VISUAL MAGNITUDES AND THE "AVERAGE OBSERVER": THE SS CYGNI FIELD EXPERIMENT

Richard H. Stanton
217 Starlane Drive
La Canada, CA 91011

Presented at the 87th Annual Meeting of the AAVSO, October 31, 1998; revised June 13, 1999

Abstract

Results of the *V-Magnitude Experiment – 1998*, designed to characterize the vision of the "average observer," are discussed. Over 650 individual observations were analyzed, submitted by 63 observers worldwide. These data lead to summary conclusions in two important areas: (1) effect of star color on observed star magnitude, and (2) the range of observation scatter expected for observers using perfect comparison star sequences. Quantitative color response measurements for 48 individuals, when compared to the "average observer," show variations generally consistent with measured levels of observation scatter. Based on these data, a color coefficient value of $b = 0.21$ is recommended for translating Johnson magnitudes (UBV) into visual magnitudes (m_v).

1. Background

It has long been known that Johnson V-magnitudes measured by various instruments (photomultipliers or CCDs) often do not accurately represent what an "average" visual observer sees (visual magnitude m_v). Even an accurately measured photoelectric sequence will typically have one or two stars that "just don't look right" to the observer. This problem has become increasingly important in recent years. Visual observations are in more demand than ever before (Mattei 1998). Significant strides have been made in the production of charts that more accurately portray what is seen in the telescope (Scovil and Leitner 1991), and more observers than ever are becoming active due to the electronic distribution of charts (AAVSO site: http://www.aavso.org). But a modern visual magnitude system, accurately representing what the "average observer" sees, has not yet been established. This lack affects not only comparisons of variable star observations, but also hampers the establishment of comparison star sequences vitally needed for all charts. To address this problem, a system is proposed based on the observations of a large number of observers. Their efforts will enable the full utilization of both the increasingly available Johnson UBV magnitudes and the two-color photometry in the *Hipparcos and Tycho Catalogues* (ESA 1997).

Defining this system by characterizing what the "average observer" sees was the primary motivation for the *V-Magnitude Experiment – 1998*. This experiment, developed jointly by the AAVSO Chart Committee and the author, followed a smaller effort in the preceding year. Participants were asked to observe as many of 23 "unknowns" in the SS Cyg field as they could reliably see, estimating their brightness relative to a sequence of "known" comparison stars on the same chart. Observations were recorded in "steps" relative to the comparison stars, much as standard variable star measurements are made. The "steps" became the raw input data for the analysis described below. Observers were asked to use their normal observing techniques so that the final results would apply as closely as possible to modern techniques and equipment.

The final output is a visual magnitude (m_v) system defined in terms of magnitudes and colors measured in the Johnson UBV system, and having the following characteristics:

1. Matching the color response of the average modern observer;

2. "Pogson," that is, a one-magnitude change represents exactly a $100^{1/5}$ change in brightness;

3. m_v = V when B–V = 0 (A0V star).

Once this transformation is defined in terms of UBV magnitudes (more precisely, V and B–V), magnitudes and colors measured in UBV, *or any other calibrated system*, can be transformed into the m_v scale and directly compared to visual estimates.

2. Formulation

The general problem of deriving a transformation for all stars can be greatly simplified by restricting the color range to –0.3 < (B–V) < 1.9. While many interesting variables fall outside of this range (e.g., Miras), it encompasses the vast majority of stars. Transforming between systems typically becomes non-linear and multi-valued for very red stars (B–V > 2.0). The *Hipparcos and Tycho Catalogues* (ESA 1997) discuss many of these complexities, such as differences between dwarf and giant branches. Having a system that gives star magnitudes consistent with the average eye is particularly valuable for establishing comparison star sequences, even if that system gives less accurate values for very red variables.

Conversion formulas that are linear in star color can be written in the following form:

$$m_v = aV + b(B - V) + c, \quad [-0.3 < (B - V) < 1.9] \qquad (1)$$

where B and V are magnitudes in the Johnson UBV system, and *a*, *b*, and *c* are constants. The adequacy of a linear formula will be seen in the following sections. If instead one uses magnitudes from the Tycho catalogue, they can be easily transformed *in this limited color range*, according to the linear relations

$$V = V_T - 0.090\,(B - V)_T \qquad (2)$$

and

$$B - V = 0.850\,(B - V)_T, \qquad (3)$$

where V_T and $(B–V)_T$ are the magnitude and color, respectively, measured in the Tycho system (ESA 1997). Measurements made by CCDs, other silicon detectors, or photomultipliers can similarly be transformed into V and B prior to the application of equation (1).

Values for *a* and *c* in equation (1) are easily derived. The requirement that the m_v scale follow the same Pogson relationship with intensity as V implies that *a* must be unity. Note that non-Pogson deviations from linearity were a significant problem in visual magnitude scales developed prior to the perfection of modern linear photometers (Zissell 1998). Defining the zero point in the traditional way, that all systems give the same magnitude for stars of zero color (spectral class A0V), forces *c* to be zero.* Thus we can write

$$m_v = V + b\,(B - V). \qquad (4)$$

Similarly, substituting (2) and (3) into (4)

$$m_v = V_T + b'\,(B - V)_T, \qquad (5)$$

* In an earlier article by the author a value of c = –0.15 was suggested (Stanton 1981). This was done as a stopgap measure to minimize the average offset between magnitudes converted to m_v and existing sequences.

where

$$b' = 0.850\, b - 0.090\ . \qquad (6)$$

All that remains is to find a *b* for the "average observer"!

