This study assessed the readability of popular print material (newspapers, magazines, and books) published in 1920. Selected passages from these books and articles were analyzed using the Flesch Reading Ease test and other, more theoretically interesting and relevant, readability formulas. Results indicated that "highbrow" publications such as the "New York Times" and the "Atlantic Monthly" contained more difficult prose than "middlebrow" publications such as the "Milwaukee Journal" and the "Saturday Evening Post." Results also indicated that detailed readability evaluation reversed anomalous ratings on the Flesch scale. Findings suggest that "high literacy" has demonstrated continued complexity, influence, and apparent inaccessibility to the average reader. (Four figures of data are included; 51 references and four appendixes of data are attached.) (RS)
DIFFICULTY OF TEXT AS A FACTOR IN THE HISTORY OF READING

William V. Trollinger and Carl F. Kaestle

Report from the Project on a Social History of the American Reading Public, 1880-1980

Carl F. Kaestle, Principal Investigator

Wisconsin Center for Education Research
School of Education
University of Wisconsin
Madison, Wisconsin

August 1986
A report for the project, A Social History of the American Reading Public, Carl Kaestle, Director. Supported by the Spencer Foundation and by the National Institute of Education (Grant No. NIE-G-84-0008), through the Wisconsin Center for Education Research.
The mission of the Wisconsin Center for Education Research is to improve the quality of American education for all students. Our goal is that future generations achieve the knowledge, tolerance, and complex thinking skills necessary to ensure a productive and enlightened democratic society. We are willing to explore solutions to major educational problems, recognizing that radical change may be necessary to solve these problems.

Our approach is interdisciplinary because the problems of education go far beyond pedagogy. We therefore draw on the knowledge of scholars in psychology, sociology, history, economics, philosophy, and law as well as experts in teacher education, curriculum, and administration to arrive at a deeper understanding of schooling.

Work of the Center clusters in four broad areas:

- **Learning and Development** focuses on individuals, in particular on their variability in basic learning and development processes.

- **Classroom Processes** seeks to adapt psychological constructs to the improvement of classroom learning and instruction.

- **School Processes** focuses on schoolwide issues and variables, seeking to identify administrative and organizational practices that are particularly effective.

- **Social Policy** is directed toward delineating the conditions affecting the success of social policy, the ends it can most readily achieve, and the constraints it faces.

The Wisconsin Center for Education Research is a noninstructional unit of the University of Wisconsin-Madison School of Education. The Center is supported primarily with funds from the Office of Educational Research and Improvement/Department of Education, the National Science Foundation, and other governmental and nongovernmental sources in the U.S.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Figures</td>
<td>vii</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>History of Readability Formulas</td>
<td>2</td>
</tr>
<tr>
<td>The 1920 Reading Public</td>
<td>13</td>
</tr>
<tr>
<td>Readability of Popular Print Material</td>
<td>14</td>
</tr>
<tr>
<td>Conclusion</td>
<td>28</td>
</tr>
<tr>
<td>References</td>
<td>31</td>
</tr>
<tr>
<td>Appendix A: Application of Flesch Formula</td>
<td>35</td>
</tr>
<tr>
<td>Appendix B: The Stories</td>
<td>41</td>
</tr>
<tr>
<td>Appendix C: Propositional Structure of the Two Stories</td>
<td>43</td>
</tr>
<tr>
<td>Appendix D: Clauses and Their Categories</td>
<td>49</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>&quot;The Play&quot;: Tree Structure</td>
<td>22</td>
</tr>
<tr>
<td>2</td>
<td>&quot;The Ghost&quot;: Tree Structure</td>
<td>23</td>
</tr>
<tr>
<td>3</td>
<td>&quot;The Play&quot;: Event Chain</td>
<td>26</td>
</tr>
<tr>
<td>4</td>
<td>&quot;The Ghost&quot;: Event Chain</td>
<td>27</td>
</tr>
</tbody>
</table>
Introduction

Historians of literacy have traditionally focussed their attention on people's ability to comprehend written language at a rudimentary level and on the socioeconomic status of individuals who possessed this crude level of literacy. Recently, however, historians of literacy have become more interested in the adult reading activities of the majority of the population who, in industrial nations, can read at some level. Together with historians of print culture, historians of literacy are attempting to learn more about the reading public or, more accurately, the several partially overlapping reading publics that exist in societies with diverse print forms and diverse education levels. What have literate people read in the past? How have they chosen what to read?

To answer questions such as these, the historian must begin with circulation figures for books, magazines, and newspapers, along with other information about the publishing industry. Once the historian has identified the most widely read material among different groups, a careful analysis of the texts themselves is necessary. Much of this analysis must focus on the content of the prose--its themes, information, and purposes--in order to understand the appeal and the function of popular publications. But some members of the reading public, now and in the past, read at a more sophisticated level than
others, and some prose is more difficult to read. Perhaps readability has been a factor in the selection of print matter in the past. A thorough analysis should therefore include attention to the more technical aspects of prose, vocabulary, syntax, and rhetorical structure.

History of Readability Formulas

At first glance this might not seem to be a terribly difficult enterpris. For years publishers, journalists, and teachers have used "readability formulas" to determine the reading ease or difficulty of certain passages. While there are well over sixty such formulas today (Klare, 1974-75, 1984), the majority involve a combination of three variables: sentence length, word length, and word familiarity. Sample passages, usually of 100 words, are pulled from the text in question, and a count is made of sentence length and/or word length and/or unfamiliar words, i.e., the number of words in the passage that do not appear on a list of commonly used English words. Each of these figures is then multiplied by weighted coefficients, and the resulting numbers are added together. This total score indicates the relative difficulty of the text.

The most important individual in the development of readability formulas is Rudolf Flesch. Flesch views himself as the missionary of clear, direct prose, a crusader who has sought to instruct people to write less like John Dewey and more like Erle Stanley Gardner. At times he presents himself as a prophet without honor, as in his recent lament that American schools have failed to adopt the phonic reading method he
has long championed (Flesch, 1985). But he also has not hesitated to
proclaim that his work on readability measurement has revolutionized
journalistic and business writing in America (Flesch, 1954, 1960). He
makes this boast with some justification. Klare (1984) has observed
that Flesch's 1948 Reading Ease formula has "become the most widely used
of all readability formulas" (p. 686). While in some measure this is
due to Flesch's promotional talents, the formula's popularity has
stemmed in great part from its simplicity: R.E. (Reading Ease) =
206.835 - .845 (number of syllables per 100 words) - 1.015 (average
number of words per sentence), with 0 being practically unreadable and
100 being easy for any literate person (Flesch, 1948).

Flesch's Reading Ease formula would seem to be a godsend to
historians of literacy. Not only is it easily understood and
implemented, but it also does not include the variable of vocabulary
familiarity. Hence, the historian is not faced with the intractable
problem that what might have been a familiar word in 1980 might not have
been familiar in 1880, and vice-versa. Using the Reading Ease formula a
scholar can quickly compute the reading difficulty of printed material
in the past, and with a few computations can chart changes over time.

Unfortunately, this is all too good to be true. Over the past
fifteen years, readability researchers, particularly cognitive
psychologists, have mounted an increasingly heavy attack on the Flesch
measuring device and other readability formulas. One basic criticism
involves the procedures used to verify that a particular formula does
indeed indicate the reading difficulty or ease of a written passage. A
perusal of Klare's lengthy discussion of current readability formulas
reveals that some have never been validated, or have been validated by comparing results with older formulas (Klare, 1974-75; Davison, Lutz, & Roalef, 1981). Of those formulas that have been independently validated, some of the most popular formulas, including the Flesch method, have used the McCall-Crabbs selection of readings (1925, 1950, 1961). These are short prose passages, each of which is followed by a set of multiple-choice questions. These passages and questions have been given to large numbers of children. Thus, it is known what percentage could answer how many questions for each passage. Formula-makers have simply taken certain textual features, such as sentence length, and then used regression analysis to ascertain to what degree each textual feature predicts comprehension of the passage (Kintsch & Vipond, 1979).

