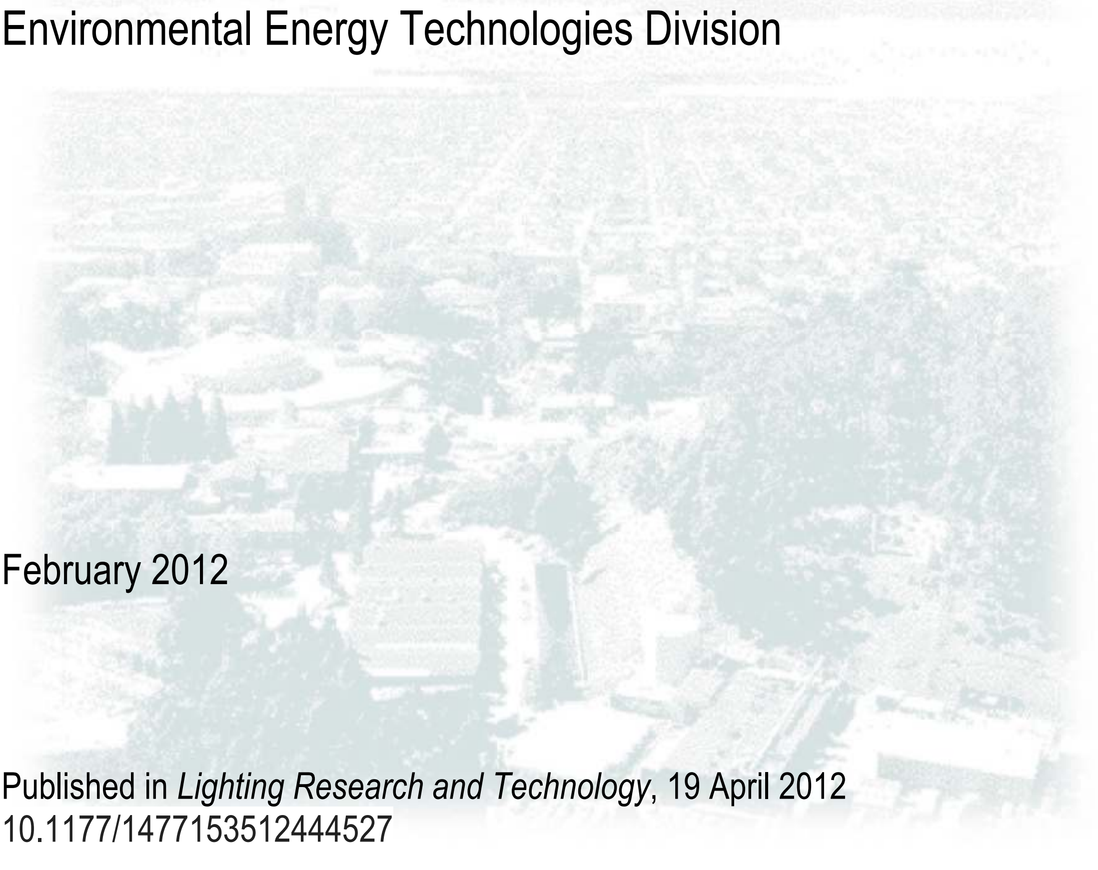
LBNL-5584E

ERNEST ORLANDO LAWRENCE
BERKELEY NATIONAL LABORATORY

Discomfort Glare: What Do We Actually Know?

R.D. Clear
Lawrence Berkeley National Laboratory

Windows and Envelope Materials Group
Building Technology and Urban Systems Department
Environmental Energy Technologies Division



February 2012

Published in *Lighting Research and Technology*, 19 April 2012
10.1177/1477153512444527

Published in *Lighting Research and Technology*, April 2012

Discomfort Glare: What Do We Actually Know?

R.D. Clear

Building Technologies Program, Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory, Mailstop 90-3111, One Cyclotron Road, Berkeley, CA 94720, USA
Phone 1-510-486-4286, Fax 1-510-486-4094, email rdclear@lbl.gov

---

Abstract

Glare models were reviewed with an eye for missing conditions or inconsistencies. We found ambiguities as to when to use small source versus large source models, and as to what constitutes a glare source in a complex scene. We also found surprisingly little information validating the assumed independence of the factors driving glare.

A barrier to progress in glare research is the lack of a standardized dependent measure of glare. We inverted the glare models to predict luminance, and compared model predictions against the 1949 Luckiesh & Guth data that form the basis of many of them. The models perform surprisingly poorly, particularly with regards to the luminance-size relationship and additivity. Evaluating glare in complex scenes may require fundamental changes to form of the glare models.

---

1. Introduction

Discomfort glare is a major issue for lighting and daylighting design. Although glare is a subjective response, engineers have developed formulas to quantify it. In fact, they have developed a number of formulas to quantify it, and new formulas are still being developed. Admittedly, some of the formulas only apply to a limited range of sizes, with different formulas applying to different size ranges. However, even within a size range there is more than one formula, and several formulas supposedly apply to both small and large targets.

We do not have a theoretical understanding of discomfort glare, and every formula evaluates glare differently. The latter factor makes it difficult to objectively compare different models for glare against each other. As a result, we have agreement on the factors that cause glare and the basic trends, but no real measure of how well the different models fit the data.

In this paper we review the early work of Luckiesh and Guth (L&G), and of Hopkinson with an eye towards understanding the basic assumptions behind the models that arose from their work, along with an exploration of the reasons for differences between them [1,2]. The simplest models based on their work were small source models, but extensions to these models soon arose to allow them to treat large glare sources. More recently, some explicitly large glare models have been developed. We present, and then have applied, a procedure to objectively compare all the models to data, such as the L&G data, where there is a fixed glare level. These comparisons illustrate some significant shortcomings in the models, and complement the more general review in suggesting important avenues for future research.

The addendum at the end of the paper lists the formulas for the models that we reviewed. These models are the British Glare index (BGI), the Discomfort Glare Index (DGI), the Cornell Glare Index (CGI), the Unified Glare Rating (UGR), the Visual Comfort Probability (VCP), the Discomfort Glare Probability (DGP), the Predicted Glare Sensation Vote (PGSV), and Osterhaus' Subjective Rating (SR) rating. The last two models were designed to model glare from large sources. We did not expect them to fit the L&G data very well, but we were interested in their forms. The DGP model has both a large and a small glare source term, so we were uncertain as to how well it might fit the L&G data. It should not be a surprise that even none of the small sources models were completely successful at fitting the data.

2. General Issues

The basic factors common to glare models are glare source intensity, which is determined by the source luminance, size, and location, and the adaptation state of the observer. Observer variability is addressed in two of the models (VCP and DGP) by expressing the degree of glare in terms of the percentile of observers who consider it at or above a fixed reference level. The remaining models express glare by the mean or median rating of the observers. Context is partly covered by one model (DGI) in that the interpretation of the numerical value of the glare rating in terms of subjective impression differs based on whether the glare is from interior lighting, or daylight. There is no correction for possible cultural differences in any of the models. Luckiesh and Guth informally reported that glare source exposure duration was not a significant factor in the glare evaluation, and none of the above models explicitly include duration as a parameter.

Glare in the L&G experiment was measured in terms of the BCD, which is the luminance that a subject reports is on the border between comfort and discomfort. Each of the common factors mentioned above was varied over a range of values, while the remaining factors were held fixed at a reference level. The experimental design assumes factor independence as only one reference level was used, so there is no information on possible interactions between factors. This assumption of independence is reflected in the functional forms of the small glare source models. To some extent this appears to be a deliberate policy to maintain simplicity in the calculation of glare. Hopkinson notes that at the 1951 CIE meeting it was agreed that discomfort glare could be assessed as a set of multiplicative functions of a single parameter, and further notes that this is an approximation [2]. In particular, Hopkinson notes that the relationships found at moderate levels of luminance and size break down at extreme values of these parameters.

The BGI and UGR models have independent terms, but they are explicitly small glare source models, and are presumably only valid for such sources. The VCP, DGI, and CGI models have luminance adaptation terms that are functions of both the surrounding luminance, and the source luminance, so they can at least approximately handle larger sources. These models have the CIE multiplicative form, but the luminance adaptation is a function of both the source and background luminance, and therefore all the parameters do not act completely independently of each other. The SR rating model has independent terms, but it was based on a study of large glare sources, and its form is very different from that of the above models. The PGSV model, which was also developed for large glare sources, has a term in source luminance times a term that is dependent upon both size and background. Finally, DGP has two additive terms. Although the terms individually show independence, their relative importance varies with values of the parameters, with the result that the relative significance of any one factor does depend on the other factors.

