Typography papers 9

Readability: discovery and disputation

William Berkson & Peter Enneson

Typography papers 9 was edited, designed, and prepared for press in the Department of Typography & Graphic Communication, University of Reading (www.reading.ac.uk/typography) and published by Hyphen Press, London (www.hyphenpress.co.uk).

This file is compatible with Adobe Acrobat 7.0 or higher.

Colour and greyscale images downsampled to 240 ppi with 'high' image quality JPEG compression; monochrome images downsampled to 400 ppi. Printing is enabled to 'high resolution'.

Minor copy-editing faults that may have occurred in the article, as printed, have been silently corrected. Substantive corrections, if any, are listed below.

Corrections

p. 134, n. 47: 2nd sentence deleted, which repeats the 2nd sentence of n. 38 (p. 131).

William Berkson &
Peter Enneson

Readability: discovery and disputation

This essay recovers the breakthrough work on readability done by Matthew Luckiesh in collaboration with the Mergenthaler Linotype company. In the late 1930s, Luckiesh developed the concept of readability as ease of reading continuous text; he also discovered that he could measure readability by studying blink rate during reading. After describing how Luckiesh began research in typography, his collaboration with Linotype is explained. The views of, and reaction to, this work by the typeface designer W. A. Dwiggins are also presented. The validity of Luckiesh's work was attacked by psychologist Miles Tinker. Tinker's own background and work are presented, along with the several dimensions of his dispute with Luckiesh. An argument is put forward that, in the light of methodological standards then and now, Tinker's case against the work of Luckiesh is not sound. The essay concludes with reflections on the reception and value of both Tinker and Luckiesh today, and on the insights Luckiesh's work in particular may bring to the practice of typography and research into reading.

> *Between the research scientist and the technician in the graphic arts there has been a degree of common interest in readability, but a colossal absence of mutual understanding.*[1] – Harry L. Gage (1942)

Since the late 1920s, when psychologists began systematically studying the impact of typography on reading, there has been a divide in understanding between them and practising typographers. Psychologists have generally been wary of typographers' expert craft knowledge, regarding their practices as the product of introspection, conditioning, and preference that lacked scientific validation or objective measures. Typographers, for their part, have often regarded psychologists as naïve about type, and their experimental results consequently weak, obvious, or misleading.

An important indicator of these differing perspectives is the term 'readability'. Many books on typography written in the last fifty years have defined the term as the ease of reading extended text. 'Legibility', also referred to in books on typography, is generally defined as how quickly readers can accurately identify individual letters. Authors generally explain the distinction at the outset, then describe how both contribute to good typography. In psychology, by contrast, a distinction between readability and legibility is not made. Instead, the concept of legibility alone – not defined, but as a rule measured by reading speed – is used to assess and describe readers' experience of type in print and on screen.

How did the concept of 'readability' come about, and why was it separated from 'legibility'? And, if the distinction is valid, as many typographers believe, why has it not been adopted by psychologists? The concept of readability of type, defined as the ease of reading extended text, was developed in the late 1930s, during a collaboration between Matthew Luckiesh, a prominent researcher for the General Electric Company, and Harry L. Gage, a vice president at the Mergenthaler Linotype Company.[2] Luckiesh found that reading strain and fatigue caused by factors such as length of time reading, low

1. Luckiesh and Moss (1942), p. ix.

2. Berkson came across the work of Matthew Luckiesh while investigating the distinction between readability and legibility made by Lieberman (1967) and others; see Berkson (2011). The subsequent narrative has been assembled by Berkson and Enneson collaboratively. We would like to thank: Kent Lew for providing documents relating to Luckiesh, which he uncovered in the Chauncey Hawley Griffith papers, Special Collections, University of Kentucky; Ann Sindelar, Research Center, Western Reserve Historical Society, Cleveland, Ohio, where the Luckiesh archive is located; Peggy Luckiesh Kundtz and John Kundtz for sharing recollections of their father and grandfather, respectively, and for help with the Luckiesh archive; and Terry McGowan for information on Luckiesh and Nela Park, Cleveland.

lighting, and small type size could be measured by increases in blink rate while reading. He proposed using the term 'readable' for text that is easier (or requires less effort) to read, and so generates less strain or fatigue.

Though simple in outline, the story of the origins of readability is a fascinating one in detail, involving an ambitious research program, a remarkable collaboration, an important discovery, an acrimonious dispute, and fundamental differences over research methods. The story also includes a penetrating appraisal of Luckiesh's work by the type designer W. A. Dwiggins. Despite recognition of Luckiesh's work on readability at the time, it subsequently fell from view, in large part the result of a dispute with a fellow researcher, Miles Tinker. So while readability as a concept has endured among typographers, there is little if any awareness of its origins in Luckiesh's work. Nor did psychologists continue his work on readability. But in recent decades, researchers interested in fatigue have returned to Luckiesh's research and confirmed its general validity. This has been especially true of blink rate as an indicator of fatigue, currently of interest to neuroscientists studying brain function.

In this essay, we will argue that a mostly one-dimensional notion of 'legibility' used by Miles Tinker against Luckiesh's readability has contributed to the divide between typographers and psychologists. Good typography can enhance the reading experience in many ways, and Luckiesh's focus on ease of reading was a promising step towards revealing the numerous dimensions of reader experience. For this reason, renewed engagement with Luckiesh's research is of value both as a recovery of an important historical episode in typography and psychology, and in the insights it may bring to issues that concern typographers, psychologists, and other scientists seeking to understand how we read.

Matthew Luckiesh

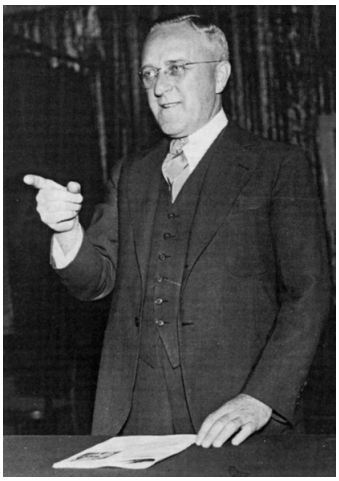


Figure 1. Matthew Luckiesh lecturing, probably 1940s.

Matthew Luckiesh (1883–1967; figure 1) rose to distinction from modest beginnings.[3] His father, orphaned in Austria, emigrated to the small town of Maquoketa, Iowa, where he worked as a school janitor. Luckiesh taught himself to play trombone in high school, and afterwards performed in one of the last touring minstrel shows in the United States, earning enough money to pay his tuition at Purdue University, where he studied electrical engineering. In 1910 he was hired by the National Electric Lamp Association (NELA) in Cleveland, Ohio, which became part of Thomas Edison's General Electric Company (GE) the following year.[4] Electric lighting was at this time a new, exciting, and socially transformative technology, and initially Luckiesh invented specialty light bulbs. But he soon turned to questions concerning the optimal conditions of electric lighting for visual tasks, part of what is now the field of human factors.

3. 'Luckiesh' is pronounced LOO-kish (rhymes with 'dish'). Unless otherwise indicated, personal information about Luckiesh is taken from Covington (1992).

4. The National Electric Lamp Association was formed in 1901 to rival General Electric. In 1913, after its acquisition by General Electric, NELA moved to its present location, Nela Park, Cleveland, Ohio. Nela Park was the first industrial research park in the world and remains the headquarters of GE Lighting. See Covington (1992), ch. 7.

Luckiesh's career at GE proved highly successful. Outgoing, confident and energetic, indeed something of a showman, Luckiesh was exceptionally industrious. By 1924 he was head of the Lighting Research Laboratory at Nela Park, a position he held until his retirement in 1949. Colleagues were apparently awestruck by his brilliance, take-charge manner, and powers of persuasion.[5] Luckiesh was well supported by GE and given almost everything he asked for, including facilities designed to his specifications. He also wrote and published extensively: over the course of his career, some 28 books and more than 600 articles dealing with subjects relating to light and lighting, ranging from color and optical illusions to camouflage and reading.[6] One likely reason GE supported Luckiesh so fully was his ability to promote light bulbs, in particular by recommending high illumination levels for tasks such as reading. General Electric's 'Better Light – Better Sight' campaign, for example, which lasted from 1933 until 1979, was based on research work at Nela Park. By showing customers how to light their homes and offices, the campaign helped to sell a lot of light bulbs.

Early work and first dispute

Matthew Luckiesh began his career when the first and second generations of experimental psychologists were active in the USA. Luckiesh did not, however, learn about methods of human factors research from psychologists, but from a physician, Percy Wells Cobb. Cobb had studied mechanical engineering before training as a medical doctor. Soon after receiving his medical qualification, he began teaching and doing research in experimental psychology, first at Western Reserve University, and then at Nela Park where he was head of the Lighting Research Laboratory during the early years of Luckiesh's career. Cobb adhered to the research methodology pioneered by Claude Bernard, the French founder of experimental medicine. In his *Introduction to the study of experimental medicine* (1865), Bernard described an approach to research that encouraged the search for the 'immediate causes' of any phenomenon. Bernard also emphasized the need to avoid being bound by past theories; one should instead constantly attempt to refute them, as well as one's own hypotheses, to discover the truth. Bernard's view of scientific method was prevalent in experimental medicine and, probably via Cobb, became a decisive influence on Luckiesh.

When Luckiesh succeeded Cobb as head of GE's Lighting Research Laboratory, his principal task was to study the effect of different lighting conditions on the eye. Two important developments in the study of illumination would lead him towards research in typography and reading. The first development was an insight that Cobb had arrived at while investigating appropriate illumination levels for various tasks. Cobb realized that Fechner's Law, one of the earliest hypotheses in experimental psychology, was inadequate as a description of the eye's response to illumination. Fechner had proposed a simple logarithmic relationship between the level of a stimulus, such as brightness, and a 'just noticeable difference' from that level. Working together with Frank K. Moss,[7] Cobb showed that the immediate causes of the eye's response to illumination were more complex.

5. This characterization of Luckiesh is based in part on the personal recollections of Terry McGowan, who worked at Nela Park, from 1961 to 1998 and met the retired Luckiesh. John Kundtz, Luckiesh's grandson, recalls playing under his grandfather's desk as a boy while Luckiesh worked furiously above him. Personal communications from Terry McGowan (June 2011) and John Kundtz (June 2011 and October 2012).

6. A further indicator of Luckiesh's stature and reputation was the commission to design the lighting for the White House during the Franklin D. Roosevelt administration.

7. For remarks on Frank K. Moss, see n. 15.

Not only was the brightness of an object involved, but also the its size, the amount of time the object was looked at, and the contrast between the object and its background. This approach of looking at multiple causal factors would later be decisive in Luckiesh's work.[8]

The second development that provided a major impetus for Luckiesh involved the work of two other researchers on illumination, Clarence Ferree and Gertrude Rand.[9] In 1911, at a joint meeting of the American Medical Association and the Illuminating Engineering Society, a physician had issued a challenge: find a way of accurately measuring the effect of various lighting conditions on the eye. The normal test for visual acuity, the optometrist's eye chart, was insufficiently sensitive to discriminate the effects of glare, low light, and other lighting conditions that accompanied the new technology of electric light. A new measure was needed.

In taking up the challenge, Ferree and Rand supposed that time – and specifically the power of the eye to sustain clear vision over time – would be a key test factor. Looking at an object for a much longer time than normal might increase the sensitivity of visual acuity tests. One test they developed and used extensively required subjects to stare continuously at the letters 'li' for three minutes under various illumination levels; the subject pressed a lever to indicate when the i became blurred and confused with the l, or when the dot of the i fused with its stem. Ferree and Rand used the length of time the eye could sustain clear vision (acuity) as an indication of its efficiency under specific lighting conditions. The test was repeated before and after periods of fixed duration, during which visual work was done continuously under varying lighting conditions. They also compared the ratio of blurred to unblurred time before and after extended work at different lighting configurations and levels, to see how 'time-on-task' and lighting affected visual acuity. The effects of time-on-task were then attributed to fatigue and reading comfort.

Cobb, on the other hand, had begun to place great importance on the fact that normally the eye is constantly in motion, with short fixation times. Following this idea, he studied the impact of contrast, size, and brightness on how *quickly* objects are recognized. He was sceptical that Ferree and Rand's 'li' test of fixed, staring eyes indicated anything important or valuable about sight; commenting on a paper published by Ferree and Rand in 1915, he pointed out that their data derived from subjective reporting by participants who were carefully chosen, trained, and aware of the goal of the experiments. Cobb was therefore unable to trust the results as an accurate measure of the eye's loss of efficiency. Some years later, in 1927, Cobb, now with Luckiesh and Moss, proposed a way to check the validity of the 'li' test.[10] They repeated the test but without any rest interval between the three-minute periods when subjects stared at the letters. Over the course of successive three-minute periods, they were unable to establish a consistent baseline value for, or drop-off in, visual acuity. From this they concluded that the test was in fact useless as an indicator of suitable illumination levels for visual tasks.

In a lengthy reply appended to the 1927 article by Luckiesh, Cobb, and Moss, Ferree and Rand objected that their tests had not been repeated with the same controls. They dismissed the critique as

8. See Cobb and Moss (1928). Cobb's insight was also the foundation of what Luckiesh called 'the science of seeing'; see Luckiesh and Moss (1937a).

9. Clarence Ferree and Gertrude Rand, husband and wife, were researchers at Bryn Mawr College and from 1928 at Johns Hopkins University. Rand is noted as one of the early, outstanding women of twentieth-century science.

based on a 'lack of understanding', 'confusion', and a 'lack of knowledge', and declared the results of the repeated experiment 'irrelevant'. Ferree and Rand also put forward additional data to support their objections. In a concluding note, Luckiesh, Cobb, and Moss replied, in turn, that they remained unimpressed by the additional data collected using the same questionable methods they were criticizing.

Aside from what had proved a heated exchange, it is notable that Luckiesh, Cobb, and Moss nevertheless agreed with Ferree and Rand's original supposition: that testing the effect periods of continuous work had on vision offered a viable way to measure the impact of lighting, from which good lighting standards could be established. Indeed, Luckiesh and and his colleagues had singled out at the outset of the exchange the idea that 'a test which would give a reliable measure of fatigue or other temporary impairment of vision after a period of work would be of immense value in lighting practice'.[11] Finding such a reliable measure, or measures, of fatigue would soon became the focus of Luckiesh's research.

Visual effort, fatigue, and ease

After the critique of Ferree and Rand's work, Luckiesh became convinced that measures of visual *performance*, such as that for acuity, were not sensitive enough to adequately understand the interaction of light and seeing. The human body's ability to compensate for the effects of low light and thereby sustain performance confused matters. The ability to accommodate for various effects did take a physiological and psychological toll on the body in the form of strain and fatigue, and it was these *costs of performance* that Luckeish thought might be the key to a better understanding of seeing. In order to measure the impact of lighting on seeing, Luckiesh would now look not only at the performance of visual work, but also at its costs.[12]

In 1929, Luckiesh published a chart that would form the conceptual basis of his research for the next twenty years (figure 2).[13] The chart distinguishes *effectiveness* in performing a visual task from the *efficiency* of that performance. Efficiency involves not only the level of performance, such as the ability to sustain clear-seeing or to sustain reading speed, but also the *resources* or *effort* needed to achieve that level of performance. While the body may sustain high performance throughout periods of sustained work, the fatigued performance requires increased effort. That greater effort correlates with changes in the body, such as increased muscular tension or weakness, indicating an increased demand on resources. The premise of Luckiesh's ongoing work was that the fatigue from periods of continuous work,

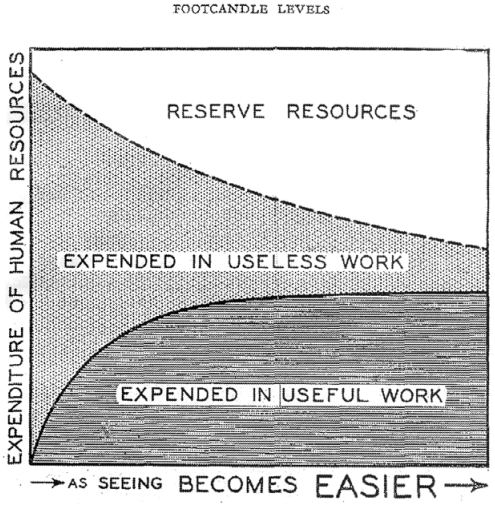


Figure 2. Graph showing the relation of any aid to seeing to the expenditure of human resources in the performance of any task of seeing. Luckiesh (1948), p. 403.