Many attempts at measuring or calculating *b* have appeared in the literature. Landis (1977) applied data from seven observers to demonstrate that the eye sees red stars as systematically fainter than their V magnitudes suggest (*b*>0). Steffey (1978) calculated m_v–V for a variety of red stars using approximations of stellar flux distribution and the assumption that the eye response is scotopic when observing faint stars. His results also show a strong systematic increase of m_v–V for redder stars (b~0.3). Howarth (1977, 1979) and Stanton (1981) showed a similar trend by comparing historical visual data with modern photoelectric measurements. Eugenio *et al.* (1959) performed regression analyses of both the Potsdam and Harvard visual catalogues compared to Johnson V. Their results demonstrated important differences between the two catalogues. While the Harvard data generally yielded *b* ~ +0.15, Potsdam fits actually gave small negative values! Since these differences could not be due to random error, they must reflect differences between observers, or more likely, between measurement equipment and techniques. Significant variations were also detected as functions of right ascension and declination, illustrating the difficulty of doing all-sky photometry using the eye. Differences between the Potsdam and Harvard results pose an interesting dilemma for those trying to characterize the modern observer's vision: Which catalogue should one use?

More recently, Hallett (1998) has shown that equating variable star observing with scotopic eye response is incorrect since both cones and rods play a role in telescopic observations. The two types of receptors are involved to various degrees, depending on distance of the image from the foveal center, brightness above threshold, and background effects. Each of these factors is very difficult to characterize for the "average observer," suggesting that an accurate computation of *b* using response curves must be considered all but impossible. Apparently two of the three approaches used above in computing *b* have serious flaws:

1. Using eye response data and star irradiance curves to calculate color differences between V and m_v is thwarted by our imperfect knowledge of what the average observer's response function really is at the telescope.

2. Relying on early visual catalogue data has the flaw that the observing conditions and techniques used were significantly different than those employed by modern observers. The fact that the comparison with V gives substantially different results for the Potsdam and Harvard catalogues clearly suggests caution.

The third approach, following Landis (1977), relies on comparing modern observations with corresponding B and V magnitudes. In order to truly represent the "average observer," this approach must include a large number of observers and observations, and must include a range of instruments and techniques that truly represent today's observer. Only then can the statistical variance inherent in visual observations be overcome, and the range of variation from observer to observer be quantified.

3. The observers

The call for observers to participate in *V-Magnitude Experiment – 1998* was both mailed to observers and placed on the AAVSO web page. Maps of the test field, SS Cyg, were included in the call, identifying both the lettered "unknowns" and the comparison stars that were to be used for the observations (Figure 1). Observers were instructed to ignore existing sequences in making their estimates. Responses were received from 63 participants, yielding a total of over 750 individual observations. Table 1 lists participating observers. These observers, from around the world, represent a cross-section of those

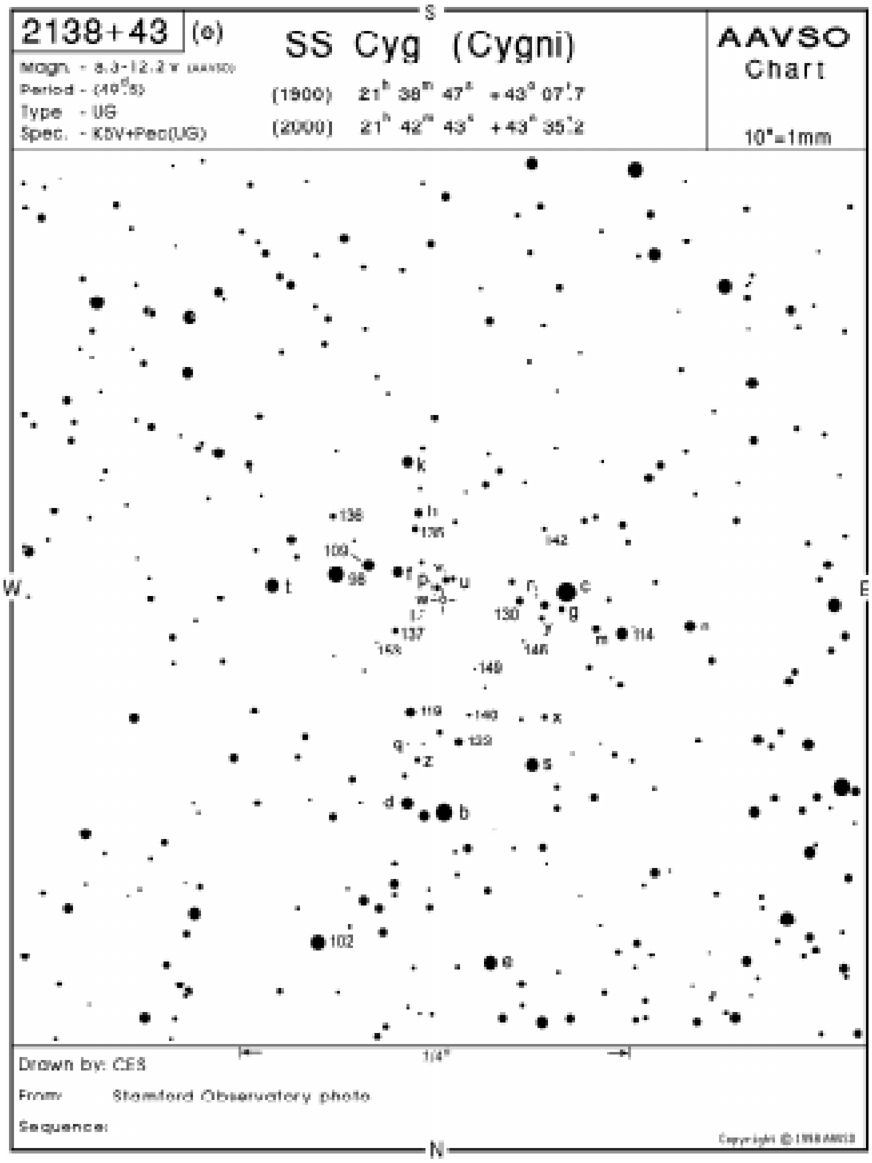

Figure 1. Modified AAVSO e-scale chart showing most of the "unknowns" and comparison stars of the *V-Magnitude Experiment.* Stars outside of this field were marked on a b-scale chart. *[Ed. note: The values of comparison star magnitudes on this chart do not necessarily match those on the chart to be used for estimating the magnitude of SS Cyg. Do not use this chart to make observations of SS Cyg.]*

contributing to the AAVSO. (The author regrets that the selected field precluded the participation of most Southern Hemisphere observers. The SS Cyg field was selected due to its rich assortment of well-measured red and blue stars.) Some individuals are among the most prolific and experienced observers in the organization. Others indicated being relatively new to variable stars. While a few are young, the average age of all participants was 47.1. If this is indicative of the AAVSO in general, we all need to recruit more young members!