In an illuminating piece on the validation of readability formulas, Ramsay Selden (1981) has observed that there are definite advantages in using the McCall-Crabbs passages (p. 20): "They were normed on a large population, and there seems to be some empirical basis for grade placement of the passages" (e.g., passage C is at a fifth-grade level). But as Selden also pointed out, this method of validating readability formulas has serious flaws. There seems to have been no theory-grounded method for choosing the particular questions. It is hence questionable whether they are as "readable" as the passages they deal with. Moreover, the results of the comprehension tests are not controlled for a reader's previous knowledge, fatigue, and the like. Finally, creators of all formulas make the unwarranted assumption that their formulas, derived from particular passages, are applicable to other unrelated passages.
The validation problem should be enough to make most historians hesitate before using traditional readability formulas, such as Flesch's Reading Ease Measurement. But a more devastating criticism is that readability formulas are atheoretical. That is to say, the standard variables used to measure readability, such as word length, are only tangentially related to what makes a text more or less understandable. Kintsch and Vipond (1979) drive the point home: formula variables "are concerned with word and sentence properties at a superficial level; at best, they are correlated with whatever makes a text easy or hard, but they are not the causes themselves" (p. 337). That word and sentence length and vocabulary familiarity are not what make a text understandable is readily apparent. One could quite easily construct a sentence or passage that is quite "readable" in its use of short and common words, and yet scramble the words and sentences to make it completely incoherent. Of course, this is an extreme example. More telling are Davison and Kantor's research findings (1982). They examined original and revised versions of four texts. The revised versions were presumably easier to read because they had been adapted in keeping with traditional readability formulas. But Davison and Kantor discovered that in many cases the adaptations had actually made the texts less readable. They concluded that "readability formulas, in our opinion, fail to give any adequate characterization of readability, except in a purely statistical sense from which no particular valid conclusions can be drawn for creating readable texts" (p. 207).

All of this said, readability formulas are not designed to explain what makes a text readable. Their purpose is to predict reading
difficulty (Klare, 1984; Selden, 1981). Even taking into account the enormous problems with validation, traditional readability formulas do seem to correlate well with scores on comprehension tests. Klare even claims that correlation "values in the low .90" have become possible. For this reason, and because of their simplicity, readability formulas are still widely used in education. So, it does seem plausible that historians could use traditional readability formulas to get some indication about how difficult or easy readers would have found a particular text.

But just because one can make this limited defense of traditional readability measurements does not mean that the historian should stop here. In the past decade readability research has moved far beyond surface features of the text. Most current work in the field is grounded in the idea that reading involves much more than simply the decoding of words and sentences (Amiran & Jones, 1982). As Robert J. Tierney and James Mosenthal have observed (1982), "consistent with contemporary psycholinguistic and cognitive viewpoints is the notion that . . . discourse comprehension and discourse production involve a complex interaction among the cognitive structures of the author, the text, the cognitive structures of the reader, and the communicative situation" (p. 55). Reading is an interactive process, and thus readability is not simply a property of the text. Instead, James R. Miller and Walter Kintsch have pointed out (1980), "the readability of a text is determined by the ways that certain text properties . . . interact with the reader's processing strategies and resources" (p. 348).
This movement beyond text to the reader has an analog in literary criticism. Just as cognitive psychologists now insist that readability is a conjunction of text features and the reader's characteristics, an influential group of literary critics have recently asserted that a literary work of art comes about only when a reader interacts with a text. Denying the possibility of universal meanings or of objective standards of literary quality, these reader-response critics focus steadfastly upon the experience of the reader (Tompkins, 1980). This phenomenological approach opens up fascinating questions. Like interactionist models of reading comprehension, reader-response criticism asserts the uniqueness of individuals and the creativity of the reading process. Both the interactionist psychologists and the reader-response critics pose analytical problems for those who wish to study readers in the mass, not a single individual, and who cannot grab hold of their readers as research subjects.

Even the leading reader-response critics, however, do not generally do empirical research on living readers. Rather, they think hard about the process of reading for an "implied" or an "intended" reader. Some text-based models of comprehension do the same thing: they consider carefully and in detail the different mental processes involved in understanding a text. Drawing on linguistics and computer science as well as cognitive psychology, reading researchers have constructed models of how the content and organization of texts affect the hypothetical average reader.

This effort to measure the impact of textual content and organization on reader comprehension has led readability theorists in a
variety of directions. Some of the early critics of traditional formulas complained that the standard measuring devices completely ignored syntax. Out of this concern came efforts to rate sentences and passages according to the degree of syntactic complexity (Botel, Dawkins, & Granowsky, 1973; Dawkins, 1975; Endicott, 1973; Richek, 1976). Other researchers have asserted that the number of different concepts in a text has a direct bearing on its level of difficulty and have proposed ways to measure content density (more on this below). Borrowing from Halliday and Hasan's work on cohesion (1976), Judith Irwin has demonstrated that the greater the number of cohesive ties, i.e., textual links such as conjunctions, the easier readers find the text (1980). In the same vein, a group of scholars (Britton, Glynn, Meyer, & Penland, 1982) has emphasized the importance of textual signals, or, those "special words and phrases [which] signal ... how ideas are related and which ideas are most important" (p. 56). Along with simple syntax, the presence of cues such as summary sentences and title words help readers understand the text. Finally, in some of the most interesting work in the field of readability, Thomas Trabasso and others have observed that the degree to which elements in a story are tightly connected in a causal chain directly affects reader recall. Those stories that are more coherent are better understood and remembered than those that are less coherent (Amiran & Jones, 1982; Armbruster & Anderson, 1982; Stein & Trabasso, 1982; Trabasso, Secco, & Van Der Broek, 1983).

This new crop of readability researchers has performed complex analyses and empirical research on the comprehensibility of selected
short prose passages, underlying the complexity of the reading process. But historians of literacy, being weak in the flesh, yearn for a simple readability formula, something readily applicable to a wide variety of texts while also capturing some of the grammatical and rhetorical features of prose. Susan Kemper (1983) has provided one such tool, with her "inference load formula." Along with Trabasso and others, Kemper views texts as consisting "of chains of causally and temporally connected events, states, and actions" (p. 392). These chains are connected either by explicit statements within the text or by inferences made by the reader. The greater the number of inferences required of the reader, the more difficult the text. Using multiple regression techniques on the always popular McCall-Crabbs passages, Kemper has developed a new readability formula for providing the inference load score. It is calculated after labelling every clause as an action, a physical state, or a mental state, and then identifying necessary inferences not given in the text. According to Kemper, the inference load score correlates significantly with the reported grade level of the McCall-Crabbs passages.

While the Inference Load formula is harder to implement than Flesch's Reading Ease formula, it is not inordinately difficult. Moreover, in contrast with traditional measuring devices, Kemper's formula is grounded in current readability theory. All of this is to the good. But Kemper has limited herself to measuring causal coherence. Historians would, of course, like a readability tool that evaluates other rhetorical and organizational aspects of written passages. Walter Kintsch and his associates have created a more complicated model. In
the mid-1970s Kintsch and Keenan (1973) and Kintsch et al. (1975)
established that the number of propositions, defined as a predicate with
one or more concepts, and the number of different concepts were
important factors in determining a text's level of readability. A few
years later Kintsch and Van Dijk (1978), Kintsch and Vipond (1979), and
Miller and Kintsch (1980) constructed a readability model that also
included textual coherence. In this model coherence is defined as
argument or concept repetition, i.e., the degree to which ideas in a
passage are linked one to another. According to Miller and Kintsch
(1980), "if a segment of text is read that is not related to the current
contents of working memory, long-term memory must be searched to locate
a part of the text that can interrelate what has been read previously
with the current" material (p. 336). If this search is successful, the
reader reinstates the earlier material in working memory and relates the
new segment to it. If the search is unsuccessful, the reader must make
an inference about how the new segment relates to the rest of the text.
Both activities take time and effort. Texts with numerous demands for
reinstatements and inferences are more difficult.

