Meta-Analysis of Eye Movements Between Children With and Without Dyslexia or Reading Disability

Stanley W. Hatch, OD, MPH, FAAO
Salus University College of Optometry
Philadelphia, Pennsylvania

ABSTRACT

Background
Hundreds of studies have compared eye movement variables in subjects with and without dyslexia or reading disability. Most studied only small sample sizes and the eye movement tasks and targets varied. The aims of this study are to determine which, if any, eye movement variable(s) differ between children with and without dyslexia or reading disability, and, if differences are found, quantify the amount.

Methods
Search engines PubMed and Salus/EBSCO Discover Database for key words eye movements OR saccades OR fixation AND dyslexia OR reading disability yielded 728 titles. Following initial study eligibility criteria (objective eye movement variable measures of children age 6-15.5 years in defined case and control groups), 43 studies qualified for in depth review. Eleven studies qualified for data synthesis. Data were extracted, tested for normality by Kolmogorov-Smirnov statistic, standardized, weighted by sample size, tested for homogeneity by Q test, pooled and measured for combined effect.

Results
Combined relative risk effect revealed fixation duration, number of fixations, and number of regressions when reading words to be 2.33 (95% CI: 2.12-2.54) times longer, 1.58 (95% CI: 1.52-1.65) times higher, and 1.83 (95% CI: 1.68-1.97) times higher respectively in children with dyslexia or reading disability compared to age normal readers. Differences reading pseudowords were not statistically significant.

Conclusions
Significant differences in fixation duration, number of fixations, and number of regressions were found during word reading. Because most reading time is during the fixation duration, children with dyslexia or reading disability need, on average, 2.33 times longer. The results provide objective data to support reading time accommodations for Individual Education Plans. Systematic review suggests that oculomotility ability depends on the amount of cognitive processing rather than purely on extra ocular muscle control.

INTRODUCTION
Oculomotility (eye tracking, saccades, eye movements, reading eye movements) represents the process where the eyes fixate on a target, usually a word or series of words, then move to fixate on another target, usually the next word or series. During fixation, words are processed by phonetic decoding and/or orthographic sight word recognition. Reading...
ability, fixation stability, context and subject knowledge determine the length of fixation and distance to the next fixation.\textsuperscript{1}

Oculomotility is measured clinically by three general methods. The first is observation where the examiner asks the subject to look back and forth between targets. The examiner grades the accuracy of saccades by the number of under shoots, over shoots, and head movements.\textsuperscript{2,3} The second is standardized visual-verbal naming tests such as the King-Devick (King-Devick Technologies, Oakbrook Terrace, Illinois) or the Developmental Eye Movement Test™ (Bernell Corp, South Bend, Indiana). Criteria for oculomotility or tracking dysfunction are one to two standard deviations below the age mean for speed and errors after controlling for cognitive speed. The third is objective instrument recordings during reading or other eye tracking tasks (RightEye™, Bethesda, Maryland; Visagraph™, ReadingPlus, Winooski, Vermont, optical coherence tomo graphy, and custom devices). Objective instruments measure time and distance of eye movements in milliseconds and milliseconds per degree.

The optometric, medical, and psychoeducation literatures contain thousands of papers on eye movements, especially on eye movements and reading. An early interest was whether children with dyslexia or other reading disability demonstrated abnormal eye movements and whether simple eye movement tests could identify them. Studies in the late 1970s and early 1980s reported contradictory findings. Pavlidis published works showing that children with dyslexia performed more fixations and more regressions when reading words as well as tracking a sequence of lights.\textsuperscript{4,5} Lefton et al.\textsuperscript{6} found similar results, but at least four other studies found no differences between subjects with dyslexia and controls.\textsuperscript{7-10} Simons concluded there to be insufficient evidence to consider oculomotor dysfunction a unique entity.\textsuperscript{11}

These contrasting results and opinions highlight the competing theories: are eye movements simply a reflection of reading ability or does oculomotor control influence reading?

It is important to note that the above studies define dyslexia as simply two grades below expected grade level on standardized reading tests in the presence of normal intelligence and education exposure. They did not use or investigate for the option of dyslexia as one of several types of reading disabilities, e.g. dyseidesia, orthographic problem, or dysgraphia. To illustrate this conundrum, Blythe et al.\textsuperscript{12} debated Stein\textsuperscript{13,14} on the role of visual pathways, particularly the magnocellular pathway, in the behavioral, physiological, and etiological aspects of developmental dyslexia. They both discuss dyslexia as one problem instead of considering that a reading disorder could be due to a decoding issue and/or magnocellular pathway issue. It might be better, therefore, to substitute the term reading disability for dyslexia in the present paper.