Another significant issue is that the small source models suffer from a problem with ambiguity. These models assume well defined sources, with uniform luminances for both source and background. When the scene becomes complex, it is not clear what actually constitutes source, or background. The ambiguities become especially obvious when there is an attempt to assess glare from simulated Radiance images, or high dynamic range (HDR) luminance maps:

A) What relative value of luminance makes a pixel count as a glare source, rather than background?
B) When can a "glare" pixel be combined with nearby pixels as a combined source, with an average or effective luminance, a combined area and a single position factor, and when does it count as a separate source?
C) Can the effect of the non-uniformity of luminance of a source and the background be ignored?

We do not have validation for the rules that have been adopted to handle these issues in complex scenes. Waters has reported that non-uniformity of the glare source does affect the perceived degree of glare, but we do not a full understanding of how it affects it [3]. Similarly, Hopkinson reported that the local surround affected the adaptation differently than the far surround, but again we do not have a quantitative formulation for this effect [2]. The CIE specifies that the background luminance is calculated as the average luminance of the field of view, not counting the glare sources, while VCP counts all sources [4,5].

A second form of ambiguity becomes obvious when comparing some of the large glare source models to the small glare source models. The small glare source models are solely concerned with luminances and solid angles. The SR model is based on illuminances, and the DGP model uses both luminances and illuminances. For small sources the distinction has no practical significance. Illuminance is just the luminance times a configuration factor, but the difference between the solid angle and the configuration factor for a circle with a diameter of 30° is less than 2%. The configuration factor has a $\cos(\Theta)$ term for off-axis sources, but this simply means that a new position factor would have to be defined that was equal to the old one divided by $\cos(\Theta)$ to the appropriate power. The issue here is whether reformulating glare in terms of illuminances would help lead to a model that can bridge the range from small sources to large sources.

A final basic issue at the heart of all the glare models is that of "additivity". Additivity can have a number of meanings for glare. The BGI and UGR models exhibit adaptation additivity, in that the glare from n identical sources is equal to the glare from just one of the sources if the background luminance is reduced by n. The SR model exhibits an area additivity, in that the source can be arbitrarily split into smaller sources without changing the glare rating. The UGR and CGI models also exhibit area additivity, but only if the position factor for an area source is interpreted as being the weighted integral of the position factor as a function of position over the area. Area additivity appears fundamental, whereas adaptation additivity does not. One can conceive of the splitting of a glare source with an infinitesimally thin line, so that there is no visual difference between the two sources, and a combined source without a line. The infinitesimal line should not change the glare level. This analysis suggests that the use of the centroid of an area for the position factor has to be an approximation, and further suggests that any model that does not exhibit this type of additivity is at best an approximation.

BGI and UGR exhibit one further form of additivity. Both models have the form of a glare rating, GR, where $GR = A \times \log B \times \sum f(L_s,L_b,\omega,PF)$, A and B are constants, and $\sum f(L_s,L_b, \omega,PF)$ is the sum of glare functions for each individual source in terms source and background luminance, solid angle and position factor. This can be recast as the formula $e^{(GR/A)}/B = \sum f(L_s,L_b, \omega,PF)$. This suggests that subjects respond to the logarithm of a glare stimulus, f, with f being additive. This is certainly an attractive property, but it is not fundamental in the sense of area additivity. Indeed, DGI and CGI exhibit this form of additivity in the limit of small sources with limited luminance, but deviate from it for larger, intense sources, because of the effect of the source on adaptation. Spatial inhibition is another mechanism that could cause this type of physiological additivity to fail.

3. Evaluating the Data - General issues

As the summary review above illustrates, the different glare models are based on different approximations, and presumably have different strengths and weaknesses. Unfortunately, it is not possible to make a direct comparison of the different models against data, because the dependent variable, the discomfort glare level, is a subjective measure that is measured differently by each model. The L&G data point to a way around this problem. Luckiesh and Guth measured the glare luminance that subjects evaluated as being at the borderline between comfort and discomfort (BCD). The important point is not the glare level chosen (BCD), but the fact that *all the measurements are at the same level*. The BCD glare luminance is a physical measurement for which averages and standard deviations are meaningful. We do not need to know the a priori relationship between the subjective BCD level and a particular glare rating or probability for a given glare model. Instead, we just require that the glare rating or probability be a constant over the different conditions. To do the actual model comparisons on an equal basis across all models, we invert the models to compute the glare source luminance as a function of the other physical parameters, and a subjective glare rating. We then use a least-squares criterion to determine the value of the subjective glare rating; BGI, DGI, and so on, that produces the least-sum of square errors in the luminances relative to the measured glare luminances as a particular given factor, such as size, was varied.

Luckiesh and Guth reported the individual BCD values over subjects for a reference condition, and the mean values for all remaining conditions. This is very far from ideal, although it is better than just having graphed data, or even just fit lines, which are the two methods of presenting data that Hopkinson used. The underlying factors causing variability in glare measurements should be the same across conditions, so the variability under one condition should be a reasonable guide to the uncertainty for other conditions. Since the BCD values cover a very wide range, we expect that the relative error, rather than the absolute error, should be approximately constant. The calculated logarithmic standard deviation for L&G's reference condition was 0.4. Ten subjects were used for the non-reference conditions, so the uncertainty estimate for the data points is 0.125 (= 0.4/√10). This measurement is based on the uncertainty over subjects, which should be large relative to the uncertainty within subjects. We can use it with a chi-square test to judge the validity of the fit over the population as whole, but it is not ideal. Since we only have the average values, the degrees of freedom for the test is the number of conditions, minus the number of parameters. This is far fewer than we would have had if the raw data was available. This in turn, makes the test much less able to correctly reject an inappropriate model, than when using a within-subjects model with all the data being available. Nonetheless, it does provide a procedure for rejecting models that deviate sufficiently strongly from the average data. Conversely, if the model is correct, the chi-square values can be spuriously small and appear to indicate that the model is too good. This can happen for functions that are linear in their parameters, as the variability in the average values around the average function is solely due to within subject variability, while our estimate of the uncertainty of the points includes between subject variability.

4. Examination of the glare data and the glare models with respect to the L&G data

4.1 Background Luminance

The L&G paper describes the results of several experiments, each of which examines the effect of a single factor on glare. The first factor examined was that of adaptation. L&G measured the BCD of the reference source (a glare source subtending 0.0011 steradians viewed on-axis) at three different background luminance levels. This data provides information on the relationship between the relative importance of the glare and background luminances, but only for the fixed conditions studied.

There are significant differences among the various models on how adaptation is handled. The two small glare source models, BGI and UGR, assume that the glare source is sufficiently small that adaptation is determined solely by the background luminance. The CIE specifies that the background luminance is calculated as the average luminance of the field of view, not counting the glare sources [4]. Hopkinson developed a formula that includes a possible immediate surround term, but the background luminance for BGI formula is again just the average luminance of the surround [2]. Neither reference provides an explicit definition of what luminance and area represents a source, instead of background.

The VCP model is applicable to electric lighting that may cover a significant solid angle. It assumes that the source affects adaptation in the same manner as the background, and simply averages all luminances in the field of view.

DGI and CGI are extensions of the BGI and UGR formulas, respectively, to large area sources. In both of these models the glare source affects adaption, albeit in different ways. For DGI the adaptation depends upon the average luminances of the background plus the source luminance times 0.07 times the square-root of the solid angle. For small sources, adaptation is dominated by the background luminance, and DGI asymptotes to the BGI formula. For large sources and large luminance ratios, DGI asymptotes to a function of the source luminances alone. For CGI, the adaptation is assumed to depend upon illuminance at the eye. When the illuminance from the source is small, CGI asymptotes to a value 3.25 units larger than UGR. When source illuminance exceeds both 500 lux and the background illuminance, the CGI adaptation term asymptotes to 0.004. Since both DGI and CGI distinguish between source and background, it is necessary to know what constitutes a source as distinct from the background. The CIE discusses CGI in the context of electric light sources, where presumably there is a large and fairly unambiguous distinction between source and background, at least in terms of illuminance [6]. DGI was designed to be used for the evaluation of daylight glare. Chauvel et. al. explicitly describes the background as being due to the luminances of the interior, while a window is assumed to represent three sources: the sky, obstructions, and ground [7]. Neither reference provides an explicit definition in terms of absolute levels, or ratios.

The DGP model is based on a study of large window glare sources, but it includes a small source term to handle small bright sources in the window view that uses an extension of the L&G position factor [8]. The large source term of the DGP model has no explicit adaptation term. The small source term uses the vertical illuminance at the eye from all sources as the adaptation term.

