Advances in modelling visual symptoms and visual skills

M K Powers¹ and W P Fisher²

¹Gemstone Foundation, Tarzana, CA, USA,
²University of California, Berkeley, CA, USA

E-mail: maureenpowers@gemstonefoundation.org

Abstract. Though it seems to be a natural process of physical development, not everyone learns on their own how to use their eyes effectively. Poor binocular vision skills negatively impact a wide range of daily life activities, including reading. This study extends previous work using existing data in multidimensional causal Rasch models of vision, Optometry, survey, and reading assessment data are evaluated in a common frame of reference. Thirty-three students performed vergence and tracking exercises in about 12 to 38 sessions each involving a sequence of five computerized training modules. The number of training modules experienced in any given session ranged from one to five. There were 4,064 total measured cases, with about 123 measures per student. Paired comparison t-tests of early and late measures show significant gains in vision quality. Two regression models predict reading scores on the basis of vision data. Results support further study of how training in Functional Binocular Vision might contribute to improved outcomes in reading education.

1. Introduction

Functional Binocular Vision (FBV) is a construct we have previously shown to be related to a set of quantifiable items on a symptom questionnaire (the Convergence Insufficiency Symptom Survey, CISS) [1, 2]. Our findings have established FBV as a separate, measurable component of the visual process that is not related to measures of acuity (the ability to see the eye chart [2]. Because many decisions about one’s visual ability are currently based on measures of acuity alone, further investigation of the relationship between FBV and additional outcome measures becomes important.

Reading is a fundamental life skill that impacts nearly everyone but has been difficult to relate directly to binocular vision. Reading is not an automatic skill: Both the neuromuscular and cognitive systems must develop over time to provide optimal input to the brain for comprehension of text, and not all people achieve good reading performance. In particular, we know that many students with poor binocular skills also read poorly [3, 4], and at least one study shows that degrading binocular vision by inserting prisms between the eyes of good readers and their reading material (thus stressing the convergence system) can reduce reading rate [5]. One hypothesis that has developed from observations like these is that training visual skills can improve reading performance.

In this work, we take advantage of an existing data set where both vision and reading variables were measured in grade school students over time during a computerized visual skills training course [6]. By examining changes in vision variables over time (during training) and comparing their change to that in

¹ To whom any correspondence should be addressed.
reading outcomes (measured pre- and post- training), we hoped to provide new insight into the nature and strength of the relationship between binocular vision and reading.

2. Study Design
This study explores possibilities for productively modelling aspects of vision using a probabilistic approach prioritizing the estimation and comparison of linear measures [7-9] explained by causal theory [10], which may be combined together in a more complex model of multiple constructs [11]. Visual optometric measurements, survey ratings, and reading assessment data are evaluated in a common frame of reference, with the aim of determining the extent to which an overall model of FBV might be feasible. Even though the data available for study were not assembled with this explicit aim in mind, they appear to justify a preliminary exploration of the possibility that further study is warranted.

2.1. Sample and training modules
Thirty-three students with binocular vision issues were enrolled in a visual skills course to train their eyes in vergence and tracking exercises hypothesized to strengthen the neuromuscular coordination needed for FBV. Each session of the course involved a sequence of five computerized training modules (Table 1). Students participated in from 1 to 50 sessions. Setting aside two students participating in three or fewer sessions, and one participating in 50, the remaining 30 participated in 12 to 38 sessions, with means of about 25 and standard deviations of 6 across the five modules.

| Number | Module name                | Module purpose                        |
|--------|----------------------------|---------------------------------------|
| 1      | Fast Focusing              | Improve accommodative facility        |
| 2      | Smooth Tracking            | Improve smooth pursuit speed and accuracy |
| 3      | Jump Tracking              | Improve saccadic eye movement speed and accuracy |
| 4      | Cross-Eyed Fusion          | Improve convergence range             |
| 5      | Wall-Eyed Fusion           | Improve divergence range              |

The number of training modules experienced in any given session ranged from one to five, with the sequence always adhering to the same order of administration. If only one module was administered, it was always the same one, and so on. Of 862 administrations of the first module, the second module was administered 827 times. Of those 827, the third was administered 810 times; of those, the fourth was administered 792 times; and the fifth, 773 times. Thus, for the 33 students across all sessions and all modules, there were 4,064 lines of data.

2.2. Reading measures
The ability to read words aloud for a sustained period of 1 minute was measured using a standardized grade-level-based battery of 1-page stories where each grade level had several stories, in order to prevent memory interference. Stories were sufficiently long for each grade that no student could finish in 1 minute. Students were told to read out loud as rapidly as possible but to try not to make mistakes. Use of fingers was not allowed. A stopwatch was used to start and stop reading, and the total number of words and the number of errors (words added, omitted, or misread) was recorded.

