-2}.

8.2 Railway Signal light levels

Typical supplier characteristics for LED signals (200mm type) are as shown in table 10.

These values, based on the Westinghouse range of LED signals, have been found usable in practical signalling applications in Victoria.

8.2.1 Aligning signal "luminous intensity" with luminance of background

To calculate contrast between signal and background, it is necessary to convert them to the common currency of "luminance".

The figure quoted by the manufacturers is "luminous intensity". The same intensity can be delivered by sources of differing sizes. The apparent brightness/ contrast of the light against the background varies according to the area of the source. That is, a LED signal with a uniform light output and larger diameter will appear less intense than the same luminous intensity than one with a smaller diameter or a point source (eg point source for some incandescents).

For a 200mm (0.0314 m^2 area) LED signal with 1,000 Cd output, the luminance is 32,000 Cd/m^2. For the same signal with 300 cd output, luminance is 10,000 cdm^{-2}.

Mashour comments regarding appropriate signal size as follows [8]:

"The relationship between signal size at a given distance and its detection has been investigated recently by Cole and Brown.

"These authors ... found ... (a) that optimal signal intensity is independent of signal size, and (b) that if two signals of different luminous areas have the same but less than optimum intensity, the smaller signal will be more effective. On this basis, Cole and Brown recommend the use of a smaller signals of sufficient intensity rather than large signals. ...

"This corroborates that size is an effective factor in the detection of signals with low luminance levels. This effect decreases, however, with increasing signal intensity".[32]

The above is written in the context of incandescent signals. Where LEDs are required for close viewing, increasing signal size with same light output can reduce the potential for glare at shorter ranges. The studies show that long range characteristics are largely unaffected by doing this.

The principle involved is similar to that used in frosted light globes for the same effect.

8.3 Signal Light – prior work with incandescent signals

To put the above figures for LEDs into context, it is useful to review some of the prior work carried out concerning appropriate signal intensity.

This work has generally involved incandescent lamps.

---

[32] Human factors in Signalling (Mashour) p 171 (table 9-1)

8.3.1 Mashour (1974) discussion of practical luminous intensity

Mashour (1974)[8] reviewed some prior studies on the topic of luminance and signal visibility in the context of contemporary signals.

"Masaki and Tanaka [7] have ... investigated the important problem of how the visual range required for the detection of signals varies with the extinction rate of light through the atmosphere in different background luminances mentioned above and weather conditions except thick fog. According to these determinations, the detection range of signal lights (red, yellow and green) of 1000 cd intensity, on a clear day, a slightly hazy and overcast day, and a rainy day, will be 1.2-3.4 km (for red), 1.1-2.8 km (for yellow) and 0.9-2.1 km (for green). These values should be judged against the background that (1) they were based on threshold values obtained in static laboratory conditions and (2) the assumption of an equal intensity emanating from the luminous surface ...".[33]

At the time of these studies, signal lights were incandescent and used colour filters of varying efficiency to display colours. A result of this was that, with a lamp with basic (unfiltered) intensity of 5500 cd, actual signal colour intensities were much less. This relationship for incandescent lamps is discussed as follows [8]:

"Signal intensity is dependent upon two factors: the light source and the filter's transmittance. The light source (usually an incandescent lamp) is for all practical purposes the same for almost all signal colours within a particular railway network. Therefore the intensity of a signal – its luminance – varies in practice only with transmittance, being considerably lower for red and green and highest for yellow filters. For example, the range of transmittance of the national standard filters in the U.S.A is 0.06-0.17 for red, 0.17-0.27 for green, and 0.58-0.71 for yellow (illuminant 2,854 K)".[34]

For typical colour aspects for incandescent signals at that time, table 11 gives the typical values used by Mashour.

It can be seen that these values for red and green are less than for the typical LED "intermediate" range signal today, as well as being less than the value thought to be optimum at the time.

In recommending the appropriate light level for a signal light, Mashour [8] concludes as follows:

"Increasing a signal's luminance improves its detection. The question is how much it should be increased. The theoretical upper limit of luminosity of a signal is the luminance required for the detection of a signal is the luminance required for the detection of a signal against the highest background luminance, that is, the luminance of new snow on a clear day or the upper surface of the clouds at noon; this being lower than, but in the vicinity of 10,000 mL. But a luminance this high and even much lower creates dazzle with

---

[33] Mashour (1974), p 174

[34] Mashour (1974), p 75.

decreasing background luminance. Usually, a range of about 300 – 2900 fL (1,000 – 10,000 cd/m^2 is considered to represent the luminance of natural backgrounds in daytime, excluding a snow surface. The upper limit of this range – 2900 fL (10,000 cdm^{-2}) – can be taken for the criterion for the selection of signal luminance in such a way that the relative contrast is positive.

"A reasonable recommendation for the upper limit of signal luminance is, in my view, to increase the luminance as much as possible but not so much as to create dazzle (and colour confusion)" .[35]

8.4 LEDs in other "signal" contexts

We as members of the public (as distinct from train drivers) encounter signal lights focussed on us at short range routinely in our day to day experience. The sources of these are traffic lights and flashing lights at level crossings.