The PGSV and SR models are both large glare source models. The SR model was based on an experiment with a single fixed task with a fixed background luminance [9]. The experiment provided no information on the possible variation of the glare sensation with changes in adaptation, and there is no explicit adaptation term in the SR model. Although this looks like a significant limitation of the model, one should note that the DGP mentioned above tends towards to this limit when the source size is large, as long as the luminance ratio between source and background is not huge. The PGSV model was designed to model the glare from windows. Adaptation in the PGSV model is assumed to be equal to the luminance of the background, however the magnitude of the effect of the background luminance, or adaptation, is dependent upon the size of the glare source [10]. For large glare sources the effect of the explicit adaptation term goes to zero, in agreement with the SR model, and to a lesser extent, the DGP model. For the PGSV model, the source appears to be defined as the window, which is presumably markedly brighter than its background.

As there does seem to be agreement that the small and large glare source regimes are not the same, we do not expect that the large glare source models will match the background luminance dependence measured by Luckiesh and Guth. There are, however, differences even between the small glare source models. Figure 1 shows a comparison of the data versus the model predictions.

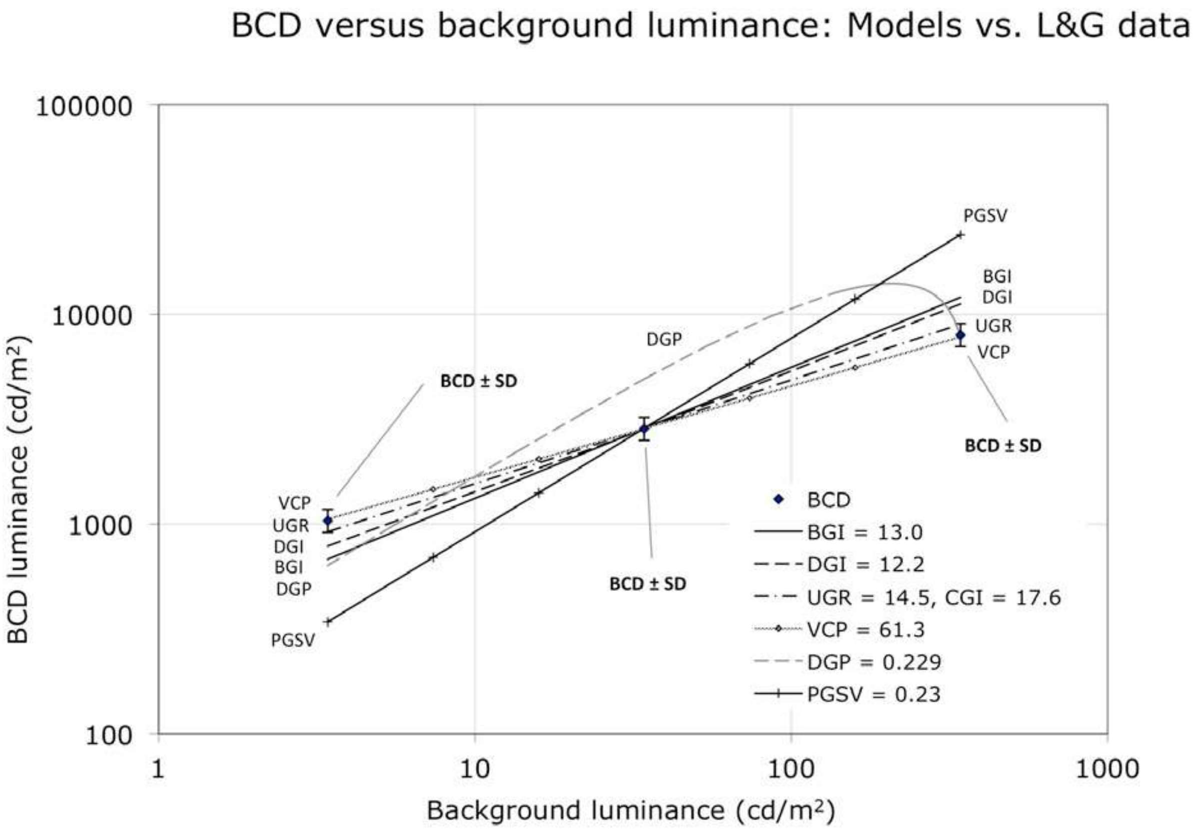


Figure 1. Measured BCD data from Luckiesh & Guth for BCD versus background luminance shown with estimated ± 1 standard error bars. Model fits (lines) are based on finding the glare rating or probability that leads to the least-square error for the BCD prediction. The legend gives the glare rating or probability for each model that gave the best least-squares fit to the BCD data. Only a single line is shown for UGR and CGI, as the separate fit lines cannot be distinguished from each other on this graph. The SR model is not plotted, as it does not have a background term.

The error bars in figure 1 are equal to our estimated standard error (0.125 from above). Table 1 shows the chi-square probability, based on this estimated standard error, for each of the models. Four of the models fit the data. The fit to DGI is close, although this model might have failed if we had had the individual data instead of just averages. The fit to VCP is actually too good, and is probably an example of a spuriously good fit as described earlier.

Table 1. Chi-square probability of fit to background luminance data

Formula	probability
PGSV	0.0%
DGP	0.3%
BGI	2.4%
DGI	12.9%
UGR	69.5%
CGI	78.6%
VCP	99.1%

The BGI and PGSV formulas show significantly higher sensitivity to background luminance than is supported by this data. BGI is a small glare source model, but it is not based on the L&G data set, and the data used in its derivation was collected with a different presentation protocol. The failure of the PGSV model can be attributed to it being a large glare source model. This is also probably the reason for the failure of DGP against this data. The strange shape of the curve is due to the fact that DGP has both a large glare source term, and a small glare source term, and they respond differently to background luminance. Although DGP has a small glare source term, it appears that it is not appropriate to use this model in the absence of a large glare source.

The chi-square probabilities listed in Table 1 imply that the PGSV, DGP, BGI and, to a lesser extent, DGI glare models do not fit the relationship between background luminance and BCD that L&G determined in their experiments.

4.2 Size

Luckiesh and Guth's second experiment examined the effect of size at a fixed background level of 34.3 cd/m^2, and a fixed location at the center of the field of view. Five sizes were measured, ranging from 0.0001 to 0.122 steradians. We again have to be careful to note that our analysis only covers this restricted set of conditions.

All of the glare models analyzed here, except PGSV, assume that the size factor is independent of background luminance, location, duration and the number of glare sources. The PGSV model has a size-background luminance interaction that can be interpreted as a size effect on the sensitivity of glare to background luminance, or a background luminance effect on the sensitivity of glare to size. The PGSV model becomes insensitive to size when the background luminance is 6.5 cd/m^2, and actually goes to an inverse relationship if the background luminance is lower. It has the same size sensitivity as BGI, DGI, UGR and CGI ($L^2\omega$) at a luminance of 685 cd/m^2, and reaches parity with source luminance ($L\omega$) at 73,000 cd/m^2. Interestingly, Hankins and Waters reference two studies that appear to indicate that, for very small off-axis sources, glare depends upon intensity, not area, and the CIE has proposed a modification to UGR that becomes independent of source size for sources less than 0.005 steradian [11,12].

Several of the models have explicit limitations on the size range in which they are valid, but this is generally due to their treatment of the influence of the source luminance on adaptation, and is not directly related to background luminance, location or stimulus duration. A more direct clue that the size effect depends upon size itself is that small glare source models, and models based on extending the small glare models (DGI and CGI), have a fundamentally different size dependence than the large glare source models, DGP and SR. With the exception of VCP, the small glare source models have a size-glare luminance dependence of the form $L^2\omega$ = constant. The large glare source models have the form L x CF = E, where CF is the configuration factor, and E is the illuminance at the eye.

Figure 2 shows a comparison of the five data points measured by Luckiesh and Guth, and all but the SR model. The SR model does not fit this small glare source data even remotely correctly, and the lines do not fit on the graph. The integrated position factor for the various size targets ranges from 1.007 to 1.18. Using the line of sight value position factor of 1, makes a slight, but perceptible difference in the graph, but does not change any of the statistical conclusions regarding the degree of fit.

