THE BALTIMORE MYOPIA STUDY 40 YEARS LATER

Joseph N. Trachtman, O.D./Vincent Giambalvo, Ph.D.

ABSTRACT

The Baltimore Myopia Project (BMP) is one of the most frequently cited studies concerning the efficacy of vision training. As a result of the study two sets of publications appeared in the literature between 1946 and 1947. One set was written by an ophthalmologist and the second set by optometrists. The ophthalmological articles stated that the results of the study showed that vision training was not efficacious in the treatment of myopia. The optometric articles reported results supporting efficacy of vision training in the treatment of myopia. After reviewing each of the articles in detail, the apparent contradiction became understandable; the two sets of articles did not utilize the same sets of data. We performed a post hoc analysis of the available data using modern statistical methods. Our conclusions are that there were statistically significant positive changes in visual acuity and that the ophthalmological opinion that the BMP indicates a lack of support for the efficacy of myopia reduction vision training is unfounded.

KEY WORDS

accommodation, myopia, vision training, optometry, ophthalmology

"With the possible exceptions of educating some patients to interpret blurred retinal images more carefully and of convincing some others that they could see better even though there was no actual improvement, this study indicates that the visual training used on these patients was of no value for the treatment of myopia."[1] These are the concluding remarks by Alan Woods, author of the Baltimore Myopia Project (BMP), published in 1946 in the *American Journal of Ophthalmology* and in the *Archives of Ophthalmology*. This conclusion has placed a stigma on vision training since its publication. The BMP is typically invoked in negative articles about vision training or in cases where there were visual acuity changes, but no reported changes in refractive error.[2,3,4,5] An illustrative example has recently been published by Grosvenor and Flom.[6] In their discussion of the BMP, they quote a negative statement of Shepard.[7] Grosvenor and Flom are apparently unaware of Shepard's comment on the previous page; "We may, therefore, properly claim that 90 per cent of the myopes may enjoy a significant improved vision through vision training."[7]

Although, there have been optometric reports of the same study published in the *Journal of the American Optometric Association* by H. Ward Ewalt,[8] and subsequent publications in the *Journal of the American Optometric Association* and *Optometric Weekly* by Carl Shepard[7] and Emmett Betts,[9] Woods' conclusion appears to ring the loudest. Even though a similar study was conducted concurrently in St. Louis and published in the *Transactions of the American Academy of Ophthalmology and Otolaryngology* by Hildreth et al.[10] with conclusions more positive toward vision training, Woods' conclusion appears to have gained the most attention.

The primary purpose of this paper is to analyze all the data of the BMP using the appropriate statistical methods. To the best of our knowledge, this task has not yet been performed. Hackman[11] did report a full analysis of the data, but his statistical procedure utilized per cent visual acuity as interval level data (see Appendix A). We are not critical of Hackman, since the differentiation of data into parametric and non-parametric levels was not elaborated until 10 years later.[12] Secondly, we felt that some of the issues raised by Hackman,[11] Betts,[9] and Shepard[7] are worth repeating 40 years later. Thirdly, we became interested in obtaining optometric refractive data to compare to the ophthalmological data. We also sought a description of the vision training procedures used, since they do not appear in either Woods' or Ewalt's publications. Finally, we wish to share this re-analysis with the optometric and ophthalmological communities.

BACKGROUND

The BMP was sponsored by the Optometric Extension Program and was a joint venture between optometry and ophthalmology. The purpose of the BMP was to determine the efficacy of vision training to reduce myopia. The BMP began in 1944 at the Johns' Hopkins University, Wilmer Eye Institute, Maryland. A protocol was developed, whereby ophthalmologists performed the pre- and post- training visual acuity and retinoscopic measurements, and optometrists conducted the vision training portion of the study.