Three contributors were color-blind, and the vast majority (89%) used red lights for reading charts. An exception to the claim of diversity was the almost complete absence of female observers. The author is grateful to Barbara Adams for being the only participant to buck this trend. Her observations line up very well with the male average, indicating that sex may not be a major discriminant in color sensitivity.

Table 1. Participating observers.

Observer	*No. Obs.*	*Julian Date (+2451000)*	*Instrument*
Patrick Abbott	4	47	12.5" Cassegrain
Barbara Adams	7	74	Celestron 8
Rudie Allison	12	80	12" LX200
Ray Berg	11	49	3" and 8" Catadioptric
Mark Biesmans	20	138	10"Cassegrain; 8cm refractor
John Bortle	20	54	16" Dobsonian; 10x50 binocs
Eric Broens	18	48;78	14" Dobsonian
Alain Bruno	15	50	Celestron 9
Wayne Clark	5	80	17.5" Dobsonian
Peter Collins	7	185	4.25" reflector
Robert Crumrine	14	54	10" f/7; Celestron 8
Frank Dempsey	10	51	10" Newtonian
Alfons Diepvens	14	140	15cm refractor
William Dillon	9	53	12" Newtonian
Pavol Dubovsky	11	78	20cm Newtonian
Shawn Dvorak	12	—	Celestron 8
Sergio Foglia	9	57;66	11.4 cm Newtonian
Marino Fonovich	9	51	Celestron 8
Mark Gable	4	75	Meade 8 f/10
Patrick Garey	8	48	13.1" Newtonian
Keith Graham	4	65;66	12" Meade
Bjorn Granslo	22	53;54	Celestron 8
Gene Hanson	12	126	18" reflector
Richard Harvan	19	69	44.5; 7.6cm reflector
Robert Hays, Jr.	18	56	15 cm reflector
G.Wyckliffe Hoffler	7	71	12.5" reflector
Robert Johanns	19	43	25cm Newtonian
Walter Kaminski	12	42	10" Newtonian
Kiyoshi Kasai	6	55	15cm Dobsonian
Attila Kosa-Kiss	10	49	6.3cm refractor; 7x50 binocs
Thomas Lazuka	9	46	6 cm refractor
Daniel Loring	12	40;70	20cm Newtonian

(Table 1 continues on following page.)

Table 1. Participating observers, continued.

Observer	*No. Obs.*	*Julian Date (+2451000)*	*Instrument*
Michael Lyons	3	41	11cm reflector
Tyler MacKenzie	4	72	10" Newtonian
Miguel Marco	10	45	21cm Dobsonian
Dmitry Matsnev	3	50	4.4" Newtonian
George Mavrofridis	29	67-69	16" Dobsonian
Tom McCague	9	45	10" Newtonian
Patrick McDonald	7	80	14" Celestron
Jerome McKenna	16	45-70	11" Celestron
Robert Modic	20	50	20" Dobsonian
James Molnar	7	62;67	10" Cave Newtonian
Warren Morrison	9	54;74	6 and 15cm refractors
Soumen Mukherjee	9	53	7.75" Newtonian
Mark Munkacsy	20	61;68	6" Dobsonian
Steve O'Connor	15	66;74	8" Cave Newtonian
Noel Peattie	1	41	4" Astroscan
Giorgio Pozzi	2	72	11.4cm Newtonian
Ronald Royer	23	45;51	12.5"&18" Newtonian
Hugh Rumball-Petre	3	56;73	3", 6" reflector
Sei-ichi Sakuma	4	67;73	16" Newtonian
Gerd-Lutz Schott	9	—	8" Celestron
Alan Sharpless	3	58	10" reflector
Jerzy Speil	10	79	20x80 binocs
Richard Stanton	26	21;110	16"Cassegrain;10"Newtonian
Philip Steffey	10	50	8" Newtonian
Robert Stewart	7	47	17.5" Dobsonian
Scott Tracy	10	143	11" SCT
Charles Trefzger	11	52	Photovisual
Daniel Troiani	13	71	10" Newtonian
Odd Trondal	2	51	24" Cassegrain
Richard Wend	8	38	Criterion 11
Frederick West	3	81	8cm binocs

4. Analysis

Each observer's measurements were entered into separate Microsoft Excel spreadsheets in which "steps" were automatically converted to magnitudes and summary parameters for that observer calculated. Each observation was tagged with the observer's initials so it could always be identified, regardless of where it was used. Data from all observers were eventually combined in a single spreadsheet and analyses run on the entire data set.

Each observation submitted had the form, $V_1 – S_1$ x $S_2 – V_2$, where the observer judged that the unknown star x was S_1 steps brighter than star V_1 and S_2 steps fainter than star V_2 (V_1 and V_2 refer to the Johnson V magnitudes of these comparison stars). Based only on these steps, and measured magnitudes and colors of comparison stars, a magnitude, V_f, *and color*, C_f, of a fictitious star was calculated using linear interpolation

$$V_f = V_1 + (V_2 - V_1)\,[S_1/(S_1 + S_2)] \qquad (7)$$

and

$$C_f = C_1 + (C_2 - C_1)\,[S_1/(S_1 + S_2)]\,, \qquad (8)$$

where V_i is the V magnitude and $C_i = B_i – V_i$ for the *i*th comparison star. Note that implicit in (7) is the assumption that the observation interpolation produces a magnitude in the Johnson V system, which is not strictly true unless the comparison stars have identical colors. For this reason, comparison stars were selected to minimize the color differences between consecutive magnitudes. Of course, the observers have no way of measuring the color of either the comparison stars or the unknown stars. But their observations are equivalent to saying that, if a star of magnitude V_f and color C_f were placed beside the unknown star, they would judge both stars to be equal. Note that the total number of steps that the observers choose to use is not important in this formulation. Substituting into (4), we can write

$$m_v = V_f + b\,C_f \qquad (9)$$

for the fictitious star that appears exactly equal to the unknown. But since we actually know the magnitude, V_u, and color, C_u, of the "unknown,"

$$m_v = V_u + b\,C_u\,, \qquad (10)$$

which implies

$$V_f + b\,C_f = V_u + b\,C_u \qquad (11)$$

or

$$V_f - V_u = b\,(C_u - C_f)\,. \qquad (12)$$

Since everything in equation (12) is known except *b*, one should be able to solve directly for this coefficient. But the high noise content in V_f (and to a lesser extent in C_f) dominates the results for a small number of observations. In order to confidently estimate *b* to 10% or better, at least 30 independent observations are needed for a single observer, and several hundred points should be used to obtain a good average for many observers.