10. Luckiesh, Cobb, and Moss (1927); Cobb's critical remarks of 1915 were first published in this article.

11. Luckiesh, Cobb, and Moss (1927), p. 77.

12. Luckiesh (1930). This paragraph draws on the terminology of David DiLaura to describe Luckiesh's concept. To define the efficiency of doing a task, DiLaura, an illuminating engineer and historian, distinguishes between *performance* and the *psycho-physical cost* of performance. For example, an office worker might be able to perform a job equally well in a good office chair or a bad one, but the cost in aches and pains would be much lower in the good one. See DiLaura (2005).

13. Luckiesh (1948), p. 403. The earliest version of the chart occurs in Luckiesh (1929), p. 39; modified versions appear in subsequent publications by Luckiesh.

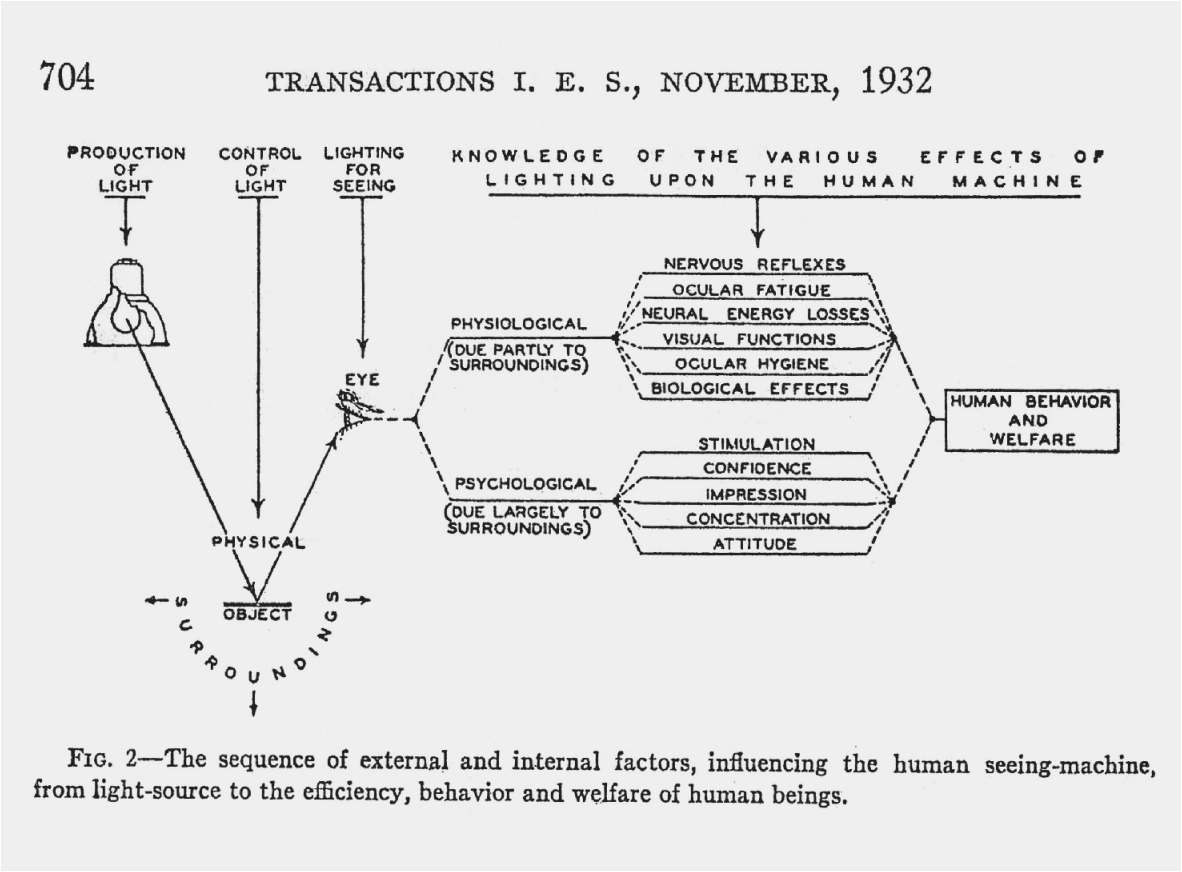


FIG. 2—The sequence of external and internal factors, influencing the human seeing-machine, from light-source to the efficiency, behavior and welfare of human beings.

Figure 3. Factors involved in visual performance. Luckiesh (1932), p. 704.

and its inverse, ease, were as important for assessing lighting or typography as the level of performance. His experimental approach was to systematically identify and employ better, objective, measures of effort. He called his research program 'humanistic' because he focused not just on output, but also on the human costs of effort and work.

From the outset, Luckiesh saw that many physiological and psychological factors were involved in visual performance (figure 3), and so he sought out good ways to measure how these different factors were affected by fatigue in the 'human seeing-machine'. Like Cobb, Luckiesh wanted to identify underlying causes. He studied how physiological and psychological resources were used in vision, and conducted experiments to determine how time-on-task and adverse visual conditions affected the use of these resources. Luckiesh also followed Cobb's example by introducing new variables to the study of reading. Where Cobb had added speed, contrast, and time to the variable of illumination level, Luckiesh added the consideration of effort, and the use of resources involved in visual tasks. Luckiesh introduced three measures in particular: general muscular tension, heart rate, and the strength of eye muscles used to maintain binocular vision. He found that each was affected by time-on-task and low illumination levels. Based on this new research, Luckiesh made his own recommendations for illumination levels for reading and for other demanding visual tasks, levels that were significantly higher than those implied by the research of Ferree and Rand.[14]

14. See, for example Luckiesh and Moss (1937a), p. 303 ff.

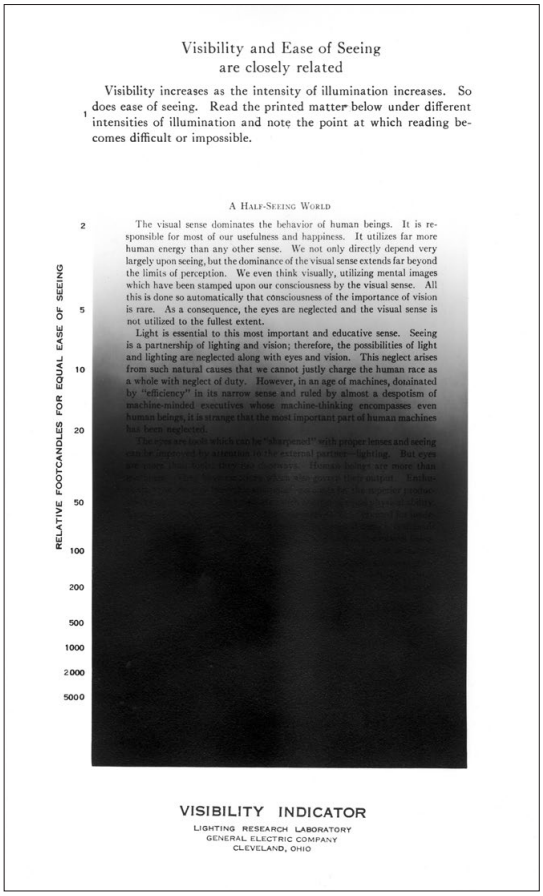


Figure 4: 'Visibility and ease of seeing', plate from *The science of seeing* (1937) by Matthew Luckiesh and Frank K. Moss.

In 1935 and 1936, Luckiesh, now working solely with Moss,[15] made two further advances in this research. The first was the invention of the Luckiesh-Moss visibility meter, which measured the visibility of objects.[16] It consisted of spectacles fitted with two filters that rotated over each eye to decrease contrast and brightness until an object could no longer be identified (figure 4). The second advance was the adoption of blink rate as a measure of fatigue in seeing. It appears that Luckiesh recognized the potential of blink rate as a measure in the mid 1930s, having encountered the idea in a 1934 article in a Russian journal; he probably then became aware of an article on blinking by Eric Ponder and W. P. Kennedy published some years earlier.[17] Ponder and Kennedy drew on a number of tests and observations to argue that spontaneous blinking was neither a reflex action nor solely a means to cleanse and moisten the eye, since blinks were too frequent for that alone; instead, blinking served to relieve 'mental tension'.

Luckiesh and Moss, with their new interest in 'non-performance' impacts of periods of continuous work, began testing how blink rate was affected by different conditions, including visibility.[18] For their tests, they turned to one of the most demanding visual tasks: reading. They found that blink rate, which is normally considerably slower during reading, in fact increased the longer a subject was reading. Similarly, they found that under conditions normally considered adverse for reading, such as the use of very small type or low levels of illumination, blink rate increased as well. Luckiesh and Moss

15. Frank K. Moss worked as an assistant to Luckiesh from 1929 until his death in 1943, aged 45. Moss never published solely under his own name, making it difficult to judge his contribution to works co-authored with Luckiesh. Based on later comments by Luckiesh, it seems that Moss was a very helpful lieutenant but not an innovator.

16. Luckiesh (1935).

17. G. A. Litinsky, 'Recording of winking as a method of study of ocular fatigue in children resulting from reading', *Sovietskii Viestnik Opht.*, 4 (1934), p. 275; cited in Luckiesh and Moss (1937), and Luckiesh and Moss (1939). See also Ponder and Kennedy (1927).

18. It is worth noting that Ferree and Rand had already been aware that visual fatigue was related to 'ocular discomfort' and that this discomfort was relieved by blinking. But they regarded blinking only as a possible confounding factor whose influence on their test of visual acuity under stress needed to be eliminated.

also used their new visibility meter to determine which typefaces were more and less visible. They evidently hoped that their meter's ability to measure visibility quantitatively would prove crucial in understanding the impact of the physical conditions of reading on reader experience. But they were not able to find a simple rule characterizing the relationships of visibility and reader experience.

Luckiesh was convinced that typographic factors were only weakly discriminated by existing performance tests that involved speed and accuracy of reading. He also dismissed introspective reports of 'esthetic comfort' by readers, or indeed by expert typographers, as 'of little value' as a measure since 'this attribute is not one which yields to standardization'. The only clear correlation he had discovered between the physical conditions of reading (size of type, lighting, and so on) and the experience of effort, or ease, was the relationship of blink rate to illumination. But he and Moss remained hopeful that 'refinements in technique' might yet establish clearer relationships between quantitatively measured visibility and ease of reading.[19]

Luckiesh and Moss published the results of their research in scientific journals,[20] printing trade journals, and more widely in magazines and newspapers with a general readership.[21] It was probably through some element of this publicity that their work came to the attention of the Mergenthaler Linotype Company, at that time the largest manufacturer of typesetting machinery in the world.

Breakthrough

In 1937, Mergenthaler Linotype appointed Matthew Luckiesh and Frank K. Moss as 'consulting research staff' who would undertake research on typographic factors in reading.[22] Their collaborator throughout this project, and probably the project's initiator, was Harry L. Gage, a vice president at Linotype.[23] Gage, who was also a painter (figure 5), had already served as president of the American Institute of Graphic Arts and had written a manual of typographic design for apprentices in the printing trade. He had also recruited the book designer and illustrator William Addison Dwiggins to design typefaces for Linotype. Gage's knowledge of typography was thus extensive.

Figure 5. Harry L. Gage as a young painter, probably 1920s.

In an article in *Linotype News* announcing the collaboration, Gage described the research Luckiesh and Moss would do for Linotype.

19. Luckiesh and Moss (1937a), p. 455.

20. Luckiesh and Moss (1935) and (1937).

21. 'Measuring type visibility', *The Trade Compositor*, December 1936. Public relations staff at Nela Park were apparently able to secure notice of Luckiesh's work in magazines and newspapers, which were then clipped out and collected in scrap books. These scrap books, containing many hundreds of clippings, are preserved in the Luckiesh archive, Research Center, Western Reserve Historical Society.

22. *American Printer and Lithographer*, vol. 105, 1937. 'Mergenthaler Linotype Co. announces that Dr. Matthew Luckiesh and Frank K. Moss, scientists, will be associated as consultants with the company's present research and development department, to specialize in research projects in type legibility. Dr. Luckiesh and Dr. Moss are widely-known for their work in the science of seeing.' The appointments were also announced in an article praising Luckiesh's work in *Linotype News,* May–June, 1937. Below, for convenience, the Mergenthaler Linotype Company will be referred to simply as 'Linotype'.

23. Harry L. Gage, foreword to Luckiesh and Moss (1942), p. ix.

In general, they would seek answers to practical questions about typography that concerned printers and publishers. This would be done by bringing typography and science together to resolve long-standing typographic issues.[24] Gage was impressed by the number and variety of techniques Luckiesh and Moss had already used to study seeing, and that these techniques had proven scientifically robust. He was cautiously hopeful about where the research might lead:

> Whether this research may open a new approach to further developments in legibility remains to be seen. Such progress would be welcome. But we shall feel well rewarded if our work clarifies and defines the many beliefs, hazy traditions and mere habits of type image which characterize modern practice in the graphic arts.[25]

Luckiesh and Moss would apply scientific method to the study of common typographic matters, such as the ideal line length and leading for a given size of type. Gage also wanted them to test the 'speed of fatigue' experienced by subjects when reading texts typeset in different ways: in uppercase only, lowercase only, in a range of type sizes and styles, as white type reversed out of black, and as printed on various coloured papers. Gage named a number of factors that Luckiesh and Moss would consider, including visibility, as measured by the Luckiesh-Moss visibility meter, and degree of muscular fatigue, as measured by temporarily weakened eye muscles.

Although Gage did not mention blink rate in his article about the new collaboration, six months later, in a first progress report in *Linotype News*, he stated that Luckiesh and Moss believed that blink rate was 'the most important factor yet appraised in these studies.'[26] According to Gage, Luckiesh regarded blink rate as the best indicator of readability, though Luckiesh still felt that visibility (as measured by the Luckiesh-Moss visibility meter) was a good indicator, too. At this point Luckiesh was also still including both performance (speed and accuracy) and cost-of-performance (ease) measures in the concept of 'readability'.[27]

The fourth of Gage's progress reports, in July 1938, documents a turning point in the collaboration. Luckiesh and Moss had compared Caslon Old Face, Textype, and Memphis Medium[28] (figure 7, overleaf) using three tests: *visibility,* a measure of contrast thresholds for reading printed type, as ascertained by the Luckiesh-Moss visibility meter; *readability*, ease of reading, as measured by blink rate in

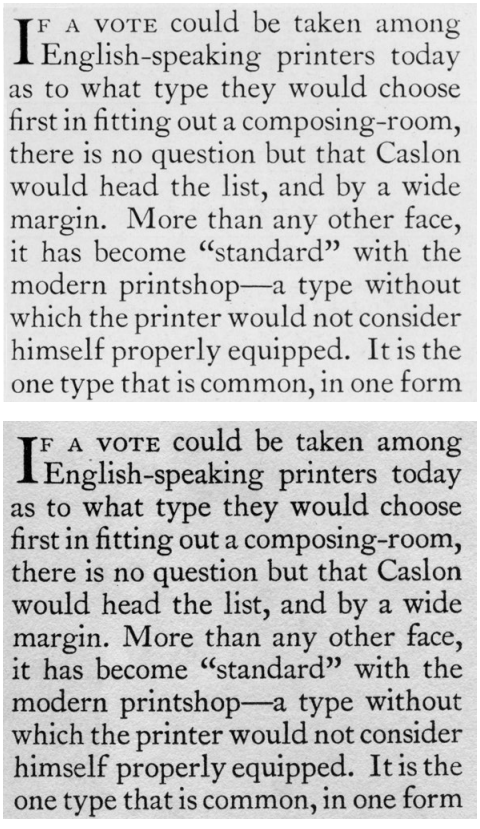


Figure 6. Comparison of Linotype Caslon Old Face on smooth (top) and rough paper. Linotype specimen booklet (1928). (See n. 28, at right.)