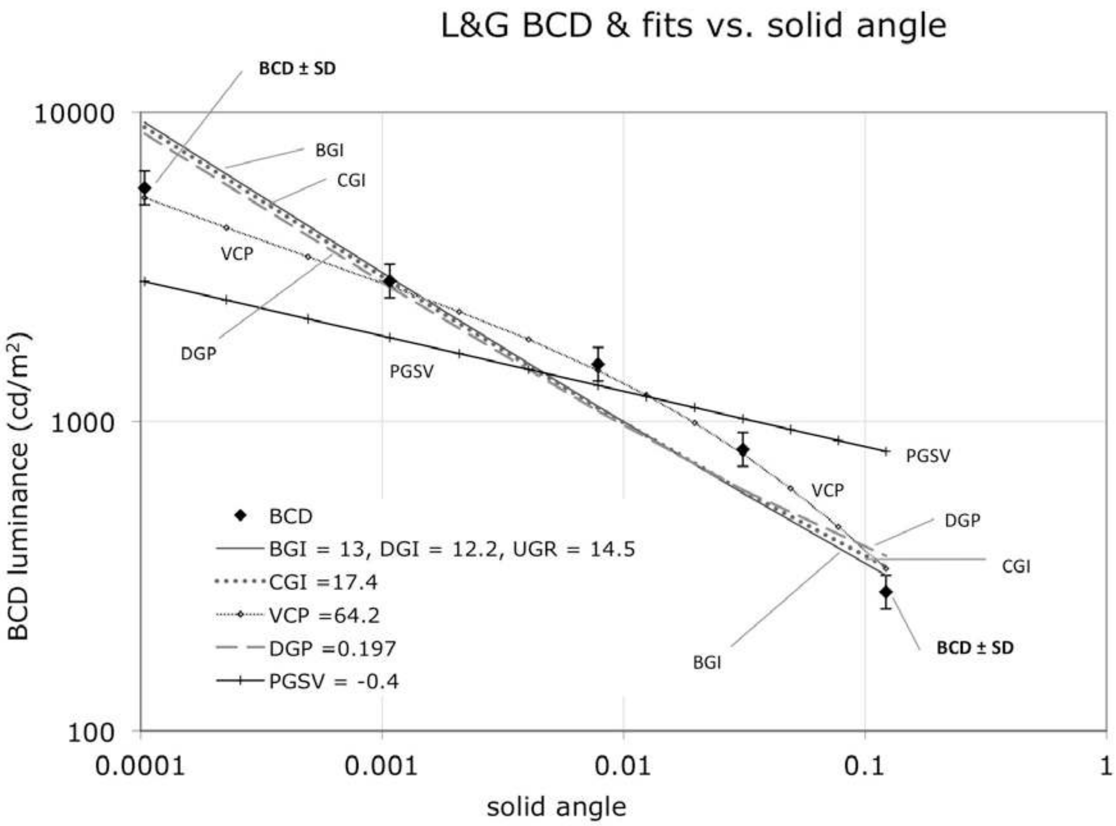


Figure 2. Measured BCD data from Luckiesh & Guth for BCD versus glare source size shown with estimated ± 1 standard error bars. Model fits (lines) are based on finding the glare rating or probability that leads to the least-square error for the BCD prediction. The SR model is not plotted, as the lines exceed the ordinates of the graph. The legend gives the glare rating or probability for each model that gave the best least-squares fit to the BCD data. Only a single line is shown for BGI, UGR and CGI, as they are otherwise almost impossible to distinquish apart in this figure.

Before we discuss the fit curves themselves, we note that the glare ratings or probabilities that led to the best fits for the size data are very close to the values that led to the best fits for the background luminance data (see legends). This indicates that the data is basically self-consistent.

The most obvious feature of the plot itself is the high degree of fit of the VCP model, and the low degree of fit for all the remaining models. The VCP model was based on the L&G data, and so it should be no surprise that it fits it. It should be similarly no surprise how poorly PGSV fits this data, as it is mostly considerably below the size range within which the model is supposed to be valid. What is a bit surprising is how poorly the other models fit the data. The chi-square test rejects these models at a probability of about 0.0012% or below. There was no significant difference in the degree of fit between the small glare source models, BGI and UGR, and their larger source extensions: DGI and CGI. The extensions to these models only affects the way in which they handle the impact of glare source luminance on adaptation.

The data in the figure does not look like it comes from a fixed exponent. As a check, the data was fit against the basic $L^x\omega$ form. The best fit gave an exponent of 2.4. The fit was rejected at the 2.2% probability level, which, while better than the p-values for x=2, still suggests that the

L&G data do not support the assumption inherent in most of the models that there is a fixed size-luminance dependence over the size range examined. The failure of a fixed exponent is consistent with the studies mentioned earlier, that indicate that the standard small glare functional form is not appropriate for very small or very large sources. The fixed luminance exponent of two, used in the small glare source models, appears to fit the data only in the limited range of around 0.005 to 0.05 steradians.

Currently, the only model that appears to fit the data well is the strictly empirical VCP model. It is tempting to ask whether the model extrapolates well to smaller or larger sizes. For large sources the VCP model extrapolates to an exponent near 1, which is similar to what is used in the DGP and SR models. This suggests that extrapolation to larger sizes should be approximately correct. For small sources the empirical nature of the VCP model becomes evident, as the model is not defined below 2.92×10^{-7} steradians. It reaches a minimum sensitivity to size at about 0.0003 steradians, and then becomes increasingly sensitive as size is reduced. This is in contradiction with the findings that glare is independent of source size for small sources (note however, that glare luminance must be inversely related to size once the glare source drops below the resolution limit of the eye$\approx 7 \times 10^{-8}$ steradians) [11,12].

4.3 Position Factor

The position factor, PF, is the ratio of luminance at an arbitrary position to the luminance on the line of sight that causes the same glare sensation. The experimental conditions were limited in a manner similar to that of their other experiments: fixed uniform background luminance of 34.3 cd/m^2, and a fixed size of 0.0011 steradians. One difference in this experiment relative to the two experiments above, is that subjects alternately saw two glare sources; a reference glare source at the line of sight set to the group average BCD, and a test glare source set at a location off the line of sight. The subjects adjusted the luminance of the off-axis glare source until they felt it matched the glare sensation of the reference source.

The values of the position factor, and even its existence as an independent multiplicative factor, is critically dependent upon all other factors being independent of position. For example, given a glare source with L = 2 840 cd/m^2, and a size of 0.0011 steradians, the formulas $(L/PF)^2\omega$ and $L^x\omega$ have the same value when $x = 1.9$ and $PF \approx 1.48$. This ambiguity cannot be addressed with the L&G data.

As a practical matter, there is not a lot of analysis that can be done with just the L&G data, as the small glare source models all use the same empirical fit to this data. However, recently, Kim et. al. reported on an experiment that is very similar to the position factor experiment performed by Luckiesh and Guth [13]. Kim graciously provided us with the data for the individual subjects for this experiment. Kim's data provides a direct check of the generalizability of the L&G data, and also provides independent information on the variability of the data as a function of location. This is particularly useful, as, Luckiesh and Guth reported that the variability of the BCDs increase as a function of the angle from the line of sight, but provide only a vague, and somewhat inconsistent, description of the degree of increase. Kim's data shows about a 50% rise in the logarithmic standard deviation of the position factors from about 10° to 50° (0.045 to 0.07), with a constant variability at higher angles. This is not a large change, and suggests that the fitted uncertainty to the L&G data, 0.062 log units, is probably a reasonable estimate for the uncertainty over the data set.

Before proceeding, we need to note that there are some minor differences from the data we analyzed and the data reported in the published paper. The data we have from Kim includes four other subjects whose values were not reported in the journal article, and does not include two subjects who were. We threw out the data for one subject (identified as # 20 in the spreadsheet, & #24 in the journal article) who had position factor values that were significantly lower than the other subjects (t-test with P = 5%) for 30 of the 33 conditions where there was data. This data is sufficiently far from normal that it is likely to skew any analysis that includes it, and it was therefore excluded, leaving 28 subjects. Two other extreme data points were thrown out, one of which was four times any of the values near it, and the other over twenty times the values nearest it (subjects # 6 and #8 (spreadsheet) at polar angles 60 and 10 degrees, and azimuths of -90 and -45).