Two features of this formulation should be noted. First, plots of $(V_f – V_u)$ vs. $(C_u – C_f)$ should pass through the origin with a slope *b*. This fact was used in estimating the color coefficient of individual observers. Forcing the regression fit to pass through the origin avoids mapping observation errors into a spurious constant term. Regression fits of data submitted were performed for each observer submitting more than six observations.

Typical results are illustrated by the author's response curve (Figure 2). Note that in this case there is not a large difference in *b* between including a constant term or forcing the fit through (0,0). Note also that the data scatter seems quite large. The RMS scatter (Standard Error) about the best-fit line for these observations was about 0.2 magnitude, implying a significant random error component, even after correction for the observer's color coefficient. As will be seen, this is a fairly typical value for the "average observer."

A second feature shows up when data from a large number of individuals are plotted together. A comparison star being judged equal to an unknown will result in a single point on a graph of $(V_f - V_u)$ vs. $(C_u - C_f)$. A second comparison star will give a different point for the same unknown. All observations of this unknown using these two comparison stars will fall on a straight line connecting these points. When a large number of observations using the same comparison stars are combined, these lines become very evident (Figure 5).

5. The stars

Table 2 lists the photoelectric and CCD magnitudes and colors of the unknown stars and the comparison stars. Also listed is V_{error}, the standard deviation of the independent measurements for each star (where there is only one published observation of a star, the V_{error} and $(B–V)_{error}$ values refer to published errors for that observation). The comparison stars were selected from a relatively narrow color range, most being in the range of $0.4 < B–V < 0.86$. An essential element of this experiment is the accuracy and stability of the magnitudes and colors of the unknown stars and comparison stars. As can be seen from the error columns, standard deviation for most V and B–V values, *at least for the times these observations were made,* is below 0.03 except for the faintest stars. This is taken as a good indication, but not as proof, that the star is constant. Additional evidence is derived from the visual observations, discussed below. The final column in the "Unknown" Stars section of Table 2 lists the number of visual observations received for each of the unknown stars. Note that no one was able to detect star *l*. From this we can conclude that either its position is misplotted or that it is fainter than indicated (V = 15.2).

Before combining all data points in an estimation of *b*, observers' data for individual stars should be examined to uncover possible variability. In Figure 3, the standard deviation, σ_i, of all visual observations for each unknown star is plotted as a function of the star's color. Each star is identified by its letter. Several conclusions can be reached from this figure:

1. When data from many observers are combined, a standard deviation not much smaller than 0.2 magnitude can be expected, even with "perfect" comparison star magnitudes.

2. There is very little color dependence in this parameter. Therefore, one should expect roughly the same data scatter for visual observations of any star in this color range.

3. Only star *w* has an σ-value significantly above the average value (σ_{ave} = 0.22 mag.). None of the other stars seem to be significantly variable, at least on the magnitude and time scales of the observations. Star *w* might be considered suspect in that it is difficult to separate from its companion, star *p*, in many telescopes. This fact probably explains the higher variance in the observations, and may be grounds for its deletion from the data set.

4. Since the variance associated with these visual observations is roughly a hundred times larger than the variance of photoelectric and CCD measurements (0.22^2 vs 0.02^2), the latter can be ignored.

5. The variance of visual observations is also significantly larger than that arising from rounding all chart magnitudes to the nearest tenth (variance for a uniform distribution of –0.05 to 0.05 is 0.029^2). Therefore, there is no need to clutter charts with comparison star magnitudes of greater precision than one decimal place.

Table 2. Magnitudes and colors of experiment stars.

Star	*V*	*B-V*	*N*	V_{error}	$(B–V)_{error}$	*Observers**	*Number of Visual Obs.*
Comparison Stars:							
75	7.396	0.143	1	—	—	S	
80	7.953	0.526	2	0.007	0.012	S,T	
85	8.502	0.432	2	0.004	0.007	S	
90	9.037	0.315	2	0.001	0.007	S	
98	9.768	0.409	2	0.009	0.010	H,S	
102	10.237	0.453	2	0.009	0.015	S	
109	10.885	0.549	4	0.008	0.006	H,S	
114	11.408	0.629	2	0.011	0.007	S	
119	11.863	0.636	4	0.007	0.003	H,HH,S	
123	12.278	0.633	4	0.004	0.012	H,HH,S	
130	12.983	0.635	3	0.015	0.014	H,HH,S	
135	13.478	0.745	2	0.001	0.023	H,S	
136	13.562	0.763	1	0.007	0.029	H	
137	13.644	0.763	2	0.001	0.003	H,HH	
140	13.998	0.610	2	0.070	0.007	H,S	
142	14.190	0.860	1	0.032	0.040	H	
146	14.608	0.788	1	0.018	0.089	H	
149	14.943	0.770	1	0.048	0.060	H	
153	15.258	0.850	1	0.029	0.060	H	
"Unknown" Stars:							
a	7.549	1.734	3	0.023	0.015	S,T	43
b	9.559	1.021	2	0.002	0.007	S	54
c	8.469	1.331	3	0.002	0.013	S	47
d	10.813	0.872	2	0.000	0.001	S	57
e	10.824	0.220	2	0.004	0.004	S	41
f	12.071	1.869	5	0.010	0.010	H,HH,M,S	44
g	13.574	0.907	3	0.016	0.030	H,HH,S	20
h	13.022	1.801	3	0.017	0.011	S	26
k	11.816	1.509	4	0.019	0.013	H,HH,S	41
l	15.236	1.480	1	0.025	0.055	H	0
m	13.203	1.753	2	0.017	0.027	H,S	20
n	11.261	1.068	2	0.004	0.001	S	46
p	14.819	1.377	2	0.009	0.022	H,M	3
q	15.251	1.819	1	0.016	0.030	H	2
r	14.254	1.850	1	0.014	0.060	H	10
s	10.950	1.160	3	0.007	0.012	S	44
t	11.001	1.652	3	0.002	0.006	S	45
u	13.680	0.398	2	0.051	0.018	M,S	13
v	13.382	0.459	3	0.024	0.003	H,HH,M	19
w	13.776	1.614	1	0.023	0.107	H	11
x	13.699	1.383	3	0.009	0.023	H,HH,S	17
y	14.064	1.529	2	0.037	0.009	H,S	14
z	13.918	1.746	1	—	—	S	14