Miller and Kintsch (1980) formalized their model, incorporating
within it a count of propositions, different concepts or arguments,
inferences, reinstatements, unfamiliar words, and long sentences. They
tested the model using 20 short prose passages, each read by 120
individuals. The test was quite successful: "multiple regression
predictions of readability (reading time per proposition recalled),
reading time, and recall ranged from $r = .8$ to $.9$" (p. 335). But Van
Dijk and Kintsch (1978) had also asserted that macroprocesses, or those

1,
processes that organize the semantic structure of an entire discourse or
text, play a key role in readability. While Miller and Kintsch's model
is geared toward the analysis of microprocesses, such as argument
repetition, one of Kintsch's students (Young, 1984) has created a model
which incorporates macroprocesses. This also allows the evaluation of
somewhat larger pieces of prose.

Kintsch's and Young's work provides historians with the best
substitute for a summary readability formula. But before enthusiastic
scholars commit themselves to using this model, they should take note of
some major problems. First, this model, particularly in its later
manifestations, is both difficult and expensive to implement. The Young
version, which analyzes macrostructures and is computerized, would,
nonetheless, be much too costly to apply to a lengthy piece of prose or
a great number of shorter passages. More important, theoretical
problems remain with both the Kintsch/Young and Kemper models. Both
claim to be operating from an assumption that reading involves an
interaction between reader and text. But these formulations are
developed on the basis of one hypothetical reader. There is no room for
variation. In commenting on the effort to create formulas based on
textual features that affect reading comprehension, Alice Davison (1984)
wrote, "These have some promise as models of one or two aspects of
reading comprehension, but to the extent that they isolate a few
variables, they share the defects of formulas in general for measuring
the difficulty of all texts and readers" (p. 141). These formulas do
not, indeed cannot, take into account that individual readers have
different cognitive skills, different motivations, and different situations, all of which play a crucial role in the reading process.

The theme of recent reading research is that reading is an interaction between one reader and one text. It is difficult to see how historians could implement such a method, except in the rare case where an individual has left detailed reactions to a particular piece of prose (e.g., Darnton, 1984). Moreover, like reading researchers, historians can use the best of the text-based models, those that imagine a reader processing a text, not just pondering big words or large sentences. Indeed, the Kintsch/Young, Kemper, and even Flesch models yield ratings that predict with high accuracy the difficulty real readers will encounter with texts, on the average. Of course, historians using these text-based readability measurements should be quite aware of their flaws and should avoid making sweeping claims about the "readability" of printed materials in the past. But imperfect analytical tools beat no analytical tools at all, and in this regard the Kintsch/Young model, the Kemper formula, and perhaps even the Flesch formula can be of some use.

As a first step in studying changes in popular prose from 1880 to 1980, we decided to explore the readability of selected prose pieces from widely circulated books, magazines, and newspapers in 1920. Before reporting how we selected the prose passages and what our readability analysis told us, a short sketch of the potential reading public in 1920 is in order. Who could read? How much education had they had? How many people could read difficult texts? How many could read the daily newspaper? How many did?
The 1920 Reading Public

America in 1920 reported low levels of outright illiteracy. Among whites the self-reported illiteracy rate was 4 percent; 23 percent of American blacks reported that they could not read, down from 30 percent a decade earlier. Education levels had risen steadily in the late nineteenth and early twentieth centuries. Thus, among those in their late fifties in 1920, about 2.5 percent had graduated from high school, while among those in their late thirties, 6 percent had. Among 17-year-olds in 1920, about 16 percent graduated from high school. College graduation was much rarer; of 23-year-olds in 1920, 2.6 percent had received a bachelors degree. While the higher reaches of education were thus restricted, going to school up through the junior high years was very common. A large national sample of people who were between 13 and 22 years old in 1920 later reported an average 9.4 total years of schooling, ranging from an average of 7.5 years among those whose fathers were in the lowest occupational status to an average 13 years among the children of the highest status fathers.¹ The United States, like many other industrial countries 65 years ago, had widespread basic literacy training, and a much more restricted group trained in higher literacy skills. The difficulty of a printed text could therefore have a dramatic effect on its potential audience (Census, 1920).

Book readers tend to be people with at least a high school education. A study of reading habits in 1923 confirmed that the

audience for books was more limited than that for magazines and newspapers. Parsons surveyed 314 adults in the Chicago area, an urban area with greater access to books and with higher levels of education than the national average. While 97 percent of the respondents said that they spend some time each day reading a newspaper, and 76 percent read in a magazine daily, only 53 percent reported reading in a book each day (Parsons, 1923).

Readability of Popular Print Material

Given this range of education levels and reading habits, we wished to assess the readability of popular print material in different categories. Were books in general more difficult to read than magazines and newspapers? Were highbrow publications like the New York Times and Atlantic Monthly written at a higher level of difficulty than local newspapers and the immensely popular Saturday Evening Post?

Of newspapers, we examined the New York Times; its rival, The World; a midwestern urban newspaper, the Milwaukee Journal; and a midwestern small-town newspaper, the Beaver Dam Daily Citizen (WI). We looked at one story as it was carried in all four newspapers. This was a more difficult task than it might appear. The Beaver Dam paper frequently reported on national news in blurbs so short that readability analysis was virtually impossible. The story we finally settled on was the April 1920 expulsion of socialists from the New York state assembly.

Our choices among magazines were the highbrow Atlantic Monthly, the middlebrow Saturday Evening Post, and the lowbrow Argosy. In keeping with the dates of our newspaper articles, we analyzed a fiction article
from the April 1920 issues of these magazines. We also analyzed a nonfiction article from the same issues of the Atlantic and the Post; Argosy did not carry any nonfiction pieces. Regarding our nonfiction choices, we had originally wanted to follow the same news story as it moved from the newspapers into magazines. This might have worked if our year was 1940 and we had a Time or Newsweek from which to choose. But our year was 1920, and because news stories in the Atlantic and the Post did not involve the reporting on a specific event, our scheme proved unworkable. We did, however, manage to look at two articles dealing with the same topic: profiteering in post-World War I America.

Finally, for books we simply chose the nonfiction and fiction best sellers of 1920. Actually, there was nothing simple about attempting to ascertain which books were the most popular. Publishers and booksellers traditionally have not divulged the precise number of copies that have been sold of a particular book. Best-seller lists typically have been constructed from calls to selected bookstores across the country, sometimes no more than 15-25 stores in all. Compiling lists from a limited and perhaps unrepresentative sample is not the only problem. If, say, twenty small stores reported that book A was number 1 in sales and five large stores reported that book B was number 1 in sales, A will rate higher on the best-seller list even though more copies of B may have been sold. And yearly best-seller lists just consist of an averaging of these possibly inaccurate weekly or monthly lists (Frase, 1953).

Despite all of these problems, a historian of popular print material in twentieth-century America is dependent on these lists. In
choosing the most popular nonfiction books for 1920 we went with Alice Payne Hackett, the monarch of best-seller lists (Hackett, 1945). Hackett developed her 1920 list from the monthly lists published in The Publishers' Weekly and Books of the Month. Her top two in 1920 were: Philip Gibbs, Now It Can Be Told, which was an account of World War I; and, John Maynard Keynes, The Economic Consequences of the Peace. But Hackett's list has an additional problem not mentioned above. A best-seller list for a particular year fails to reflect the fact that a book's sales will quite often be spread over two, even three, years. Fortunately, we had an alternative to Hackett when it came to fiction books: Irving Harlow Hart's "Best Sellers in Fiction During the First Quarter of the Twentieth Century" (1925). In constructing his 25-year list Hart noted all of the months a book was on the standard best-seller lists. Using Hart's list, we chose to examine the top four fiction books that included 1920 as one of the primary sales years: Sinclair Lewis, Main Street; V. Blasco-Ibanez, Four Horsemen of the Apocalypse; Harold Bell Wright, Re-Creation of Brian Kent; and, Zane Grey, Man of the Forest. We decided to look at four novels because we wanted to compare an acclaimed classic, Main Street, with the Wright and Grey pulp novels.