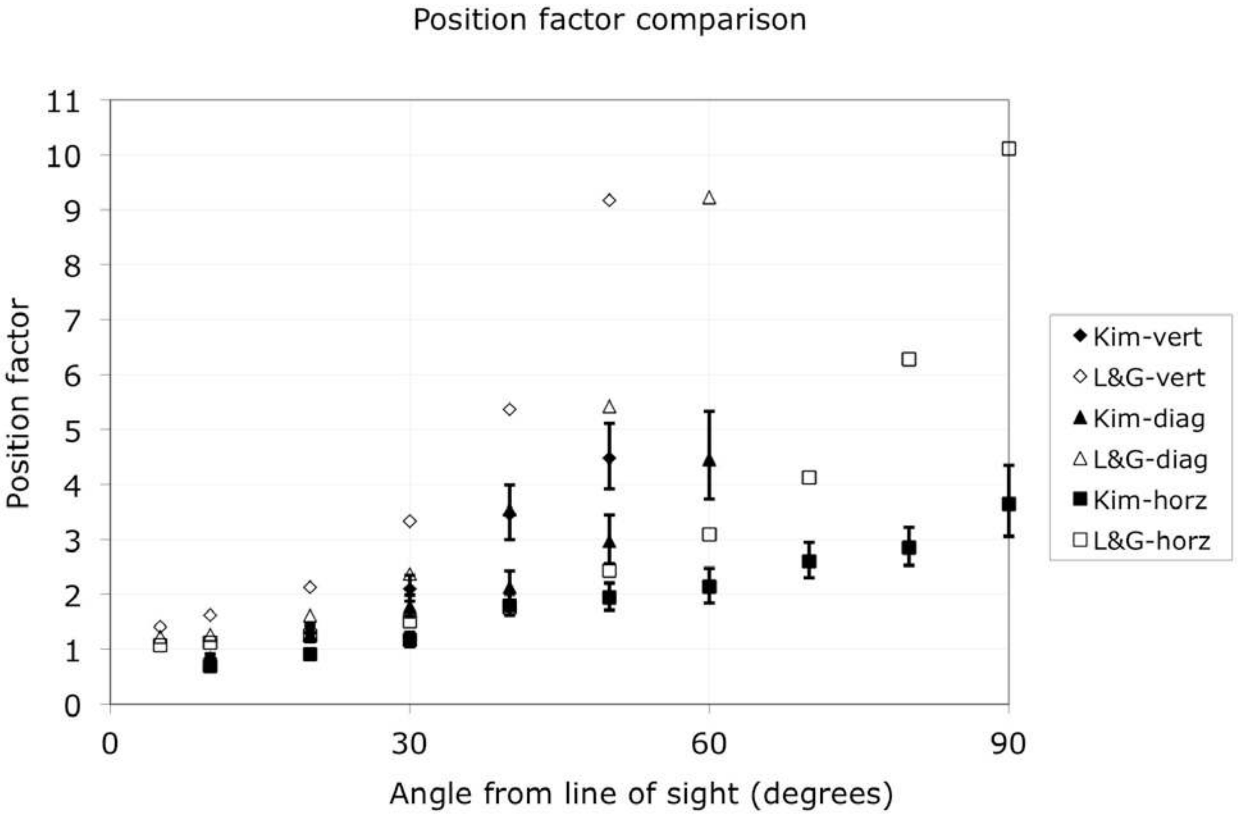


Figure 3. Position factors (BCD(θ,ϕ)/BCD(0,0)) from L&G experiment (open symbols) and Kim et. al. experiment (filled symbols). Kim et. al. data shown with ±1 standard error bars.

Luckiesh and Guth only measured position factors for azimuth angles above the horizontal (0° to 90°), while Kim et. al. measured them over the whole visual field. Figure 3 shows a plot of the L&G data, and the data from Kim, with standard errors, for azimuth angles above the horizontal. The position factors measured by Kim et. al. are lower than those measured by Luckiesh and Guth for 19 of the 20 conditions that are common to both experiments, and they are further statistically significantly different (two-sided t-test at P = 5%) for 18 of those conditions. The protocol for Kim's experiment is essentially the same as for L&G. The only obviously significant difference is that Kim's subjects were all Korean students, while L&G merely note that the 10 subjects selected for the position factor and other studies were selected on the basis of their availability for the extended set of measurements - which suggests that they were company employees of an American company. Hopkinson has noted significant differences between experienced and naive subjects, and this may well be another example of such an effect [2]. Alternatively, there may be a cultural difference.

Possible evidence for there being a significant difference between the two subject pools is that while the within subject variability for the Kim et. al. data was 0.366, the between subject variability is 0.777, or almost twice the 0.40 value for the L&G subject pool. The 0.40 value is outside the 95% confidence limits for this quantity in the Kim et. al. experiment (0.61 to 1.06) [14]. In addition, the position factors for individual subjects were highly correlated across the different view directions, which means that different subject samples could have significantly different position factor patterns. The results suggest that this is an area that needs further data even over the existing conditions, and it also indicates that any analysis of L&G results that involve off-axis viewing has to be done with the L&G position factors.

4.4 Additivity

Luckiesh and Guth (1949) describe four experiments which examined the issue of the additivity of glare sources. The first two experiments compared a single source (size = 0.0011 steradian) located on the line of sight, to two half size sources (size = 0.00055 steradian) located symmetrically about the line of sight. The third and fourth experiments compared a reference glare source to one, two, three, or four glare sources of the same size (size = 0.0011 steradian) located at increasing angles from the line of sight. The first experiment is a direct test of additivity, as the test and reference sources have the same total area (to within 0.3%), and the BCD luminance of full sized sources at the test locations was measured in the previous position factor experiment. In the second experiment the BCD level for a full size source at either of the test locations has to be estimated from the position factor formula, as it was not measured in this experiment or the previous position factor experiment. In the third experiment the BCD for the reference size single glare source at the test locations is again known, but in the test case there are two or more sources each with the same area of the reference size sources, so the experiment mixes size and additivity. Finally the fourth experiments mixes the issues of additivity, size, and position factor together. Note that in all these experiments the multiple sources have a significant physical separation. This is a stronger test of additivity than arises from the mathematical concern that two sources separated by a vanishingly small distance should be additive.

The analysis of the additivity experiment differs from that of the size and background experiments in that the experiment uses a comparison of two sources, so the analysis requires no unknowns. In the analysis, the glare level is computed for the reference condition, and then used to compute the expected BCD luminance for the test condition.

Because all four experiments compare a test source to a reference source, the between subject variability for the reference source should cancel from the measurement, leaving within subject variabilities, and in experiments two and four, the uncertainty of the fitted position factor estimate (from the previous section). We do not have any data specific to the within-subject variability for the L&G data, but we do have the value, 0.366, for the Kim et. al. data. This value is almost as large as the combined within and between subject variability for the one condition we have for the L&G data, and thus it is very likely to represent an upper bound for the within subject variability for the L&G subjects. This leads to a standard error estimate of 0.116 (0.366/√10) for experiments one and three, and 0.131 (√[$0.116^2 + 0.062^2$], where 0.062 is the standard deviation of the position factor fit) for experiments two and four.

Figure 4, below, shows the L&G data for experiment one; the computed values assuming additivity, and the computed values from the listed non-additive glare formulas. The chi-square test is fairly weak, as the uncertainty estimate is an upper bound, and it rejects only the fit to the VCP model (P= 0.6%). Note, however, that the non-additive models show a bias error. The data points scatter to both sides of the additive formulas, but are consistently larger than the curves for the non-additive models. With only four data points (P = 6.25%) this is not enough to reject the non-additive models, but it does suggest that there may be a problem.