METHODS

Subjects

One hundred and three (111 according to Ewalt[6]) patients started the study and 103 patients completed the training and testing protocols. The age range of the completed patients was 9 to 32 years. No further details on age could be located in any of the references. The refractive error of the subjects ranged from: -0.50D to -9.00D, with one case, O.D.: - 2.50 and O.S.: +5.00. The subjects were selected according to the following criteria: first come basis with a broad range of socioeconomic status and general health. However, this protocol was not followed. An inordinate number of individuals were from a low socioeconomic status. Further, "A considerable number of the trainees were not desirable candidates for vision training on the basis of standards established in most private practices. This is particularly true of those trainees who may have had general health problems." [8]

Procedure

Before the vision training began, each subject had distant visual acuity measurements OD, OS and OU using Snellen Letters, Snellen Numbers, Tumbling E's, Landolt C's, and cycloplegic retinoscopy. Woods[1] states that the reason for the four different sets of visual acuity measurements was to average the fluctuations usually obtained among the different charts. The cycloplegic agent was homatropine 5% and two instillations were administered to subjects 9 years and older. "Atropine sulfate, 1% solution three times a day for two days was used in younger individuals."[1] The end point for the retinoscopy was not defined in the text; the protocol did not specify the end point from the high or low end. There were no measurements of subjective refractive errors. Visual acuity and retinoscopy were the dependent variables. Vision training, which was the independent variable, was described by Woods in the following way; "cylinders, prisms, and targets were employed in this training program with the idea of reorganizing the visual behavior patterns (so) that the visual skills, including acuity, can be improved significantly".[1] The training was given from September, 1944 to the first week of December, 1944. Three sessions per week for 13 weeks were planned, but the actual average number of training sessions was only 25.

RESULTS

We used the STATS Computer Program (Statsoft, Inc.) on an ATandT 6300 Plus microcomputer to analyze the data reported by Woods.

All the data from the visual acuity measurements (from Woods,[1] Table IX, N = 103) were analyzed using the Correlated T-Test program. The results showed that there was a significant statistical improvement from the pre-test to the post-test (see Table 1).

The same type of analysis was performed on the refractive error measurements (from Woods,[1] Table VIII, N= 67) using the Wilcoxon T Testa program. Before we evaluated the data for statistical significance, we saw, on an overview, extremely large changes, both increases and decreases between the pre- and post-training retinoscopic measurements. These large shifts were sometimes larger than the value of the initial refractive error. Examples of the large variability may be seen by looking at Figure 2, where: one patient had a pre-training retinoscopy measurement of -0.50D and a post-training measurement of -2.25D; a second patient had a pre-training retinoscopy measurement of -3.87D and a post-training measurement of -0.25D; a third patient had a pre-training retinoscopy measurement of -3.25D and a post-training measurement of -4.75D. The actual differences in retinoscopy are graphically shown in Figure 1 for the right eye and Figure 2 for the left eye.

According to proper statistical procedure,[12] the variability of the retinoscopic changes reported by Woods[1] were statistically too large and upon inspection[13] too great to perform an accurate analysis. We maintain that position even though there was a statistically significant change toward less myopia in retinoscopy for

VISUAL ACUITY RESULTS
from Woods Table IX; N = 103

TARGET	EYE	t-TEST	PROBABILITY
LETTERS	RIGHT	6.294	>.01*
LETTERS	LEFT	7.445	>.01*
LANDOLT	RIGHT	6.016	>.01*
LANDOLT	LEFT	5.631	>.01*
E	RIGHT	7.285	>.01*
E	LEFT	7.429	>.01*
NUMBERS	RIGHT	6.362	>.01*
NUMBERS	LEFT	5.210	>.01*

*statistically significant

Table 1.