24. Gage (1937).

25. Gage (1937); cf. n. 27, below, for subsequent remarks by Gage on 'legibility'.

26. Gage (1938), a. These and subsequent progress reports published in *Linotype News* were principally distributed to the printing trade.

27. Gage (1938), b. 'We prefer the term readability, rather than legibility since, (1) it is descriptive of the act of reading and (2) it is not so likely to be confused with visibility. Obviously, readability is a function of visibility and, in specific cases, the two may be synonymous for empirical reasons. In general, the readability of printed or written material may be defined as that characteristic which determines the speed, accuracy and ease with which it may be read.' The identical passage occurs in Luckiesh (1937a), p. 455.

28. Textype was designed 'primarily for printing textbooks and related works requiring intensive study and prolonged reading'. In figure 6, and in the Luckiesh and Moss test material, Caslon Old Face was printed on smooth paper, which produces a weaker, thinner type 'image' than when the typeface is (letterpress) printed on rough paper – its traditional use – where ink gain (or squash) produces a stronger, sturdier image.

CASLON OLD FACE, 11-point, 3 points leaded

This type represents a faithful reproduction of the original type of William Caslon of 1734. In selecting seven type-faces for a research, this 11-point Caslon Old Face was estimated to be approximately equal in face-size to the other 10-point faces of the series. The length of the lower-case alphabet is 122 points and the average number of characters per pica is 2.75. The visibility of 11-point Caslon Old Face is 91 percent of that of 10-point Textype.

TEXTYPE, 10-point, 3 points leaded

Readability of this size of type for a line-length of 21 picas increases markedly as the leading is increased from solid-set to 3 points. Apparently there is some gain from a further increase in leading but the increase is so little compared with the unsatisfactory appearance of the page that no further increase in leading appears to be desirable for 10-point type. It is particularly interesting to compare this paragraph with the preceding one.

MEMPHIS MEDIUM, 10-point, 3 points leaded

This type is the second of four weights or degrees of boldness of the square-serif family. The average length of the lower-case alphabet is 135 points and the average number of characters per pica is 2.5. The visibility of 10-point Memphis Medium is 106 percent of that of 10-point Textype.

Figure 7. Specimens of Caslon Old Face, Textype, and Memphis Medium, printed on coated paper. Luckiesh and Moss (1942) pp. 402–4.

reading extended text; and *reading speed*, in this case the unhurried 'natural' reading speed of passages set in the different typefaces. On the basis of these tests, Luckiesh and Moss rated Memphis Medium highest, followed by Textype, with Caslon Old Face rated much lower.

Gage, however, thought that these results were not credible. In reference to the apparently low rating of Caslon Old Face, Gage quotes the eminent printer Daniel Berkeley Updike's encomium of Caslon as the ultimately readable type. He also criticized the assumption made by Luckiesh and Moss that because, on average, readers completed the fewest lines of texts set in Caslon Old Face, it was the slowest to read. On the basis of character count, which is much higher for Caslon than for the other typefaces, more characters were in fact read, even if fewer lines of text were involved. Subjects therefore read Caslon faster than reported. Gage was also sceptical about the inclusion of Memphis in the comparison, since it was designed for display setting, not text. That a South American newspaper (unnamed) had recently used it for text merely reflected the special printing conditions of newspapers, in which ink might appear thin and grey and was applied to poor-quality paper. In his opinion, Textype was much more readable in text than Memphis Medium. Pointing out that Memphis Medium is distinctly blacker than Textype, Gage raised the question of whether the visibility meter was simply measuring

weight, a measure that he believed was not in itself a reliable indicator of readability:

> If we were to be finally bound by these tests, we would say that 10-point Memphis is more readable than 10-point Textype, which is in turn superior in legibility to 11-point Caslon Old Face – provided that each is set 21 picas measure, leaded 3 points, and printed on white book paper.
>
> But we do not yet know further factors which must control any sweeping comment on the general legibility traits of a type family as a whole. Would these tests made with 8-point or 14-point sizes bring similar results?
>
> If the favorable showing of Memphis Medium is due to its weight, then how about the still heavier weights of the same family? The Bold and Extra Bold are far higher in visibility rating. Are we to have blackness of type the determining factor for ease of reading? Common sense says 'no.' But how far does habit control our judgment? Has tradition stopped our use of types that might be easier to read than those normally selected?
>
> Such are the questions we have tossed back to our collaborators, and the further development of this research must inevitably seek the answers.[29]

Luckiesh's response to the questions raised by Gage in his report was based on his own experience in lighting research, as well as his doubts about the role of aesthetics.

> The widespread conflict between utilitarian and esthetic considerations emphasizes the need for scientific data. Often when these become available, esthetic demands are tempered or diminished; and design and practice are more equitably influenced by utilitarian ends. In lighting practice only a decade ago lighting fixtures were designed, sold and bought on the basis of 'beauty,' regardless of the penalties human beings paid for living and working under glaring and inadequate light. Incontrovertible data from sound researches have greatly changed the design of lighting equipment. In a similar way, type-design, typography, etc., must strike a proper balance, and this can only be achieved with certainty if the data are available pertaining to the utilitarian aspects.[30]

Luckiesh did, though, accede to testing the impact of boldness on reading. Linotype supplied him with samples of text in 10-point Memphis, in four weights: light, medium, bold and extra bold (figure 8, overleaf). The results of the tests conducted with them, as plotted graphically, were surprising (figure 9).

The key result was the *divergence* among the measures of visibility, readability (blink rate), and reading speed. Crucially, between medium and bold, there was little difference in reading speed, some difference in visibility, but a significant difference in blink rate. Before these tests, Luckiesh evidently believed that while blink rate was the most sensitive measure associated with readability, visibility and reading speed would still correlate: thus low blink rate would go together with high visibility and higher reading speed. But the tests in fact showed the different measures diverging. While bold showed higher visibility and reading speed than medium, it also showed a significantly higher – not lower – blink rate. Gage's suspicions about the too-simple correlation of weight with readability were confirmed. Luckiesh immediately understood that this implied a separation of

29. Gage (1939), c.
30. Gage (1939), c.

MEMPHIS LIGHT

Meanwhile the plough does its work year by year, the harvests are gathered, the builders build as they are told, the tradesmen work and acquire fresh devices; the knowledge of writing spreads; novel things, the horse and wheeled vehicles and iron, are introduced and become part of the permanent inheritance of mankind; the

MEMPHIS MEDIUM

At a remote period in geological time there is good reason for supposing that there was land where now the Atlantic waters roll, but there is no evidence for and much against any westward extension of Europe or Asia since the Miocene period. But civilization is a matter of at most the last twenty thousand years and probably

MEMPHIS BOLD

Very soon the seafaring men must have realized the peculiar freedom and opportunities the ship gave them. They could get away to islands; no chief nor king could pursue a boat or ship with any certainty; every captain was a king. The seamen would find it easy to make nests upon islands and in strong positions on the main-

MEMPHIS EXTRA BOLD

The trade that was going on in the ancient world before the sixth or seventh century B.C. was almost entirely a barter trade. There was little or no credit or coined money. The early empires got along without coin altogether. The ordinary standard of value with the early

Figure 8. Specimen showing four weights of Memphis: Light, Medium, Bold, and Extra Bold. Luckiesh and Moss (1942), p. 168.

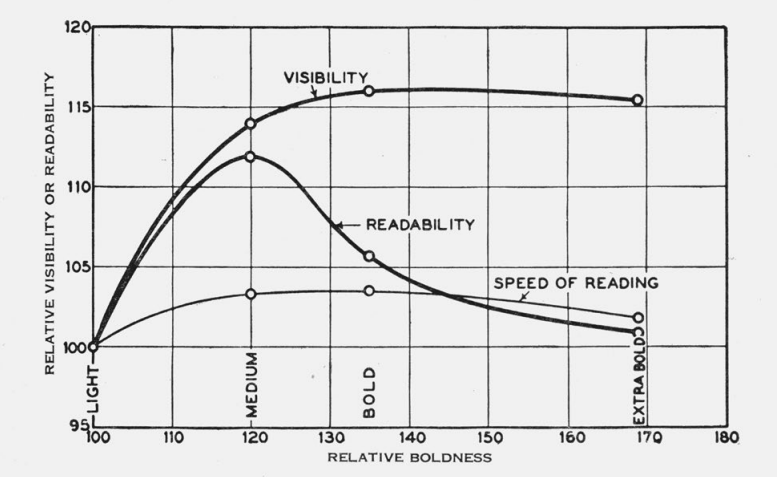


Figure 9. Graph comparing the visibility, readability, and reading speed of four weights of Memphis (Light, Medium, Bold, Extra Bold). Luckiesh and Moss (1942), p. 171.

ease of reading from *speed* of reading. Previously, he had included reading speed and accuracy as aspects of readability; from now on he would reserve the term 'readability' for ease of reading alone, and would identify blink rate as the single most reliable measure of readability. Luckiesh henceforth used the term readability to express 'the integral effect of physical factors which influence ease of reading' (in 1939), and 'that attribute of reading materials that governs the relative ease with which different materials may be read by subjects possessing normal vision and exhibiting normal responses' (in 1942).[31]

In addition to the breakthrough to a new definition of readability (figure 10, opposite), Luckiesh was also excited to note that 'the outstandingly important fact revealed by the visibility-boldness relationship is *the definite indication of an optimum boldness*.'[32] The existence of some optimum was suggested by the chart just discussed (see figure 9), and it tempted Luckiesh to read more into it. Undeterred by Gage's scepticism over his claim that the bolder Memphis Medium was more readable than Textype, Luckiesh now asserted that optimum boldness was *the* key factor in readability, and that other stylistic variations were functionally less important, and concerned aesthetics more than utility. The lower blink rate associated with

31. Luckiesh and Moss (1939), p. 652; Luckiesh and Moss (1942), p. 93. 'Ease' is the low end of a scale involving effort, with low effort being 'ease' in reading and high effort 'difficulty' in reading. Luckiesh's definition is apt because the subjective experience readers have of text of poor readability is of having to struggle, to expend more effort in achieving or trying to achieve a normal reading rate. The experience is of ease or difficulty in achieving and sustaining normal reading speed. However, the measures Luckiesh used are of 'strain' or 'fatigue' resulting from heightened efforts in reading. Ease of reading, or high readability, then, is inversely correlated to resulting strain or fatigue – more ease, less resulting strain or fatigue.

32. From the dossier compiled by Chauncey Griffith on Matthew Luckiesh, Margaret I. King Library, University of Kentucky (Lexington), box 8, folder 4; italics here represent underlining in the typescript.

Figure 10. Report on the importance of blink rate in measuring eyestrain, published in the *Chicago Daily Times*, 19 January 1939. The information and photograph were sent to newspapers by the GE Lighting promotional staff at Nela Park.

Memphis Medium indicated that most text types should be bolder. Luckiesh felt that indeed Memphis Medium hit that level of optimal boldness. Further, he saw that in the range of boldness of normal text types, he might use the visibility meter as a proxy for the blink rate tests involving numerous subjects. Using the visibility meter one could, he believed, quickly measure the visibility of a typeface, and then rule on whether its boldness matched the standard of Memphis Medium.[33]

These assertions were included in what appears to be a final report to Linotype on the results of the research conducted by Luckiesh and Moss. In the 'Conclusions' of the report, Luckiesh wrote:

> 1. A marked enhancement in the readability of the printed page can be obtained by augmenting the boldness of many types which are now being recommended for body text.
>
> 2. An enhancement in readability is decisively less promising by means of alterations in the configuration of modern type-faces than by utilizing the optimum degree of boldness.
>
> 3. The design of a type of optimum readability may now be guided with reliability and exactness by measurements of visibility within a range which is now fairly definitely known.
>
> It is emphasized again that all these conclusions are based on the intrinsic visibility of the types studied and upon the facility with which they are actually read, and not upon introspective appraisals of the appearance of the printed page.
>
> . . . .
>
> The design of a superior type – a 'Super Textype' seems to us to be an important future step in our cooperative research program. We are preparing a separate report on the immediate possibilities of further research and of the development of an ideal type-face *in the light of our current knowledge.*[34]

The reaction of Linotype

Although Harry Gage was apparently the guiding hand at Linotype in the collaboration with Luckiesh and Moss, the only record of the company's reaction to Luckiesh's last report is from Gage's colleague, Chauncy H. Griffith, vice president in charge of type design.[35] Griffith was a formidable figure who had, since 1915, brought quality and efficiency to the design and production of Linotype typefaces. His successes included the 'Legibility Group' of types for newspapers. For help in arriving at a considered view of the Luckiesh and Moss research, Griffith turned to his friend and Linotype's foremost type designer, W. A. Dwiggins.

Dwiggins's assessment of the Luckiesh and Moss work was both admiring and critical.[36] He felt they had made important discoveries, but that their conclusions over-reached their test results. Dwiggins began his assessment by agreeing with Luckiesh that 'in the case of any given size of letter, there is certainly an optimum weight for that letter and size, and it's good to have a way for finding it.' But Dwiggins also observed that there were optimums for other variables, too, and that the problem therefore was 'how to hitch it [boldness] up with other "optimums" and make a team – that is what is needed

33. Luckiesh used his visibility meter to test various newspaper types, as he had done with Memphis Medium; several newspapers published letters from him to demonstrate how the visibility and readability of their types had been improved. Luckiesh archives, box 6 (clippings of letters), Western Reserve Historical Society Library.

34. Griffith dossier on Luckiesh, cited above, n. 32.

35. Griffith dossier on Luckiesh, cited above, n. 32.

36. The assessment by W. A. Dwiggins is contained in the Griffith dossier on Luckiesh, cited above. Quotations that follow are from this assessment.

to make the boldness findings valuable in the case of new designs.' Dwiggins then listed the other four other variables, recommending that their optimums should also be determined:

– the ratio of stem weight (breadth of the vertical stroke) to white space (i.e. the combined total area of white space of a letter's counter(s) and its side bearings);
– the triple ratio of areas: stem (boldness), to counter, to side bearings;
– the side bearings; and
– the thick-thin contrast, or 'modeling' ('One would like to know: whether monotone and no modeling; or whether modeling, and if so how much or how little.')

For Dwiggins, none of these variables, boldness included, operated independently; instead, 'all these factors interplay; and the investigator has to keep all the balls in the air at once, as I see it, because each variable influences all the others. … Optimum weight alone is not enough to go ahead with'. He thought that the research should continue since the basis for an ideal text type was far from clear.

> I don't want any of this to make it seem that I am blowing cold on the laboratory end of the game. I'm for getting all the facts *via* eye-blinks that a feller can get together. …
>
> It's simply that I feel a little shaky about the L.& M. findings because I find the investigators so eager and willing to find a Super-Textype on 'boldness' alone: '*Now* we've got a sure basis to work on!' They haven't. They've built one corner of the foundation very nicely.

Despite its insightful analysis, it seems that Dwiggins's assessment was not shared with Luckiesh and Moss; indeed they seem to have been quite unaware of the typographic variables Dwiggins listed and discussed, or their influence on a readers' subjective impression of boldness. However, Luckiesh was aware that a complex combination of factors influenced readability, and so likely would have been receptive to Dwiggins's ideas. These circumstances alone certainly warranted continuation of the research, but that never took place, as at this point the collaboration between Linotype and Luckiesh and Moss ended.