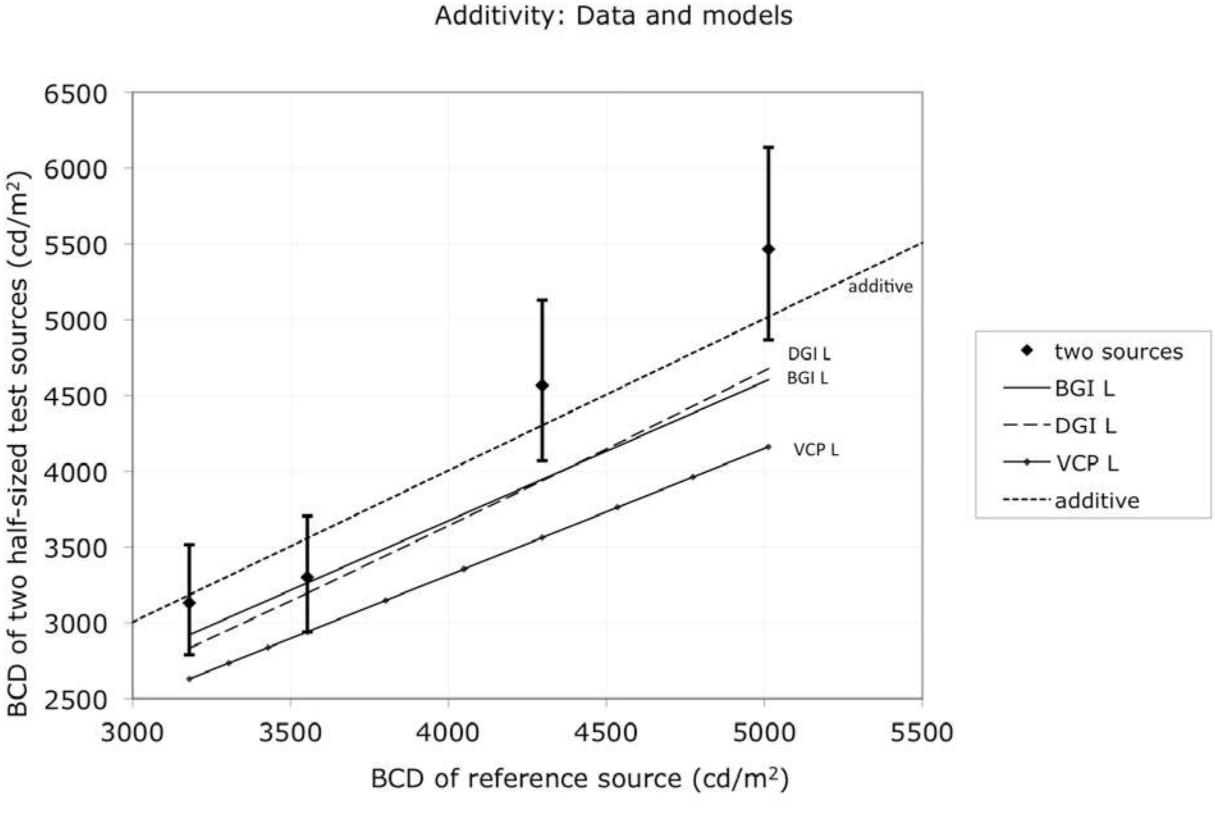


Figure 4. Data from L&G's first additivity experiment. The x-axis gives the BCD of the reference 0.0011 steradian source, located at one of four angles from the line of sight. The y-axis gives the measured or fitted BCD of the two half-size (to within 0.3%) sources located symmetrically about the line of sight as the reference source. The error bars are an estimated upper limit for the standard errors calculated from the within subject standard deviation from Kim et. al. (0.366) divided by the square-root of the number of subjects in the L&G experiment (see text). The additive models have slight differences from each other due to non-additivity in the adaptation term (CGI), differences in how they handle the slight size discrepancy (UGR versus SR), or combinations of these effects (DGP). The curve shown is the UGR curve, but the differences versus the other models are difficult to see on the scale of the figure.

The results from experiment two are similar to those of experiment one, despite this experiment being a test of both additivity and position. VCP is rejected even more roundly (P = 0.3%), and DGI is rejected as well (P = 2.4%), but none of the remaining models are rejected. However, the non-additive models are again biased. Since the additive models do fit this data, it seems reasonable to assume that the position factor is not a critical issue in judging the fits. This suggests that we can combine the results of the two experiments. The combined results do reject the non-additive models, as the probability of all eight points being biased is only 0.4% (binomial test). Despite the lack of individual error estimates for the L&G data, the data is strongly consistent with additivity, at least for these small sources, and appears to be inconsistent with the degree of non-additivity found in the BGI, DGI and VCP formulas.

In L&G's third additivity experiment, the reference glare source is located along the line of sight, and the test glare sources are presented in an array of one to four sources starting 10 degrees above the reference source and extending to 40 degrees above the source when there are four sources. The reference glare source is used to compute the reference glare level for each of the glare formulas, and the measured position factors are used with the formulas to compute the predicted luminance levels for the test sources. The SR model does not included position factors, and therefore does not properly correct the small glare source BCDs for location. Figure 5, below, shows that the SR model is not able to fit this data at all. The remaining models do include position factors, and therefore this experiment becomes just a test of size and additivity. From the analyses we have performed above we know that only the VCP model appears to properly handle size, at least over a large range of sizes, but that BGI, DGI and VCP do not exhibit additivity. Given the natural perversity of expectations it should therefore come as no surprise that only BGI and DGI appear to fit the data in experiment three.

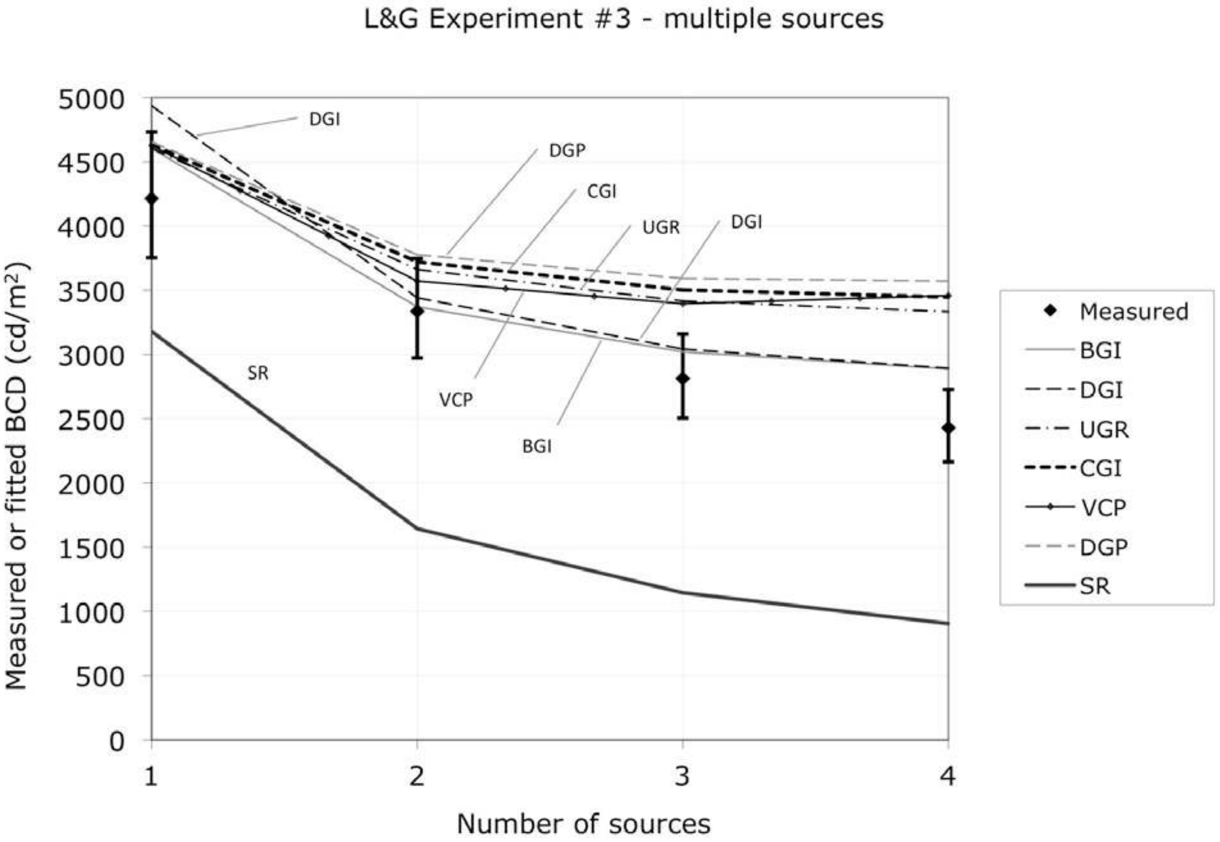


Figure 5. BCD for multiple glare sources: data from L&G versus model fits. The error bars are again estimated standard errors from Kim et. al. (see figure 4 caption). The BCD values are the luminances of glare sources from 10 to 40 degrees above the line of sight that match a single reference source located along the line of sight. The SR model does not use position factors, and fits very badly. The DGP model also has a term that does not include a position factor, and fits moderately badly. The remaining models have the position factor built in, so deviations from the data are due to chance, or a failure of the size/additivity relationship in the model.

In this L&G experiment, additivity effects only apply to the situations where there are two or more glare sources. This leaves only three data points. BGI and DGI appear consistent with the data (P = 45% and 42% respectively). The VCP model is soundly rejected by the chi-square test (0.6%). CGI, UGR and DGP are also significantly rejected by the chi-square test (≤ 1.2%). It appears that the size failure in the models is more significant for these off-axis points than would have been thought, and is evidently in the opposite direction of the additivity failures in the BGI and DGI models.