Typical specifications for Level Crossing Flashing Lights are given in table 12.

Typical specifications for Road Traffic Light Heads are given in table 13.

These values (200 – 700 Cd) are viewed by car drivers with adaption levels provided by headlights and typical street lighting. As noted above, good street lighting gives luminance of 10 Cd m^{-2}. Car headlights raise this level to (perhaps) 50 Cd m^{-2}.

At the other end of the scale, these lights are regarded as sufficient for viewing in sunlight in the road traffic context.

The "medium" signals refered to in table 10 above have luminance roghly twice that of the traffic signals referred to in table 13.

9 Discussion of signal light requirement

9.1 General context

General levels of background illumination were discussed in section 8.1 above..

For the 1,000 Cd signal (30,000 cdm^{-2}) considered, contrast in normal daylight will vary from 0.67 to 0.97[36]. The internal contrast ratio against grass on a standard sunny day (2900 cdm^{-2}) is about 10, close to the ideal quoted in the Handbook of Traffic Engineering. The outcome using the lower bound of contrast can be gauged by the casual observer by looking at an LED signal on a sunny day. It can be seen to stand out, but not always shine against the background.

For the 350 Cd signal (11,000 cdm^{-2}), the internal contrast ratio against grass on a standard sunny day (2900 cdm^{-2}) is about 4, which is within the acceptable range around the lower bound.

The background luminance at night is much less than during the day.

The full moon, for instance, generates a background luminance in the order of 0.1 Cd/m^2. For a vehicle (eg train) with headlight, the light within the headlight beam can illuminate the environment to above 50 Cd/m^2.

---

[35] Human factors in Signalling (Mashour) p 173

[36] For the 300 cd signal, contrast is 0.00 – 0.90, so brightness is of same order as background. Colour is then relied on for contrast (eg against black background, or red against natural backgrounds).

While this dimmer background provides a much higher contrast for the signal than daylight, it is within the range of normal “non-adaptive” vision for a person. A person can generally perceive dark and bright objects together over a range of 1000x luminance (bright object to dark object).

The internal contrast ratios for both 350 Cd signals and 1000 Cd signals is high at night, but would seem acceptable against a background adaption level of 50 cdm^{-2}. It is confirmed by very wide usage that 350 Cd signals can be viewed without excess glare. There is possibly room for demonstration that the same conclusion can be reached for the 1000 Cd signals.

9.2 LED light levels

9.2.1 General review

Taking driver night adaption level as suitable for headlight illumination (minimum 22 Cd m^{-2})[37], signal luminance up to 1,000 times that should be discernable. This is 22,000 Cd m^{-2}.

Converting this to signal light level gives about 700 Cd output for a 200mm^2 diameter signal.

Such a light level would also give a positive contrast against snow.

There is a degree of uncertainty in each of these values. In general, the 1000 Cd level recommended for signal lamps in 1970s studies can be seen to be at the high end for night time viewing.

This level is consistent with light outputs found in typical intermediate range LED signals (1000 Cd). The values for “long range” (2000 – 4000 Cd) appear to be above the ideal level.

It is noted that while “600m range” (300 – 350 Cd – consistent with levels selected for traffic lights) signals are selected in NSW, “1500m range” (800 – 1100 Cd with 1800 – 2000 Cd for yellow) are often selected in Victoria. If drivers are experiencing glare with the red aspects, it may be indicative that the yellow aspects are producing luminance above the glare threshold and not being seen reliably at night with respect to colour.

Practical light levels for incandescent signals (red and green aspects) fall in the 500 Cd range to (yellow aspect) 2500 Cd range. These light levels have not been reported to cause glare.

9.2.2 What is important?

It is clearly important that all colours can be seen by drivers at all light levels. Thus all signal lights must have light levels below the glare threshold for night viewing to allow their colours to be accurately identified.

For a train standing at a signal, the subjective glare experienced from the red aspect is the most significant. While green and yellow signals will be passed at speed, red signal will be stopped at and viewed for some time. Of more importance than this subjective experience of glare is the objective ability of the driver to judge colour at night. This may be the effect if the signal is above the glare threshold but not viewed for an extended period.

---

[37] The adaption level is taken from the paper: *Age and Glare Recovery Time for Low-Contrast Stimuli*; Frank Schieber (

Where glare is perceived to be a problem at shorter ranges, the approach of increasing physical size for same light output can also be considered.

The strategy of moving the signal to peripheral vision to avoid glare is least successful with the red aspect. As discussed in section 7.1 above, detection of red light drops off rapidly outside the fovea region of the eye and red is not detectable at all using the scotopic vision available at the peripheries of sight.

9.2.3 The Victorian yellow aspect

It was previously identified that the Victorian drivers perceived a need for the yellow aspect to be brighter than either the red or green aspect in a signal.