Though many studies have investigated this topic, most used small sample sizes and measured many variables. Meta-analysis is a statistical technique for combining data from similar studies to increase power. The specific aims of this study were to determine which measurements of eye movement variables are available for data pooling and perform meta-analysis to quantify effect size and test for statistical significance.

**METHODS**

The study was approved by the Salus University Institutional Review Board as category 104(d.4.i) exempt. This study limited eye movement variables to fixations and saccades since they are the primary oculomotor activity in reading. Objective instrument measure was chosen because fixation duration, number of saccades, number of regressions, and associated variables move at the millisecond or millimeter level and cannot be accurately assessed without laser or infrared tracking devices. To pool data for meta-analysis, several intermediate steps were necessary. First, a literature review was
conducted to determine the instrumentation and methods used by different investigators. Once identified, statistical tests for homogeneity were performed to determine if the measures were likely the same variables and populations. Second, specific literature search protocols sampled key words, title words, and abstracts to collect data for secondary analysis. Articles meeting subject and study design but not data criteria were assessed qualitatively. Articles meeting inclusion and exclusion criteria for data synthesis or meta-analysis provided data for quantitative analysis.

Meta-analysis was conducted by standard methods recommended by PRISMA and others.15-17 Multiple electronic search engines were utilized during the study period from March to September 2019. The two primary search engines, National Library of Medicine (U.S.) PubMed and Salus/EBSCO Discovery Database, were accessed March 22, 2019, between 3:00 and 3:38 PM U.S. Eastern time. The Salus/EBSCO search engine includes over 160 journals. It searches journals in optometry, audiology, speech and language, education, and occupational therapy some of which are not included in PubMed. All fields key word criteria were eye movements, saccades, dyslexia, and reading disability. Search settings required at least one of eye movements OR saccades AND at least one of dyslexia OR reading disability. Later searches added the term fixation. Both singular and plural, e.g. saccade and saccades, yielded the same results. Additional searches with Educational Resource Information Center (ERIC) and PsycINFO were used in June and September 2019 to corroborate and expand the search to include the possibility of psychology, education, and other journals that might not have been indexed in the first two search engines. In addition, the Journal of Behavioral Optometry, Optometry and Visual Performance, and Vision Development and Rehabilitation were searched manually via each journals’ website index function or by the author reviewing each digitally available annual index.

Studies were coded with the Newcastle-Ottawa Scale for assessing the quality of nonrandomized case-control studies in meta-analyses.18 The case-control scale was chosen since the design is the closest equivalent to cross sectional comparison studies. A check list (Table 1) was developed to assure uniform inclusion and exclusion criteria and then quality ranking.

<table>
<thead>
<tr>
<th>Table 1. Study assessment form with inclusion criteria based on the Newcastle-Ottawa Scale (NOS).18</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Methods/Exposure (Maximum 4 stars)</strong></td>
</tr>
<tr>
<td>□ Objective eye movement measure by automated instrument*</td>
</tr>
<tr>
<td>□ Same for cases and controls</td>
</tr>
<tr>
<td>□ Variable means and standard deviations or data which allowed conversion published*</td>
</tr>
<tr>
<td>□ All cases and controls completed each measure*</td>
</tr>
<tr>
<td>□ Description of test targets sufficient to categorize as reading words, pseudowords, or objects*</td>
</tr>
<tr>
<td><strong>Selection (Maximum 4 stars)</strong></td>
</tr>
<tr>
<td><strong>Case definition</strong></td>
</tr>
<tr>
<td>□ Reading level at least 1 standard deviation below age expected on standardized reading test performed in the study*</td>
</tr>
<tr>
<td>□ Normal IQ tested or confirmed by records to be within 2 standard deviations of mean on standardized test*</td>
</tr>
<tr>
<td><strong>Control definition</strong></td>
</tr>
<tr>
<td>□ Controls selected from same community. Note: if cases selected from a school for learning disabled children, controls would have to be selected from the same geographic area.*</td>
</tr>
<tr>
<td>□ Reading level at or above mean age expected on standardized reading test performed in the study*</td>
</tr>
<tr>
<td>□ Normal IQ tested or confirmed by records to be within 2 standard deviations of mean on standardized test*</td>
</tr>
<tr>
<td><strong>Comparability (Maximum 2 stars)</strong></td>
</tr>
<tr>
<td>□ Study controls for vision conditions by optometrists or ophthalmologists performing eye exams and excluding subjects with clinically significant uncorrected vision problems*</td>
</tr>
<tr>
<td>□ Study controls for other factors which could influence reading ability such as developmental deficits, neurologic conditions, or psychiatric conditions by reviewing medical records*</td>
</tr>
<tr>
<td><strong>Study Assessment</strong></td>
</tr>
<tr>
<td>□ At least 9 stars (*) = strictest criteria</td>
</tr>
<tr>
<td>□ At least 5 stars (*) = relaxed criteria per NOS.</td>
</tr>
</tbody>
</table>
RESULTS