The fourth experiment actually varies three factors: multiplicity of targets, size, and location. With all three factors varying it appears that the errors are partially canceling, and only the SR model is rejected by the chi-square test. Unlike in the third experiment where all but the SR model overestimate the BCD luminance, the BGI and DGI models now underestimate the BCD luminance, and UGR and VCP only overestimate two of the points. If we combine the results of the third and fourth experiments together we find almost the exact opposite of the results from experiments one and two. BGI and DGI do a reasonable job of fitting the data in the latter two experiments (P ≥ 47%), while VCP and the additive models all fail (P≤ 2.4%).

There are a number of possibilities for why additivity holds in experiments one and two, but fails in experiments three and four, but the most obvious concern is that the source-luminance versus size relationship used in the models is not adequate. In the limit that source luminance does not affect adaptation, and the position factors are approximately equal, all of the small source models except for VCP assume a relationship of the form $L^2 \times \omega$ for the reference source = $L_1^2 \times \Sigma\omega$ for the test sources. The exponent has no effect on additivity for two half size sources. For multiple full size sources $L_1 = L/N^{(1/2)}$. If the luminance exponent is incorrect, then the formula will not predict the correct luminance for multiple sources, even if additivity is correct. The calculation is more complex if the position factor is not constant, but the basic message that the exponent affects the expected luminance when there are multiple sources remains the same. If the exponent is the problem, it affects both the size and the additivity data.

If the problem is just a matter of the wrong exponent, then the inner term of the glare model has to have the form $L^n \times \omega$ (or $L^n \times CF$, where CF is the configuration factor), with n to be determined in order for it to fit experiment three (and to a lesser extent, experiment four). This type of model does actually work for these experimental conditions. With n fit to the value 1.38, it returns a p-value of about 80% with the chi-square test for experiment three. The fit to experiment four is considerably less good, but is still accepted at the $P = 40\%$ level. Unfortunately, the best fitting exponent in the size experiment over the three smallest sizes (which are centered on the sizes used in the additivity experiment) is 3.3 ± 0.05, which range does not overlap the value found in experiments three and four. In short, although this is a very appealing explanation for the additivity failure it appears that it is inconsistent with the data on the relationship between size and luminance.

There are other possible explanations for the observed additivity failure. Hopkinson has observed that in the limit of many sources the ambiguity between source and background becomes significant, but this seems remote from the situation above [2]. It is possible that additivity of the stimulus doesn't hold because of inhibition or feedback effects. It is also possible that the effect of size on glare depends upon location. Figure 6 shows how the size and additivity data fails to fit the CIE CGI model (chi-square probability about 0.0001%). Figure 7 shows a fit of the L&G data to a model based on configuration factors, and the assumption that L&G's measured position factors are solely due to a change in the luminance exponent. The model uses the position factor fit to instead calculate an equivalent exponent from a guess to the asymptotic exponent along the line of sight. The fitted exponents vary only by a small amount over the size range studied by Luckiesh & Guth, but tend towards one for large glare sources. The expression should therefore be close to a vertical illuminance model for large or off-axis glare sources, which would make it compatible with the large glare source models. This model returns an acceptable chi-square probability of 13%, but we have no data that confirms that the size effect actually varies with position, so the model, while suggestive, is purely speculative.

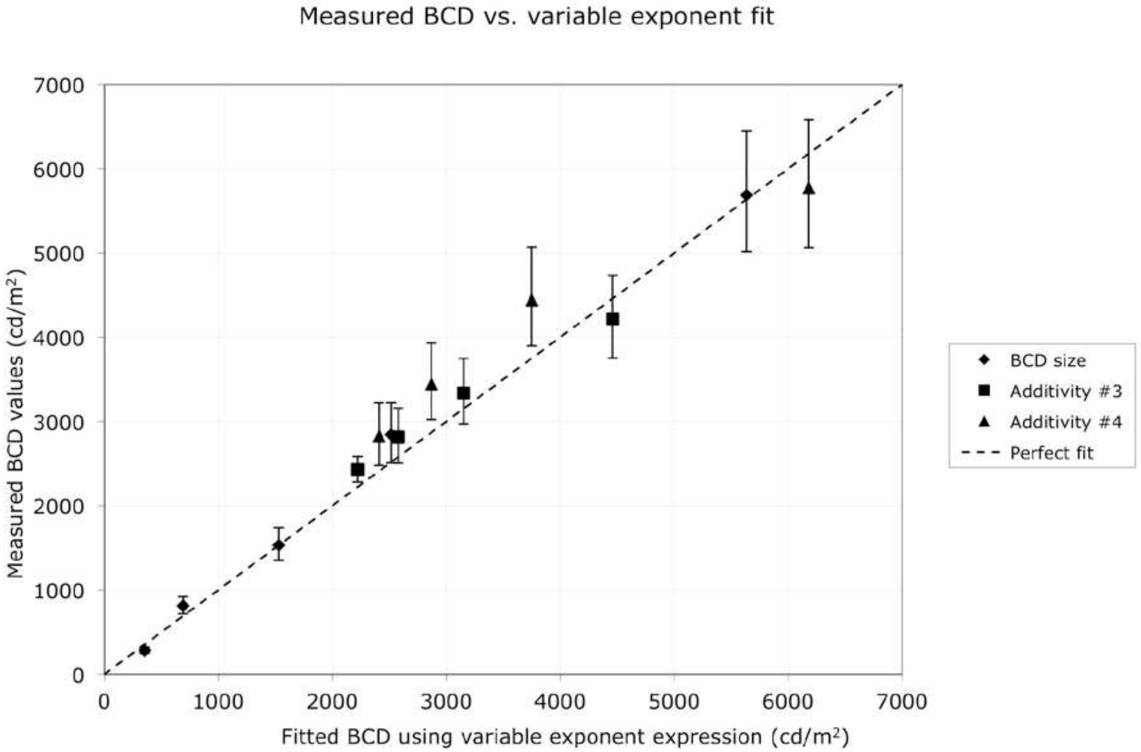


Figure 6. Measured BCD values with estimated standard error bars for the size experiment and the last two additivity experiments, plotted against values predicted using the CGI glare model.

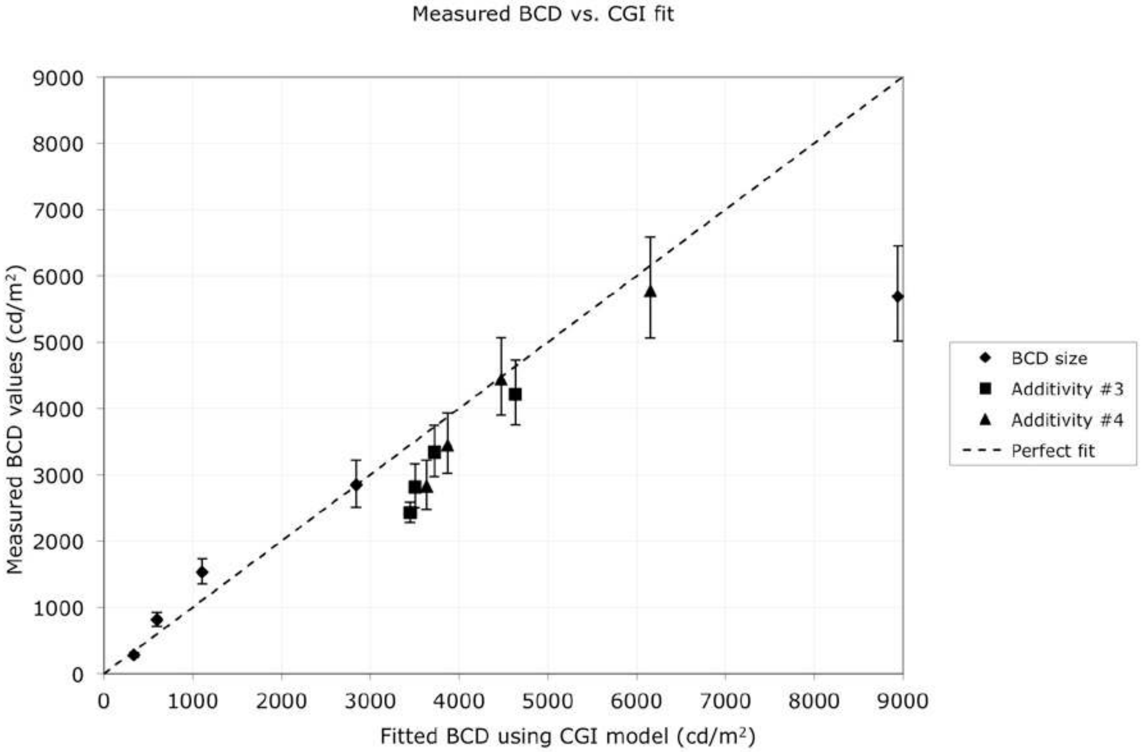


Figure 7. Measured BCD values with estimated standard error bars for the size experiment and the last two additivity experiments, plotted against values predicted using a variable exponent expression (see text).

