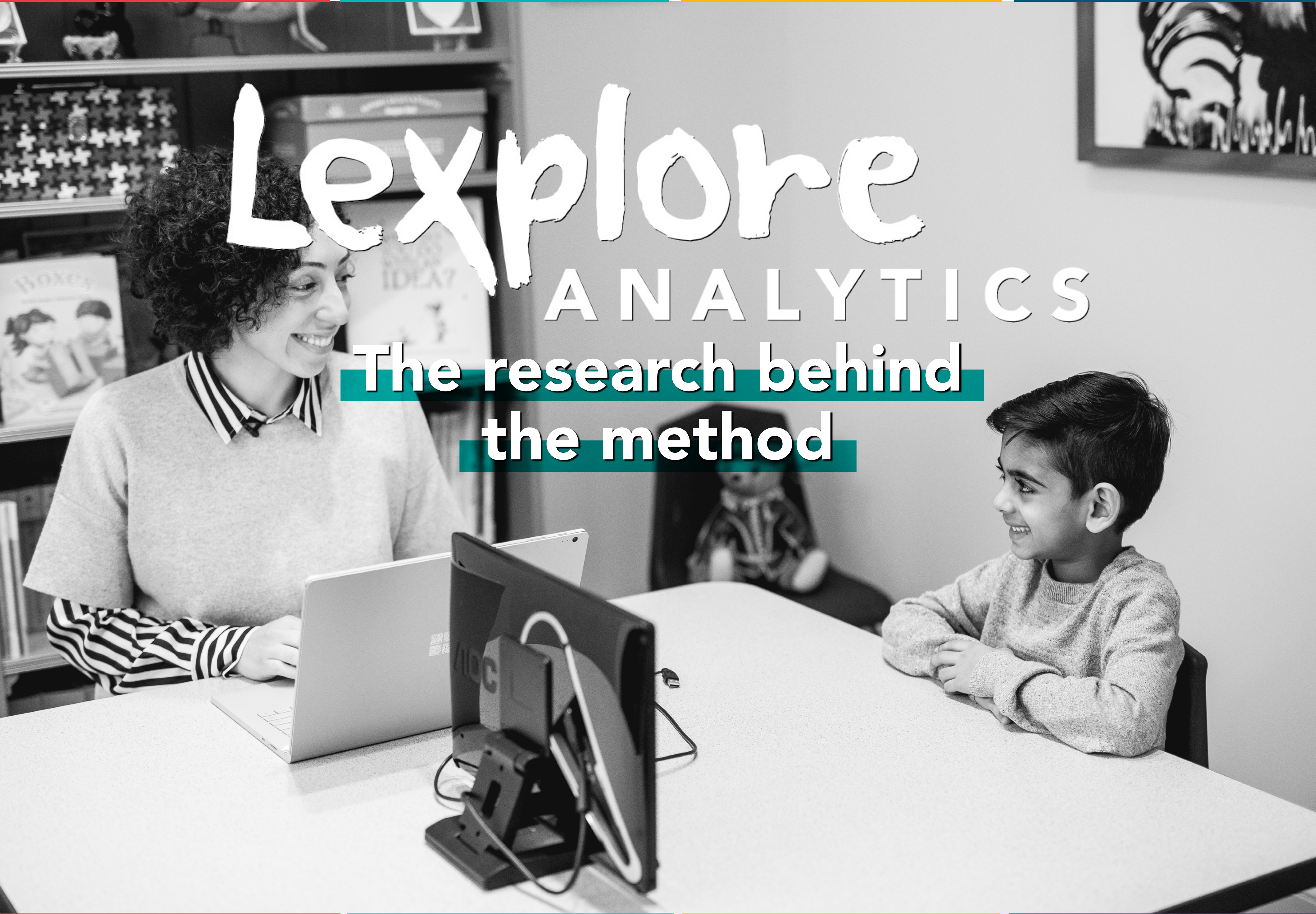
# Lexplore ANALYTICS

## The research behind the method



# Research overview

• Lexplore Analytics was founded following a research project, beginning in 2007, which was conducted at the Marianne Bernadotte Cenre, part of the Karolinska Institute in Stockholm.