This can be seen to mimic the relationship in the incandescent signal where the effective output of the yellow aspect was very much higher than either the red or the green aspect. But why does this lead to a different perceived need in Victoria compared with other states?

One possible answer is the issue of the long range viewing. When viewed from very long distance, the “green over red” or the “red over green” aspects are seen as yellow due to the properties of additive colours. Making the yellow light brighter may help to distinguish the “yellow over red” (or “red over yellow”) from the green/red combination at distance.

It should be pointed out that a similar effect could be obtained by reducing the light output of the red “b arm” in 3 aspect signals.

10 Practical issues in selecting “range”

There appear to be significant differences in practice between jurisdictions.

While these differences may be accommodatable within the normal range of human vision within the practical environment, additional work is recommended to confirm that this is the case.

Red aspects as provided in NSW (600m range, 350 cd) do not appear to be at a light level which should cause glare. The light levels provided in Victoria (1500m range, 1000 cd) appear consistent with European practice and also should not cause glare. If glare is perceived in the Victorian signals the situation could be improved by providing signals with the light levels specified in NSW.

Both types appear to fall within acceptable range, though additional confirmation may be sensible for the Victorian signals.

Provision of light levels higher than these may be problematic at night.

10.1 Variation in light intensity with viewing angle

The optics of LED signals generally involves light being directed by lenses fitted as part of the LED assembly itself. Light outputs quoted are achieved by a combination of LED output and lenses.

Changes in light levels between “medium” and “intermediate”[38] can be

---

[38] Adopting terminology used in Westinghouse documents

achieved by increasing LED output (as in UGI signals) or by a combination of increasing LED output and narrowing the beam by lense effects (as in Westinghouse signals)

The step from "intermediate" to "long range" is achieved by narrowing the beam using lenses.

More detailed "beam spread" data was obtained from Westinghouse. This is presented in diagrams 4 and 5.

With the signals located 3m from nearest track and focussed to 2-300m (assume straight track), the light reaching the driver's eye at closer range (20 – 50m short of signal) will be below the maximum light level. Indicative calculations applicable for the Westinghouse signal range (based on diagrams 4 and 5 – note that peak figures shown appear to be approximately double the values quoted in the specification sheets) are provided in table 14.

From this table it can be seen that each signal type provides similar light intensity at approximately 50m distance.

At closer approach, light levels drop off significantly for the intermediate and long range signal types. Sighting in poor sighting conditions (eg heavy fog) will thus be adversely affected by selecting the longer range types.

At longer distances, light levels are higher depending on the type.

It can be concluded from this data:

- Medium range signals provide best performance in poor visibility (eg fog) without the risk of dazzling in darkness.
- Intermediate range signals provide inferior performance in poor visibility (eg fog) and carry some risk of dazzling in the darkness. The difference in performance is not large.
- Long range signals are only usable where short range viewing is not important.

11 Miscellaneous Issues

11.1 Banner Indicators, Railtrack practice

Banner Indicators have been suggested as a solution to signal sighting problems generally. The following is the view expressed in the Railtrack guideline on this topic [13]:[39]

"It can be argued that the presence of the banner repeater signal can act to prime the driver of the approaching signal and as such could help overcome a deficiency in reading time.

"Recent human factors evidence suggests that no more than half a second of the required reading time can be deducted as a result of the driver being "primed" (this figure is not dependent on how long the banner repeater signal is visible for).

"The additional cost of providing a banner repeater signal to "prime" the driver would be assessed against a quite small benefit, unless other benefits have also been identified.

<u>GE/RT 8037 Cl 5.3.2 [12]</u>*: "Usually the main signal shall become visible when the driver loses sight of the banner repeater. However, it is permissible to have a gap. … The gap shall be as short*

---

[39] GE/GN 8537, P44-5

as possible, subject to achieving good sighting of the banner signal (typically no more than 3 seconds)".

The following examples of use are provided:

- A "start on yellow" SPAD. … Depending on the speed that the train has reached and the achievable reading distance, a banner repeater signal may provide additional warning of the stop aspect so as to avoid the SPAD.
- If a risk of a driver reading the wrong signal has been identified (perhaps due to a curved approach to a gantry) the provision of the banner may help a driver avoid observing the wrong signal.
- A signal at the exit from a tunnel or other very dark section may be difficult to read as a result of the sudden impact of daylight. A banner repeater positioned within the tunnel may assist the driver if the signal is not easily observed in the bright sunlight.
- Banner repeaters provided for performance enhancement (rather than to achieve signal sighting requirement) are normally associated with locations whee an early warning of a signal having cleared from stop to proceed is considered to be worth the additional cost involved.

It can be concluded that the application of placing a Banner Indicator close to a platform for viewing by a departing train would not be considered good practice by Railtrack. If a Banner Indicator is thought to be needed to cover a "depart on yellow" risk, it would be placed ahead of the geographic feature obscuring the view of the main signal in accordance with the Railtrack Guideline.