2.3. Original data structure and basis for ratings
Ten indicators were tracked by the computer during the administration of the training modules (Table 2). Data were stored as numeric counts of seconds, as percentage fractions to four decimal places, and as other kinds of computed scores occurring in various numeric ranges. Frequency charts and histograms were produced for each of these, with each distribution broken out into eight equipercentile ranges. The values in each range were then assigned ratings from 1 to 8, where 1 was associated with the value indicating the worst possible performance, and 8, the best. One item's top percentile range was confined in a range between 0.99990 and 1.00000; scoring went to four decimal places, so these are set at 7.
2.4. Form of the measurement model

The measurement model formalizes the expectation of a unit quantity, hypothesizing an invariant relation between visual function and challenges:

\[ \ln\left(\frac{P_{nij}}{(P_{nij} - 1)}\right) = B_n - D_{ij} \] (1)

which says the log-odds of any student responding within the range of seconds represented by any given category on any visual challenge is equal to the difference between the visual function \( B \) of student \( n \) and the visual challenge \( D \) of item \( i \) at the level of category \( j \), relative to the level of category \( j-1 \). [8-9]. Data quality is evaluated relative to the model via information-weighted and outlier-sensitive mean square statistics, both with expected values of 1.00, and in terms of principal components analyses of the standardized residuals [12-13].

2.5. Expected reliability

With nine eight-category items, and one seven-category item, Rasch generalizability theory [14] predicts a measurement uncertainty of about 0.3. Reliability is, of course, a function of the ratio of true variation to uncertainty [15]. However, these data are not a calibration sample chosen with careful attention to variation in the construct of interest. The range of variation is therefore restricted by the study's focus on students with identified vision and reading problems. A low ratio of variation to uncertainty, and low reliability coefficients, can therefore be expected.

As shown in Table 2, the first three items are included in all five modules, and so they serve to connect the 4,064 by-module-by-session by student measures in a common frame of reference.

Table 2. Items by module.

| Number | Code | Focus                | M1 | M2 | M3 | M4 | M5 |
|--------|------|----------------------|----|----|----|----|----|
| 1      | AHT2 | Average Hit Time     | X  | X  | X  | X  | X  |
| 2      | BHT2 | Best Hit Time        | X  | X  | X  | X  | X  |
| 3      | OHP2 | Overall Hit Percent  | X  | X  | X  | X  | X  |
| 4      | THR2 | Tracking Hit Rate    |    |    |    |    |    |
| 5      | RTN2 | Red Transition Number| X  |    |    |    |    |
| 6      | BTN2 | Blue Transition Number|   | X  |    |    |    |
| 7      | RTA2 | Red Transition Average|  |    | X  |    |    |
| 8      | BTA2 | Blue Transition Average|  |    |    | X  |    |
| 9      | MB2  | Maximum Break        |    |    |    | X  | X  |
| 10     | MR2  | Maximum Recovery     |    |    |    |    | X  |

3. Results

3.1. Measurement analysis

Three measures were at the minimum score extreme; there were none at the maximum. Model fit statistics appeared satisfactory, and the mean measurement uncertainty was 0.29, near the expected 0.3. The true standard deviation (0.41) and measurement separation reliability were low (0.67), as expected.

Principal Components Analysis (PCA) of the standardized residuals shows more unexplained variance captured in the first three contrasts (41%) than is explained by the measures (38%). Contrast eigenvalues and loadings are high, and the measures implied by the different subsets of contrasting items have low disattenuated correlations. The analysis separates the items into contrasting groups by the module they appear in, suggesting quite different constructs are being measured [12, 13].

Anm’;lsdms’esdkmother analysis racking the data across modules within sessions was also completed, but given the by-module significance of the PCA results, these five groups of measures were used.
3.2. Regression analyses

The repeated measures from each of the five module administrations were aggregated into four sets of mean values by averaging each module's measures within ranges of 1-7, 8-12, 13-19, and 20 or more sessions. The number of measures summarized in the 20 cells of the five modules by the four session groups ranged from about 150 to about 260, with about 750 to 1,250 measures within a session group (Figure 1).

![Bar Chart](Image)

**Figure 1.** Numbers of measures by module by session group.

The 20 by-session, by-module means per student were then used in conjunction with previously reported [1, 2] pre- and post-intervention optometric measures of the same 33 students, and from the Convergence Insufficiency Symptom Survey, to predict oral reading fluency. The latter was measured via a log-odds transformations of counts of words read aloud correctly and incorrectly in one minute.