• The research project was undertaken by two scientists; Gustaf Öqvist Seimyr and Mattias Nilsson Benfatto. They are also the founders of the company Lexplore and still participate in product development and management of the company today.

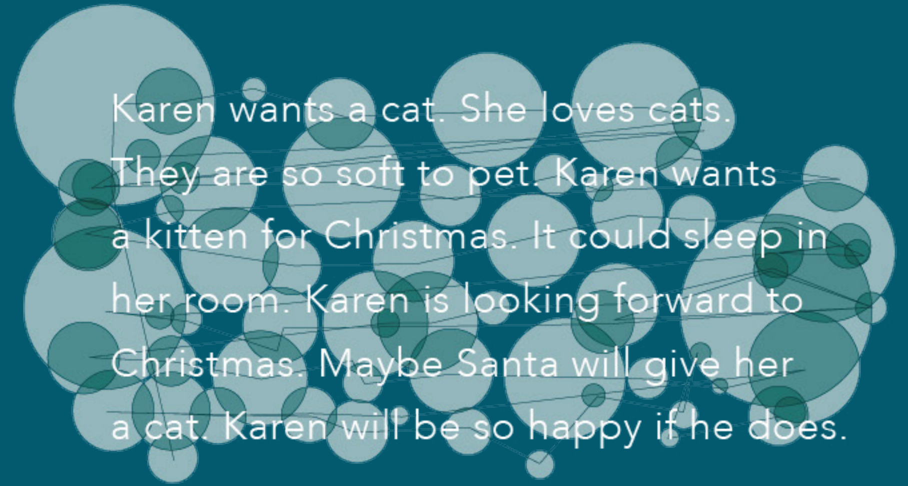
• The research behind Lexplore's unique method of assessing reading also extends much further back in time to the Kronoberg Study, undertaken more than 25 years ago by Christer Jacobsson and other scientists from the Karolinska Institute.

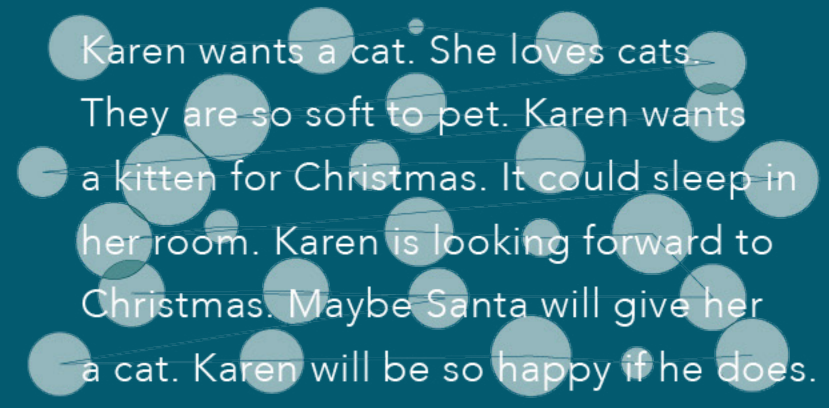
Karolinska Institutet

# What did the research involve?

During the 2007 research project, eye movement recordings were taken from a large number of pupils (almost 3500 in years 2 - 5). Each student performed a letter and word segmentation task, a sight-word and non-word reading task, rapid naming of letters, and read two texts at their year level. The average standard scores for the rapid naming, sight word, non-word, as well as the word segmentation tasks were used as references for training and testing machine learning (ML) models for each year group.

These machine learning models are now able to make accurate correlations between new and existing eye movement data in order to determine reading attainment to a high degree of accuracy.



# Lexplore Researchers

The researchers behind Lexplore recently featured on the 2019 Royal Academy of Science and Engineering's list of the most prominent 100 Swedish scientists.