11.2 Signal Backgrounds

The effectiveness of large signal background compared with smaller signal backgrounds has been discussed, as has the use of white borders. This is of particular relevance to ARTC since Searchlight/ LED tri-colour backgrounds are 900mm while LED multi-light backgrounds are 600mm.

Railtrack guidelines make the following comments on this issue [13]:[40]

"The use of matt black sighting boards can enhance the contrast between signal aspects and light, or cluttered backgrounds. Research has shown (no ref given) that larger backplates (1.5 times the current standard size of a 4-aspect signal backplate [this is 900mm instead of 600mm]) can improve the visibility of signal aspects, due to the increase in contrast between the signal and its background.

"White backplate borders have been used in the past to try to draw attention to problem signals. However, unless the approach speed is slow (15mph or under) and the view uncluttered, borders have the opposite effect as they merely serve to reduce the apparent size of the black backplate, thereby reducing contrast and visibility."

---

[40] GE/GN 8537 (Appendix 3), P62

Based on data referred to in section 6.2, it is found that the 900mm backgrounds can provide improved viewing compared with 600mm backgrounds at approach distances between 400m and 600m. There may also be a benefit at closer distances. This benefit is reduced when white borders are used.

11.3 Other Visual Affects – Atmospheric Optics

11.3.1 Mirage and atmospheric lensing affects

Air is not a uniform optical medium. The refractive index of air varies depending on its moisture content and absolute temperature (which determines density). The usual pattern is for the air to reduce in temperature and density with altitude. However, close to ground level, cold air can form a stable layer below warm air due to the fact that “hot air rises”. This is known as “inversion”.

Combined with these layering affects with height, vertical “fronts” can occur where hot and cold air masses collide, separating air masses across a sharp interface. These are well known in Melbourne with the experience of the fast moving “cool change”. At other locations these fronts can be much more complex, slower moving or even subject to retrograde motion.

At particular types of locations, such as near the sea or in deserts, large temperature gradients can occur in the air over very short distances. These can cause lensing, total internal reflection between air layers (so that the air layer acts akin to optic fibre where the effect persists) and other atmospheric effects. These cause distant objects to appear much closer, or in a different location than they would normally be seen. Conversely distant objects may not be visible from location where normally they would.

The following is quoted from a text on light in the outdoors [9]:[41]

“There are days when everything can be seen with extraordinary clarity, and a faraway town or lighthouse suddenly becomes visible which in ordinary circumstances would be impossible to see at all because it lies below the horizon. Very often it gives the impression of being surprisingly near.

“Two very striking cases of this kind were once observed along the English Channel. Once, the whole of the French coast opposite Hastings could be seen from the beach there with the naked eye, whereas in ordinary circumstances it cannot even be seen with good binoculars. Another time, the whole of Dover Castle was seen from Ramsgate to appear from behind the hill that usually covers the greater part of it.

“And, conversely, there are cases where distant objects that usually project above the horizon disappear as if they lay below it. These conditions too give the strong impression of proximity.”

A further example can be cited from the Australian outback. In this case a glowing ball was observed close to a campsite at nighttime. It appeared to be moving in space, appearing and disappearing at intervals. A colleague of the observer was in a car more than 50

---

[41] Minneart (1992): Light and Colour in the Outdoors, p62.

miles distant and in contact by radio. After some discussion concerning the object, it was realised that the mysterious object could be made to disappear and appear again by switching the car headlights off and then back on again.

This was an example of atmospheric lensing.[42]

The Nullabor in Australia is particularly suited for phenomena such as those described above. It is important to note that these effects are physical, not perceptual. They cannot be influenced by changing the height of the object (in this case the signal), or by increasing or decreasing the brightness of the light or altering its colour.

It is also important to note that these effects involve distant objects, not objects at the distances between 0 and 800m.

11.3.2 Sighting signals in the absence of landmarks

Driver route knowledge is often relied on to ensure that the driver knows when to expect signals and when best to apply the brakes having passed a warning signal.

Some rail authorities provide Advance Waring Boards to assist the driver to know when best to apply the brakes for various classes of train at various speeds to stop at particular signals. These have been found to be of use to the extent that they do not contribute to "clutter of signage".

At locations such as on the Nullabor, the problem is not too many signs, but too few landmarks. In this context, there can be value in adding landmark and approach signs to fill the gap caused by the otherwise featureless landscape.

11.3.3 Colour shift when viewing over distance or through haze

When viewed through haze or dust over long distances, light sources tend to be shifted towards the red end of the spectrum. Thus, yellow lights can appear to be red.

Note that this affect cannot be avoided by mounting the signals on higher masts since typically the distances involved involve viewing over the horizon.

It should also be noted that this effect is quite different from that where a "green over red" signal appears as a shimmering yellow over distance. This second is a fundamental optical effect caused by the merging of the red and green lights at distances of 1-2km and more. It is not caused by haze.

12 Need for consultation with other groups re "siting" requirements

Various other parties apart from drivers and operators have an interest in the location of signals. Some of these are discussed in the following sections.