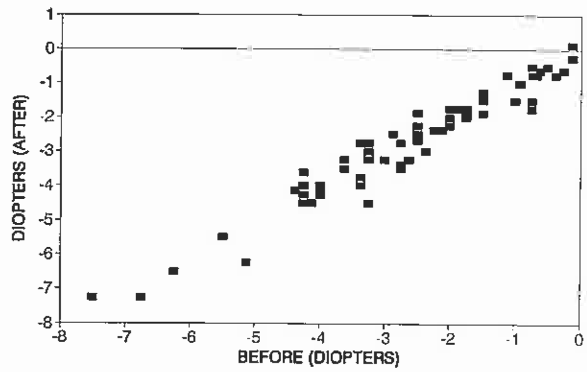

Figure 1. PRE- VS. POST-TRAINING RETINOSCOPY - O.D. from Woods Table VII, N = 67

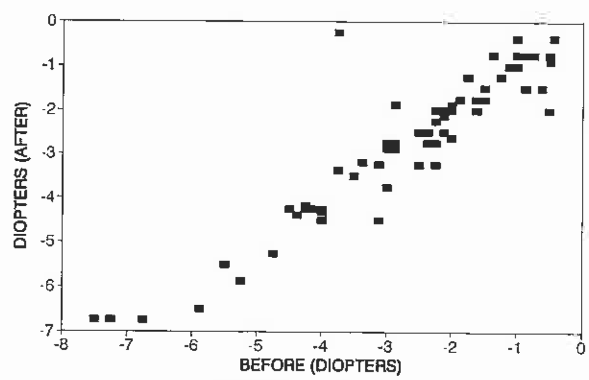

Figure 2. PRE- VS. POST-TRAINING RETINOSCOPY - O.S. from Woods Table VII, N = 67

REFRACTIVE ERROR DATA
from Woods' Table VIII; N = 67

EYE	WILCOXON T	Z SCORE	PROBABILITY
RIGHT	T = 742	Z = 2.48	p = .0126*
LEFT	T = 888	Z = 1.568	p = .1127

*statistically significant

Table 2.

right eyes according to the Wilcoxin T Test (see Table 2).

We then examined the data reported by Ewalt[8] and found *different* visual acuity data. These data showed an even greater improvement than that shown by the ophthalmologists' data. The optometric visual acuities were determined by the Clason Visual Acuity Meter, which allows for a more precise measurement of visual acuity by employing the psychophysical procedure of ascending and descending method of limits.[14] We could not analyze the optometric retinoscopy data since none had been reported.

DISCUSSION

After reviewing the data, we can conclude that there was a significant improvement in visual acuity, whether using Woods' or the Ewalt's data. Our evaluation of the retinoscopic data led us to conclude that Woods' data were incomplete since only 67 out of 103 cases were reported for pre- and post-training measurements. Further, the measurements were too variable to analyze properly. We could not obtain any optometric measurements of refractive error.

We then considered several possible alternate hypotheses, for the variability in the ophthalmological retinoscopic data. They are:

1. Measurement error.

The error in retinoscopy has subsequently been discussed by Hyams, Safir, and Philpot[15] and found to be +/-0.50D using an Analysis of Variance for 10 subjects (20 eyes). In addition, there was no statement about inter-observer reliability in the BMP. The lack of control for measurement error would, therefore, contribute to a certain amount of variability.

2. Bias in the measurements.

The BMP was conducted without a control group since this was not a double blind experiment.[1] This fact is most notable in Hackman's[11] analysis comparing the visual acuity data from the ophthalmologists to the data from the optometrists. The average improvement for the ophthalmologists was one line and two lines for the optometrists. Although, these biases can not be translated into diopters, the direction is clear. It is plausible to assume that if the bias were convertible to diopters, proper retinoscopic data may have been significant.

3. Actual results of the training.

Since all subjects did not receive the same amount of vision training and since motivation across subjects was also variable, the actual results of the training could account for some of the variability of the retinoscopic data.

4. Artifact of the cycloplegic refraction.

Ewalt commented quite specifically on the problem of the cycloplegic by noting reports of its long lasting effects on the pre-training visual acuity measurements.[8] He describes abnormally large pupil sizes persisting so long that ; "Judging by the prolonged effect, a powerful cycloplegic was used."[8] According to authoritative pharmacological textbooks both in medicine[16] and optometry,[17] 5% homatropine has the following effects: A. The pupil can be dilated for 5 or more days after instillation; B. The accommodative amplitude may not return to normal for more than 15 days after instillation; and C. It does not produce complete cycloplegia. Residual accommodation ranges from 0.5 to 2.00D. While atropine does produce complete cycloplegia, it effects are even longer lasting than that of homatropine, with both pupil and accommodation affected for 7 to 12 days after instillation.