The abrupt termination of the collaboration may have resulted from Griffith's negative reaction to the final report on the research, or possibly from a change in Gage's role at Linotype. In any case, the report's dismissal of design factors other than boldness as irrelevant to readability must have struck Griffith as not only lacking an understanding of type design, but also presumptuous. Writing some years later, in 1956, Griffith described Luckiesh and Moss's results as 'disappointing'. The pair, he felt, displayed a 'lack of understanding of the basic principles of type designing which contribute to the ease and pleasure of reading. Their Conclusions … were weighted with theories so abstract and impractical in their broad application to the subject under consideration as to become of little value in our work, and the research project was discontinued.'[37]

After its conclusion, Luckiesh nevertheless remained enthusiastic about the collaboration with Linotype. It represented a culmination of the work he and Moss had done on ease of reading, which had begun more than ten years earlier when they began investigating

37. Griffith dossier on Luckiesh, cited above. Alternatively, it seems equally fair to suggest that as an applied scientist, Luckiesh, in his urgency to find practical applications, was too quick to reach for single answers and practical measures, thereby lessening the value his research might have had for typography.

non-performance indicators of fatigue. Over this period, Luckiesh and Moss had published numerous articles on their work, which they now summarized in the book, *Reading as a visual task* (1942). The book's most important contribution was to present in one place the authors' advances of the previous ten years, namely, the development of a specific, scientifically constructed and validated concept of readability. This presentation was, in turn, supported by the authors' clear sense and understanding of the realities of type and typography. The relative sophistication of the book was no doubt partly enabled by Gage's contributions to the research of Luckiesh and Moss, including preparation of the test samples supplied by Linotype. The analysis of typography that resulted was well informed, notwithstanding the absence of those issues raised by Dwiggins. The book's sophistication is paired with the authors' inventiveness in devising new tests that consider the reader's experience from different angles. Luckiesh and Moss produced two different measures of reading speed, and measures of reflectance, visibility, fatigue in muscles surrounding the eye, blink rate, and many others. These tests and their measures all contribute to an understanding reading as a visual task, both its performance and the costs of that performance.

In retrospect, *Reading as a visual task* holds its own as indeed sophisticated, inventive, well-informed, and of practical as well as theoretical value. But its publication was not entirely well-received. The most important critical review at the time appeared in the *Journal of Applied Psychology*, edited by Donald Paterson. The reviewer was Miles A. Tinker, another leading researcher on typography and reading. In his own work, Tinker relied predominantly on tests of reading speed, tests which he carried out in collaboration with Paterson. Apparently threatened by the minimal importance assigned to reading speed by Luckiesh and Moss, Tinker's review of *Reading as a visual task* dismissed its methodology as unscientific and its results on reading as invalid. It would be the start of a dispute with Luckiesh that would preoccupy Tinker for the next 25 years.

Miles Tinker

Miles Tinker (1893–1969), like Luckiesh, grew up from modest beginnings. One of nine children of a poor Massachusetts farmer, he was the only member of his family to go on to higher education.[38] Tinker completed an undergraduate degree at Clark University in Worchester, at the time one of the leading centers for psychology in the United States; a PhD followed at Stanford University under the leading psychologist Miles Terman. Tinker, the student, struck his teachers as intelligent, academically ambitious, hard working and agreeable, but not particularly original. Terman found Tinker's farm-boy manner of speaking disconcerting and gave him a list of words with which to practice correct pronunciation, so he would not be mistaken as ignorant; Tinker was not offended but grateful. Tinker went on to spend the whole of his career at the University of Minnesota where he was highly productive. There his research focused on measuring eye movements in reading and on the impact of typographic variables on reading speed.

38. Biographical information about Tinker is taken from Sandra Wright Sutherland, 'Miles Tinker and the zone of optimal typography' (PhD thesis, University of Washington, 1989). Here and below, our account of Tinker's family history, education, attitudes to Luckiesh, and reception among typographers (and others) is indebted to Sutherland.

Methodology

Tinker's research methodology followed the philosophy of his teachers, who were doing 'psychometrics', attempting to objectively measure mental qualities, and other human abilities and qualities. One of the pioneers of psychometrics had been his teacher Miles Terman. Terman belonged to the second generation of research psychologists in the USA, many of whom were anxious to show how their field could be of social benefit. His approach was to define a characteristic – most famously, intelligent quotient (IQ) – measure it reliably, and correlate it with other measures such as academic and life achievement.[39] Terman treated the mind as a 'black box' within which he did not seek to understand the causal mechanisms. Instead, he used statistical methods in an effort to ensure that a given trait was reliably measured (i.e. that tests consistently produced the same result for the same person) before doing correlations with other variables.

Tinker's approach to research was also influenced by Paterson, who was a colleague at the University of Minnesota and partner in much of Tinker's work.[40] Like Terman, Paterson had been inspired by Alfred Binet's measurement of intelligence and was similarly devoted to psychometrics. His own research method was criticized by one former student as 'dust-bowl empiricism' – implying that data about empirical relationships, amassed in large quantities but lacking any unifying theory, was barren and unenlightening. Another of Paterson's students named the approach Paterson advocated 'the Minnesota point of view' and characterized it more positively:

> Concepts should be defined; definitions should be operational so that they can be measured; questions should be approached through empirical research; the measurement of individual differences is central; conclusions should be based on objective data rather than on subjective surmise; and research should focus on the search for results that can be applied.[41]

In valuing a narrow focus on the measurement of empirical data, Paterson's method appeared sober and careful. But the aversion to theory meant the method was in fact susceptible to bias and confounding factors. This is because theory unavoidably enters into the choice of what data should be collected and studied, and therefore the effort to avoid a theoretical framework increases the likelihood that one is simply uncritical of the theory that one has (unconsciously) adopted.[42]

In a manner similar to Terman, Tinker would focus on an apparently singular attribute, legibility, and use just one measure of it, reading speed, to determine differences among typefaces and text settings. And although in his career Tinker did do eye movement studies, and occasionally discussed theories of the reading process, in his legibility tests he followed the a-theoretical methodology of Paterson. At the urging of Paterson (who hired him at Minnesota), Tinker embarked on what he felt was a rigorous and methodologically sound program of work to test reading speed against typographic variables.

39. Terman, notably, was the 'Stanford' in the 'Stanford-Binet' intelligence test, which built on the work of Alfred Binet. Terman also defined and measured other qualities such as 'masculinity' and 'femininity'.

40. Paterson's role as 'partner' seems to have been as instigator and guide to Tinker, who then did the work. Tinker published many articles on typography as sole author, including all his articles debating with Luckiesh (discussed below). Paterson published nothing on typography as sole author; rather, he was a frequent collaborator with others on a variety of topics, published many co-authored articles, and supervised numerous graduate students. See Sutherland (1989), ch. 3.

41. Lofquist (1991), p. ix.

42. The measurement of IQ provides an example, where scientists have questioned the concept as too one-dimensional in that it fails to capture a full expression of intelligence, i.e. that the test can measure intelligence in certain areas, such as in mathematics, but not in others, such as social skills. A good theoretical framework could provide a richer view. One troubling outcome of this methodology was Miles Terman's initially racist interpretations of IQ.

Tests, results, reactions

Miles Tinker was one of the earliest researchers to study type and reading systematically. Prior to his work, experiments in legibility had been mostly confined to discerning letters. One previous study focused on reading had used speed of reading as a measure. This study had been carried out by an advertiser, Daniel Starch, in an effort to understand what made advertising copy effective; the study, however, had been done without any systematic controls.[43] Tinker's plan was to adopt Starch's approach (i.e. measure speed of reading), but introduce controls as advocated by Paterson. Tinker would systematically test numerous aspects of typography, from type size and line width in text, to the tabular setting of numbers.

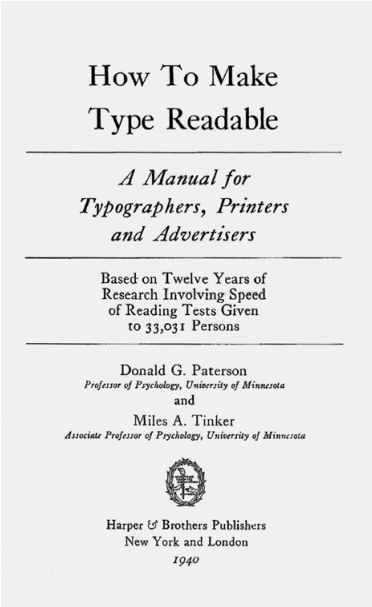
Tinker's first step was to employ a test for reading speed that also controlled for comprehension. To do this, he adopted an existing test, the Chapman-Cook. It involved the measurement of both reading speed and comprehension by inserting into a two-sentence text a 'rogue' word whose meaning was inconsistent with the overall meaning of the text; the reader was required to cross out the rogue word to demonstrate comprehension. For example:

> When I am enjoying anything very much, time seems to go very quickly. I noticed this the other day, when I spent the whole afternoon reading a very uninteresting book.[44]

(The rogue word is 'uninteresting', which should be 'interesting'.) By ensuring that the text was being read for comprehension, Tinker felt he could then isolate and vary typographic factors to assess and compare their effect on reading speed. To ensure that these factors were indeed isolated, subjects underwent two Chapman-Cook tests, each with a different text, which Tinker then compared. Finding no differences of statistical significance in the results, the tests were declared generally 'reliable'. But at no point did Tinker compare the Chapman-Cook test results against other tests of reading speed – a serious lapse in view of later methodological standards.

This research into reading speed, undertaken by Tinker and Paterson in the late 1920s and throughout the 1930s formed the basis for their co-authored book *How to make type readable: a manual for typographers, printers and advertisers* (1940). On the title page they declared that it was 'based on twelve years of research involving 33,031 persons' (figure 11, overleaf). In their introduction they explained that the book made no reference to earlier research: 'since the bulk of previous investigations are seriously deficient or misleading, due to defects in methodology, references to them would not have been helpful.'[45] They also dismissed as irrelevant the subjective preferences of typographers, since the results of reading tests demonstrated that any variation in the choice of common roman typefaces (of the kind used by typographers) made little difference to 'legibility'.[46] Tinker and Paterson's message, in effect, was a claim of ownership over their subject, a claim that displaced other researchers and typographers, and which was based on supposedly superior methodology incorporating a validated test (Chapman-Cook) and a large number of test subjects.

43. Tinker and Paterson (1928), p. 359.
44. Tinker (1963), p. 21.
45. Paterson and Tinker (1940), p. xvi.
46. In *How to make type readable*, Tinker and Paterson refer to 'legibility' and 'readability' as synonyms; previously, Tinker had referred to 'legibility' alone before adopting 'readability' as the preferred term. The change was possibly influenced by *The science of seeing* (1937), which Tinker reviewed, in which Luckiesh and Moss argue that the term 'readability' was preferable because it would not be confused with visibility (p. 455). Both Luckiesh and Moss, and Tinker and Paterson, claimed responsibility for developing the term. In the late 1940s Tinker reverted back to using 'legibility' alone, ostensibly because 'readability' had also begun to be widely used to characterize the contribution good writing (rhetoric) could make to ease of reading. Tinker's reversion in effect served to distinguish him from Luckiesh, and suppress the specific association of readability with physical factors affecting ease of reading.



How To Make
Type Readable

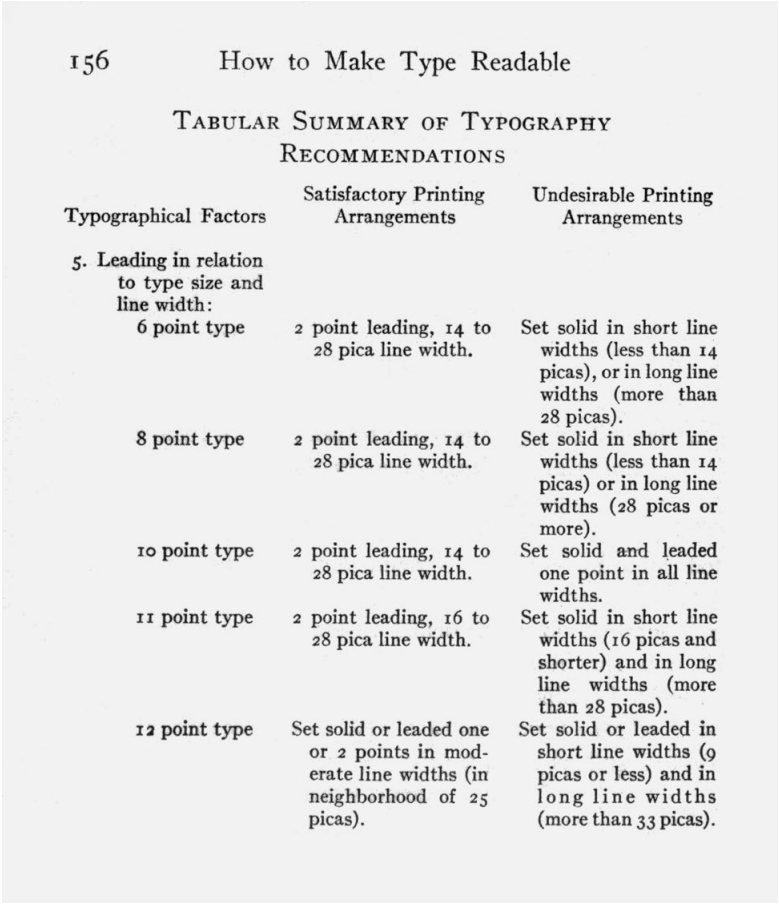
A Manual for Typographers, Printers and Advertisers

Based on Twelve Years of Research Involving Speed of Reading Tests Given to 33,031 Persons

Donald G. Paterson
Professor of Psychology, University of Minnesota
and
Miles A. Tinker
Associate Professor of Psychology, University of Minnesota

Harper & Brothers Publishers
New York and London
1940

Figure 11. Title page of *How to make type readable* (1940) by Donald G. Paterson and Miles A. Tinker.



156 How to Make Type Readable

TABULAR SUMMARY OF TYPOGRAPHY RECOMMENDATIONS

Typographical Factors	Satisfactory Printing Arrangements	Undesirable Printing Arrangements
5. Leading in relation to type size and line width:		
6 point type	2 point leading, 14 to 28 pica line width.	Set solid in short line widths (less than 14 picas), or in long line widths (more than 28 picas).
8 point type	2 point leading, 14 to 28 pica line width.	Set solid in short line widths (less than 14 picas) or in long line widths (28 picas or more).
10 point type	2 point leading, 14 to 28 pica line width.	Set solid and leaded one point in all line widths.
11 point type	2 point leading, 16 to 28 pica line width.	Set solid in short line widths (16 picas and shorter) and in long line widths (more than 28 picas).
12 point type	Set solid or leaded one or 2 points in moderate line widths (in neighborhood of 25 picas).	Set solid or leaded in short line widths (9 picas or less) and in long line widths (more than 33 picas).

Figure 12. Excerpt from the summary of typography recommendations. *How to make type readable* (1940), p. 156.

A notable feature of Tinker and Paterson's book is a table of recommendations for (readable) typography. At first glance, the recommendations appear impressively comprehensive, covering kinds of type, size of type, width (i.e. length) of line, size of type in relation to width of line, leading, leading and line width in relation to type size, spatial arrangements of the printed page, black print versus white print, colour of print and background, and printing surfaces. Yet on closer inspection, the recommendations are odd and come across as arbitrary, unsupported, or simply ill-advised from the standpoint of typographic practice.

One example may suffice to demonstrate the character of the recommendations; it concerns the arrangement of type that relates leading to type size and line length (figure 12). There is no reference to the impact of typeface variables such as x-height and set width (the length of a lower case alphabet); the recommendations give as 'satisfactory' line lengths of up to 28 picas for any type size between 6- and 11-point; line lengths shorter that 14 picas are (bizarrely) given as unsatisfactory for 6- and 8-point sizes; and so on. Throughout, the recommendations seem to reflect a wish to fix 'optimal' ranges of typographic variables, but without demonstrating the dynamic relationships between them, as might have been shown (for example) by graphing data comparatively (as Luckiesh had done).[47]

How to make type readable was intended by Tinker and Paterson as a manual for practitioners to improve their work. But it appears

47. Sutherland (1989); Sutherland's thesis is that Tinker's work was animated by a search for 'the zone of optimal typography'.

that the 'Typographers, Printers and Advertisers' to whom it was addressed found its recommendations of little or no use, and its publication was met largely with indifference.[48] Among those who ignored the book and the studies that led up to it were Luckiesh and Moss. Although they had earlier quoted Tinker in *The science of seeing*, their discoveries relating to boldness and blink rate apparently led them to regard Tinker's work as of little or no importance, since they now argued that reading speed was a relatively insensitive measure of the quality of a reader's experience, showing only small variations above thresholds for fluent reading.[49] Visibility (as measured by the Luckiesh-Moss meter) and blink rate, on the other hand, showed larger variations and so were more revealing measures of a reader's experience. Luckiesh and Moss's unwillingness to take much notice of Tinker's work – which was invested wholly in reading speed as a measure – and the dismissal that this implied, probably played a significant part in spurring on Tinker's campaign against Luckiesh's work.