* H = Henden (1998); HH = Henden & Honeycutt (1997); M = Misselt (1996); S = Author's measurements. H, HH, and M are CCD measurements; S is photoelectric.

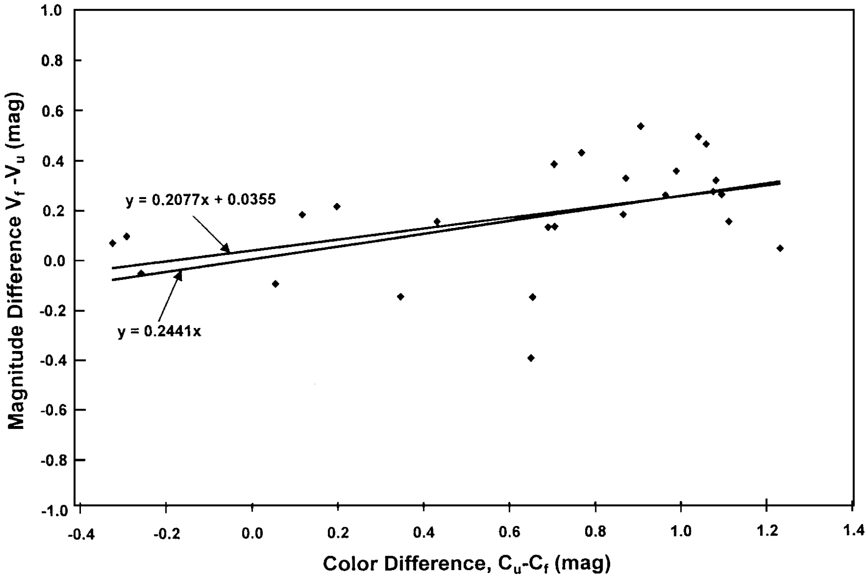

Figure 2. Typical response curve for an individual observer. The slope of the line, y = 0.24x, gives the color coefficient *b* for this observer. Note that allowing the least squares fit to include a constant term introduces only a minor change in slope.

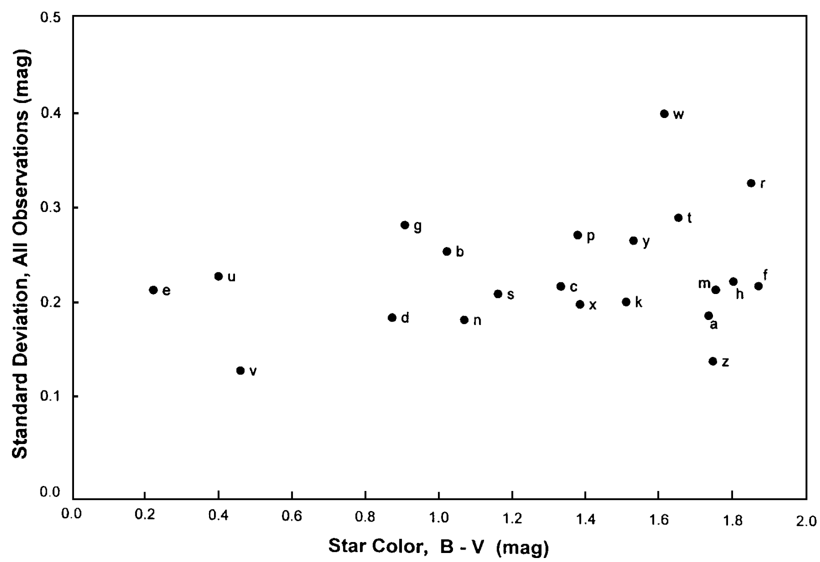

Figure 3. Standard deviation of all observations made of each "unknown" star plotted against the photoelectrically measured color (B–V) of the star. These data include the effect of differences in the color response of individual observers. The number of observations of each unknown is tabulated in Table 2. Note: star "q" is not plotted, since both observations received for this star were identical (V_I ~15.5), yielding a misleading standard deviation ($\sigma = 0$).

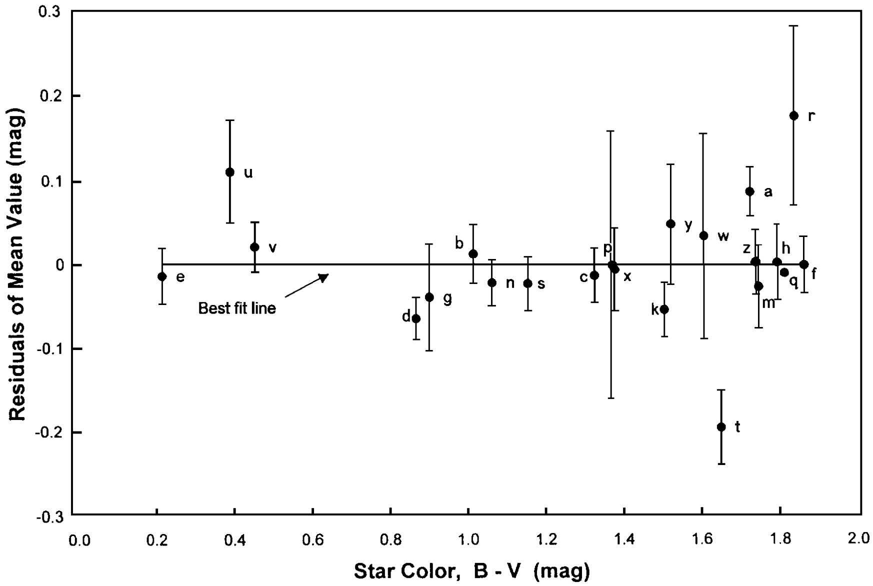

Figure 4. Residuals between the mean value of all observations of each star, and the best fit line through those means, plotted as a function of star color. Error bars show the expected standard deviation of each mean.