All fields search returned 436 citations in PubMed and 395 in Salus/EBSCO Discovery Database (Figure 1). Twenty-six duplicates were removed. An additional 77 citations designated as book chapters, symposia, transactions of societies, or missing title were not screened. ERIC, PsycINFO, and individual journal reviews yielded only duplicates of previously detected articles or studies unrelated to the screening criteria. Records screened totaled 728. Case reports, language other than English, editorials, and retrospective studies were excluded. Abstracts of remaining citations were screened for the following inclusion criteria: objective measure of at least one saccadic eye movement variable, sample size greater than 3, and presence of control group. Full text articles with abstracts meeting initial inclusion were reviewed in detail for meta-analysis eligibility or qualitative literature review.

Forty-three articles qualified for in depth study. Twenty-three met the general study design criteria. Eleven of those met meta-analysis criteria. Five of the 11 articles satisfied the Newcastle-Ottawa strongest criteria with 9 or 10 stars. Six of the remaining 11 met moderate strength criteria of at least 5 stars. These six all rated 7 stars. Reasons for not meeting strongest criteria included: controls within one standard deviation below the mean on standardized reading tests, not confirming normal IQ in the control groups, and not performing vision exams or did not specify normal neurologic health by history. These 6 studies were included in the meta-analysis for reasons described next. The first two issues involved only controls. If controls in some studies had slightly below average reading ability or lower IQ, any bias would be towards the null hypothesis. Regarding the third issue, not validating normal vision and neurologic health, it seems significant impairments would have been obvious and subjects with such impairments would not have been diagnosed with dyslexia as the primary cause of reading disability and would naturally have been excluded. However, confounding factors such as small refractive errors, binocular and accommodative problems could have been missed. These are addressed later.

The first objective was to identify the eye movement variables measured by various labs. Oculomotility variables reported by at least 5 qualifying studies included fixation duration, number of saccades, and number of regressions for word reading and pseudoword reading. Excluded from analysis due to small sample size were the following variables reported in 4 or less qualifying studies: saccadic reaction time right and/or left, saccadic angle, mean peak velocity, number of right saccades, number of left saccades, mean overlap right and/or left saccades, mean gap overlap, saccadic reaction time gap and overlap, saccadic amplitude right and/or left, predictive saccadic time right and left, predictive saccadic learning right and left, and saccadic latency right and left.

The second objective was to determine if the variables identified across studies measured the same functions. Several statistics tests are available. For normally distributed variables, the Dixon Q test is widely accepted, but before applying the Q test, the data had to be tested for normality. Outcome variables were grouped based on fixation target- words, pseudowords, or objects. Only variables where at least five studies met inclusion criteria were incorporated in the analysis. Table 2 lists the eye movement variables from qualifying studies, their tests for normality (Komologrov-Smirnov statistic and skew), and Q statistic (random effects model) for homogeneity. All variable distributions were consistent with the normal distribution and combined effects met criteria for homogeneity for meta-analysis. Some would argue that number of regressions has a higher than accepted skew but given the K-S corresponding P=.29, this author felt criteria for normality was met. The tests for homogeneity reveal there were no statistical
both a risk ratio and combined mean difference between groups. Combined effect was calculated by matched pair correlated t test- each study control group matched to its case group. Since all studies were cross sectional designs, the derived risk ratios could be treated statistically as odds ratios.

Meta-analysis proceeded for each outcome variable where reported by at least 5 studies. Each study’s mean ratio was weighted based on total number of subjects (cases + controls). Data were pooled and mean risk ratio calculated and tested for significant difference from relative risk of 1.0.

Table 3 shows the combined risk factor difference, e.g. X times higher cases to controls, with 95% confidence intervals for each variable. A risk ratio greater than 1.0 indicates X times longer time of fixation duration or X times higher number of saccades and number of regressions. Confidence intervals that do not include 1.0 indicate a statistically significant difference.