The Luckiesh-Tinker dispute

Miles Tinker's review of *Reading as a visual task* by Luckiesh and Moss, in the *Journal of Applied Psychology* (edited by Paterson), was highly critical of the book – indeed so critical that it sparked a seven-year dispute, in print, between Tinker and Luckiesh. Tinker's colleague Paterson stayed out of the debate, but other researchers would also become involved.[50]

In his review, Tinker praised the book's results on visibility but dismissed its account of readability, asserting that other researchers had found blink rate an invalid measure of visual fatigue. He also dismissed Luckiesh's critique of reading speed as an insensitive measure of readability, saying that Luckiesh's failure to control for comprehension invalidated Luckiesh's data on reading rates. He concluded by accusing Luckiesh and Moss of an 'extreme lack of experimental controls', of 'ignor[ing] certain worthwhile research contributions', and of 'inadequate appreciation of certain fundamental principles of reading'.[51] In responding to Tinker's initial attack, Luckiesh pointed out that Tinker failed to cite the evidence he claimed contradicted Luckiesh and Moss's results, and challenged Tinker to do so. As to Tinker's belief that rate of reading is a measure of readability, Luckiesh remarked 'he merely claims this to be.

48. Sutherland (1989), ch. 6, *passim*.

49. Luckiesh (1940), p. 268: 'The normal rate of reading is limited by perceptual phenomena rather than by the physical characteristics of the visual stimuli when the reading is done under the usual supra-threshold conditions.' See also Luckiesh and Moss (1942), p. 121.

50. The exchanges between Tinker and Luckiesh include Tinker (1943a), (1945), (1946), (1948), (1948a), (1949), (1950); and Luckiesh (1943), (1944), (1946), (1947), (1947a), (1948), (1948a), (1948b), (1949). A related exchange between Bitterman and Luckiesh includes Bitterman (1945), (1946), (1947), (1948a); Luckiesh (1946), (1947a), (1948a); and Wood and Bitterman (1950). Luckiesh made no further comment after retiring in 1949; Tinker summarized his case in a review article of 1950 (Tinker, 1950), then recapitulated his arguments in his widely cited book, *The legibility of print* (1963). In the discussion that follows, we are particularly indebted to the research of John Stern and his colleagues who reassessed this complex debate fifty years later, with admirable clarity and insight. See Stern (1994a).

51. Tinker (1943).

He has published no results which prove this.'[52] Luckiesh also (perhaps unwisely) returned some of the insults of Tinker's review, questioning Tinker's own competence as a researcher.

In this first acrimonious exchange, and indeed over the course of the debate that followed, there was nevertheless a lack of full engagement on issues, as Tinker and Luckiesh never agreed on precisely what the actual issues were. One example was the issue of testing actively for comprehension. In all his research Tinker had used the Chapman-Cook test, or an expanded version of it, which included repeated checks on comprehension during the measurement of reading speed. Tinker argued that one had to regularly check for minimum basic comprehension otherwise reading rate measures could be misleading because of 'dawdling' by readers. Luckiesh, on the other hand, believed that when testing for ease of reading, the reading should neither be rushed nor interrupted; instead it should proceed at a 'natural rate' if the test was to be reliable. Luckiesh had been assessing ideas such as Ponder and Kennedy's hypothesis that blinking relieves 'mental tension', and so relieves one kind of fatigue or encumbrance to reading. Rushing the 'natural rate of reading' might increase fatigue, while interrupting reading (for example to test for comprehension) might give time for recovery, thereby reducing the fatigue that had resulted from continuous reading.[53]

For Luckiesh, the consistency and convergence of his results spoke to their reliability. A large part of his work on seeing was not applied research, but rather basic research, which he hypothesized and tested for possible causes of difficulty and fatigue in reading, and the physiological impact of these causal factors. He called this basic research 'axiomatic', stating 'Certain "axiomatic" researches indicate with surprising consistency a qualitative relationship (which is sufficient for our objectives) between decreasing blink rate and increasing readability – ease of seeing. Dr. Tinker completely ignores the extensive coordination and apparent consistency of our results'.[54] In a subsequent response to Luckiesh,[55] Tinker countered by citing the blink rate experiments of McFarland, Holloway, and Hurvitch.[56] These three researchers had found no consistent increase in blink rate relative to time-on-task, and therefore rejected blink rate as an unreliable measure. But in their key experiments, they used just three test subjects who completed only a single (unrepeated) test; as Luckiesh later noted, the (small) sample was 'anything but reliable'.[57]

Throughout the dispute, Tinker never addressed Luckiesh's own claims for his results: that ease of seeing is a different variable from reading speed, and that it has a complex relation to reading speed. This is what Luckiesh was investigating using blink rate. After Luckiesh had repeatedly urged the necessity of using his experimental controls and conditions, Tinker did conduct experiments that he claimed replicated these. Tinker's goal was to determine whether blink rate was both a 'reliable' measure and a 'valid' one. By reliable Tinker meant that if a measurement was repeated, it would produce the same results consistently. In his reliability tests, Tinker repeated blink rate tests on the same individuals and got the same results, confirming that blink rate was indeed a reliable measure. As for

52. Luckiesh (1943), p. 361.

53. In regard to the mitigation of 'dawdling', Luckiesh chose skilled readers for his tests and used what he felt was interesting reading material (*A short history of the world* by H. G. Wells); test subjects were told to read as they normally would.

54. Luckiesh (1943), p. 360.

55. Tinker (1943a).

56. McFarland *et al* (1942). McFarland, Holloway and Hurvitch were researchers at Harvard Business School, which also the published their report.

57. Luckiesh (1947a), p. 267. More recently Stern *et al* (1994a, p. 4) noted that while they 'do not wish to appear unkind' to McFarland and his colleagues, their work 'probably would not have been accepted for publication by a psychological journal, since it contained no statistical evaluation of results but many conclusions.'

validity, Tinker defined it as 'fidelity to an established criterion.'[58] When he measured blink rate against his own criterion measure of reading speed, he found that blink rate gave different results from reading speed, and so concluded that blink rate is not a valid measure.[59]

In this subsequent experimental work by Tinker a variety of problems are, however, evident in the 'proof' that blink rate lacked validity as a measure. The first problem was in the posited criterion itself. Luckiesh had argued that blink rate and reading speed diverge at certain points, with blink rate being the more sensitive measure of the impact of varying typographic conditions on ease of reading. Tinker, by simply positing reading speed as the criterion measure of validity for all aspects of reading, was making an illegitimate circular argument (known as '*petitio principii*' or 'begging the question'). The issue under debate – the validity of using reading speed to measure the impact of typographical factors on ease of reading – is assumed and reasserted, rather than defended by an appeal to independent evidence or principles. Tinker's fallacious approach improperly guaranteed that wherever a measure diverged from reading speed, it would be declared invalid.

There are also specific problems with Tinker's experiments. For example, in a test comparing text set in uppercase with text that mixed upper- and lowercase, Tinker used the Chapman-Cook test, which Luckiesh had pointed out violated the controls necessary for valid results. Again, when comparing newspaper text with book text, Tinker did not control for the x-height or line length of the texts; Luckiesh had found short lines gave lower blink rates, thus undermining the validity of contradictory results Tinker claimed to have got. Finally, in Tinker's experiments on illumination levels, which most closely replicated earlier experiments by Luckiesh, Tinker got seemingly incompatible results in 1945 and 1949.[60] Did the problem lie in Tinker's experimental set-ups, or protocols? In any event, Tinker never acknowledged the problem or sought an explanation.

In addition to Tinker's repeated critical articles over several years, the work of Luckiesh and Moss was also notably challenged by L. Carmichael and W. F. Dearborn in their book *Reading and visual fatigue* (1947). They concluded that sustained reading, even for up to six hours, had no fatiguing effects. Blink rate was tested and no increase was observed, contrary to the results of Luckiesh and Moss. Other researchers, however, raised serious concerns about the validity of Carmichael and Dearborn's work. J. Brožek immediately pointed out that their research involved regular interruptions to reading to test for comprehension, thus violating Luckiesh's control conditions.[61] And later, E. C. Poulton crucially pointed out that Carmichael and Dearborn had made errors in their statistical calculations; when the calculations were done correctly, their data in fact showed increases in blink rate, roughly in line with the results of Luckiesh and Moss – even with the interruptions Carmichael and Dearborn had introduced.[62]

Apart from the contributions and contentions of other researchers, the dispute between Tinker and Luckiesh – a sequence of claim,

58. Tinker (1950).
59. Tinker (1945) and (1946).
60. Tinker (1945) and Tinker (1949)
61. Brožek (1948), p. 420. They also, it seems, did not record food and toilet breaks over a 6-hour period, which would allow significantly more time for recovery from fatigue.
62. Poulton (1958). Luckiesh never responded to Carmichael and Dearborn. There are reasons to suspect that their work did some damage to Luckiesh's subsequent reputation. These include Luckiesh's failure to address their assertions, the prestige of the the authors' association with Harvard, and the passing of ten years before Poulton published his discovery of Carmichael and Dearborn's erroneous statistical calculations. It may also be the case that, with the prestige of its Harvard researchers behind it, *Reading and visual fatigue* played some role in discouraging Linotype from making further reference to Luckiesh's work after 1947, following their publication of *Researches in readability*, which featured and praised Luckiesh's work (discussed below). An example of the persistence of Carmichael and Dearborn's apparent credibility is Sheedy and Larson (2008), who, when discussing blink rate, refer to Carmichael and Dearborn without noting the problems in their methods and subsequent corrections by other researchers.

counterclaim, explanation, and retort – continued until 1949, when Luckiesh retired, after which he made no further comment publicly. By 1950, Brožek, Simonson, and Keys put forward their moderate recommendation that researchers should exercise 'greater caution than at times was present in the controversy regarding the utility of this criterion [of blink rate]'.[63] Tinker, however, without hesitation, continued to discredit Luckiesh both publicly and in private.[64] In his last published articles before retirement, Luckiesh, though irritated by Tinker, seems much less affected by Tinker's criticisms, which he apparently viewed both with contempt and perhaps some resignation; on his part, Luckiesh was confident that the foundation of his work was 'incomplete, but not unstable'.[65] Tinker, unable to let the matter rest, persisted in discrediting Luckiesh even after both had retired, getting in the last word in his *The legibility of print* (1963). By this time Tinker had refined the presentation of himself as an objective and conscientious scientist delivering authoritative judgments based on solid evidence. The book is superficially very persuasive in its dismissal of Luckiesh's work on blink rate – and by extension, Luckiesh's notion of readability. But when one studies the evidence, as fatigue researcher John Stern and his colleagues later did, a quite different picture emerges.[66] Tinker was not actually acting as a fair-minded scientist, but rather as a lawyer assembling a case against Luckiesh, without admitting that much of the evidence cited as authoritative – that of McFarland, Holloway, and Hurvitch, or of Carmichael and Dearborn – either lacked rigour or had been convincingly contested.

Apart from studying the evidence of Tinker's critique of Luckiesh's research, as found in the published record, it is also worthwhile to look more generally at the methodological divide between them. For Tinker's arguments against Luckiesh – and indeed against earlier researchers on reading – were based on his claim that his research methods were 'valid' and others' research methods were not. To understand the issues concerning validity of concepts, we need in our next section to take a brief excursion into developments in philosophy of science, particularly in the field of psychology.

Methodological divide

Tinker and Luckiesh were divided by differing views of good scientific method, in Tinker's case his adherence to the notionally a-theoretical 'black box' approach of psychometrics, and in Luckiesh's, the creative search for immediate causes following Claude Bernard's model for experimental medicine.

For a black box approach to research into a trait or attribute to succeed, a researcher must choose as a variable for testing a factor that is actually a key part of the causal chain of events. But if the variable for testing and quantifying the trait or attribute (IQ, for example) does not correspond to a causal factor, the correlations that the researcher measures are not informative, or worse, are misleading. The weakness of a black-box-and-correlations approach is its vulnerability to 'confounding' factors or variables. If factor A correlates with factor B, it may be that there is another factor, C, that has a causal impact on both A and B, and therefore confounds the claim of a causal

63. Brožek (1950), p. 62.

64. Tinker's son, Gordon Tinker, remembers his father pacing around the house muttering 'that Luckiesh son of a bitch'; he believed his father regarded Luckiesh as a fraud. Sutherland (1989), p. 86.

65. Luckiesh (1948c), p. 931. In his final published comments, Luckiesh observed that Tinker had come to illumination issues 'as a newcomer in this complex field in 1934 [and] took a definite position, based largely upon the inadequate criterion of rate of reading.' (Luckiesh, 1948a, p. 885) Luckiesh also noted that, in correspondence with Tinker, he had made efforts to clarify Tinker's misunderstandings, but found that Tinker's 'misunderstandings have become misrepresentations.' (p. 885) Unlike Tinker's family, Luckiesh's family was unaware of the dispute; personal communication from Peggy Luckiesh Kundtz and John Kundtz, 4 October 2012.

66. Stern (1994).

connection between A and B.[67] Those using a black-box-and-correlations approach have been very aware of the danger of confounding factors, and have sought to avoid them through the use of good statistical methods, operational definitions, or other techniques that would allow them to form a truthful causal picture. But confounding factors cannot be avoided in a black box approach. Three theorists of scientific method, Karl Popper, Paul Meehl, and Donald Campbell, each studied the problem of confounding factors and all drew the same conclusion: the black box approach is fundamentally flawed because there is no substitute for hypothesizing possible causes; nor is there a routine, or 'algorithm', for avoiding confounding factors. As Claude Bernard had urged, researchers need to hypothesize causes, then rigorously test the hypotheses.

Popper supported Bernard's call for a creative search for immediate causes by pointing out that there is no purely logical path from data to reliable theory, that is to say, no algorithm. Theories, by definition, exceed any finite set of data, and so are fallible. Like Bernard, Popper argued that the only way to eliminate errors and get at the truth is to make bold hypotheses about causes, then seek out data to test them. Hypotheses are searchlights that can locate revealing tests; tests, in turn, are crucial for discriminating among theories that are nearer or more distant from the truth. To attempt an a-theoretical approach by simply gathering data and looking for correlations, which Popper called 'inductivism', is to risk working from a theory in any case, only unconsciously. The notionally a-theoretical approach thus makes a researcher uncritical of theories that may nevertheless be at work. Confounding factors may be overlooked and experimental results vitiated. And, with fewer imaginative theories of causation, fewer searchlights are in play and less of interest or practical use is likely to be discovered.

Tinker believed that by focusing on a single criterion for legibility – speed of reading – he could effectively control for problems of confounding factors entering his experiments. Paul Meehl, however, has shown that Tinker's confidence in such a single criterion was misplaced. Meehl completed his doctorate under Paterson and spent his career at the University of Minnesota, but later became a fierce critic of the Minnesota approach.[68] He pointed out that the claimed benefits of a single criterion for measuring an attribute, even when the criterion was operationally defined, were overblown. He noted that some concepts in physics, whose rigor psychologists were attempting to emulate, could indeed be measured by a single criterion and operationalized. 'Electrical resistance', for example, is defined as the ratio between the measures of voltage and current. However others concepts, such as 'electron', have no single criterion for measurement, and no associated operational definition. To show that 'electron' is a valid concept involved development of a rich theoretical framework, and a whole variety of tests of that framework.