We began our analysis of these selected books and articles by applying the Flesch Reading Ease test to all of them. The formula and Flesch's instructions on how to interpret the scores are noted in Appendix A below. In keeping with Flesch's instructions, we randomly selected three 100-word passages from each article except the abbreviated Beaver Dam Daily Citizen article, and twenty-five 100-word
passages from each book. For each piece we applied the formula to all the 100-word passages and then computed an average Reading Ease score. Our results are given in Appendix A. The beginning and ending words of each article passage are given; for reasons of space, only the page numbers are given for the book passages. The reading ease score for each passage is given, as well as the average reading ease score for the entire piece.

The analysis of newspaper articles yielded no real surprises. New York newspaper articles were more difficult than Wisconsin newspaper articles. But as might be expected, the Milwaukee Journal and the Beaver Dam Daily Citizen did not have their own reporters covering the expulsion of Socialists from the New York assembly. Both newspapers were using wire service stories. Their markedly high Reading Ease scores lead us to speculate that perhaps newspaper editors in "the hinterland" rewrite wire service stories with an eye toward simplifying the prose. Plausible as this may be, we did not investigate this possibility further.

Applying the Flesch Reading Ease test to books did produce some surprises. Certainly it was not startling to discover that Economic Consequences of the Peace, by John Maynard Keynes, rated much more difficult to read than the five other best sellers. But one would not think that Sinclair Lewis' Main Street would have a higher Reading Ease score than Harold Bell Wright's The Re-Creation of Brian Kent and Vicente Blasco-Ibanez's The Four Horsemen of the Apocalypse. The technical reason for this is simple: Lewis' sentences are much shorter, particularly in comparison with Wright's elongated sentences. This
result illustrates some of the problems with the Flesch test. No theoretically sound reason exists for assuming that short sentences are necessarily easier to read than long sentences. Still, it might be true that short sentences generally make a text more readable. If so, perhaps the heightened readability of Main Street contributed to its best-seller status.

A major surprise came in periodical fiction. We expected that the highbrow Atlantic would have the most difficult prose, followed in order by the Post and the Argosy. But, as can be seen in Appendix A, according to the Flesch formula the Post story was difficult to read, the Argosy story rated a little harder than standard, and the Atlantic story ranked between fairly easy and easy to read. Could it be true that, at least in some cases, low- and middlebrow fiction in the 1920s was much more difficult to read than highbrow fiction? To test these anomalous results, and because time and money did not allow us to perform further readability analysis on all of our selections, we decided to apply the more sophisticated Kemper and Young models to the Atlantic and the Post stories. Specifically, we examined the Atlantic passage rated by the Flesch formula to be the easiest and the Post passage assessed to be the most difficult. Because the Kemper and Young models perform better with longer selections, we extended each of the original passages by over 100 words.

The story in the Atlantic is about a man afraid of a ghost named Malcolm. Its vocabulary is relatively simple, and its sentences are relatively short. Its Flesch reading ease score is 82.0 (easy). The Saturday Evening Post story is about a successful play. It uses flowery
language and long sentences. Its reading ease score is 36.7 (difficult). The text of each story excerpt is given in Appendix B. These seemed good examples to test with a more sophisticated analysis. We wanted to determine whether a more thorough analysis of the prose would confirm the Flesch readability estimate or would tell us something different.

A program developed by Sheryl Young attempts to simulate the reading process, or at least some important aspects of it. The first step in this analysis is to "propositionalize" the text. The sentences are broken down into their simplest discrete propositions. Thus the sentence "He heard, as he waked next morning, that it was heavily raining" becomes six "micropropositions": he heard rain; he waked; he heard it when he waked; the time was morning; it was raining heavily (see Appendix C).

If an inference is required to relate one microproposition to the previous text, the human researcher provides the inference as a new, special microproposition. The computer is programmed to mimic the reading process in the following way. It reads each microproposition, storing it in short-term memory, until short-term memory is full, then it moves earlier micropropositions to long-term memory. It relates each microproposition to the preceding text by linking up identical words, by relating nouns to their modifiers, and calling up other relationships it is programmed to recognize. If it cannot relate a microproposition to the material in short-term memory, it searches long-term memory, which takes more time (that is, it slows down comprehension). This is called a "reinstatement." Inferences are also noted, for they also slow
comprehension in human readers. The product of this analysis is two-fold: a tree diagram showing the relationships between all of the micropropositions in the text, and an analysis of the macrostructure, the main ideas, and which key micropropositions belong to each macrostructure.

The difficulty of comprehending a given text is assessed by determining how many reinstatements are necessary, how many inferences are necessary, and how deep the tree structure of micropropositions is. One can also measure the density of micropropositions and the number of macropropositions for a text of a certain length. All of these would affect speed of comprehension. The emphasis, however, is on inferences, reinstatements, and depth of subtrees: Does the reader have to do a lot of searching around and figuring things out? Texts that are difficult to comprehend may be just poorly organized and incoherent, or they may be subtle, aesthetically challenging, or composed of very complex ideas. In the case of our two stories, the Young analysis clearly reverses the comprehensibility predictions of the Flesch readability test. The Atlantic Monthly article, about the ghost, proves to be more difficult, partly because the ideas are more difficult and partly, perhaps, because it is not written as clearly as it might be. The Saturday Evening Post article, on the other hand, despite its fancy words and longer sentences, is very simple to comprehend.

In terms of the Young-Miller program, the story about the play had 5 macropropositions, no reinstatements, no inferences, and 11 levels of microstructure. It had 89 micropropositions in its 211 words (or 42.2 per 100). The ghost story had 6 macropropositions, required one
reinstatement and two inferences, and had 12 levels of microstructure. It had 101 micropropositions in its 227 words (or 44.5 per 100 words).

One of the necessary inferences (that Malcolm was a ghost) proved to be particularly difficult for human readers. In an informal experiment, we gave the two stories to four office workers and asked them to write down the main ideas of each of the two passages. None had any difficulty with the passage about the play; all had some difficulty writing a good summary of the ghost story.

Figures 1 and 2 present the tree structures for the micropropositions of the two stories. The dotted lines outline the macropropositions for each story. The macropropositions for the story about the play, from the Saturday Evening Post, are easy to draw; the key proposition is related in a hierarchical way to the other main micropropositions. For Atlantic Monthly's ghost story, it is almost impossible to draw circles around the macropropositions: some of the key micropropositions are on other subtrees altogether. This is a graphic representation of the messiness the human reader has to sort out in comprehending the story about Malcolm, the ghost who doesn't show up on rainy days.

We do not claim, of course, that all seemingly simple prose in the Atlantic Monthly was incredibly difficult to read, or that everything in the Saturday Evening Post was at a homogeneously simple level in terms of its logical structure, its rhetoric, and the complexity of its ideas. What this exercise shows is that we can learn something more about the difficulty of text from the techniques developed recently by cognitive psychologists than we can from old-fashioned readability formulas. It
"THE PLAY": Tree structure of a story excerpted from the Saturday Evening Post, April 3, 1920 (Young-Miller analysis)
"THE GHOST": Tree structure of a story excerpted from the Atlantic Monthly, April 1920 [Young-Miller analysis]
just happened in this case that the analysis of macrostructure confirmed the common-sense prediction that one usually finds more sophisticated (or at least more difficult) prose in the *Atlantic* than in the *Post*. We reserve for our conclusion some discussion of why historians might be interested in the difficulty ratings of old prose.