12.1 Maintainer needs

Railtrack guidelines make the following comments regarding the needs of the maintainers for safe access [13]:[43]

"Legislation promotes safe access for maintainers. Placing the signals

---

[42] No literature reference (was reported on the radio). Further observation was made that UFOs did not visit Earth quite so often prior to the invention of the car headlight.

[43] GE/GN 8537, P12

unnecessarily between the running lines can work against this requirement.

"Access route needs to be provided."

ARTC maintainers have expressed similar views as part of a separate risk assessment process. Maintainers should be involved in the signal sighting process for this reason.

12.2 Engineering design

It is a worrying trend that the design is made "subject to signal sighting" rather than the other way around. By requiring the designer to accept the signal sighting committee's view as final for the location of signals, some of the responsibility for the correctness of the final design is removed from the designer.

This is of concern since the Signal Sighting Committee will not always be in possession of all information relevant to designing the location of the signal. As a committee, it is also less well placed to take responsibility for the impact of technical decisions it may make.

Railtrack guidelines reverse the effective roles by stating that the Signal Sighting Committee makes recommendations only. The final design remains the responsibility of the designer. It states [13]:[44]

"As a critical part of the signalling design, it is important that the signal sighting recommendations are checked in a similar manner to any other signalling drawings."

"The role of the checker is to review the forms and comments made by the signal sighting committee, looking in particular for evidence that:

1. the assessment processes have been carried out.

2. the calculations are correct and any assumptions made are reasonable.

3. the recommendation is either compliant or that departures derogations are identified.

4. any departure/ derogation proposed is adequately justified and supported with evidence, where necessary."

12.3 Stopping on sight – Reasons for seeking longer sighting distances

The material presented to date emphasises the importance of adequate viewing between short and intermediate distances. In this section, the possible contribution "long sighting" can make to safety is discussed.

The Human Factors research supports the concept of keeping the signalling as simple as possible. The simplest form of signalling is 2- aspect (Red and Green) as is used, for instance, on the London underground.

This can be an effective approach in some circumstances which depends largely on the braking characteristics of the train in question. The risk of attempting to apply this approach (eg by providing 2km sighting distances to red signals) in the context of Passing Lanes is discussed in this section.

The Victorian rule book allows that for line speeds up to 80km/hr, 2-position signalling may be put in place without the need for distant signals. The basis for this is that, for such areas, provided

---

[44] GE/GN 8537, P21

that the stop signal is sighted from a distance of 400m, a loco-hauled passenger train can be expected to be able to stop on sight of that signal.

For much of the Melbourne Metro area, where line speeds are 80km/hr or less, and sighting of 400m is typically available, "stop on sight" capacity becomes a viable backup of last resort where warning aspect signals have not been observed.

The sighting distance needed for "stop on sight" varied according to the line speed and the braking characteristics of the train. Modern EMUs and DMUs have better braking than loco-hauled passenger trains. GW40 represents a significantly lower braking rate than any passenger train.

For passenger trains, a rough guide to "stop on sight" distances needed for various line speeds and braking rates is given in table 15.

Within this range, VLP loco hauled brakes at 0.6m/s/s, EMU and Sprinters brake at 0.8 m/s/s, XPT brakes at 0.9 m/s/s, Vlocity brakes at 1.1 m/s/s Freight trains (GW40) generally brake at approx 0.23 m/s/s[45].

Taking account of the typical sighting distance, a passenger train having not responded to the warning signal (disregard) often has a good probability of stopping after responding to the red signal alone. On this basis, Mashoud (1974) acknowledges that, as a "final backup" after all other defences have failed, "stopping on sight" can provide a basis for extending sighting beyond the minimum otherwise recommended.

The same argument fails with respect to freight trains operating at any but low and medium speeds, since "stopping on sight" fails as a strategy for the required stopping distances.

13 Conclusion

Understanding the physiology of the eye and the mental processes which occur between a signal presenting itself and being responded to be a driver are crucial in designing signalling.

The requirement found in GE/RT 8037 to place the red aspect of the signal close to the track and at eye level can be seen to be solidly grounded in such an understanding.

As train speeds increase and tracks continue to be built with curves, it important to understand the types of errors which can occur in responding to signals and the potential severity of each type of error. This allows limits to acceptability on a number of dimensions to be set on a rational basis with knowledge of the human "system" which must deal with any compromises made.

Finally, it can be seen that safety often means doing things simply. Section 4 is particularly instructive in this regard. As signal complexity increases, response times can be seen to increase together with the occurrence of errors.

It is hoped that this paper can be a resource for those tasked with designing and positioning signals.

14 Appendix – Comment regarding units

It is easy to get lost in the units used for measuring different aspects of light. Not

---

[45] Various sources including VRIOG 012.0.1 – 2008 and GW braking curves.

only are the units easily confusable, the metric and imperial equivalents use similar names. In this section, some of the units will be explained.