Meehl noted that in psychology as well, many legitimate concepts do not have unitary operational definitions and should instead be regarded as 'hypothetical constructs.'[69] In psychology, these hypothetical constructs include any trait that cannot be measured simply and directly by a single operation; IQ, introversion, and depression

67. The notion of a confounding factor has been illustrated by a scenario involving ice cream and drowning. It might be proposed that eating ice cream causes cramps, and that cramps cause people to drown; therefore these two variables appear to correlate. A confounding factor, however, may be outside air temperature: in warm months, people both eat more ice cream and swim more frequently; it may be this variable that is causal, while the correlation of ice cream and drowning is spurious. See 'Confounding', en.wikipedia.org/wiki/Confounding.

68. Meehl (1989). While at Minnesota, Meehl contributed to the Minnesota Multiphasic Personality Inventory (MMPI), an instrument that is still (in revised form) one of the most widely used to assess personality traits.

69. MacCorquodale and Meehl (1948).

are examples. Meehl argued that for these hypothetical constructs 'validation' of their use in scientific experimentation and theorizing was a much more complex process.

Donald Campbell took up Meehl's idea and developed what became the most widely used approach to validation of hypothetical constructs. The problem of confounding factors, Campbell pointed out, goes beyond failure to recognize causal variables that are additional to those being tested. A construct may itself be misconceived or invalid, especially if it bundles two or more factors. For example, 'legibility' (broadly defined) bundles ease of reading and reading speed, so the correlations using that construct will be misleading.

Campbell proposed that to validate a construct, one must show that the construct stands up to different tests which *converge,* showing that the construct (such as legibility) varies in the same way in different situations. There must also be tests that *discriminate* this variable from other variables. Only when a variable that is a hypothetical construct is subjected to convergent and discriminating tests can a researcher draw conclusions about whether the variable labels just one causal factor. Campbell's methodology pushes experimental psychologists to consider competing causal hypotheses, just as Bernard and Popper had recommended.[70]

The guidelines Campbell set out in his important article of 1959 are very similar to those Luckiesh had adhered to twenty years earlier in his 'axiomatic researches'. Luckiesh did different tests of readability whose indicators of ease of reading (or low fatigue in reading) do converge. His measures of general muscular tension, of fatigue in the eye muscles involved in binocular vision, and of blink rate, were each independent and all three converged. Importantly, Luckiesh was also careful to do tests of the second kind identified by Campbell, those that discriminated ease of reading from other variables or constructs, namely visibility and reading speed; here Luckiesh found divergence. Luckiesh's concept of 'ease of reading' therefore exhibited Campbell's notion of construct validity.

In light of Campbell's standards for construct validity, the weakness in Tinker also becomes clear. In his treatment of the term 'legibility', he avoided evaluating legibility as a hypothetical construct. By choosing to use only speed of reading as legibility's measure, he was presuming without any independent check that a single criterion was enough to rank typefaces, type sizes, measures, line-feed, and so on, according to their legibility – which he also assumed is the same thing as readability.[71]

Tinker was well aware that the ability to read text depends on a number of factors being above minimum thresholds, and that these factors may continue above thresholds to have an impact on reader experience. Yet in his discussion in *The legibility of print*, he assumes (without any comment) that a single criterion is needed, then looks at candidates for that single criterion. He writes that 'to a considerable degree, legibility is defined in terms of a specific method of approach to the study of the problem.'[72] He mentions measuring speed of perception, perceptibility at a distance, perceptibility in peripheral vision, visibility, 'rate of work' (reading speed), and

70. It is no accident that Campbell and Popper take similar lines: both studied under followers of the anti-associationist Würtzberg School of psychology; Tinker's influences, on the other hand, were all associationists. For a discussion of how assumptions about human psychology influence research methodology, see Berkson (1984).

71. In 1928, when Tinker began his series of thirteen studies of typographical factors that would later become the backbone of *How to make type readable*, he focused on the impact of different typographic variables on reading speed, that is to say, on 'documenting empirical relationships' specifically (Tinker and Paterson, 1928). But by the early 1930s, he shifted from this narrow focus to the broader one of assessing 'relative legibility' (Tinker and Paterson, 1931). The shift was made without ever addressing questions about the validity of legibility as a hypothetical construct.

72. Tinker (1963), p. 5.

fatigue. This variety of factors suggests that Tinker's 'legibility' qualifies as a 'hypothetical construct', calling for the kind of (demanding) validation described by Campbell. But after characterizing legibility so expansively, Tinker then goes through a process of elimination, stating the limitations of each measure, as well as of blink rate, concluding that 'the large majority of investigators have come to depend upon some aspect of rate of work or speed … for studying the legibility of print. … It seems to have high reliability and apparently good validity.'[73] It is thus striking that Tinker makes a claim for 'good validity' in *The legibility of print* since, in his discussion of the speed of reading measure, he does not in fact demonstrate its validity as a criterion measure. Throughout *The legibility of print*, Tinker makes no formal case for the validation of reading speed as a measure of legibility, but instead remained content in his claim that because it is widely used, it is 'apparently' valid.

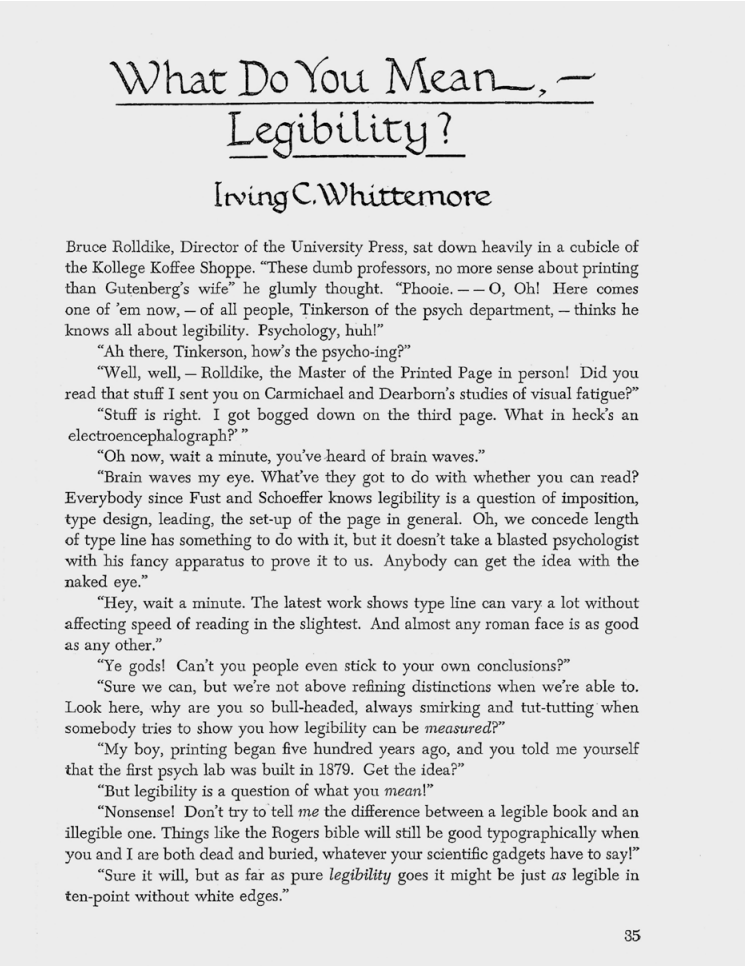
Tinker's treatment of 'legibility' was indeed muddled fundamentally. He wanted the notion of legibility to have a broad scope, including ease of reading and threshold measures. This would require the kind of complex validation Meehl and Campbell called for in the case of theoretical constructs. But Tinker settled on just a single measure and treated it much like an operational definition, whose meaning was unproblematic and therefore did not require demanding validation. While he settled on a single measure, giving 'legibility' a narrow scope, he continued to use the term in a broader sense. Tinker also failed to explore how reading speed relates functionally to visibility, distance, and other measures; nor did he discuss how ease of reading and speed of reading are related. Instead, he simply ploughed ahead with his own assessments of various typographic layouts, using only his preferred tests of reading speed; other possible approaches were ignored.

How to validate concepts was widely discussed in the 1940's and 1950's, including by Paul Meehl in Tinker's own department at Minnesota. But consideration of these discussions was consistently neglected by Tinker, with the result that his claims for 'validation' of reading speed as the sole measure of legibility were not well founded, either by the standards of his time or by today's.

Reception and value

In 1948, at the height of the dispute between Luckiesh and Tinker, a particularly astute survey of the state of reading research drew attention (by implication) to the narrowness of Tinker and Paterson's analysis. The popular graphic arts magazine, *Print*, published the article 'What do you mean, – Legibility?' by Irving C. Whittemore[74] (figure 13, overleaf). Whittemore, a psychologist, was also knowledgeable about typography and was aware of the debate. He began the article with a mock dialogue between 'Bruce Rolldike' (a conflation of Bruce Rogers, Carl Purington Rollins, Daniel Berkeley Updike) and 'Tinkerson' (i.e. Tinker and Paterson). The satire is heavy-handed, but Whittemore nevertheless concludes with a series of incisive questions, noting that the issue of 'legibility' (and reading) has many dimensions:

73. Tinker (1963), p. 30.
74. Whittemore (1948), pp. 35–7.



What Do You Mean—, —
Legibility?

Irving C. Whittemore

Bruce Rolldike, Director of the University Press, sat down heavily in a cubicle of the Kollege Koffee Shoppe. "These dumb professors, no more sense about printing than Gutenberg's wife" he glumly thought. "Phooie. — — O, Oh! Here comes one of 'em now, — of all people, Tinkerson of the psych department, — thinks he knows all about legibility. Psychology, huh!"

"Ah there, Tinkerson, how's the psycho-ing?"

"Well, well, — Rolldike, the Master of the Printed Page in person! Did you read that stuff I sent you on Carmichael and Dearborn's studies of visual fatigue?"

"Stuff is right. I got bogged down on the third page. What in heck's an electroencephalograph?' "

"Oh now, wait a minute, you've heard of brain waves."

"Brain waves my eye. What've they got to do with whether you can read? Everybody since Fust and Schoeffer knows legibility is a question of imposition, type design, leading, the set-up of the page in general. Oh, we concede length of type line has something to do with it, but it doesn't take a blasted psychologist with his fancy apparatus to prove it to us. Anybody can get the idea with the naked eye."

"Hey, wait a minute. The latest work shows type line can vary a lot without affecting speed of reading in the slightest. And almost any roman face is as good as any other."

"Ye gods! Can't you people even stick to your own conclusions?"

"Sure we can, but we're not above refining distinctions when we're able to. Look here, why are you so bull-headed, always smirking and tut-tutting when somebody tries to show you how legibility can be *measured*?"

"My boy, printing began five hundred years ago, and you told me yourself that the first psych lab was built in 1879. Get the idea?"

"But legibility is a question of what you *mean*!"

"Nonsense! Don't try to tell *me* the difference between a legible book and an illegible one. Things like the Rogers bible will still be good typographically when you and I are both dead and buried, whatever your scientific gadgets have to say!"

"Sure it will, but as far as pure *legibility* goes it might be just *as* legible in ten-point without white edges."

35

Figure 13. Satire on legibility by Irving C. Whittemore. *Print* magazine, 1948.

Legibility? What do you mean, *Legibility?* Do you mean:
(1) easy to read fast,
(2) easy to read at a distance,
(3) easy to read in dim light,
(4) easy to read when you haven't your glasses,
(5) easy on the brain,
(6) not tiring to the eyes,
(7) possible to grasp in big gulps of meaning,
(8) pleasant to read,
(9) inviting to the eye, or
(10) something else?
Before you pick your fights, Mr. Rolldike, Mr. Tinkerson – *answer the question, gentlemen!*

Whittemore's questions both identify the key issues that should be addressed by typographers and scientists, and in hindsight are suggestive of the gap between them. For master typographers (such as Rogers, Rollins, or Updike), enhancing reader experience was a multi-dimensional problem, involving many aspects of accessibility and aesthetics; for certain scientists (such as Tinker and Paterson), only one dimension of reader experience was proving of interest: efficacy of performance as measured by reading speed. Whittemore's challenge – as Dwiggins, too, had articulated earlier – was to understand the many dimensions of reader experience.

Tinker

Among typographers, there is little evidence that Tinker was accepted as an authority on the issues that concerned them. Books on typography sometimes refer to his work, but do not embrace his recommendations.[75] Among experimental psychologists, on the other hand, Tinker's *The legibility of print* is often treated as a standard reference on early research in typography, and is referred to with respect and as a trusted source.[76] But more recently, Tinker's claims to superior methodology and results have not held up well, in particular under the scrutiny of two commentators who are, in effect, the contemporary counterparts of Tinker and Luckiesh. They are Gordon Legge, an influential researcher into reading at University of Minnesota, and Mark Rea, a leading researcher in illumination at the Lighting Research Center at Resselaer Polytechnic (a contemporary equivalent of Nela Park).

In his published work, Legge acknowledges the extensive work of his predecessor, Tinker, which first defined the field of scientific testing of typography's impact on reading. But Legge also draws attention to the numerous problems in Tinker's work, and so, indirectly, throws into relief its flawed influence. The first problem Legge identifies is Tinker's exclusive reliance on the Chapman-Cook reading test, discussed above. Today, the Chapman-Cook test is apparently used only to assess impaired cognition due to brain injury, and is not used as a test of reading speed. Legge confirms the view that the Chapman-Cook, in particular its testing for comprehension, introduces confounding factors. He notes that 'Any measure of reading performance with high cognitive demand may dilute the impact of

75. This judgment is based principally on Sutherland (1989), as well as (e.g.) *The visible word* (Spencer, 1969), and more recently Lund (2004), Beier (2009), and others. Sutherland reviews graphic arts literature from the period during which Tinker and Paterson were active publishing their research, up until 1989, concluding that 'The audience which might have used this information did not embrace it.' Sutherland observes that, with the exception of Whittemore's article in *Print* (1948), the work of Tinker and Paterson on typographical factors and how to make type readable is 'not mentioned in the graphic arts literature of [the] time'; and that more recently Tinker and Paterson's work has 'survived only minimally' (see Sutherland, 1989, pp. 89–100). In *The Thames and Hudson manual of typography* (1980), McLean, who mentions Tinker as an authority, nevertheless maintained that 'no research so far published has been seriously helpful to designers concerned with the design of a straightforward reading matter for literate adults, except insofar as it has, in general, confirmed their practice. Research in legibility, even when carried out under the most "scientific" conditions, has not yet come up with anything fundamental that typographic designers did not already know – or believe – with their inherited experience of five hundred years of printing history'. (p. 47)

76. Legge (2006) refers, with apparent respect, to the work of Tinker and Paterson as 'influential', and that it 'surveyed the body of research to that time', though he follows this with the critique summarized below. Legge's critical view, however, is an exception. Elsewhere, when Tinker's results and assessment of Luckiesh are mentioned, they are typically just reported, not challenged. A passing comment by Kerfoot (1967, p. 121) – 'an authoritative review of the research on the legibility of print was prepared by Tinker (1963)' – is probably indicative of how *Legibility of print* was then and is still viewed by scientists making reference to typography. In general, Tinker's preferential reliance on reading speed as a measure of legibility, and the tacit acceptance of legibility as a simple, transparent, and unproblematic construct, seems widely shared. And while hardly conclusive, it may be of interest that Google Scholar search results give 126 citations for *How to make type readable* and 517 citations for *Legibility of print*.

visual factors including print legibility. The 'rogue' word method does seem to have substantial cognitive demand.'[77] Legge finds other problems, too, such as Tinker's ignoring of distance to reading material, which affects visual size and is therefore an important influence on reading speed. By doing so, another confounding factor was allowed to entered into his work.[78] Legge faults Tinker more summarily for his 'behaviourist' approach to research, which had 'operationalized the definition of legibility in terms of the measurement process.'[79] The views of Rea on Tinker are given in two articles published in 1986 and 1987.[80] Rea dismembers Tinker's work on illumination (though this lies outside the scope of this essay), pointing out many confounding factors in his experiments. Like Legge, he also takes issue with reading speed as the sole measure of visual experience.