Gustaf Öqvist Seimyr

Gustaf is a researcher in eye movement tracking and reading at the Marianne Bernadotte Centre. He is a computer linguist with a degree from Uppsala University, where he submitted his PhD in 2006 on evaluating readability on mobile devices. Gustaf's research focuses on measuring and analysing eye movements to create greater understanding of how vision works in health and disease under everyday circumstances such as reading, interpreting images, and in facial recognition.

Read more

Mattias Nilsson Benfatto

Mattias gained his PhD in computer linguistics at Uppsala University in 2012. As a graduate student he worked on the development of new analytical methods to give us a better understanding of eye movement control in reading. Mattias is currently a researcher at the Marianne Bernadotte Centre of the Karolinska Institute. Here he has continued his investigations into eye movements with a special focus on the relationship between eye movements and neurological disabilities.

Read more

# The Kronoberg Project Summary

Lexplore's method of assessing reading is originally based upon data from the Kronoberg project. This was an entirely unique, longitudinal study of reading and writing which began almost 30 years ago. As part of the project, eye movement recordings were taken for hundreds of children both with and without reading difficulties. Their academic and reading progress was then followed from year 3 to adulthood.

Technology at the time of the study was far from being as developed as it is now, meaning that the eye movement data collected was completely unique material. However, at the time it was not possible to carry out the same statistical analysis possible today. By analysing patterns from this project our researchers were able to show that trained statistical models could predict - with a high level of accuracy - which of the original pupils would have difficulties from their original eye movement correlations. This was possible after just 30 seconds of reading.

The overall results from this study, published in PLoS One (Benfatto et al, 2016), showed that this method of statistical analysis had the potential to be used to determine reading skills in schools.

PLOS | ONE

Screening for Dyslexia Using Eye Tracking during Reading

Abstract

Introduction

# The Dyslexia Project Summary

Additional research was then carried out as part of the Dyslexia Project in the Swedish municipalities of Järfälla and Trosa. During the two year project more than 3000 pupils were tested in years 2 to 5 with the intention of developing a method of assessment which could be used in schools.

Pupils sat a variety of other different assessments including rapid naming tests, word chains, naming words and non-words. Based upon results from these tests, researchers were able to determine each pupil's reading ability in relation to their year level. These results were then used in combination with eye movement recordings to train the machine learning algorithms to carry out a statistical mapping between the two.

By analysing the results of the Dyslexia project our researchers were able to show that the method they had developed to assess reading worked as accurately as intended. Analysis was able to show the method maintained a balance between sensitivity and specificity when determining each pupil's reading attainment level (Figure 1).

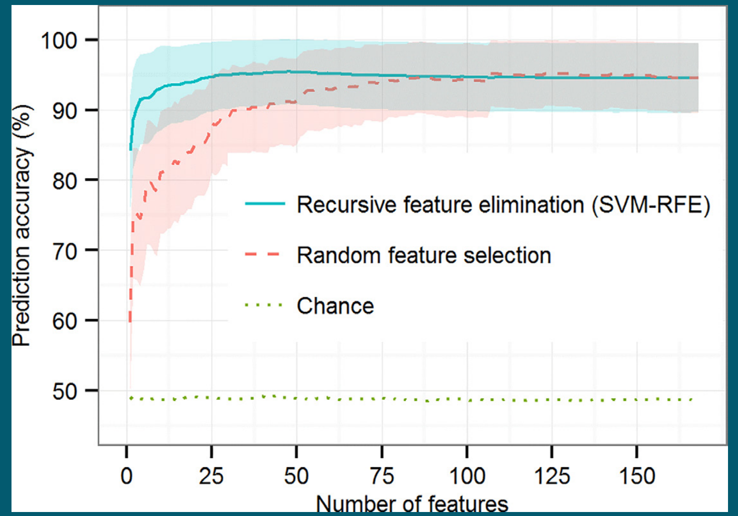



Figure 1. Accuracy as a function of the number of eye movement parameters in the classification of recordings from the Kronobergsprojektet (Benfatto et al, 2016).