The following is quoted from the Handbook of Transportation Engineering [3]:

"The … most important photometric measurements used to describe sign visibility are luminous intensity, illuminance, luminance and reflectance.[46]

"Luminous intensity, expressed as candelas (cd), is a description of a light source itself and is therefore independent of distance. That is, no matter how far away an observer is from a lamp, that lamp always has the same intensity. Luminous intensity is the photometric measurement most often specified by lamp and LED manufacturers.

"Illuminance, or incident light, is measured in lux (lx) and is a measure of the amount of light that reaches a surface from a light source. Illuminance is affected by distance and is equal to luminous intensity divided by the distance squared. …

"Luminance is expressed in candelas per square metre (cdm^{-2}). Luminance is the photometric that most closely depicts the psychological experience of 'brightness'. Luminance can refer to either the light that is emitted by or reflected from a surface, and is an expression of luminous intensity (cd) over an extended area (m^2). Like luminous intensity, a source's luminance is constant regardless of distance. …

"Reflectance is the ratio of illuminance to luminance and, as such, reflectance describes the proportion of incident light that is absorbed and the proportion that is reflected by a surface. If, for example, 100 lx hits an object's surface and that surface has a luminance level of 5 cdm^{-2}, that surface has a reflectance of 5 percent."

Luminance in the imperial system, is measured in "foot lambert" (fL). Mashour (1974) refers to this unit continually, providing the following conversion formulae:

$$1\ cdm^{-2} = 0.292\ fL.$$

$$1020\ fL = 3500\ cdm^{-2}$$

Since this paper concerns brightness and contrast, it is these measures of luminance which are of most interest.

Acknowledgements

I would like to acknowledge:

Australian Rail Track Corporation for providing permission for publication of material originally prepared for them. Trevor Moore for assisting and reviewing.

Westinghouse Signals for providing typical data on signal types. Wayne McDonald for assiting in this respect.

There are no sources in the current document.

References

[1] Boyce P R. Human factors in lighting. London: Applied Science Publishers, 2003.

---

[46] Handbook of Transport Engineering; ed M Kutz (2004), p11.3

[2] The Rt Hon Lord Cullen PC. The Ladbroke Grove rail inquiry – part 1, report. Health and Safety Commission, 2001;

[3] Garvey P M, Kuhn B T. Highway sign visibility. In: Kutz M ed. Handbook of transportation engineering. New York: McGraw-Hill.

[4] Institution of Railway Signal Engineers. IRSE Signalling philosophy review, April 2001;

[5] International Standard. ISO/FDIS 9241-303 Ergonomics of human-system interation – Requirements for electronic visual displays (final draft). Geneva: ISO, 2008.

[6] Julian W G. Lighting: basic concepts (6th edition). Sydney: University of Sydney, 2006.

[7] Masaki, H, and Tanaka, H. Visual range of signal lights (I). Quartl. Rep. (JNR), 1963, 4 (3), pp67-71.

[8] Mashour M. Human factors in signalling systems – specific applications to railway signalling. Stockholm: Almqvist and Wiksell International, 1974;

[9] Minnaert M G J. Light and colour in the outdoors (translated and revised by Len Seymour). Springer, 1992;

[10] Nikandros G, Tombs D. Measuring railway signals passed at danger. 12th Australian Conference on Safety Critical Systems and Software, Adelaide: Australian Computer Society, 2007.

[11] Office of Transport Safety Investigations. Rail safety investigations report – derailment of 4BM7, Lidcombe, 4 November 2005. Sydney: OTSI, 2006.

[12] Rail Safety and Standards Board. GE/RT 8037 Signal positioning and visibility, 2003;

[13] Rail Safety and Standards Board. GE/GN8537 Guidance on signal positioning and visibility, 2003;

[14] Railtrack Safety and Standards Directorate. GK/RT 0037 Signal Sighting (issue 3, December 1997, superseded). Railtrack, 1997.

[15] Sivak, M and Olson P L. Optimal and minimal luminance characteristics for retroreflective highway signs. In Transportation Research Record no 1027, 1985. pp 53-56.

[16] Staplin L. Older driver and highway safety literature review and synthesis. Working paper in progress. US Department of Transportation, FHWA, Washington DC, 1995.

[17] Victorian Rail Industry Group Standards. VRIOGS 012.0.1 – 2007 Standard for signalling design and documentation. Melbourne: Director of Public Transport, 2007.

Results for basic 3 aspects were as shown in table 2 [8][47].