The considerable reputation of Tinker's work grew out of the behaviourist tenor of his times. But in retrospect his work stands up neither to the generally accepted principles of methodology set out Campbell, nor to the critiques of Stern, Legge, or Rea. His reputation as an experimenter is surely unmerited: his methodology was flawed, his results questionable, and he was unjust both to his predecessors and to Luckiesh. His unfortunate influence lies not in so much in his specific claims as in his narrowing of the study of reader experience, which has probably held back progress in the understanding of typography and reading. But if Tinker's work is flawed, it also made a contribution that should be acknowledged. The performance measures he used, such as reading speed and eye movements, *are* important indicators of reader experience,[81] even if a good theory of reading should explain the impact of typographic features on *all* performance levels (reading speed, eye movements, comprehension, and so on) and non-performance costs. Given the likely importance of reading speed in particular, Tinker's data offer a starting point for further inquiry. His reliance on the Chapman-Cook test, however, means that none of his data can be relied on until his tests are repeated with better ones. Tinker's data on typography might still serve as a reference point for researchers, but should not be treated as authoritative.

77. Gordon Legge, e-mail to authors, 12 July 2011. For evidence, Legge cites Crossland (2008), an article in which he collaborated. The article documents a test comparison of reading speeds of sentences, with and without a true-false judgment being required. The true-false judgment was found to alter the reading speeds, including in tests on the influence of typographic features. Since the Chapman-Cook requires a kind of search process, in which (for example) line length might have a strong influence, it seems likely, on the evidence of Crossland, that the results of Tinker's tests would have been misleading.

78. It is worth noting two further examples of probable confounding factors in Tinker's work: he used point size rather than x-height as a measure of type size, and he did not control for illumination.

79. Legge (2006), p. 108. In this article, Legge emphasizes that research into reading requires a wider range of methods than was used by Tinker, which should include the psychophysical study of visual perception.

80. Rea (1986); and Rea (1987), p. 130 f. The latter is cited by the Illuminating Engineering Society as among the one hundred best articles in the field.

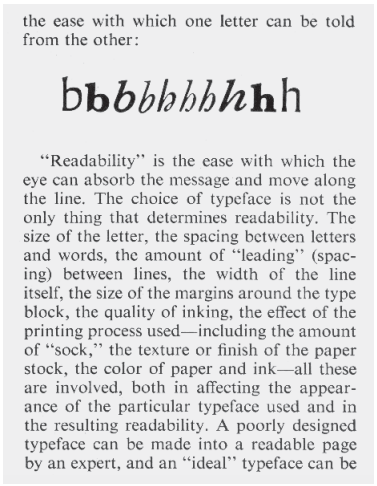
81. See, for example, Legge and Bigelow (2011).

Luckiesh

The reception and value of Luckiesh's work is, in several respects, the inverse of Tinker's. There are good reasons to conclude that the persistence of Tinker's attack on Luckiesh's work resulted in its near total disregard among experimental psychologists. Among typographers, however, Luckiesh's influence has been significant and lasting, if obscured. In English-speaking spheres, the notion of readability – ease of reading – he pioneered has been widely endorsed as a goal in the design of text and of typefaces intended for text composition. Luckiesh's name, however, and the experimental basis of readability, have been largely forgotten.

This situation arose in two stages, separated by two decades. In 1942 Harry Gage expressed admiration for Luckiesh's research, whose importance he emphasized in a laudatory foreword to *Reading as a visual task* (1942). Five years later, a handsome, illustrated pamphlet about Luckiesh's work, *Researches in readability*, was published by Linotype (also probably written by Gage).[82] This pamphlet now introduced two different concepts: Luckiesh's 'readability'; and 'legibility', which the pamphlet's author associated with 'quickness of perception' or the 'quick recognition of a word or phrase', such as in reading display type.[83] As already noted, legibility was a concept that Luckiesh had little use for, since it was generally linked to reading speed.[84] But here the two terms came together and were juxtaposed.

Much later, in 1967, the paired but distinct concepts were taken up by J. Ben Lieberman in his book *Types of typefaces*. Lieberman had consulted Paul A. Bennett who, as director of typography at Linotype between 1932 and 1962, was undoubtedly well aware of the research Luckiesh and Moss had done for the company. Lieberman probably learned from Bennett about the distinction Linotype had made between readability and legibility, but he tried to improve it. He defined legibility as 'ease with which one letter can be told from another'; as an example of a 'legibility problem' he illustrated the confusion that could occur between italic *b* and old face italic *h* (figure 14). His definition of readability drew on Luckiesh: the 'ease with which the eye can absorb the message and move along the line.'[85] But while taking up Luckiesh's definition, Lieberman did not directly cite Luckiesh and Moss. His book therefore seems to be the point at which the readability/legibility distinction was solidified, but the connection to Luckiesh lost. Many subsequent works on typography follow the distinction (with varying phrasing), and recommend


the ease with which one letter can be told from the other:

bbbhhhhhh

"Readability" is the ease with which the eye can absorb the message and move along the line. The choice of typeface is not the only thing that determines readability. The size of the letter, the spacing between letters and words, the amount of "leading" (spacing) between lines, the width of the line itself, the size of the margins around the type block, the quality of inking, the effect of the printing process used—including the amount of "sock," the texture or finish of the paper stock, the color of paper and ink—all these are involved, both in affecting the appearance of the particular typeface used and in the resulting readability. A poorly designed typeface can be made into a readable page by an expert, and an "ideal" typeface can be

Figure 14. Illustration of 'legibility'. *Types of typefaces* (1967) by J. Ben Lieberman, p. 85.

82. Linotype (1947). The pamphlet focuses on Luckiesh; Moss had died in 1943.

83. Linotype (1947). No tests on reading display type are cited in the pamphlet to support this definition of 'legibility', which seems instead to rely on the conventional view in typography that display typefaces function differently than text typefaces. Linotype sold text as well as display typefaces, and possibly wanted to employ a distinctive term when discussing the latter, in keeping with the conventional typographic distinction.

84. In Luckiesh and Moss (1942), p. 390, legibility is described as 'a term which has been narrowly defined as that characteristic of printed or written material which determines the speed and accuracy with which it may be read. As commonly used, it is indefinite and often ambiguous.'

85. Lieberman (1967), p. 65. Lieberman's definition of legibility may have been influenced by work at that time on letter differentiation; see, for example, Poulton (1965).

readability as a goal in continuous text composition, offering the testimony of practice that Luckiesh's original concept is sound.[86]

Recent and future research

Luckiesh's work can, we believe, inspire new ideas and insights both in the practice of typography and in experimental psychology. For the practice of typography an important insight of Luckiesh was that reader experience is the 'integral effect'[87] of many diverse factors affecting reading:

> 'The various factors determining the visibility and readability of reading material are always encountered in a complex combination. … As none of the factors is ever unaccompanied by others, it is impossible to rank them in a categorical manner.'[88]

This 'integral effect' suggests that different aspects of reading as a visual task may prompt different typographic decisions. Depending on the situation and the designer's goal, different design decisions can enhance one or another aspect of the reader experience.

The Luckiesh boldness chart (see figure 9) begins to capture this complexity in a way that may extend to other performance factors. Bolder types are generally used by typographers to direct the reader's attention or to orient the reader to a document's structure. Luckiesh's identification of visibility as a characteristic of bold typefaces offers a psychophysical explanation for this practice. High visibility seems to counteract the effects of crowding in peripheral vision, effects that make it difficult to isolate a word or words in a field of other words or letters. Similarly, Luckiesh's blink tests reveal that there is a psychophysical reason why medium weights of text are easier to read in continuous settings. There may also be a creative tension between pure performance measures and non-performance costs that typographers can productively exploit. Reading speed and reading comfort may at times be in conflict with one another, where a gain in one (reading speed) might outweigh a gain in the other (by minimizing effort), or vice versa. When this is the case, this can inform decisions about line lengths and column widths, where such a trade-off may be at issue.

Luckiesh's work points to the need to address the complexity of the psychophysical factors in reading, and therefore to consider carefully the intention of any piece of typographic design. Is the most important factor that a piece of text looks inviting, or has a specific visual look or aesthetic mood? Is it most important that readers can read the text with comfort? Is it most important that readers can find a part of the text readily on the page? All of these factors play against one another, and how different aspects of the text relate to each other and to the impression of the whole are of paramount importance. Awareness of the diversity of demands on the reader and their tensions with one another can help guide typographic decisions.

Reading and blinks

In the field of reading research, reading speed as a measure has proven to have more to offer than was revealed either by Tinker's

86. See, for example, Walter Tracy, *Letters of credit* (1986), James Felici, *Complete manual of typography* (2003), and Mitchell and Wightman, *Book typography: a designer's manual* (2005); among these examples, 'legibility' sometimes includes quickness in recognizing words as well as letters. For sources commenting on typeface design and readability, see for example Tracy (1986), pp. 30–2, and Berkson (2010) and (2011).

87. Luckiesh (1939), p. 652.

88. Luckiesh (1942), p. 370.

narrow approach or Luckiesh's downplaying of its significance. The recent work of Pelli, and of Legge, shows that (along with thresholds and critical values) one can identify 'fluent ranges' from graphs of reading speeds; Bigelow and Legge show fluent ranges for type size, using a variety of typefaces.[89] It seems likely, in turn, that graphing factors such as contrast, illumination, time, distance, and size, in order to observe how they compare and intersect, will add to the knowledge of how fluent ranges change. Building on the example of Luckiesh will involve studying the range and diversity of variables associated with the forms of typefaces, taking into account both performance measures that identify thresholds and ranges, and non-performance costs that identify optimums.

Research into spontaneous blinks of readers – Luckiesh's main innovation in research tools – is now more promising than ever, as recent discoveries about eye blinks open new avenues of research.[90]

To better understand the promise of new research into blink rate, it is necessary to briefly review Luckiesh's own research on blink rate. In his early work, Luckiesh tested for fatigue in the muscles controlling the eye, and in strain caused by eyeglasses with deliberately improper prescriptions; he also tested for general muscular strain and changes in heart rate. In tests on typography, he also studied fatigue by measuring blink rates from the first and last five minutes of an hour of continuous reading, and comparing them. But later, when studying typographic variables in detail, he only compared blink rates during a single 5-minute period. He never explained why he thought this shorter period equally valid to the earlier test. One plausible reason, though not stated explicitly by Luckiesh, is that he came to believe that ease of reading for the most part related not to fatigue in eye muscles but to the encumbrance of reading caused by 'mental tension'.[91] Luckiesh also acknowledged that some kinds of demands on the visual system initially reduced, rather than increased blink rate. The diversity of effects these demands produced indicates that different kinds of strain, and different factors inducing fatigue, should be separated and evaluated on their own, possibly by different types of tests involving eye blinks. Here, and in general it seems, Luckiesh was mistaken in putting all causes of strain and all factors inducing fatigue into one bundle.

The differentiation of kinds of strain and factors inducing fatigue has begun to take place in more recent research. James Sheedy and colleagues at the College of Optometry at Pacific University, for example, have studied strain in eye muscles caused by taxing physical conditions associated with reading, including reading from a screen. They have concluded that different conditions can cause strain in different eye muscles.[92] But there is reason to suspect (following Luckiesh) that the typographic issues involve the costs of brain processing rather than eye muscle strain. Further study using blink rate seems warranted.

Yu-Chi Tai and colleagues, also at the College of Optometry at Pacific University, studied the eye movements (saccades) associated with blinks in reading.[93] They found that blinks do not occur randomly, but are 'more associated with interruptions in reading:

89. Pelli (2007), Legge (2011), Legge and Bigelow (2011). Fluent range, as defined by Bigelow and Legge, is the broad but limited size range over which text can be read at the fastest presentation rate yielding a criterion reading accuracy. According to their data, for a reading distance of 40 cm, the fluent range corresponds to type with an x-height between 1.4 mm (4 points) and 14 mm (40 points).

90. Stern *et al* (1994).

91. It is this mental tension that Ponder and Kennedy (1927) claimed was relieved by blinking.

92. Sheedy and Larson (2008); Gowriskaran *et al* (2012). Sheedy and colleagues make reference to the blink rate work of Stern, and thus indirectly to Luckiesh.

93. Yu-Chi Tai (2009), slides 11, 13, 21.

corrective saccades, regressive saccades, and line change saccades.' These kinds of eye movements are interruptions of the normal reading process for 'visual difficulty' or 'interruption in visual acquisition'. The blinks associated with corrective saccades take longer than those during normal saccades, but are still shorter than the time spent in eye fixations. Such results support Luckiesh's theory that increased blink rate reflects some kind of tax on brain processing, specifically in the visual cortex.[94]

Other recent studies on blinking suggest more directly that Luckiesh was right in thinking that spontaneous blinks indeed reflect how visual processing taxes the brain. Volkmann and colleagues showed that visual processing is substantially suppressed during eye blinks.[95] Johns and colleagues showed, using reaction times on a vigilance test, that attention or vision is suppressed before and after both blinks and saccades.[96] This research of both these groups suggests that eye blink provides a physiological opportunity for some kind of recovery mechanism from visual work, in both neural processing and eye movement.

Colzato and colleagues showed an intriguing connection between eye blink and attentional blink.[97] Attentional blink (which is not an actual eye blink) occurs in tasks involving Rapid Serial Visual Presentation (RSVP). When a test subject is asked to identify a particular letter from a series of letters flashed at them, they succeed in doing so, but are unable to identify the next letter after the one identified if it is different and follows at a sufficiently short time interval. The critical interval varies with the individual; as it turns out, the attentional blink interval for an individual is highly correlated with that individual's spontaneous eye blink rate. Colzato and colleagues hypothesize that eye blink is thus connected to the attention processes in the brain and efficient processing in the visual cortex.[98] They also theorize that these processes involve dopamine, a neurotransmitter, suggested by the fact that those who suffer from Parkinson's disease, caused by low dopamine, have reduced eye blink rate, while those who suffer from schizophrenia have both increased dopamine levels and increased blink rates. This theory comes full circle to Ponder and Kennedy's original claim that blinks relieve 'mental tension'.

Colzato's studies suggest the hypothesis that eye blink is specifically connected to the temporary depletion of dopamine or other neurotransmitters, a depletion relieved by the way blinking momentarily shuts down activity in the visual cortex, allowing time for blood circulation to replenish chemicals to the temporarily inactive nerve cells. If this depletion is caused by non-optimal processing, Colzato's and Sheedy and colleagues' findings strongly suggests that Luckiesh's theory of blink rate – that the increase in blink rate over time may be a kind of summative measure of different typographic and environmental factors causing strain or fatigue in continuous reading – is ready to be revived.

For testing typographic variables, blink rate has the advantage of being exquisitely sensitive to brain activity; but it also has the disadvantage of being responsive to *many different* brain activities. It is important, therefore, to carefully separate different factors

94. The most recent work on blink rate by researchers at Pacific University (Gowrisankaran, 2012) shows a reduced total blink rate over a 30-minute period with greater cognitive demands. This work did not, however, compare initial and final blink rates over time, as Luckiesh did in tests that were more thorough-going.

95. Volkman *et al* (1980).

96. Johns *et al* (2009).

97. Colzato *et al* (2008).

98. This is reminiscent of Luckiesh and Moss (1934), which reports on efforts to supply quantitative data pertaining to the 'character or efficiency of the "cortical" integrational process' in the 'occipital' or visual cortex. Luckiesh never integrated this early work into his later research on reading, though the report is included in the bibliography of *Reading as a visual task*.

and their affects on blink rate. By holding potentially confounding factors constant, the impact of typographic variables on reading might be newly tested by blink rate and other aspects of eye blinks, checking Luckiesh's work and extending it to important questions that he did not address, such as the influence of serifs, or differences between reading from printed matter and from screens.