Lexplore ANALYTICS

# The Research Continued

Reading attainment can be described as a continuum where reading and writing difficulties appear in the lower part of the normal distribution. Exactly where to draw this line is arbitrary: we have chosen to use the 10th percentile which is common in Sweden.

| | Students | Risk | Sens. | Spec. | Accu. |
|---|---|---|---|---|---|
| Grade 1 | 860 | 9.8% | 84.4% | 87.3% | 85.9% |
| Grade 2 | 910 | 10.1% | 86.2% | 90.3% | 88.3% |
| Grade 3 | 956 | 9.8% | 82.9% | 86.2% | 84.6% |
| Total | 2 726 | 9.9% | 84.5% | 87.9% | **86.2%** |

Table 1. Classification of reading and writing difficulties based on the 10th percentile for grades 1-3 or years 2-4 (Seimyr & Benfatto, 2017).

Research presented at the International Dyslexia Association Conference (Seimyr & Benfatto,2017) showed that results from the Lexplore assessment aligned with those determined by much lengthier reference tests. The results were also able to offer an insight into reading across many different components as opposed to focusing on a singular component like reading speed or comprehension (Figure 2).

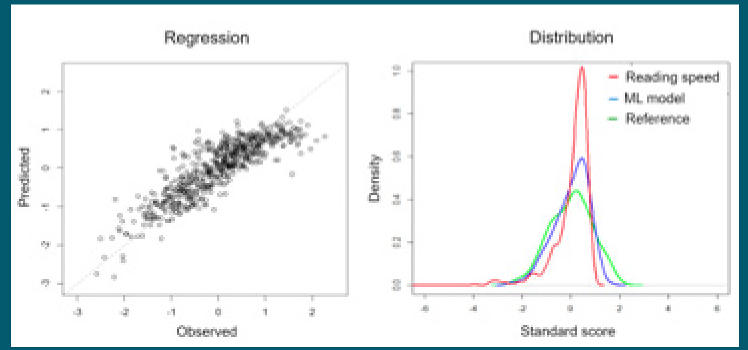



Figure 2. Accordance with the predicted and observed reading ability based on our method (R2 =. 78) compared to only using reading speed (R2 =. 48) (Seimyr & Benfatto, 2017).

Lexplore ANALYTICS

# Eye Movements - The Research

When reading - like you are now - you're probably not thinking about the complex cognitive processes which enable you to turn the words typed on a page into meaningful content. Reading fluently is the result of a complicated interplay between many linguistic and cognitive processes, which together enable children to decode text and comprehend written content. Learning these processes can be extremely difficult; children must first understand that words are composed of individual sounds – phonemes. In English we have many letter combinations that go with the same phoneme sound - for example, 'sh' can be formed in many ways. Not only is this sound found in the word 'sheep', but also in 'mission', 'sugar', 'lotion' and 'ocean'. Children who are good readers have strong phonemic awareness, vocabularies and grammatical skills; they also understand and can easily apply alphabetic principles.

There are no boundaries between the many different cognitive and linguistic processes involved in reading; all these processes form part of a much larger communication network and occur simultaneously. As an adult, these processes often occur subconsciously, meaning reading can often be confused as something which comes naturally for human beings. In reality, it involves many complicated processes which children much master for themselves.

To measure how effectively the many processes involved work together, we use the latest eye tracking technology. By measuring where, when and how the eyes move in relation to the words you read, we can gain a unique insight which immediately reflects the complex cognitive and linguistic interplay occurring. This insight is entirely spontaneous and reflects natural reading without the influence of any external factors.