Using the Young-Miller program proved to be as expensive and time-consuming as we had expected. It required the expert assistance of Sheryl Young to propositionalize the text, run the program, and advise us on its interpretation. To do this for a large sample of texts would be impossible. The program is complex; it is a research device, not a pragmatic replacement for a readability formula. As we mentioned above, Susan Kemper's work on inference load appeared to be yet another way to explore the macrostructure of these stories.

As noted earlier, Kemper's model is predicated on the notion that texts consist of chains of causally connected events, actions, and states. According to Kemper, the ease or difficulty of the text is dependent on the degree to which the events, actions, and states are explicitly connected. The more inferences a reader is required to make, the harder the selection will be to read.

For us to use her model we had to "decompose" our two stories into event chains. We began by parsing the text into clauses, classifying each clause as an action (A), a physical state (PS), or a mental state (MS). The list of clauses and their classifications are listed in Appendix D.
Thus classified, the clauses were ready to be organized into an event chain. In creating the event chain we had to follow Kemper's rules concerning causal connections. According to Kemper:

1. ACTIONS -- result in -- PHYSICAL STATES
2. ACTIONS -- initiate -- MENTAL STATES
3. PHYSICAL STATES -- initiate -- MENTAL STATES
4. PHYSICAL STATES -- (dis)enable -- ACTIONS
5. MENTAL STATES -- provide reasons for -- ACTIONS

These are the only types of causation. Kemper argues that "an action cannot cause a new action without an intervening physical or mental state, a physical state cannot lead to a new physical state without an intervening action, and mental states cannot cause new mental or physical states without intervening actions" (Kemper, 1986). In other words, a causally connected event chain can not have any action-action, physical state-physical state, mental state-mental state, and mental state-physical state sequences. All such sequences require inferred actions or states to fill the gaps in causal connection.

Following these guidelines, we created an event chain for each story, with Ms. Kemper's advice. The resultant event chains are displayed in Figures 3 and 4, with the inferences circled. The chains completed, we were ready to apply Kemper's "inference load formula." Kemper derived her formula by, first, constructing event chains from 62 passages in the 1979 McCall-Crabbs readings. She then performed a regression analysis to obtain an inference load formula for assessing textual difficulty. The resultant formula is:

\[
\text{Inference Load score} = 8.41 + .54(\text{Stated Mental States}) - .86(\text{Stated Physical States}) +
\]
Figure 3

"THE PLAY": Event chain of a story excerpted from the Saturday Evening Post, April 3, 1920 [Kemper analysis]

PS14
They were flaming in color

PS15
They were permeated with mystery

PS16
They were splendid in size

PS4
It was Bartlett's first venture

MS5
It is known as a legitimate field

PS7
He was preeminent as a producer

A6
He staged it

A9
It told of treachery

A10
It had a spectacular background

A11
It depicted India

PS1
It was a sensation

PS18
Cast was distinguished

A19
They gave versimilitude

PS20
Company included girls, animals etc.

A
This increased the excitement

PS13
It presented a series of settings

A17
Audience greeted revelations with gasps

Audience greeted revelations with gasps
FIGURE 4

"THE GHOST": Event chain of a story, excerpted from the Atlantic Monthly April, 1920 [Kemper analysis]
.23 (Inferred Mental States). The higher the score, the harder the passage is to read, with the score correlating $r = .63$ with the reported grade levels of the McCall-Crabbs passages.

Applying Kemper's formula to our two stories, we obtained inference load scores of 2.30 for the *Post* passage and 10.48 for the *Atlantic* passage. In other words, according to the Kemper model the *Post* prose was written at a second-grade level, while the *Atlantic* prose was fit only for someone with at least a tenth-grade education. In keeping with the Young formula, these results contradict the Flesch Reading Ease scores and correlate with the notion that highbrow fiction is more difficult to read than middlebrow fiction.

Conclusion

Our foray into the world of readability theory, on behalf of historical studies, suggests two points, one methodological and one substantive.

The methodological point is this: there are ways to explore the complexity and thus the probable readability of published reading material from the past. Moreover, they are ways that relate to the mental skills required of readers, not just the mechanical features of word length and sentence length. Certainly these more complex formulas are not perfect instruments. For example, neither the Kemper nor the Young-Kintsch procedure accounts for syntactical order, which clearly affects speed of comprehension; moreover, neither can account for the prior knowledge of the reader, which is the empirical problem all reading researchers are tangling with at present. But while the current
formulas are flawed and destined to be superseded, the basic point seems rather uncontroversial: for those who wish to take the time, there are now ways to analyze the readability features of a text that are more theoretically relevant and more interesting than traditional readability formulas.

The substantive point is more speculative. Our detailed readability evaluation reversed our anomalous ratings on the Flesch scale. The *Saturday Evening Post* fiction turned out to be easier to read than the story in the *Atlantic*. We make no claims that this experiment is more than suggestive, and, indeed, we shrink from what it suggests: that there were some advanced readers and other readers limited to more popular publications, because the prose read by the advanced readers differed in the difficulty of the mental processes required—the number of inferences, the density of propositions, and the demands on longer-term memory. It would be an immensely costly and time-consuming task to test this bifurcation hypothesis thoroughly. But the notion that the audience for *Atlantic* and the *New York Times* was limited in 1920 by their difficulty levels does seem quite plausible, particularly in light of recent reading achievement assessments, which have shown that few Americans read well at a level requiring critical and interpretive skills (Kirsch & Jungeblut, 1986; NAEP, 1985).

All of this smacks of the old high culture versus popular culture debate of the 1950s. However, our reaction to such a split is quite different from the high culture priests who feared its erosion by popular forms in the 1950s. On the contrary, we believe that high literacy has shown great staying power. We are more impressed with its
complexity, its influence, and its apparent inaccessibility to the average reader. In saying this we are not making a value judgment regarding average readers. There is no question that many readers of the popular press succeeded in their jobs and led virtuous, intelligent lives. But the fact remains that our educational system equips only a limited number of people with sophisticated reading skills. Perhaps it does not matter that few people can read *Atlantic*. But sophisticated reading skills also act as an entry qualification and a necessary set of skills in many commanding institutions of our society. In this regard educational stratification contributes to the stratification of American society and is reflected in the varying complexity of publications read by adults.
REFERENCES


APPENDIX A

THE APPLICATION OF FLESCH FORMULA TO SELECTED 1920 SELECTED TEXTS

Reading Ease = 206.835 - .846 (syllables/100 words) - 1.015(words/sentence)

If total is 0 to 30, style is very difficult.
If total is 30 to 50, style is difficult.
If total is 50 to 60, style is fairly difficult.
If total is 60 to 70, style is standard.
If total is 70 to 80, style is fairly easy.
If total is 80 to 90, style is easy.
If total is 90 to 100, style is very easy.