---

[47] Ibid. Summarised from section 8.2, pp 140-145

Figures and diagrams

Aspect	Average Response	Range (all systems)[48]
“G” (proceed line speed)	1.78s	1.46–2.20s
“R” (Stop)	1.89s	1.41–2.50s
“Y” (Stop next signal)	2.70s	2.22–3.36s

Table 2: response time to signal aspects

Aspect	Average Response	Range (all systems)
“R/Y” (medium speed warning)	3.46s	2.42–4.21s
“R/G” (Proceed medium speed)	2.65s	2.38–3.13s
“Y/G” (Reduce to medium speed)	3.32s	2.60–3.52s

Table 3: response time to signal aspects

Aspect	Average Recognition	Range (all systems)
“G” (proceed line speed)	89.1%	51% - 100%
“R” (Stop)	90.5%	80% - 98%
“Y” (Stop next signal)	64.7%	22% - 95%

Table 4: error rate viewing various signal aspects

---

[48] The studies investigated a number of different signal aspect sequences in use in Europe at that time. These were rated with differing aspect complexities, many more complex than that used in Australia. “all systems” means the range across all those systems included in the studies

Aspect	Average Recognition	Range (all systems)
“R/Y” (medium speed warning)	48.9%	20% - 78%
“R/G” (Proceed medium speed)	56.8%	17% - 91%
“Y/G” (Reduce to medium speed)	58.9%	21% - 90%

Table 5: error rate viewing various signal aspects

Position Label	Description of Location
L	5.63 degrees to left of central visual point. This is outside fovea region but just within central visual field.
M	0.76 degrees from visual centre. This is within fovea region
R	18.18 degrees to right of central visual point. This is outside central visual field.

Table 6: eccentricity values used in signal detection study

Signal aspect	“L” detection time	“M” detection time	“R” detection time
“G”	2.11s	0.78s	2.44s
“R”	2.27s	0.78s	3.85s
“Y”	1.25s	0.90s	3.68s

Table 7: response time to signals at various eccentricities

Signal aspect	“L” detection	“M” detection	“R” detection
“G”	31.7%	96.7%	5.0%
“R”	28.3%	99.2%	5.0%
“Y”	68.3%	93.3%	5.8%

Table 8: detection rates for signals at various eccentricities

Situation	**Illuminance on horizontal surface (lm m^{-2})**	**Typical Surface**	**Luminance (cd m^{-2})**
Clear sky in summer	150,000	Grass	2,900
Overcast sky in summer	16,000	Grass	300
Textile inspection	1,500	Light grey cloth	140
Office work	500	White paper	120
Heavy Engineering	300	Steel	20
Good street lighting	10	Concrete road surface	1.0
Moonlight	0.5	Asphalt road surface	0.01

Table 9: lighting levels in various environments

Description	**Nom Range**	**Min Peak light (Cd) New**	**Luminance (cdm^{-2})**	**I Contrast (day)**	**I Contrast (night)**
Medium range	600m	350	11,150	4	200
Intermediate range	1500m	800	25,500	9	500
Long range	2500m	2000	63,700	22	1300

Table 10: light levels for various LED signal lamps

Aspect colour	Lense transmittance	Aspect intensity (cd)
Green	0.10	550
Yellow	0.45	2,475
Red	0.10	550

Table 11: transmittance of incandescent lenses by colour

Description	Colour	Nom Range	Min Peak light (Cd) New	Luminance (cdm^{-2})
Crossing 200mm	Red	500m	300	9,550
Crossing 200mm Superbright	Red	1000m	700	22,300
Crossing 300mm	Red	1000m	400	5,660

Table 12: typical specifications for flashing lights at level crossings

Description	Colour	Min Peak light (Cd) New	Luminance (cdm^{-2})
Traffic Light 200mm	Red/amber/ green	>200	6,370
Traffic Light 300mm	Red/amber/ green	>400	5,660

Table 13: typical specifications for lights at road intersections

Distance of train from signal	Signal distance from track	Signal angle from axis	Light level - Medium type	Light level – Intermediate type	Light level – Long range type
20m	3.0m	8.5°	600-700 cd	100-200 cd	0-100 cd
20m	2.5m	7.1°	600-700 cd	200-300 cd	0-100 cd
50m	3.0m	3.4°	1000-1100 cd	1100-1200 cd	800-900 cd
50m	2.5m	2.8°	1000-1100 cd	1200-1300 cd	1200-1300 cd
300m	-	0°	1000-1100 cd	1600-1700 cd	4800-4900 cd

Table 14: intensity by lense / distance

Braking Rate ->	0.6 m/s/s	0.7 m/s/s	0.9 m/s/s	1.0 m/s/s	1.2 m/s/s
Line Speed					
LS- 80 km/hr	410m	350m	270m	250m	210m
LS - 100 km/hr	640m	550m	430m	390m	320m
LS - 115 km/hr	850m	730m	570m	510m	430m
LS - 130 km/hr	1090m	930m	720m	650m	540m
LS - 160 km/hr	1650m	1410m	1100m	990m	820m

Table 15: Braking distance to stop from various line speeds with various braking rates