To ensure that we also measure beyond the technical side of reading, we also include a number of comprehension questions to assess whether pupils have taken in and understood the content they have read (Figure 3)

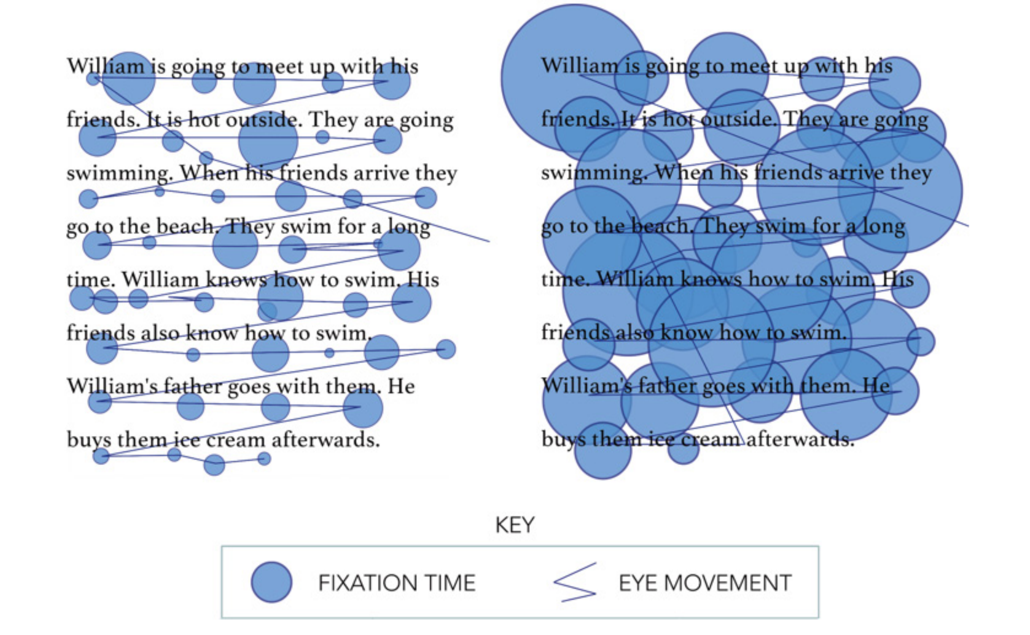



Figure 3: Visual representation of the reading process, showing the eye movements of a strong reader (left) and a weaker reader (right).

Eye tracking has become an increasingly popular research method for studying the relationship between different behaviours and underlying cognitive processes, especially because it is relatively simple and entirely objective. Further research has shown eye movement measurements provide a very accurate and useful method for us to study different aspects of the reading process (Just and Carpenter, 1980; Rayner, 1998; Clifton et al., 2007; Rayner 2009).

By varying different linguistic characters in the text, we know that eye movement measurements can pick up on minor differences in how the brain processes text on both lexical (Ehrlich and Rayner 1981; Inhoff and Rayner, 1986; Rayner and Duffy, 1986; Rayner and Well, 1996; Schilling et al., 1998), syntactic (Frazier and Rayner, 1982; Altman et al., 1992; Clifton et al., 2007; Demberg and Keller, 2008), semantic (Carol and Slowiaczek, 1986; Sereno and Rayner, 1992) and structural levels (Ehrlich and Rayner, 1983; Duffy and Rayner, 1990) .

Our knowledge of eye movements and the insight they offer into reading is continually being developed as research continues. There are now close to 5000 published studies in this area. A search for "reading" and "eye movements" in the reference database 'Web of Science' generates approximately 3000 titles.

There is, therefore, an established and well-documented correlation between the reading process and its various expressions in different eye movement patterns. Despite this peer reviewed research, Lexplore's method, which combines some of the latest AI and eye tracking technology in an objective reading assessment, remains entirely unique.

Many schools rely on the traditional combination of standardised assessments, comprehension tests and teacher observations when it comes to evaluating reading. As many of these reading tests only give one-dimensional, simple scores, teachers often struggle to collect the in-depth information they need to back their judgements, especially when deciding where those pupils performing at borderline levels lie. It is also possible for pupils using coping strategies to fall under the radar.