A. Newspapers
1. "Democrats Fought Hard to Hold up the Final Vote"
   THE WORLD, April 1, p. 1
   a. "Under the domination" 158 syllables to "and Orr, against" 54.5 words per sentence RE = 17.849
   b. "With virtually" 164 syllables to "he is obliged" 36.0 words per sentence RE = 31.551
   c. "This dispute" 179 syllables to "after the session" 35.0 words per sentence RE = 19.876
      *AVE. RE = 23.092

2. "Democrat Filibusters Delay Vote"
   NEW YORK TIMES, April 1, p. 1
   a. "At 3 o'clock" 157 syllables to "stick it out" 75.75 words per sentence RE = 47.877
   b. "The discussion of" 154 syllables to "speech from the" 39.5 words per sentence RE = 36.458
   c. "Other features" 167 syllables to "abandoned it. The" 33.0 words per sentence RE = 32.058
      *AVE. RE = 38.798

3. "Disloyalty Charge Causes 5 Men to Lose Seats in N.Y."
   MILWAUKEE JOURNAL, Thursday, April 1, 1920, p. 1
   a. "Five Socialists" 168 syllables to "and Orr" 29.6 words per sentence RE = 34.663
      (numbers excluded, in keeping with Flesch's instructions)
4. "Socialists Are Expelled from Assembly"
BEAVER DAM DAILY CITIZEN, Thursday, April 1, 1920, p. 1.
   a. "The five suspended"
      to
      "nothing in return"
      RE = 61.882
      (numbers excluded, in keeping with Flesch's instructions)
   b. "After the final"
      to
      "an extraordinary
      session" (95 words)
      RE = 34.122
      *AVE. RE = 49.502

B. Periodicals: non-fiction
1. "Profiteering and Prices"
   Melvin Copeland, ATLANTIC MONTHLY, 125 (April), 524-527.
   a. "Where are the"
      p. 524 to
      "shoe retailers"
      RE = 58.958
   b. "Such a general"
      p. 525 to
      "factors, such as"
      RE = 44.441
   c. "The amount of credit"
      p. 526 to
      "longer, and the"
      RE = 52.461
      *AVE. RE = 51.953

2. "Branding the Profiteer"
   a. "You are a bold"
      p. 18 to
      "the profiteer"
      RE = 70.124
   b. "The writer of"
      p. 18 to
      "workers are making"
      RE = 60.821
   c. "There is no"
      to
      "because of the"
      RE = 63.188
      *Ave. RE = 64.711
C. Periodicals: fiction (1 dialogue)

1. "The Third Window"
   ATLANTIC MONTHLY, 125(April), pp. 496-513.
   a. "He heard"
      p. 496 to "he went"
      130 syllables
      8.818 words per sentence
      RE = 87.905
   b. "She did not"
      p. 497 to "this bad weather"
      137 syllables
      25 words per sentence
      RE = 65.558
   c. "Well, it all"
      p. 498 to "She was silent"
      128 syllables
      25 words per sentence
      RE = 87.72
      *AVE. RE = 80.394

2. "A Prince There Wasn't"
   SATURDAY EVENING POST, 192(April), pp. 16-17,124,127,130
   a. "They're always"
      p. 16 to "three sheets"
      143 syllables
      19.25 words per sentence
      RE = 66.318
      (dialogue)
   b. "Written by"
      p. 16 to "somewhat bombastic"
      174 syllables
      29.333 words per sentence
      RE = 29.958
   c. "J. Herbert"
      p. 17 to "dressing rooms"
      160 syllables
      30.25 words per sentence
      RE = 40.771
      *AVE. RE = 45.667

3. "Anything Once"
   a. "The white dust"
      p. 449 to "ditch and the"
      148 syllables
      39 words per sentence
      RE = 42.042
   b. "When Lou awakened"
      p. 459 to "rocks where"
      134 syllables
      35 words per sentence
      RE = 57.946
   c. "It was the first"
      p. 469 to "not a mite"
      124 syllables
      20 words per sentence
      RE = 81.631
      (dialogue)
      *AVE. RE = 60.54
### D. Best-sellers: non-fiction

1. Philip Gibbs, *NOW IT CAN BE TOLD*
   - a. pg. 22 RE=68.773
   - b. pg. 44 RE=51.62
   - c. pg. 66 RE=40.246
   - d. pg. 88 RE=59.54
   - e. pg. 110 RE=76.252
   - f. pg. 132 RE=76.251
   - g. pg. 154 RE=64.206
   - h. pg. 176 RE=91.817
   - i. pg. 198 RE=52.696
   - j. pg. 220 RE=74.897
   - k. pg. 242 RE=65.951
   - l. pg. 264 RE=60.109
   - m. pg. 287 RE=16.836
   - n. pg. 308 RE=39.675
   - o. pg. 330 RE=68.773
   - p. pg. 352 RE=74.018
   - q. pg. 374 RE=93.815
   - r. pg. 396 RE=79.653
   - s. pg. 418 RE=65.951
   - t. pg. 440 RE=52.696
   - u. pg. 462 RE=62.681
   - v. pg. 484 RE=74.016
   - w. pg. 506 RE=66.738
   - x. pg. 528 RE=54.899
   - y. pg. 550 RE=62.445
   - z. pg. 582 RE=66.738
   - AVE. RE=64.573

2. J. M. Keynes, *THE ECONOMIC CONSEQUENCES OF THE PEACE*
   - a. pg. 12 RE=26.472
   - b. pg. 24 RE=32.311
   - c. pg. 36 RE=54.013
   - d. pg. 48 RE=49.485
   - e. pg. 60 RE=62.036
   - f. pg. 72 RE=20.046
   - g. pg. 84 RE=80.555
   - h. pg. 96 RE=40.183
   - i. pg. 108 RE=29.058
   - j. pg. 120 RE=43.562
   - k. pg. 132 RE=34.087
   - l. pg. 144 RE=64.007
   - m. pg. 156 RE=19.032
   - n. pg. 168 RE=36.711
   - o. pg. 180 RE=26.051
   - p. pg. 192 RE=32.902
   - q. pg. 204 RE=43.055
   - r. pg. 216 RE=19.033
   - s. pg. 228 RE=46.439
   - t. pg. 240 RE=32.062
   - u. pg. 252 RE=57.47
   - v. pg. 264 RE=50.331
   - w. pg. 276 RE=55.411
   - x. pg. 288 RE=50.668
   - y. pg. 297 RE=42.293
   - z. pg. 309 RE=42.293
   - AVE. RE=39.59

### E. Best-sellers: fiction

   - a. pg. 18 RE=96.293
   - b. pg. 36 RE=53.237
   - c. pg. 54 RE=89.243
   - d. pg. 72 RE=89.582
   - e. pg. 90 RE=86.197
   - f. pg. 108 RE=57.402
   - g. pg. 126 RE=59.094
   - h. pg. 144 RE=61.076
   - i. pg. 162 RE=88.287
   - j. pg. 180 RE=85.209
   - k. pg. 198 RE=66.741
   - l. pg. 216 RE=86.413
   - m. pg. 234 RE=56.082
   - n. pg. 252 RE=85.741
   - o. pg. 270 RE=70.758
   - p. pg. 288 RE=81.439
   - q. pg. 306 RE=79.769
   - r. pg. 324 RE=68.559
   - s. pg. 342 RE=93.377
   - t. pg. 360 RE=86.709
   - u. pg. 38 RE=90.726
   - v. pg. 396 RE=82.866
   - w. pg. 414 RE=7.091
   - x. pg. 432 RE=64.543
   - y. pg. 450 RE=68.474
   - z. pg. 462 RE=68.474
   - AVE. RE=74.224