Luckiesh's approach to the study of reading can help bridge the divide in understanding between typographers and psychologists. Seeking to understand the costs of visual performance, and not just the performance itself, opens up many avenues for understanding the reading process and its relationship to typography. Questions of ease and fatigue are probably of key importance in assessing typographic variables at levels above basic thresholds for fluent reading. Luckiesh's concern for multiple factors in reading is also rich with possibilities. Though Luckiesh tended to emphasize one factor, blink rate, he researched many. The comments of his type-aware critics, Dwiggins and Whittemore, indicate how Luckiesh's ideas could be extended to new factors and variables that may affect reading in diverse ways. Investigation of this varied interaction through measures such as blink rate, visibility, and reading speed promises new insights into legibility and readability. The fruits of the collaboration between Luckiesh and Moss and Linotype can be a source of inspiration for typography and reading research well into the future.

Bibliography

Ackerman, P. L. (ed.) (2011). *Cognitive fatigue: multidisciplinary perspectives on current research and future applications*. Washington, DC: American Psychological Association

Beier, Sophie (2009). 'Typeface legibility: towards defining familiarity'. PhD thesis, Royal College of Art

Berkson, William and John Wettersten (1984). *Learning from error: Karl Popper's psychology of learning*. La Salle Illinois: Open Court

Berkson, William (2010). 'Reviving Caslon', part 2 ('Readability, affability, authority'), ilovetypography.com/2010/11/02/reviving-caslon-part-2-readability-affability-authority/

Berkson, William (2011). 'Readability and revival: the case of Caslon', *Printing History*, vol. 10, July 2011, pp. 3–24

Bitterman, M. E. (1945). 'Heart rate and frequency of blinking as indices of visual efficiency', *Journal of Experimental Psychology*, vol. 35, no. 4, pp. 279–92

Bitterman, M. E. (1946). 'A reply to Dr. Luckiesh', *Journal of Experimental Psychology*, vol. 36, no. 2, pp. 182–4

Bitterman, M. E. (1947). 'Frequency of blinking in visual work: a reply to Dr. Luckiesh', *Journal of Experimental Psychology*, vol. 37, no. 3, pp. 269–70

Bitterman, M. E. (1948). 'Lighting and visual efficiency: the present status of research', *Illuminating Engineering*, vol. 43, pp. 906–22; discussion (Matthew Luckiesh): pp. 923–31

Brožek, Josef (1948). 'Visual fatigue; a critical comment', *The American Journal of Psychology*, vol. 61, no. 3, pp. 420–4

Brožek, J., E. Simonson, and A. Keys (1950). 'Changes in performance and in ocular functions resulting from strenuous visual inspection', *The American Journal of Psychology*, vol. 63, no. 1, p. 51–66

Campbell, Donald T. and Donald W. Fiske (1959). 'Convergent and discriminant validation by the Multitrait-Multimethod Matrix', *Psychological Bulletin*, vol. 56, no. 2, pp. 81–105

Carmichael, Leonard and Walter F. Dearborn (1947). *Reading and visual fatigue*. Boston: Houghton Mifflin

Cobb, Percy W. and Frank K. Moss (1928). 'The four variables of the visual threshold', *Journal of The Franklin Institute*, vol. 205, pp. 831–47

Colzato, L. S., H. A. Slagter, M. Spapé, and B. Hommel (2008). 'Blinks of the eye predict blinks of the mind', *Neuropsychologia*, vol. 46, issue 13, pp. 3179–83

Covington, Edward J. (1992). *A man from Maquoketa: a biography of Matthew Luckiesh*. Nela Park, Ohio: General Electric Lighting

Crossland, M. D., G. E. Legge, S. C. Dakin (2008). 'The development of an automated sentence generator for the assessment of reading speed', *Behavioral and Brain Functions*, vol. 4 (www.behavioralandbrainfunctions.com/content/4/1/14)

De Vinne, Theodore Low (1886). *Historic printing types*. New York: The Grolier Club

DiLaura, David L. (2005). *A history of light and lighting: a celebration of the centenary of the Illuminating Engineering Society of North America*. New York: Illuminating Engineering Society of North America

Felici, James (2003). *Complete manual of typography*. Berkeley: Peachpit Press

Gage, Harry L. (1937). 'Research in readability 1 – the program for research', *Linotype News*, September 1937, vol. 16, no. 2, p. 2

Gage, Harry L. (1938). *Linotype News*, a. 'Research in readability 2 – effects of leading' (March 1938, vol. 16, no. 5, p. 2); b. 'Research in readability 3 – effects of color in paper' (May 1938, p. 2); c. 'Research in readability 4 – comparison of Caslon Old Face, Textype and Memphis medium' (July 1938, p. 2); d. 'Research in readability 5 – The quantitative relationship between visibility and type size' (November 1938, p. 2)

Gage, Harry L. (1939). *Linotype News*, a. 'Research in legibility 6 – the visibility of print on various qualities of "white" paper' (Jan 1939, p. 4); b. 'Research in readability 7 – the significance of the blink test' (July 1939, p. 5)

Gowriskaran, S., N. K. Nahar, J. R. Hayes, and J. E. Sheedy (2012). 'Asthenopia and blink rate under visual and cognitive loads', *Optometry and Vision Science*, vol. 89, no. 1, pp. 97–104

Johns, M., K. Crowley, R. Chapman, A. Tucker, and C. Hocking (2009). 'The effect of blinks and saccadic eye movements on visual reaction times', *Attention, Perception, & Psychophysics*, vol. 71, no. 4, pp. 783–8

Kerfoot, James F. (1967). 'Reading in the elementary school', *Review of Educational Research*, vol. 37, no. 2, pp. 120–33

Legge, Gordon E. (2006). *Psychophysics of reading in normal and low vision*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers

Legge, Gordon E. and Charles A. Bigelow (2011). 'Does print size matter for reading? A review of findings from vision science and typography', *Journal of Vision*, vol. 11, no. 5, article 8

Lieberman, J. Ben (1967). *Types of typefaces*. New York: Sterling

Linotype, Mergenthaler (1947). *Researches in readability*. Brooklyn: Mergenthaler Linotype Co.

Lofquist, Lloyd H. (1991). *Essentials of person-environment correspondence counseling*. Minneapolis: University of Minnesota Press

Luckiesh, Matthew, Percy W. Cobb, and Frank K. Moss (1927). 'An investigation of the reliability of the li test', *Transactions of the Illuminating Engineering Society*, vol. 22, pp. 43–51; discussion: pp. 51–78

Luckiesh, Matthew and Frank K. Moss (1929). 'The new science of seeing', *Electrical World*, vol. 93, pp. 15–39; discussion: pp. 39–49

Luckiesh, Matthew and Frank K. Moss (1930). *Transactions of the Illuminating Engineering Society*, vol. 25, p. 15

Luckiesh, Matthew (1932) 'The human seeing machine', *Transactions of the Illuminating Engineering Society*, vol. 27, p. 699 ff.

Luckiesh, Matthew and Frank K. Moss (1934). 'A view of the cortical integrational process through liminal visual stimuli', *Journal of Experimental Psychology*, vol. 17, no. 3, pp. 449–61

Luckiesh, Matthew and Frank K. Moss (1935). 'Visibility: its measurement and significance in seeing', *Journal of the Franklin Institute*, vol. 220, p. 431

Luckiesh, Matthew and Frank K. Moss (1937). 'The eyelid reflex as a criterion of ocular fatigue', *Journal of Experimental Psychology*, vol. 20, no. 6, pp. 589–96

Luckiesh, Matthew and Frank K. Moss (1937a). *The science of seeing*. New York: Van Nostrand

Luckiesh, Matthew and Frank K. Moss (1939). 'The visibility and readability of printed matter', *Journal of Applied Psychology*, vol. 23, no. 6, pp. 645–59

Luckiesh, Matthew and Frank K. Moss (1939a). 'Frequency of blinking as a clinical criterions of ease of seeing', *The Australasian Journal of Optometry*, vol. 22, issue 10, pp. 501–8

Luckiesh, Matthew and Frank K. Moss (1940). 'Criteria of readability', *Journal of Experimental Psychology*, vol. 27, no. 3, pp. 256–70

Luckiesh, Matthew and Frank K. Moss (1940a). 'A summary of researches involving blink-rate as a criterion of ease of seeing', *Illuminating Engineering*, vol. 35, pp. 19–32

Luckiesh, Matthew and Frank K. Moss (1942). *Reading as a visual task*. New York: Van Nostrand

Luckiesh, Matthew (1943). 'Some comments on Dr. Tinker's review of *Reading as a visual task*', *Journal of Applied Psychology*, vol. 27, no. 4, pp. 360–2

Luckiesh, Matthew (1944). 'On the rate of involuntary blinking', *Illuminating Engineering*, vol. 39, pp. 69–71

Luckiesh, Matthew (1946). 'Comments on criteria of ease of reading', *Journal of Experimental Psychology*, vol. 36, no. 2, pp. 180–2

Luckiesh, Matthew, Sylvester K. Guth, and Frank K. Moss (1947). 'The blink-rate and ease of seeing', *Illuminating Engineering*, vol. 42, pp. 584–8

Luckiesh, Matthew (1947a). 'Discussion: reading and the rate of blinking', *Journal of Experimental Psychology*, vol. 37, no. 3, pp. 266–8

Luckiesh, Matthew (1948). 'Recommended footcandle levels for prolonged critical seeing', *Journal of the Optical Society of America*, vol. 38, issue 8, pp. 712–18.

Luckiesh, Matthew (1948a). 'Discussion: a reply to Dr. Miles A. Tinker', *Illuminating Engineering*, vol. 43, pp. 882–905

Luckiesh, Matthew (1948b). 'Footcandle levels: threshold, ideal, optimum and recommended', *Illuminating Engineering*, vol. 43, pp. 395–411; discussion: pp. 411–15

Luckiesh, Matthew (1948c). 'Lighting and visual efficiency: The present state of research', *Illuminating Engineering*, vol. 43, pp. 923–31

Luckiesh, Matthew (1949). 'Important concepts underlying lighting for critical seeing', *Illuminating Engineering*, vol. 44, pp. 226–30

Lund, Ole (1999). 'Knowledge construction in typography: the case of legibility research and the legibility of sans serif typefaces'. PhD thesis, University of Reading

Lund, Ole (2004). 'Evidence-based typography or easy-going operationalism?', in *Proceedings of the first international conference on typography and visual communication: history, theory, education: June 2002 Thessaloniki*. Thessaloniki: University of Macedonia Press, pp. 93–8

MacCorquodale, K. and Paul Meehl (1948). 'On a distinction between hypothetical constructs and intervening variables', *Psychological Review*, vol. 55, no. 2, pp. 95–107

McFarland, R. A., A. H. Holoway, and L. M. Hurvich (1942) 'Studies on visual fatigue, part 5: ordinary blinking and visual fatigue', Cambridge, MA: Graduate School of Business Administration, Harvard

Meehl, Paul (1978). 'Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology', *Journal of Consulting and Clinical Psychology*, vol. 46, no. 4, pp. 806–34

Meehl, Paul (1989). 'Paul Meehl' [autobiography], in G. Lindzey (ed.), *A history of psychology in autobiography*. Stanford: University of Stanford Press, pp. 337–89

McLean, Ruari (1980). *The Thames and Hudson manual of typography*. London: Thames and Hudson

Mitchell, Michael and Susan Wightman (2005). *Book typography: a designer's manual*. Marlborough: Libanus Press

Patterson, D. A. and Miles A. Tinker (1940). *How to make type readable*. New York: Harper Brothers

Pelli, D. G., K. A. Tillman, J. Freeman, M. Su, T. D. Berger, and N. J. Majaj (2007). 'Crowding and eccentricity determine reading rate', *Journal of Vision*, vol. 7, no. 2, article 20, pp. 1–36

Ponder, E. and W. P. Kennedy (1927). 'On the act of blinking', *Quarterly Journal of Experimental Physiology*, vol. 18, pp. 89–110

Poulton, E. C. (1958). 'On reading and visual fatigue', *The American Journal of Psychology*, vol. 71, no. 3, pp. 609–11

Poulton, E. C. (1965). 'Letter differentiation and rate of comprehension in reading', *Journal of Applied Psychology*, vol. 49, no. 5, pp. 358–62

Rea, Mark S. (1986). 'Toward a model of visual performance: foundations and data', *Journal of the Illuminating Engineering Society*, vol. 15, pp. 41–57

Rea, Mark S. (1987). 'Toward a model of visual performance: a review of methodologies', *Journal of the Illuminating Engineering Society*, vol. 16, pp. 128–42

Sheedy, J. and K. Larson (2008). 'Blink: the stress of reading', *Eye*, vol. 17, no. 67, p. 95

Spencer, Herbert (1969). *The visible word*. New York: Hastings House

Stern, John A., Donna Boyer, and David Schroeder (1994). 'Blink rate: a possible measure of fatigue', *Human Factors*, vol. 36, no. 2, pp. 285–97

Stern, John A., Donna Boyer, and David Schroeder (1994a). 'Blink rate as a measure of fatigue: a review', report no. DOT/FAA/AM-94/17, August 1994. Washington, DC: U. S. Department of Transportation, Federal Aviation Administration

Sutherland, Sandra Wright (1989). 'Miles Tinker and the zone of optimal typography'. PhD thesis, University of Washington

Tai, Yu-Chi, J. Sheedy, and J. Hayes (2009). *Blink is not a random event in reading*. http://commons.pacificu.edu/compvision/2009/thursday/11/

Tinker, Miles A. and Donald G. Patterson (1928). 'Influence of type form on speed of reading', *Journal of Applied Psychology*, vol. 12, no. 4, pp. 359–63

Tinker, Miles A. and Donald G. Patterson (1931). 'Studies of typographical factors influencing speed of reading. VII. Variations in color of print and background', *Journal of Applied Psychology*, vol. 15, no. 5, pp. 471–9

Tinker, Miles A. (1934). 'Illumination and the hygiene of reading', *Journal of Educational Psychology*, vol. 25, no. 9, pp. 669–80

Tinker, Miles A. (1935). 'Cautions concerning illumination intensities used for reading', *American Journal of Optometry*, vol. 12, pp. 43–51

Tinker, Miles A. (1943). [Review of *Reading as a visual task*], *Journal of Applied Psychology*, vol. 27, no. 1, pp. 116–18

Tinker, Miles A. (1943a). 'A reply to Dr. Luckiesh', *Journal of Applied Psychology*, vol. 27, no. 5, pp. 469–72

Tinker, Miles A. (1945). 'Reliability of blinking frequency employed as a measure of readability', *Journal of Experimental Psychology*, vol. 35, no. 5, pp. 418–24

Tinker, Miles A. (1946). 'Validity of frequency of blinking as a criterion of readability', *Journal of Experimental Psychology*, vol. 36, no. 5, pp. 453–60

Tinker, Miles A. (1948). 'Readability of book print and newsprint in terms of blink rate', *Journal of Educational Psychology*, vol. 39, no. 1, pp. 35–9

Tinker, Miles, A. (1948a). 'Trends in illumination standards', *Illuminating Engineering*, vol. 43, pp. 867–81

Tinker, Miles A. (1949). 'Involuntary blink rate and illumination intensity in visual work', *Journal of Experimental Psychology*, vol. 39, no. 4, pp. 558–60

Tinker, Miles A. (1950). 'Reliability and validity of involuntary blinking as a measure of ease of seeing', *Journal of Educational Psychology*, vol. 41, no. 7, pp. 417–27

Tinker, Miles A. (1963). *Legibility of print*. Ames: Iowa State University Press

Tracy, Walter (1986). *Letters of credit*. London: Gordon Fraser Gallery

Volkman, F. C., L. A. Riggs, and R. K. Moore (1980). 'Eye blinks and visual suppression', *Science*, 207, pp. 900–2

Whittemore, Irving C. (1948). 'What do you mean, legibility?', *Print, a quarterly journal of the graphic arts*, vol. 5, no. 4, pp. 35–7

Wood, C. L. and M. E. Bitterman (1950). 'Blinking as a measure of effort in visual work', *The American Journal of Psychology*, vol. 63, no. 4, pp. 584–8