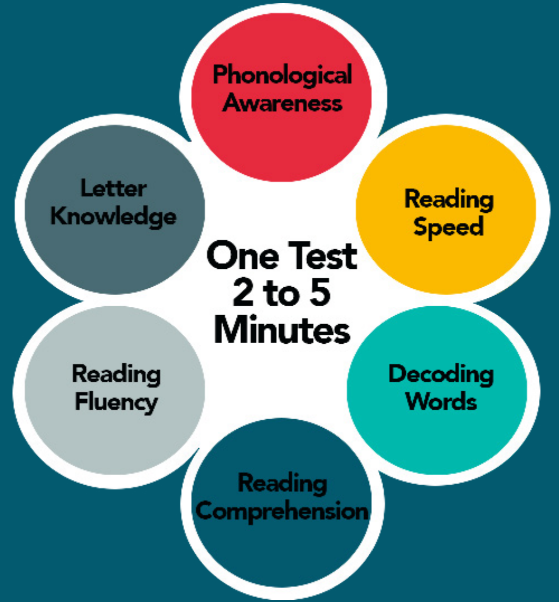
Lexplore's new technology can offer a startling insight into reading, helping pinpoint specific difficulties or problem areas in a matter of minutes from eye movement correlations. With immediate, objective and in-depth results, teachers can then make quick, evidence-based decisions to tailor support.

# Reading - The Definitions

According to "The Simple View of Reading" (Gough & Tunmer, 1986), the goal of reading is comprehension and this is defined as a product of decoding and language knowledge. Decoding refers to the process of converting letters, phonemes and graphemes into words and content which can be understood, whilst language knowledge refers to the processes involved with understanding words and being able to determine their context in written material.

Although this model of reading may seem simple, it has been shown in several studies to hold a strong empirical base (Kirby & Savage, 2008). Differences in decoding skills and language knowledge can also contribute to variations in children's comprehension skills; especially when children are still learning to read (Tilstra et al., 2009).

The relationship between comprehension and children's foundational decoding and language skills enables us to use eye movement analysis to study how the many different processes involved with reading work together for a child to read with proficiency, fluency, and comprehend content.



There are several studies which demonstrate the importance of including reading speed in the calculation of reading attainment (Deno et al., 1982; Fuchs et al., 1988; Silverman et al., 2013). Reading speed not only demonstrates how effectively the underlying processes involved with reading work, but also how they interact.

By using the insight eye movements are able to offer, Lexplore's method is able to look beyond reading as simply a product of comprehension, or a measure of speed, and instead as a very fine and complex interplay between many different processes. These processes are often difficult to assess using traditional, one-dimensional paper-based tests.

At primary school level, reading is often described in reference to a child's fluency and comprehension skills (The National Agency for Education, 2016). Reading with fluency means that pupils can effectively use their skills of decoding and language knowledge to move effectively and proficiently through a text. Reading with fluency can also be the result of a child developing strong sight reading skills.

As children move into secondary education, more emphasis is placed upon comprehension; especially when it comes to assessment. The new GCSE English specification, for example, now involves reading a larger number of texts which deal with complicated ideas in what can sometimes be an unfamiliar language. Many other subjects have also become much more content heavy – the lion's share of resources for history, geography, and even languages require children to have strong foundational reading skills. Literacy issues need to be spotted at the earliest opportunity so children can receive the support they need to succeed in the classroom.

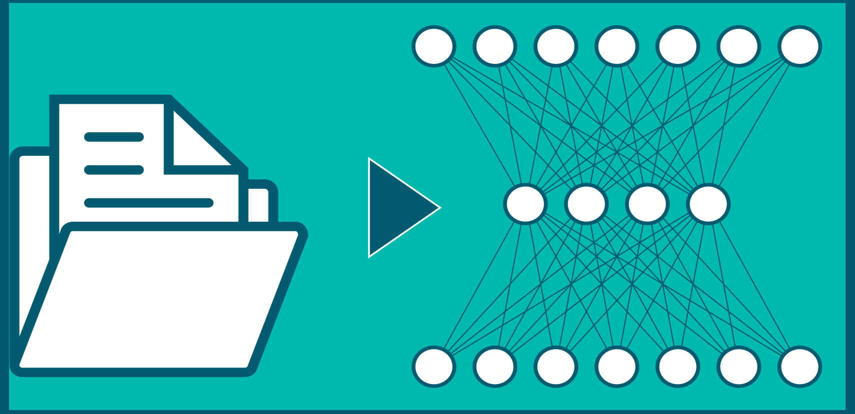
Definitions of reading from PISA and PIRLS also stress the importance of children's comprehension skills and ability to reflect on the content they have read (The National Agency for Education, 2017). Once children can reflect on texts, they can read to achieve, develop their knowledge and participate in society.