<table>
<thead>
<tr>
<th>Vol.</th>
<th>Page</th>
<th>RE</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>19</td>
<td>63.119</td>
</tr>
<tr>
<td>b</td>
<td>38</td>
<td>61.327</td>
</tr>
<tr>
<td>c</td>
<td>57</td>
<td>54.017</td>
</tr>
<tr>
<td>d</td>
<td>76</td>
<td>55.984</td>
</tr>
<tr>
<td>e</td>
<td>95</td>
<td>76.537</td>
</tr>
<tr>
<td>f</td>
<td>114</td>
<td>68.142</td>
</tr>
<tr>
<td>g</td>
<td>133</td>
<td>57.131</td>
</tr>
<tr>
<td>h</td>
<td>156</td>
<td>71.097</td>
</tr>
<tr>
<td>i</td>
<td>171</td>
<td>58.452</td>
</tr>
<tr>
<td>j</td>
<td>190</td>
<td>71.353</td>
</tr>
<tr>
<td>k</td>
<td>209</td>
<td>72.325</td>
</tr>
<tr>
<td>l</td>
<td>228</td>
<td>61.269</td>
</tr>
<tr>
<td>m</td>
<td>247</td>
<td>77.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Vol.</th>
<th>Page</th>
<th>RE</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>266</td>
<td>16.169</td>
</tr>
<tr>
<td>o</td>
<td>285</td>
<td>78.64</td>
</tr>
<tr>
<td>p</td>
<td>304</td>
<td>73.847</td>
</tr>
<tr>
<td>q</td>
<td>323</td>
<td>71.731</td>
</tr>
<tr>
<td>r</td>
<td>342</td>
<td>77.908</td>
</tr>
<tr>
<td>s</td>
<td>361</td>
<td>75.829</td>
</tr>
<tr>
<td>t</td>
<td>380</td>
<td>75.878</td>
</tr>
<tr>
<td>u</td>
<td>399</td>
<td>75.438</td>
</tr>
<tr>
<td>v</td>
<td>418</td>
<td>75.878</td>
</tr>
<tr>
<td>w</td>
<td>437</td>
<td>62.41</td>
</tr>
<tr>
<td>x</td>
<td>456</td>
<td>77.929</td>
</tr>
<tr>
<td>y</td>
<td>475</td>
<td>75.083</td>
</tr>
<tr>
<td>z</td>
<td>494</td>
<td>64.327</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ave.</th>
<th>RE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>64.327</td>
</tr>
</tbody>
</table>

3. **Zane Grey, MAN OF THE FOREST**

<table>
<thead>
<tr>
<th>Vol.</th>
<th>Page</th>
<th>RE</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>15</td>
<td>76.652</td>
</tr>
<tr>
<td>b</td>
<td>30</td>
<td>84.026</td>
</tr>
<tr>
<td>c</td>
<td>45</td>
<td>80.21</td>
</tr>
<tr>
<td>d</td>
<td>60</td>
<td>79.364</td>
</tr>
<tr>
<td>e</td>
<td>75</td>
<td>62.169</td>
</tr>
<tr>
<td>f</td>
<td>90</td>
<td>79.091</td>
</tr>
<tr>
<td>g</td>
<td>105</td>
<td>87.758</td>
</tr>
<tr>
<td>h</td>
<td>120</td>
<td>93.117</td>
</tr>
<tr>
<td>i</td>
<td>135</td>
<td>76.554</td>
</tr>
<tr>
<td>j</td>
<td>150</td>
<td>81.464</td>
</tr>
<tr>
<td>k</td>
<td>165</td>
<td>71.353</td>
</tr>
<tr>
<td>l</td>
<td>180</td>
<td>84.107</td>
</tr>
<tr>
<td>m</td>
<td>195</td>
<td>84.738</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Vol.</th>
<th>Page</th>
<th>RE</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>210</td>
<td>64.373</td>
</tr>
<tr>
<td>o</td>
<td>225</td>
<td>91.784</td>
</tr>
<tr>
<td>p</td>
<td>240</td>
<td>85.963</td>
</tr>
<tr>
<td>q</td>
<td>255</td>
<td>89.666</td>
</tr>
<tr>
<td>r</td>
<td>270</td>
<td>95.097</td>
</tr>
<tr>
<td>s</td>
<td>285</td>
<td>82.035</td>
</tr>
<tr>
<td>t</td>
<td>300</td>
<td>92.627</td>
</tr>
<tr>
<td>u</td>
<td>315</td>
<td>92.522</td>
</tr>
<tr>
<td>v</td>
<td>330</td>
<td>94.444</td>
</tr>
<tr>
<td>w</td>
<td>345</td>
<td>76.358</td>
</tr>
<tr>
<td>x</td>
<td>360</td>
<td>98.844</td>
</tr>
<tr>
<td>y</td>
<td>375</td>
<td>76.083</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ave.</th>
<th>RE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>83.14</td>
</tr>
</tbody>
</table>

4. **H. B. Wright, THE RE-CREATION OF BRIAN KENT**

<table>
<thead>
<tr>
<th>Vol.</th>
<th>Page</th>
<th>RE</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>15</td>
<td>64.035</td>
</tr>
<tr>
<td>b</td>
<td>28</td>
<td>49.147</td>
</tr>
<tr>
<td>c</td>
<td>41</td>
<td>86.24</td>
</tr>
<tr>
<td>d</td>
<td>54</td>
<td>76.723</td>
</tr>
<tr>
<td>e</td>
<td>67</td>
<td>63.612</td>
</tr>
<tr>
<td>f</td>
<td>80</td>
<td>65.558</td>
</tr>
<tr>
<td>g</td>
<td>93</td>
<td>73.339</td>
</tr>
<tr>
<td>h</td>
<td>100</td>
<td>77.69</td>
</tr>
<tr>
<td>i</td>
<td>119</td>
<td>58.123</td>
</tr>
<tr>
<td>j</td>
<td>132</td>
<td>85.049</td>
</tr>
<tr>
<td>k</td>
<td>145</td>
<td>58.96</td>
</tr>
<tr>
<td>l</td>
<td>158</td>
<td>73.341</td>
</tr>
<tr>
<td>m</td>
<td>171</td>
<td>74.442</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Vol.</th>
<th>Page</th>
<th>RE</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>184</td>
<td>79.261</td>
</tr>
<tr>
<td>o</td>
<td>197</td>
<td>34.088</td>
</tr>
<tr>
<td>p</td>
<td>211</td>
<td>72.664</td>
</tr>
<tr>
<td>q</td>
<td>223</td>
<td>83.575</td>
</tr>
<tr>
<td>r</td>
<td>236</td>
<td>30.538</td>
</tr>
<tr>
<td>s</td>
<td>249</td>
<td>42.214</td>
</tr>
<tr>
<td>t</td>
<td>262</td>
<td>52.531</td>
</tr>
<tr>
<td>u</td>
<td>275</td>
<td>67.317</td>
</tr>
<tr>
<td>v</td>
<td>288</td>
<td>45.338</td>
</tr>
<tr>
<td>w</td>
<td>301</td>
<td>47.963</td>
</tr>
<tr>
<td>x</td>
<td>304</td>
<td>84.627</td>
</tr>
<tr>
<td>y</td>
<td>327</td>
<td>56.422</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ave.</th>
<th>RE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>64.112</td>
</tr>
</tbody>
</table>
APPENDIX B

THE STORIES


The Ganges Princess was the dramatic sensation of a decade. It had been running for a solid year at the huge Hendrik Hudson Theater in New York, having weathered a hot summer with hardly a noticeable failing off of receipts. It was Chester Bartlett's first venture into what is technically known as the legitimate field, and he had staged it with that lavish disregard for expense and with that keen sense of the artistic which had given him preeminence as a producer of light musical entertainment.

Written by one of America's most flamboyant playwrights, it told a turgid story of Oriental passion and treachery set against a spectacular background depicting scenes in ancient India. As sheer spectacle it quite transcended anything hitherto attempted in the United States. It presented a series of settings which were so flaming in their color, so permeated with the mystery of the East and so splendid in their suggestion of great size and vast distances that each new revelation was invariably greeted with gasps of amazement from the audiences. A cast bristling with distinguished names gave verisimilitude to the somewhat bombastic dialogue, and purely incidental members of the company included a troupe of fifty real nautch girls, six elephants, five camels and a flock of sheep.

He heard, as he waked next morning, that it was heavily raining. When he looked out, the trees stood still in gray sheets of straightly falling rain. There was no wind.

The mournful, obliterated scene did not oppress him. The weather was all to the good, he thought. He had always liked a rainy day in the country; and ghosts don't walk in the rain. If Malcolm hadn't come in the moonlight, he wouldn't come now. He felt sunken, exhausted, and rather sick; yet his spirits were not bad. He was fit for the encounter with Antonia.