While the discussion about the side effects of the cycloplegia may be speculative, they could have interfered with aspects of the training dealing with accommodative flexibility. Consequently to obtain a more complete perspective on the problem with the cycloplegic, we contacted Mr. Robert Williams, Executive Director of the Optometric Extension Program for his advice on individuals who were knowledgeable about BMP. Drs. Gerry Getman, H.Ward Ewalt and Daniel Woolf were suggested. Each was sent a letter (see Appendix B) and all promptly replied. Each one, independent of the others, felt that the cycloplegic agent was the cause of the variability of the retinoscopic measurements. Woods also commented "It is immediately apparent that the visual training had produced no change in the basic myopic error. The occasional minor changes noted are those which would result in variations of the action of the cycloplegic."[1]

We concluded, therefore, that any of the first three alternate hypotheses could produce variability in the data, although the most likely explanation was the problem with the cycloplegia. Two studies[18,19] showed that in conducting a pre- and post-test on subjects using cycloplegia, 1/3 will remain the same, 1/3 will increase in refractive error, and 1/3 will decrease in refractive error. The refractive error data from Woods shows the same distribution. The changes in refractive error due to the cycloplegia, therefore, could account for all the variability in the refractive error measurements. Consequently these data from the BMP can not be used in a statistical analysis because of the excessive variability.

While our primary purpose in writing this paper was to re-analyze the data in a proper statistical manner, we felt it is also important to point out some of the discrepancies between the ophthalmologic data and the optometric data as discussed by Shepard,[7] Betts,[9] and Hackman.[11] We were unable to obtain any additional optometric data on either refractive error or the training program, other than the letters from Drs. Ewalt, Getman, and Woolf stating their concern with the problem of the cycloplegia. We hope that this paper will be helpful in obtaining additional comments.

CONCLUSION

After the analysis of the data we can conclude that there were: A. Statistically significant changes in visual acuity; B. Variability difficulties in the refractive error data, most likely due to the cycloplegic effects, which make the data impossible to properly analyze; C. A lack of adequate protocol for the retinoscopy which could bias the measurements, and D. No consideration was given to the measurement error. This could be particularly significant for the low myopes.

In light of this we can state: Woods' conclusion that vision training is not efficacious in the treatment of myopia is not supported by the data that can be analyzed, and thereby, must be based on something other than the data. Woods' conclusion, therefore, is not valid.

APPENDIX A
EXPLANATION OF DATA SCALES

The concept that different data are on different scales is not new to optometry.[20] However, we think that it is important that it be reviewed. There are four data scales: nominal, ordinal, interval and ratio.[21] An example of a nominal scale is iris color. You would not take an average of 2 brown, 3 green and 4 blue eyes and come up with 3 gray eyes. For nominal data you can only report frequency. An example of a ordinal scale is exophoria. On a behavioral basis, we do not say that someone with a 10 p.d. exophoria need be twice as exophoric as someone with a 5 p.d. exophoria. For ordinal scales, typically the mode and frequency are reported. Interval scales have equal differences between units, but a relative zero point. An example of an interval scale is visual acuity when based on the log of the minimum angle of resolution. The arbitrary zero point is the equivalent of 20/20 Snellen and the differences between a value of 1,2 and 3 have been found to be equal. Ratio scales have an absolute zero point and equal intervals between units. An example of a ratio scale is intra-ocular pressure, where zero pressure is zero mm/Hg and a pressure of 10 mm/Hg is twice as much as 5 and half of 20.

Nominal and ordinal scales are known as non-parametric, while the last two scales, interval and ratio, are known as parametric. Non-parametric data are typically either given frequencies and/or are ranked. Parametric data can be added, squared, and multiplied. While this discussion may seem very theoretical, we have taken non-parametric data and found that if they are used inappropriately many errors of interpretation can result.[22]

After consideration of the above laws of data scales, we decided that the visual acuity data converted to minimum angle of resolution obeyed the properties of interval data,[23] and the retinoscopy data obeyed the properties of ordinal data. We, therefore, selected the Correlated T-Test for the visual acuity analysis, and the Wilcoxon Test for retinoscopy analysis because it allows a non-parametric comparison between matching pairs of data.