5. Discussion and Conclusion

There are three parts to this paper. The first part is a summary of some of the basic assumptions and limitations of existing models. Some of this material, such as Hopkinson's work on the immediate versus distant surround, was known 60 years ago, but it is not addressed in the existing models [2]. Similarly, more recent work by Waters et. al. on non-uniform sources has not yet been reduced to practice [3]. Finally, some of it, such as the questions regarding interactions between the various factors, seems to a large extent to have slipped past people's notice.

The second part addresses the difficulty in actually comparing different models against each other. It is difficult to compare subjective measures against each other. This difficulty can be avoided, if an experiment measures multiple points at the same level of glare, by inverting the subjective measure to give an objective measure, such as source luminance, which is common to all glare models. The glare rating for each model is treated as an unknown, which can be fit to give the least-square error in the source luminances over the set of points in the experiment. This procedure can be used for both past and future experiments, and therefore may be the most significant result reported.

The third part of the paper compared the position factor studies of Kim and L&G, and applied the procedure developed in part two to the data presented by L&G in their 1949 paper [13,1]. The L&G data was unfortunately primarily presented as mean values, without any error bars, or individual values, and unfortunately many papers still only present mean values, or summary conclusions, so that the researcher interested in further understanding the work has to try to get the data from the authors. I wish to interject a plea here. We are now in the computer and internet age, and there is really no excuse not to take advantage of it. Data should be posted on the internet. Ideally this should be done at central repositories, so that it doesn't get lost, and is freely available.

The L&G analysis came to the rather unsurprising conclusion that large glare source models do not fit the data for small glare sources. Other results were less mundane. Studies have shown that the glare source exponent is probably not valid for very small or large sources, but our analysis showed that it is even an approximation over the range of sizes in the L&G study [12]. The L&G and Kim data for position factors do not agree, which is rather alarming. Finally we also looked at the question of area additivity. The L&G data are partially consistent with additivity. Additivity is an explicit feature of the CIE glare models, UGR and CGI, and its lack is considered a flaw in the older BGI and DGI models. It was therefore a surprise to discover that BGI and DGI fit the L&G data that mixed both additivity and size, while the UGR and CGI models did not. It appears that the problem is that none of the models fit size properly, and that for this set of data, and possibly for many real situations, the size error for BGI and DGI partially cancels their additivity errors, with the result that these models provide a better fit than the CIE models. This was not expected. On a speculative note, it was found that the size and additivity data could be fit by replacing the position factor with a source exponent that varies with the angle from the line of sight. There is no data explicitly supporting an inhibition term, or this position factor - luminance source exponent trade-off, so this may be an avenue that is ripe for further study.

Addenda: Glare formulas

BGI (British Glare Index)[8]:

$$BGI = 10log(0.4784\sum_{i=1}^{n} L_{si}^{1.6}\omega_{si}^{0.8} / L_b P_i^{1.6})$$

$\omega < 0.027$

DGI (Discomfort Glare Index)[8]:

$$DGI = 10log(0.4784\sum_{i=1}^{n} L_{si}^{1.6}\omega_{si}^{0.8} / (L_b + .07\omega^{0.5} L_{win} P_i^{1.6})$$

DGP (Discomfort Glare Probability)[8]:

$$DGP = 5.87\times10^{-5}E_v + 0.0918log(1+\sum_i \frac{L_{s,i}^2\omega_{s,i}}{E_v^{1.87}P_i^2}) + 0.16$$

UGR (Unified Glare Rating)[8]:

$$UGR = 8log(0.25 / L_b)\sum_{i=1}^{n} L_{si}^2\omega_{si} / P_i^2$$

$0.0003 < \omega < 0.1$

CGI (Cornell Glare Index)[8]:

$$CGI = 8log((2(1+E_d / 500)/(E_d + E_i))\sum_{i=1}^{n} L_{si}^2\omega_{si} / P_i^2$$

PGSV (Predicted Glare Sensation Vote)[10}:

$$PGSV = 3.2logL_s - 0.64log\omega + (0.79log\omega - 0.61)logL_b - 8.2,$$

$0.021 < \omega$

SR (Subjective Rating)[9]:

$$SR = 0.1909E_v^{0.31}$$

$0.03 < \omega$

VCP (Visual Comfort Probability)[5]:

$$VCP = 100 / \sqrt{2\pi}\int_{-\infty}^{6.374-1.3227ln(DGR)} e^{-t^2/2}dt$$

$$DGR = [\sum_{i=1}^{n} M_i]^{n-0.0914}$$

$$M = 0.50L_{si}Q / P_i F^{0.44}$$

$Q = 20.4\omega_{si} + 1.52\omega_{si}^{0.2} - 0.075$ (solid angle term)

$F = [L_w\omega_w + Lf\omega_f + L_c\omega_c + \sum_{i=1}^{n} L_{si}\omega_{si}]/5$ Average luminance assuming that visual field of view is 5 steradians.

Definitions:
E_v: Vertical luminance at the eye

L_{si}: Luminance of source i

L_b: Background luminance

L_{win}: Average luminance of the window (average of sources)

L_w: Average luminance of walls

L_f: Average luminance of floor

L_c: Average luminance of ceiling

P_i: Position factor of source i

ω: Total solid angle of sources

ω_{si}: Solid angle of source i.

ω_w: Solid angle of walls

ω_f: Solid angle of floor

ω_c: Solid angle of ceiling

Addenda: Position factor formula

$lnP = (A + B\tau + Ce^{D\tau})\sigma + (E + F\tau + G\tau^2)\sigma^2$
where
τ = the "angle from vertical of plane containing source and line of sight" in degrees, $\tau \leq 90°$.
σ = the angle between the line of sight and the line from the observer to the source in degrees,

A = 0.0352
B = -0.00031889
C = -0.00122
D =-2/9
E = 2.1×10^{-4}

$F = 2.6667 \times 10^{-6}$
$G = -2.963 \times 10^{-8}$

In the DGP model P is extended to below 90° using Einhorn's analytical extension [8].

Acknowledgments

This work was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Office of Building Technology, State and Community Programs, Office of Building Research and Standards of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Francis Rubinstein made helped make the figures and text readable.

References

[1] Luckiesh M, Guth SK. Brightnesses in Visual Field at Borderline Between Comfort and Discomfort (BCD). Illuminating Engineering Nov 1949; 44: 650-670.

[2] Hopkinson RG. Architectural Physics: Lighting.
London: HMSO, 1963: 99.

[3] Waters CE, Mistrick RG, Bernecker CA. Discomfort glare from sources of non-uniform luminance. J. Illum Eng. Society 1995; 24(2): 73-85.

[4] CIE. Discomfort Glare in Interior Lighting. Technical Report CIE 117 1995.

[5] IESNA. IES Lighting Handbook, 1984, Reference Volume. 345 East 47th Street, New York, N.Y, 10017, 1984: 9-47

[6] CIE. Discomfort Glare in the Interior Working Environment. CIE Publication 55 (TC-3.4), 1983.

[7] Chauvel P, Collins JB, Dogniaux R, Longmore J. Glare from windows: current views of the problem. Lighting Research & Technology 1982; 14(1): 31-46.

[8] Wienold J, Christoffersen J. Evaluation methods and development of a new glare prediction model for daylight environments with the use of CCD cameras. Energy and Buildings 2006; 38(7): 743-757.

[9] Osterhaus WKE. Brightness as a Simple Indicator for Discomfort Glare from Large Area Glare Sources: Proceedings of the first CIE Symposium on Lighting Quality. Vienna, Austria, 1998: 113-120.

[10] Iwata T, Tokura M. Examination of the limitations of predicted glare sensation vote (PGSV) as a glare index for a large source: Towards a comprehensive development of discomfort glare evaluation. Lighting Research & Technology 1998; 30(2): 81-88.

[11] CIE. Glare from small, large, and complex sources. Vienna, TC 3-01 Report, 2002.

[12] Eble-Hankins ML, Waters CE. VCP and UGR Glare Evaluation Systems: A Look Back and A Way Forward. Leukos 2004; 1(2): 7-38.

[13] Kim W, Han H, Kim JT. The position index of a glare source at the borderline between comfort and discomfort (BCD) in the whole visual field. Building and Environment 2009; 44: 1017-1023.

[14] Zar JH. Biostatistical Analysis. Confidence Limits for the Population Variance. 2nd Edition, Prentice Hall, 1984: 115