The residual of each star's mean value with respect to the best fit for all 22 stars is plotted in Figure 4. The apparent outliers in this plot are stars *r*, *t*, and possibly *u* and *a*. In a perfect world one would expect the size of these residuals to approximate

$$\sigma_{mean} = \sigma_i / N^{0.5}, \qquad (13)$$

where N is the number of measurements used in calculating the mean and σ_i is the standard deviation of those measurements. The error bars displayed for each star correspond to this σ_{mean} calculated for that star. From these bars it is clear that both *r* and *u* are well within 2σ of their expected values, and *a* is probably acceptable at < 3 σ. But at 4.4 times σ_{mean}, and 43 observations, the residual for star *t* is deemed excessive. Whether this residual is due to measurement error in V, an unusual spectrum, or star variability, is not known. Before implicating star *t* itself, the possibility must be considered that one of the comparison stars was causing this residual. However, this possibility can be dismissed since most observers also used the same comparison stars (10.9 and 11.4) for stars *s* and *n*, which behave normally in Figure 4. As a result, star *t* was removed from the final data set. Conversely, although star *w* exhibits a relatively large data scatter (Figure 3), it is retained in the data set due to its small residual error.

6. The "average observer"

It is now possible to analyze the entire data set of 21 stars (excluding stars *l* and *t*). In order to find a value for *b* in (4) that applies to *all* observers, observations were included regardless of such parameters as weather, color of chart illumination, or color

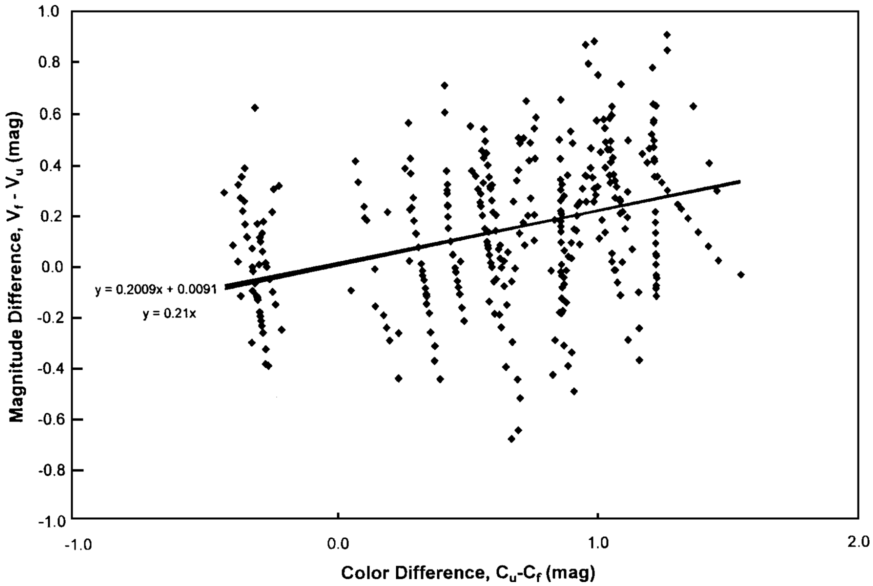

Figure 5. All 594 observations used in calculating *b* are plotted, along with two straight-line fits of the data. The ordinate is the difference between the magnitude, V_f, of a fictitious star that matches the "unknown" (calculated from the observer's "steps"), and the measured Johnson V for the "unknown". Similarly, the abscissa is the difference in color between Johnson B–V for the "unknown" and the fictitious star's color, C_f, also calculated using the observer's "steps." Refer to equation (12). The conspicuous "dotted lines" result from the use of the same two comparison stars by many observers. Many of the dots represent several identical observations. Note the slight difference in slope *b,* depending on whether the constant term (offset at zero color difference) is estimated. Since this offset *should* be zero, *b* = 0.210 is the desired composite value for the "average observer."

blindness (the last involving only 11 observations). Figure 5 shows this summary, plotting all 594 data points of $(V_f - V_u)$ vs. $(C_u - C_f)$. From (12), the parameter *b* is calculated as the slope of a linear least-squares fit though all these points, which turns out to be almost exactly 0.21. Several points should be noted relative to Figure 5:

1. As described above, all data points fall on straight lines corresponding to linear interpolation of color and magnitude between two comparison stars. This is particularly evident where most observers used the same two comparison stars.

2. When the data are fit allowing an additive constant, the zero intercept is only 0.0091 magnitude, well within one standard error (0.016) of the expected null value.

3. Despite the large number of participating observers, each with unique eyes and instrumentation, the standard error calculated for *b* is only 0.0116, or approximately 5%. This value is consistent with the very strong correlation between star color and magnitude error.

4. The validity of equation (7), using the eye to interpolate Johnson V magnitudes, was tested by rerunning the entire estimation with all magnitudes translated to m_v using a color coefficient of *b* = 0.21. The resulting coefficient was *b* = 0.0003, which is completely negligible.