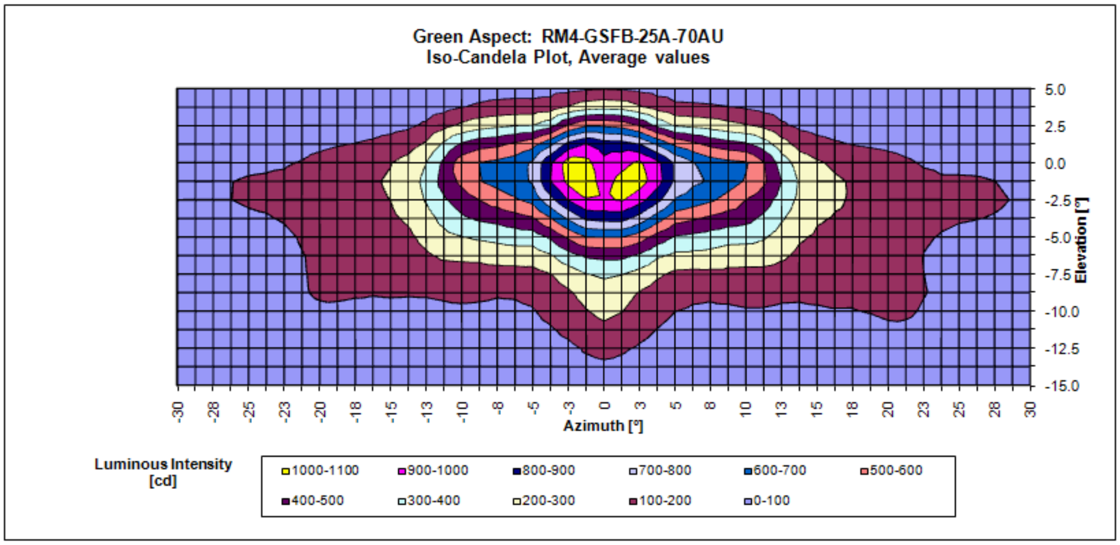

Diagram 4: intensity and spread for Medium Range signal

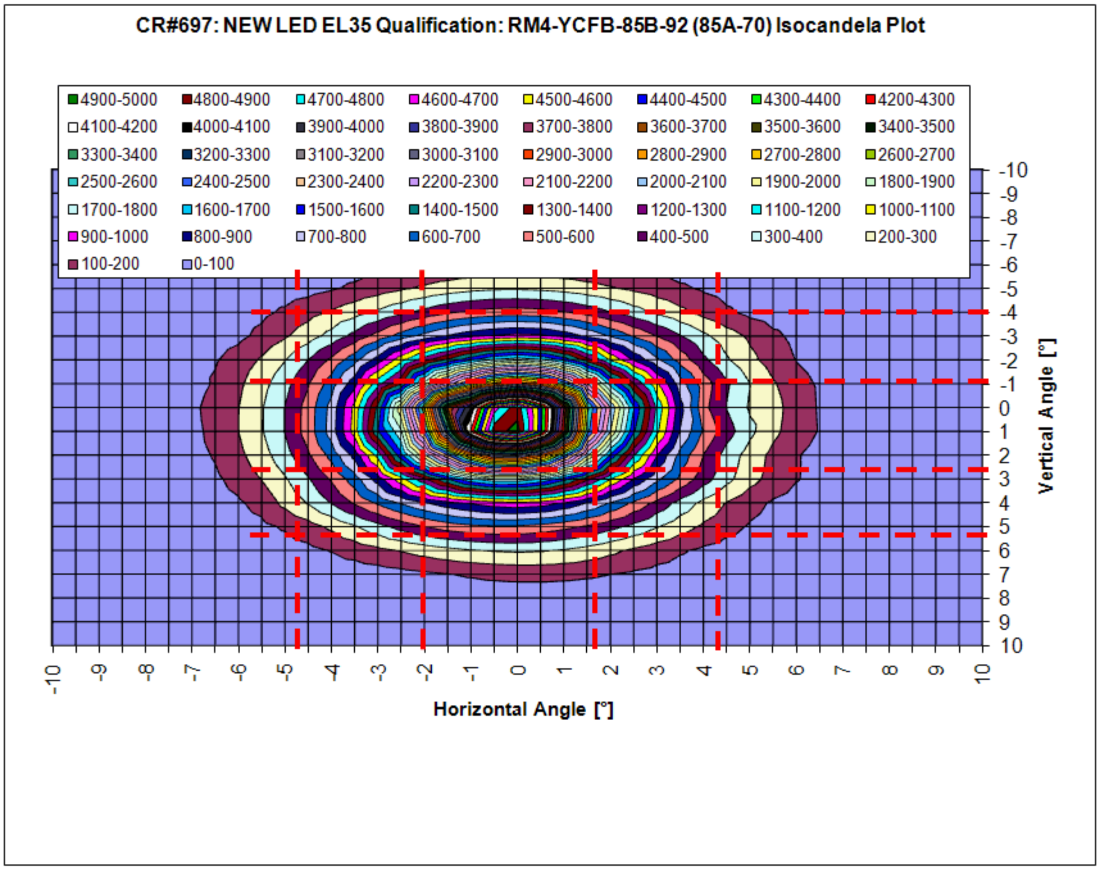

Diagram 5: intensity and spread for Long Range signal