By including simple comprehension questions, the assessment is able to provide a truly unique insight into the many different components which make a strong reader. Teachers can then use results to tailor support quickly to the needs of every child in the classroom.

# Artificial Intelligence and Machine Learning

At Lexplore, we combine the latest in both artificial intelligence and eye tracking technology with the "human touch" of good teachers.

Our machine learning algorithms are able to recognise correlations and patterns in eye movement data, so they can quickly and objectively determine a reading attainment. Some of these patterns might even elude most human beings or other computer programmes. Therefore, their introduction into the classroom can help to offer teachers an entirely different perspective when it comes to reading, as well provide the valuable information teachers need in minutes.



The phrase 'Artificial Intelligence' was first adopted in 1956 by American computer scientist John McCarthy, but the idea of intelligent machines has fascinated scientists and philosophers for centuries. Although today artificial intelligence is all around us, it is still a dynamic and ever improving field of scientific research. Today, it has become a much more umbrella term, which encompasses multi-disciplinary fields of research with many different subareas.

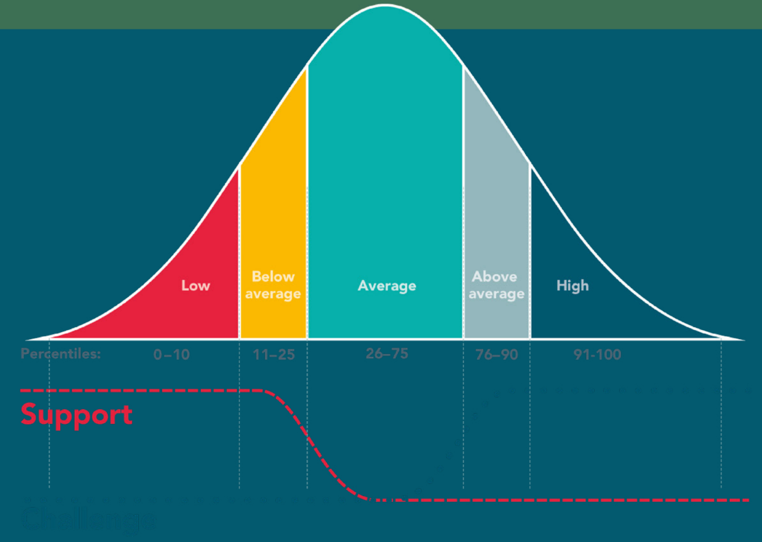
Machine learning is not only one of these branches, but perhaps the one we actually interact with most. In fact, it is a common misunderstanding that AI and ML as phrases are interchangeable. Machine Learning is actually a sub area within the field of AI that refers to machines which can learn from their interpretation of data and draw conclusions for themselves rather than following set rules, as in AI. Together, AI and Machine Learning enable us to instantly learn from the analysis of large amounts of data, in a capacity which exceeds what is humanly possible. IBM's Watson computer system is able to mine and perform a complex analysis on 200 million pages of text in just three seconds.


But, how can machines be trained to determine reading attainment from eye movement correlations? Quite simply, they use algorithms, which are sequences of instructions almost like recipes. They describe what needs to be inputted to produce a specific result. Algorithms need to be trained to perform the variety of complicated calculations required to answer different questions. At Lexplore, we ask the question, "What is the reading level of this pupil based upon their eye movement correlations?".

Lexplore machine learning models then carry out a specialised programme of analysis to assess a child's eye movement data following steps to determine reading attainment. These have been trained based upon the original data from around 3000 pupils. Their eye movements and reading ability were measured through using a variety of traditional tests. Our algorithms have been trained to automatically make links between specific eye movement patterns and reading attainment based upon these original results. Such training involves lengthy processes of identifying and quantifying connections.