In addition to measuring the risk ratio, fixation duration, the one variable independent of reading task length, was also treated as a continuous variable for which means from each study could be pooled. Weighted pooled mean difference for fixation duration was

<table>
<thead>
<tr>
<th>EYE MOVEMENT VARIABLE</th>
<th>Word reading</th>
<th>Pseudo-word reading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of fixations</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Fixation duration</td>
<td>8</td>
<td>&lt;5</td>
</tr>
<tr>
<td>Number of regressions</td>
<td>8</td>
<td>&lt;5</td>
</tr>
</tbody>
</table>

Table 2. Variables, number of studies (n), Komologov-Smirnov (K-S) statistic (associated P value), skew, and Q statistic (associated chi square P value). P values >.05 were considered consistent with the normal distribution (K-S, Skew) and homogeneity (Q) across studies. NA = not applicable since n<5.
calculated to be 17.72 Msec between cases and controls reading words (n=8 studies). Matched pair correlated t test for weighted pooled mean difference satisfied the criteria for statistical significance (t=2.50, P=.04).

**DISCUSSION**

Pooled data demonstrated a statistically significant 2.33 times longer average fixation duration in children with dyslexia or reading disability compared to age expected controls when reading words. Combined mean difference of fixation duration, raw data when analyzed as continuous, was also significantly different. Children with dyslexia or reading disability had, on average, a statistically significant 1.58 times higher number of fixations and 1.83 times higher number of regressions during word reading. Pseudoword reading did not elicit statistically significant differences. Object or light tracking were not reported by enough studies to include in meta-analysis.

The majority of reading time is spent during the fixation. Saccade eye movements between fixations last about 20 Msec while fixation duration times range from 300 to 1100 Msec depending on reading ability. Patients whose fixation duration is two times longer are at significant disadvantage. On average they would require more than 60 minutes to read the same material their classmates complete in 30 minutes. When viewed from this perspective, an Individualized Education Plan which grants a student with reading disability time and half for reading, may not be adequate. The average student with reading disability needs at least two times longer and fifty percent of students may require longer.

Twelve studies with valid designs and statistical analysis could not be included in the meta-analysis due to different test statistics reported. Most of these studies reported analysis of variance summary table without the necessary descriptive data. Five of 6 studies using word text found significant differences in eye movement variables though one noted differences in only low frequency words. Compared to studies analyzing object or light tracking, only 2 of 7 found significant differences in eye movements. One study found a difference above age 9 but no difference below age 9. These results and the data presented in the present study support the view that eye movements differ when the visual task requires reading but not when tracking pseudo words, objects or lights.

Returning now to the question, are eye movements primarily a reflection of reading ability or are there cases where oculomotor dysfunction impairs reading? Overall, the evidence suggests that reading ability leads reading eye movement skill, but several exceptions have been cited. The exceptions might be explained by variability of subjects. The papers reviewed for this study almost exclusively used the term dyslexia in the general sense— for any healthy child with normal intelligence and educational exposure who tested two grades below expected in reading. The optometric literature generally separates dyslexia as a language based neurologic dysfunction impeding the decoding of visual symbols into their phonetic components. This acknowledges the potential for other subclasses of reading disability including visual perceptual causes, memory deficits, and magnocellular pathway dysfunction to name a few. Neuropsychologic, electrodiagnostic, anatomic, and imaging studies have found differences in perceptual abilities, magnocellular pathway, as well as language processing areas. It is unknown if these occur as a continuum of one condition, or they represent separate and often comorbid conditions. If reading disability subtypes are eventually delineated, it will be interesting to see if eye movement patterns correlate with subtype. In theory, if dyslexia (language based decoding impairment) and orthographic dysfunction (visual recognition based impairment) exist separately, then visual confusion during saccades might be the basis...
of erratic eye movements and poor reading in an orthographic impaired group.

Limitations

Meta-analyses need to address the potential of publication bias. Publication bias exists when negative studies, e.g. results which support the null hypothesis, are not published as frequently since a new etiology or effect was not found. Though impossible to quantify what is not known, publication bias is probably very low for this topic. The relationship between vision and reading disability, especially eye movements and dyslexia, has been controversial between and within medicine, optometry, education, neuroscience, and psychology for over 50 years. As previously discussed, the number of studies reporting positive and negative results were equivalent. Without a preponderance of studies published in favor of or against the theories of interest, publication bias is unlikely.

Overall, these efforts should satisfy the standards for minimizing sampling error and addressing publication bias. However, it is acknowledged that the exclusion of non-English language articles and articles in non-peer reviewed publications could introduce some, hopefully negligible, selection bias.