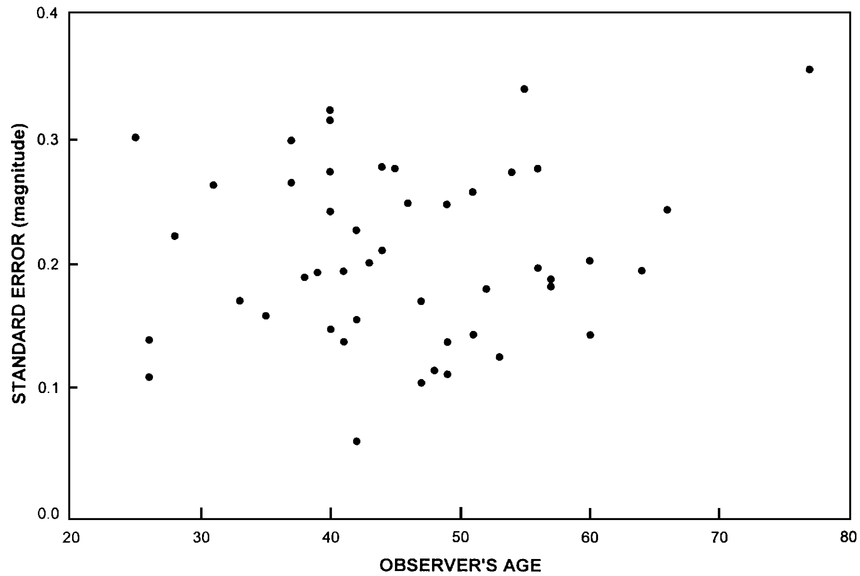

Figure 6. Standard error of observations submitted by individual observers versus their age. Since this error is measured relative to a best-fit value of *b* for each observer, average effects of color have been removed.

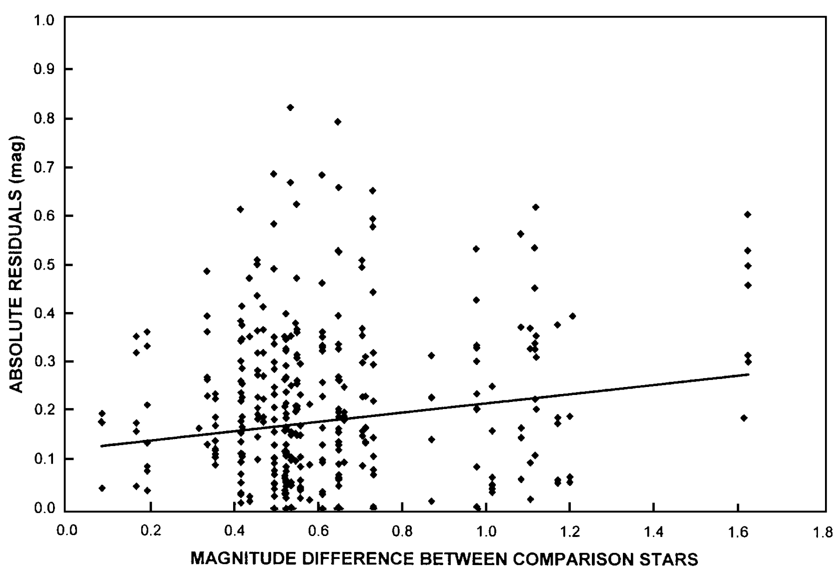

Figure 7. The absolute value of the difference between each observation and the corresponding best fit value, $m_v = V + 0.210\ (B–V)$, plotted versus the difference in magnitude between the comparison stars used. As one would expect, using comparison stars widely separated in magnitude increases measurement error.

5. Multiple regression fits with other parameters, such as observer age or star magnitude, found no significant correlations.

6. The data of all observers using white light to read charts yielded $b = 0.23$. The standard error for this measurement was 0.06 (due to fewer data points), which is consistent with using $b = 0.21$ for everyone.

7. Only 12 observations were received from color-blind observers. This limited set suggests that a larger coefficient, perhaps $b = 0.3$, would be appropriate for this group.

7. Individual observers

Having characterized the "average observer," it is important to understand the variations of individuals relative to this average. As mentioned above, a regression was performed to determine a color coefficient for each observer who submitted a sufficient number of data points (usually seven or more). Figure 6 plots the standard error calculated for the data points used in these fits against the observers' ages. Since this error is measured with respect to the best-fit color response curve for that observer, systematic color differences between individuals are eliminated. It is clear that while some observers approach the "one-tenth magnitude accuracy" often attributed to visual observations, many others fall in the range of 0.2 to 0.3 magnitude.

One way to reduce some of this scatter would be for observers to more carefully select comparison stars that are closest (in magnitude) to the variable. A number of observers used the same set of stars to compare with several unknowns, even though much closer comparisons were available. A graph of the effect of this approach is shown in Figure 7, plotting the absolute residuals of each observation versus the *magnitude difference* between the comparison stars used. If observers always selected comparison stars closest to the perceived brightness of the "unknown," the maximum difference on this plot would be 0.73 magnitude. The significant upward trend in the best-fit line illustrates the importance of choosing the "best" comparison stars for each observation.

Graphed in Figure 8 are the color coefficients for the 48 observers for whom linear regression fits were performed, plotted as a function of the observer's age. Error bars show the standard error in the estimate of b for each observer. Standard error in b for an individual observer is related both to the *scatter* and the *number* of observations used in the fit. Given an observation scatter of 0.2 magnitude, it requires *on the order of 100 independent observations* to measure an observer's coefficient to better than 10%. Since most participants submitted fewer than 15 observations, it is not surprising that the standard deviation in individual color coefficients is as large as $\sigma_b = 0.11$ (about a mean of 0.206).

But this large standard deviation for individual observers does not imply that the average b for all observers is inaccurate. As we have seen, when observations from all observers are combined, the standard error in b is approximately 5%. This strongly suggests that, if this experiment were repeated with a different set of 63 observers, the results would yield the same b value within a few percent.