Of the 33 students, 28 had module 1, 2, and 3 mean measures in both the earliest and latest session groups (1 and 4). These were significantly different at a .001 level in a paired comparison t-test for module 1 (t=-6.05, 27 df, p=.000), but not for module 2 (t=-2.41, 27 df, p=.023), or module 3 (t=-.30, 27 df, p=.76). Paired comparison t-tests also showed significant differences for module 4 (t= 5.46, 26 df, p=.000), and module 5 (t=-4.47, 25 df, p=.000). This shows that changes in mean optometric and symptom measures occurred during the training course.

Of the 33 students, 23 had reading measures. In the first regression model, measures from module 1 in session groups 1 and 2, and modules 4 and 5 in session group 1, predicted reading fluency (R=0.825, Rsq=0.68, adj. Rsq=0.61, SE=0.60, F change = 9.56, 4 df1, 18 df2, p=0.000). Figure 2 shows a scatterplot of reading fluency in logits and the regression standardized predicted value.
Figure 2. First model, 23 students' measures from module 1 in session groups 1 and 2, and modules 4 and 5 in session group 1 predicting reading fluency.

In the second model, for 16 students with visual acuity and reading measures, changes from the first to the fourth groups of sessions in modules 1, 4, and 5, along with changes in visual acuity, predicted reading fluency ($R=0.767$, $R^2=0.589$, adj. $R^2=0.44$, $SE=0.74$, $F$ change $=3.94$, 4 df1, 11 df2, $p=0.032$). Figure 3 shows a scatterplot of reading fluency in logits and the regression standardized predicted value.

Figure 3. Second model, 16 students' changes from the first to the fourth groups of sessions in modules 1, 4, and 5, along with changes in visual acuity, predicting reading fluency.
4. Discussion
As expected from the results of the paired comparison t-test, the larger changes in visual capacity indicated in modules 1, 4, and 5 were associated with reading fluency, while the measures from modules 2 and 3 were not. Further study will be required to understand these differences; however they suggest that training accommodation and vergence skill may be more important for improving reading fluency than training tracking skills. To the extent that symptoms of vergence problems (on the CISS) involve blurring due to offset visual images when the ability to coordinate the eyes during reading, these results may indicate that poor vergence skill can have a larger impact on reading fluency than poor tracking skills. Whatever the precise mechanism, these results suggest that training in Functional Binocular Vision (FBV) can impact outcomes in reading education.

Broader interpretation suggests implications as well for current practices of vision screening in schools. Most states in the USA still use some form of the eye chart as their primary screening tool [16]. These instruments can be very good at identifying students with refractive error [17]. However, they are administered monocularly at a distance and thus cannot detect binocular problems. New instruments, perhaps similar to the CISS symptom survey, which are easy to administer and interpret, could be developed with more knowledge of their relationship to FBV.

Finally, a major benefit of the current approach is its quantitative nature. Vision measures taken in the clinic or field can vary widely due to different definitions and practices. When the measures are carefully defined and uniformly applied, not only are the characteristics of any particular variable easier to determine, but valid relationships among them can emerge.

5. Conclusion
The results overall indicate that continued investment of resources in the design and execution of a calibration study focused on better powered tests of an overall model of FBV will pay significant returns.

Acknowledgments
The data analyzed here were collected with support from National Institutes of Health research grant R01 EY07414 to MP.

References
[1] Powers M, Fisher W P Jr and Massof R W 2016 J. Phys.: Conf. Ser. 772 012045
[2] Powers M and Fisher W P Jr 2018 J. Phys.: Conf. Ser. 1065 132009
[3] Grisham D, Powers M and Riles P 2007 Optometry 78 542-9
[4] Powers M, Grisham D and Riles P 2008 Optometry 79 228-34
[5] Powers M, Miner G L, Morita Y, and Tyler C W 2010 Invest. Ophth.Vis. Sci. 51 3068
[6] Powers M 2006 J. Mod. Optic. 53 1313-23
[7] Rasch G 1980 Probabilistic models (Chicago: University of Chicago Press)
[8] Wright B D and Masters G N 1982 Rating scale analysis (Chicago: MESA Press)
[9] Andrich D 1978 Psychometrika 43 561-73
[10] Stenner A J, Fisher W P Jr, Stone M H and Burdick D S 2013 Front. Psychol. 4 1-14
[11] Briggs D and Wilson M 2003 J. Appl. Meas. 4 87-100
[12] Smith R M 1996 Struct. Equ. Modeling 3(1) 25-40
[13] Linacre J M 2008 Rasch Meas. Trans. 22 1164
[14] Linacre J M 1993 Rasch Meas. Trans. 7 283-4
[15] Fisher W P Jr 2008 Rasch Meas. Trans. 22 1160-3
[16] https://nationalcenter.preventblindness.org/school-requirements-childrens-vision March 7 2019
[17] Tong L, Saw S-H, Tan D, Chia K-D et al. 2002 Optometry Vision Sci. 79 650-7