APPENDIX B
REQUEST FOR INFORMATION

Dear Dr.

I am writing this letter to request information regarding the Baltimore Myopia Project. Recently, Vincent Giambalvo, Ph.D. and I have re-evaluated the study and hope to have the data published in the OEP Curriculum II papers.

The analysis of the visual acuity data showed a statistically significant improvement in all the visual acuity measurements. Before analyzing the refractive error data we noticed large changes in refractive error for some patients, both increases and decreases. We suspect that some of the variance in these data are due to artifacts from the use of cycloplegic agents.

We would appreciate any insights that you may have regarding the variability of the refractive error data. Please let me know if you have any questions.

REFERENCES

1. Woods A. Report from the Wilmer Institute on the results obtained in the treatment of myopia by visual training. Am J Ophthalmol, 1946;29(1):28-57.
2. Balliet R, Clay A, Blook, K. The training of visual acuity in myopia. J Am Optom Assoc, 1982;53:719-24.
3. Grosvenor T, Maslovitz B, Perrigin D, Perrigin J. The Houston Myopia Control Study: A preliminary report by the patient care team. J Am Optom Assoc, 1985;56(8):636-43.
4. Graham C, Leibowitz H. The effect of suggestion on visual acuity. Int J Clin Exp Hypnosis, 1972;20(3):169-86.
5. Press L. Myopia - Annual Review of the Literature, 1986. J Opt Vis Dev, 1987;18:1-17.
6. Grosvenor T, Flom MC. Refractive anomalies. Boston:Butterworth-Heinemann, 1990:348.
7. Shepard C. The Baltimore Project. Optom Week, 1946;37(5):133- 35.
8. Ewalt H. The Baltimore Myopia Control Project. J Am Optom Assoc, 1946;17(6):1-22.
9. Betts E. An evaluation of the Baltimore Myopia Control Project--Part A: Experimental Procedures. J Am Optom Assoc, 1947;18(4):481-85.
10. Hildreth H, Meinberg W, Milder B, Post L, Sanders T. The effect of visual training on existing myopia. Transactions of the Am Acad of Ophthalmology and Otolaryngology, 1947;30:260-77.
11. Hackman R. An evaluation of the Baltimore Myopia Control Project--Part B: Statistical Procedures. J Am Optom Assoc, 1947;18(3):416-26.
12. Siegal S. Nonparametric statistics for the behavioral sciences. New York:McGraw-Hill, 1956.
13. Sidman M. Tactics of Scientific Research. New York:Basic Books, 1960.
14. Borish I. Clinical refraction. 3rd ed, Chicago:Professional Press, 1970.
15. Hyams L, Safir A, Philpot J. Studies in refraction. Arch Ophthalmol, 1971;85(1):33-41.
16. Gilman AG, Goodman LS, Rall TW, Murad F. Goodman and Gilman's The pharmacological basis of therapeutics. 7th ed, New York:MacMillan Pub Co, 1985.
17. Bartlett JD, Jaanus SD. Clinical ocular pharmacology. 2nd ed, Stoneham, MA:Butterworths, 1989.
18. Rengstroff R. Observed effects of cycloplegia on refractive findings. J Am Optom Assoc, 1966;37(4):358-59.
19. Ludlam W, Weinberg S, Twaroski C, Ludlam D. Comparison of cycloplegic and non-cycloplegic ocular component measurement in children. Am J Optom, 1972;49(10):805-17.
20. Pierce J. Research design, statistics, and interpretation. Duncan, OK:Optom Extension Prog, 1968.
21. Stevens S. On the psychophysical law. Psychol Rev, 1957;64:153-81.
22. Trachtman J, Giambalvo V, Dippner R. On the assumptions concerning the assumptions of a t-Test. J Gen Psychol, 1978;99:107-16.
23. Westheimer G. Scaling of visual acuity measurement. Arch Ophthalmol, 1946;29(1):28-57.

Corresponding author:
Joseph N. Trachtman, O.D., Ph.D.
Institute for Advanced Vision Technology™
26 Schermerhorn Street
Brooklyn Heights, NY 11201
Date Accepted for publication:
January 16, 1991