8. Summary

A detailed analysis of observations submitted in response to the *V-Magnitude Experiment – 1998* has yielded quantitative insight into many characteristics of observations generated by modern observers:

1. The vision of the "average observer" is more sensitive to blue stars than the V-magnitude scale indicates. This is consistent with Hallett's observation that the Johnson V scale is blue-blind relative to the human eye. The visual magnitude scale can now be accurately defined using a conversion from Johnson V and B magnitudes by the linear formula

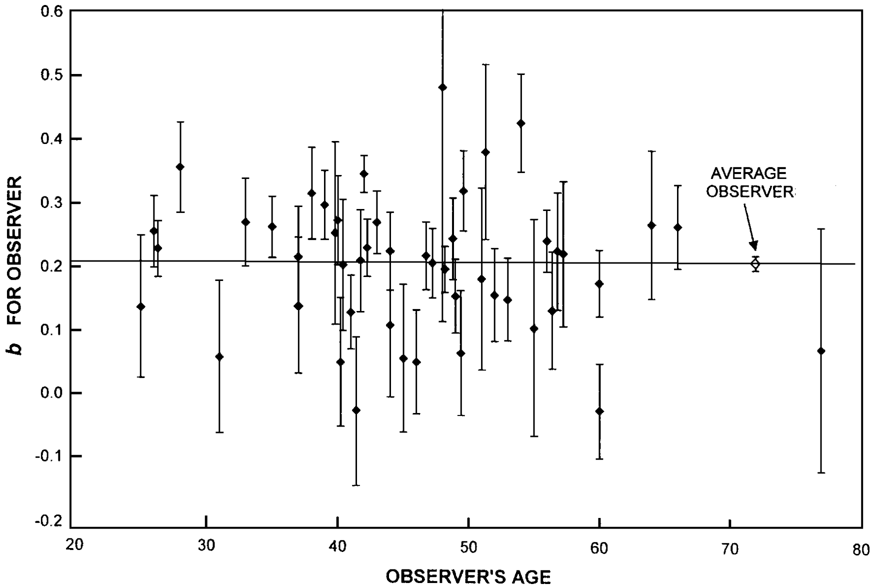

Figure 8. Color coefficient, *b*, of individual observers contributing measurements of seven or more stars. Error bars are the standard error *in the estimate of b,* due to scatter of individual observations and the limited number of observations available from any observer. The point for the "average observer" illustrates the small statistical error achieved when data from all observers are combined. This point is not plotted at the average age of participating observers (47.1 years) in order to clearly separate it from other data points.

$$m_v = V + 0.210\,(B - V)\,. \qquad (14)$$

The same magnitude scale can be expressed in terms of Tycho magnitudes:

$$m_v = V_T + 0.089\,(B - V)_T\,. \qquad (15)$$

2. Although individuals may achieve measurement consistency of 0.1 magnitude or better under controlled circumstances, random errors on the order of 0.2 magnitude can be expected when data from many observers are combined. Note that these errors are present *after the effects of star color have been corrected for each observer* (Figure 2).

3. The factor most strongly correlated with measurement error (after star color effects are removed) seems to be the selection of comparison stars (Figure 7). Observers who based their estimates on stars selected closest in magnitude to the unknown star fared better than those who used the "one pair fits all" approach.

4. Several factors do not appear to be particularly correlated to the color coefficient or observation error. Plots of measurement residuals for observers using red or white light gave similar results. Age (Figure 8) does not seem to be a significant factor. Even plots of the brightness of the "unknown" above the observer's threshold did not show a strong effect, except that errors increased for stars near threshold, as one would expect. Apparently most observers have developed techniques for handling very bright stars, such as defocusing or stopping down the aperture.

What does all this mean to the AAVSO? We can now confidently use comparison star magnitudes measured by any accurate photometric instrument, and map them into visual magnitudes for the "average observer." To accomplish this, the instrument magnitudes need satisfy only two conditions:

1. All data must include measurement of star color. Measurements made only with a single filter, or without a filter, should *not* be used.

2. The measurements must be accurately transformed to the Johnson UBV or Tycho photometric systems (enabling application of equation (14) or (15)).

Complex questions such as how to coordinate a switch to a revised magnitude scale, or whether to revise existing chart sequences, are now being carefully considered by the AAVSO Chart Committee and collaborators around the world. Please note that any revisions to chart magnitudes according to the revised scale discussed in this paper should *not* be used until a final decision is reached.

9. Acknowledgements

The author wishes to thank Janet Mattei and members of the AAVSO Chart Committee for many helpful recommendations and assistance in conducting the *V-Magnitude* survey. He is indebted to Arne Henden for his sequences and suggestions that led to the final form of the survey. Finally, a special thanks goes to all the observers who took the time to contribute their individual observations.

References

Eugenio, E., Mendoza, V., and Gomez, T. 1959, *Bol. de Obs. de Tonantzintla y Tacubaya*, **33**, 111.
Hallett, P. E. 1998, *J. Amer. Assoc. Var. Star Obs.*, **26,** 139.
Henden, A. A. 1998, private communication.
Henden, A. A., and Honeycutt, R. K. 1997, *Publ. Astron. Soc. Pacific*, **109,** 441.
Howarth, I. D. 1977, *J. Amer. Assoc. Var. Star Obs.*, **6,** 88.
Howarth, I. D. 1979, *J. Amer. Assoc. Var. Star Obs.*, **8,** 26.
Landis, H. J. 1977, *J. Amer. Assoc. Var. Star Obs.*, **6,** 4.
Mattei, J. A. 1998, *J. Amer. Assoc. Var. Star Obs.*, **26,** 170.
Misselt, K. A.1996, *Publ. Astron. Soc. Pacific*, **108,** 146.
ESA 1997, *The Hipparcos and Tycho Catalogues*, SP-1200, Vol.1, 57, European Space Agency, Publications Division, Noordwijk, Netherlands.
Scovil, C. E., and Leitner, R. A. 1991, *J. Amer. Assoc. Var. Star Obs.*, **20,** 1.
Stanton, R. H. 1981, *J. Amer. Assoc. Var. Star Obs.*, **10,** 1.
Steffey, P. C. 1978, *J. Amer. Assoc. Var. Star Obs.*, **7,** 10.
Zissell, R. E. 1998, *J. Amer. Assoc. Var. Star Obs.*, **26,** 151.