Following careful validation and properly executed training, the algorithms are able to identify these correlations and immediately determine reading attainment to a high degree of accuracy. With quick, objective and in-depth results, teachers can then dedicate their time to providing the support that all children need for their reading journey, instead of spending large amounts of time manually evaluating reading themselves.

> "200 children went through the Lexplore assessment process in a week. Within five minutes of each assessment we had the results of their reading activity and an accurate attainment level. Previously, it took teachers an entire term to do accurate assessments on their whole class."
>
> **Adam Luxford**
> **Freemantle Church of England Community Academy**



# Lexplore Analytics Today

Despite varying backgrounds and experiences in business, education and research, our team are united by the same overarching commitment to education upon which the company was founded many years ago.

Today, our work is still driven by that desire to help all children discover the magic of books and the belief that reading and writing difficulties should not prevent children from making the most of their education.

Fundamentally, teachers can only address literacy issues if they know what they are in the first place. It's our new technology that paves the way exactly for this as it is able to look beyond reading as a standardised or percentile score, and instead as complex interplay between many linguistic and cognitive processes.


**www.lexplore-analytics.co.uk**

hello@lexplore.com

+44(0)161 697 4166

@LexploreAnalyticsUK


Lexplore ANALYTICS

# References


Altman, G. T., Garnham, A., and Dennis, Y. I. L. (1992). Avoiding the garden path: Eye movements in context. Journal of Memory and Language, 31:685–712.

Benfatto, M. N., Seimyr, G. Ö., Ygge, J., Pansell, T., Rydberg, A., and Jacobson, C. (2016). Screening for Dyslexia Using Eye Tracking during Reading. PloS one, 11(12), e0165508.

Carroll, P. and Slowiaczek (1986). Constraints on semantic priming in reading: A fixation time analysis.Memory and Cognition, 14:509–522.

Clifton, C., Staub, A., and Rayner, K. (2007). Eye movements in reading words and sentences. In van Gompel, R., editor, Eye movements: A window on mind and brain, pages 341–372. Amsterdam: Elsevier.

Demberg, V. and Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. Cognition, 109:193–210.

Deno, S. L., Mirkin, P. K., and Chiang, B. (1982). Identifying valid measures of reading. Exceptional Children, 49(1), 36-45.

Duffy, S. A. and Rayner, K. (1990). Eye movements and anaphor resolution: Effects of antedecent typicality and distance. Language and Speech, 33:103–119.

Ehrlich, S. F. and Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. Journal of Verbal Learning and Verbal Behavior,20:641–655.

Ehrlich, S. F. and Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. Journal of Verbal Learning and Verbal behavior, 22:75–87.

Frazier, L. and Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. Cognitive Psychology, 14:178–210.

Fuchs, L. S., Fuchs, D., and Maxwell, L. (1988). The validity of informal reading comprehension measures.RASE: Remedial and Special Education, 9(2), 20-28.

Gough, P. B., and Tunmer, W. E. (1986). Decoding, reading, and reading disability. Remedial and special education, 7(1), 6-10.

Inhoff, A. W. and Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. Perception & Psychophsics, 40:431–439.

Just, M. A., and Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. Psychological review, 87(4), 329.

Kirby, J. A. and Savage, R. S. (2008). Can the simple view deal with the complexities of reading? Literacy, 42(2), 75-82.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. Psychological Bulletin, 124:372–422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. The Quarterly Journal of Experimental Psychology, 62:1457–1506.

Rayner, K. and Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. Memory & Cognition, 14:191–201.

Rayner, K. and Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. Psychonomic Bulletin & Review, 3:504–509.

Schilling, H. E. H., Rayner, K., and Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. Memory & Cognition, 26:1270–1281.

Seimyr, G.Ö. and Benfatto, M. N. (2017). Screening for Reading Deficits using Eye Tracking and Machine Learning. Poster presented at the annual International Dyslexia Association Conference. Atlanta, GA.


Lexplore ANALYTICS