Regarding measurement variability, each study had different numbers of trials per subject and different number of words or objects to read or track. Longer reading demand could confound eye movement outcomes. Weighting factors were based on number of subjects. Thus, given equal number of subjects, a study with 4 trials was weighted the same as a study with 20 trials. Instruments and instructions could have varied across studies as could print size, text or object spacing. Studies using words varied in reading material. Some presented individual words, others presented grade level text, and others compared performance between common words and uncommon words. Whether these differences lead to differential misclassification (bias towards one result) or only contributed to random misclassification cannot be determined. However, if some study methods inadvertently biased towards a positive result, it seems just as likely others would bias towards a negative result.

The decision to use studies meeting the NOS moderate strength (at least 5 stars) rather than only the highest quality studies (at least 9 stars) deserves attention. All six studies categorized as moderate quality achieved 7 stars. Reasons for not achieving 9 stars were: not confirming normal IQ in controls, including controls with reading levels within one standard deviation below the mean, or assuming normal vision and neurologic history rather than confirmation by examination. The inclusion of some readers slightly below the mean expected reading level might bias the results towards the null and thus would have diluted rather than exaggerated these results. However, it is conceivable that the 3 studies that did not confirm IQ in controls could have had above average IQ readers in the control group. This might have lead to greater differences in eye movements. It can be pointed out, though, that when classifying a child as having a reading disability, the comparison is made to all other children in the grade, not just those with an IQ between 70 and 130. Thus, not evaluating IQ in the control group is more reflective of a school population because at or above expected students do not routinely receive IQ testing. The third issue, studies not performing eye exams or confirming normal neurologic health, could introduce confounding. None of the 6 studies ranked 7 out of 10 stars included complete eye exams. Two simply confirmed normal corrected visual acuity. One screened by parental history. The other three did not report any vision measures. It is fairly well accepted that children with reading disabilities have higher prevalence of binocular and accommodative dysfunctions, though there are exceptions. If vision problems were disproportionately present in some dyslexia or reading disability
groups, the results found in this study could be due in part to other vision problems.

CONCLUSIONS
These results can be applied directly to the education system in addition to providing scientific insights. Fixation duration and thus reading time was found to average 2.33 times longer for children with reading disability. This suggests the average child with reading disability needs an accommodation for reading time of at least two times, and others may need more. The objective measures provide experimental evidence for educators to justify reading time accommodations in Individual Education Plans. Scientifically, these data support the view that most differences in eye movements between children with reading disability and age expected readers occur after visual inputs connect to brain centers involved in symbol recognition, decoding, and integration, whether phonologic or orthographic, rather than direct visual motor processing. Oculomotor therapy, therefore, should be integrated with phonologic and orthographic remediation if the goal is improved reading ability.

REFERENCES
AUTHOR BIOGRAPHY:

Stanley W. Hatch, OD
Philadelphia, Pennsylvania

Dr. Hatch currently serves as Chief of the Pediatrics and Binocular Vision service at Salus University, Pennsylvania College of Optometry. He received his Doctor of Optometry from Michigan College of Optometry in 1990 and completed his residency in pediatrics and vision therapy at the Pennsylvania College of Optometry under Dr. Mitchell Scheiman in 1991. He received a Master of Public Health from Harvard Chan School of Public Health in 1995, he is a Diplomate in the American Academy of Optometry and a recent Fellow of the College of Optometrists in Vision Development.
VT on the Edge Youtube

What do you get when you combine a developmental optometrist, an optometric vision therapist, and a professional film crew? The world’s first vision therapy youtube show!

VT on the Edge is designed to be a place for the VT community. We discuss the hows and why’s of vision therapy, guest host VT’s brightest minds, discuss what is happening in the world of vision therapy, and have a blast during the process.

Other things to look forward to:
• Vision therapy activity demonstrations and pearls
• Vision Therapist tricks of the trade
• Practice Management tips
• Dr. Winters losing a competition and having to eat the hottest Buffalo Wild Wings while trying to do the Hart Chart

Please tune in by subscribing to our Youtube channel - Emergent VT

Online Vision Therapist Training Course

Designed to provide foundational concepts of vision therapy. This curriculum is ideal for new Vision Therapists or therapists preparing for their COVT boards.

The course has over 50 videos with 45 comprehension checks.

The course also includes:
• Access to a COVT mentor
• A final written exam
• A final oral review with a COVT
• A printed End of Course Report for the supervising doctor that includes:
  1. Course/exam scores
  2. Proof of credit hours
  3. Curriculum covered
  4. Strengths/weaknesses of the student
  5. Personalized notes regarding the student’s performance.
• A printed Emergent certificate (if the student receives a passing grade)

info@EmergentVT.com • www.EmergentVT.com