When he went down to the dark dining-room, darker than ever today, he found only one place laid. The maid told him that both the ladies were breakfasting in their rooms. This was unexpected and disconcerting. But he made the best of it, and drank his coffee and ate kedgeree and toast with not too bad an appetite. A little coal-fire had been lighted in the library, and he went in there after breakfast and read the papers and wrote some letters, and the morning passed not too heavily.

But at luncheon-time his heart sank, almost to the qualm of the night before, when he found still only one place laid. After half an hour of indecision over his cigarette, he wrote a note and sent it up to Antonia.
APPENDIX C

PROPOSITIONAL STRUCTURE OF THE TWO STORIES

(Young-Miller Analysis)

1. Micropropositions for "The Play" excerpt.
   P1  Is a Ganges-Princess sensation
   P2  Type-of sensation drama
   P3  Of sensation decade
   *SENTENCE*
   P4  Running Ganges-Princess year
   P5  Mod year solid
   P6  Where P4 theater
   P7  Mod theater huge-Hendrik-Hudson
   P8  Location theater New-York
   *CLAUSE*
   P9  Weathered Ganges-Princess summer
   P10 Mod summer hot
   P11 How P9 P13
   P12 Falling-off receipts
   P13 Negate P12 hardly
   *SENTENCE*
   P14 Was Ganges-Princess venture
   P15 For venture Chester-Bartlett
   P16 Mod P15 first
   P17 Into venture field
   P18 Type-of field legitimate
   P19 Mod P18 known technically
   *CLAUSE*
   P20 Staged Chester-Bartlett Ganges-Princess P21 P23
   P21 With disregard expense
   P22 Mod disregard lavish
   P23 With sense
   P24 Mod sense keen
   P25 Type-of sense artistic
   P26 Given P25; Chester-Bartlett preeminence
   P27 As preeminence producer
   P28 Type-of producer entertainment
   P29 Mod entertainment musical
   P30 Mod musical light
   *SENTENCE*
   P31 Written Ganges-Princess playwright
   P32 Mod playwright flamboyant
   P33 Type-of playwright American
   P34 Told Ganges-Princess story
   P35 Type-of story turgid
   P36 Of story passion
   P37 Of story treachery
Type-of passion oriental
Set Story P40
Against background
Mod background spectacular
Depict background scenes
Location scenes India
Mod India ancient

Was a story spectacular
Mod spectacular sheer
Transcended story anything
Mod anything hitherto attempted
Where attempted America

Presented story settings
Number--of settings series
Were settings flaming
In flaming color
Were settings P55
With permeated mystery
Of mystery East
Were settings splendid
In P57 suggestion
Of suggestion size
Of suggestion distance
Mod size great
Mod distance vast
In scenes revelations
Mod revelations new
Greeted revelations gasps
Of gasps amazement
From gasps audience

Bristling cast names
Of cast Ganges-princess
Type-of names distinguished
Gave names verisimilitude
To verisimilitude dialogue
Mod dialogue bombastic
Qualify bombastic somewhat

Has cast members
Of members company
Mod members incidental
Mod incidental purely
Included P77 elephants
Included P77 camels
Number-of elephants six
Number-of camels five
Included P77 sheep
2. Macropropositions for "Play" excerpt

M1 = (Integrate) P1, P2, P3
   /M1 idea= The Ganges-Princess was a sensation

M2 = (Construct from) P4, P8, P6
   /M2 idea= The Ganges-Princess ran a year in a New York theater

M3 = (Integrate) P15, P14, P16
   /M3 idea= The Ganges-Princess was Chester-Bartlett's first venture

M4 = (Integrate) P43, P44, P40
   /M4 idea= The play is set in India

M5 = (Generalize) P68, P70, P69
   /M5 idea= The cast was good

3. Micropropositions for "Ghost" excerpt

P1 Heard he rain
P2 When P1 P3
P3 Waked he
P4 Time-of P3 morning
P5 Mod morning next
P6 Mod rain heavily
*SENTENCE*
P7 When P8 P9
P8 Looked he out
P9 Stood trees still
P10 In P9 rain
P11 Mod rain falling
P12 Mod falling straightly
P13 Type-of rain sheets
P14 Mod sheets gray
*SENTENCE*
P14.1 INFER: When P8 P16
P15 Exist wind
P16 Negate P15
*SENTENCE*
P17 Oppress scene he
P18 Not P17
P19 Mod scene mournful
P20 Mod scene obliterated
*SENTENCE*

P21 Thought he P22
P22 Was weather good
P23 Mod good all-to-the

*SENTENCE*

P24 Like he day
P25 Mod day rain
P26 Mod P24 always
P27 In P25 country
P28 Walk ghost; rain
P29 Not P28
P30 Liked he P29

*SENTENCE*

P31 If P33 P35
P32 Come Malcolm moonlight
P33 Not P32
P34 Come Malcolm now
P35 Not P34
P36 INFER: Malcolm ghost

*SENTENCE*

P37 Felt he sunken
P38 Felt he exhausted
P39 Felt he sick
P40 Mod sick rather
P41 Not P42
P42 Were spirits bad

*SENTENCE*

P43 Belong spirits he

*SENTENCE*

P44 Was he fit
P45 For fit encounter
P46 With encounter Antonia

*SENTENCE*

P47 Where P49 P54
P48 Went he down
P49 Went he diningroom
P50 Mod diningroom dark
P51 Was diningroom darker
P52 Mod darker than-ever
P53 When P52 today
P54 Found he place
P55 Was laid place
P56 Number-of place one

*SENTENCE*

P57 Told maid he P58
P58 Were ladies breakfasting
P59 Where P58 rooms
P60 Belong rooms ladies

*SENTENCE*

P61 Was P57 disconcerting
P62 Was P57 unexpected
P63 Relation P61 P62 P64  
P64 Made he best-of-it  
P65 How P64 P66 P67 P68  
P66 Drank he coffee  
P67 Ate he kedgree  
P68 Ate he toast  
P69 With P67 P68 appetite  
P70 Mod appetite bad  
P71 Negate P70  
*SENTENCE*  
P72 Lighted fire  
P73 Type-of fire coal  
P74 Size-of fire little  
P75 Location fire library  
P76 Went he library  
P77 When P76 after breakfast  
P78 Read he papers  
P79 Writes he letters  
P80 Mod letters some  
P81 While P76 P82  
P82 Passed morning  
P83 How P82 heavily  
P84 Not P83  
*SENTENCE*  
P85 Sank heart he  
P86 When P85 lunch-time  
P87 Mod P85 P88  
P88 To qualm of night  
P89 Mod P88 almost  
P90 Mod night before  
P91 Why P85 P92  
P92 Found he place  
P93 Number-of place one  
P94 Mod P93 still  
P95 Mod P93 only  
*SENTENCE*  
P96 After indecision P100 P101  
P97 When indecision P98  
P98 Over cigarette he  
P99 Amount-of indecision half-hour  
P100 Wrote he note  
P101 Sent he note Antonia  

4. Macropropositions for "Ghost" excerpt  
M1= (Integrate) P2, P1, P3, P4  
/ M1 idea= When he woke up the next morning, he heard the heavy rain 7  
M2= (Construct) P29, P28, P30, P24, P25  
/ M2 idea= Ghosts don't walk in the rain, so he liked the rainy day 7
M3 = (Generalize) P49, P48, P54, P55, P56
    /M3 idea= He found one place laid in the diningroom/

M4 = (Integrate) P78, P82, P77
    /M4 idea= After breakfast he read papers/

M5 = (Integrate) P86, P85, P92, P93
    /M5 idea= His heart sank when he found only one place/

M6 = (Generalize) P100, P101
    /M6 idea= He wrote a note to Antonia/
APPENDIX P

CLAUSES AND THEIR CATEGORIES (KEMPER ANALYSIS)

Abbreviations:

- A = action
- MS = mental state
- PS = physical state

1. "The Play" (Saturday Evening Post)

<table>
<thead>
<tr>
<th>Number</th>
<th>Type</th>
<th>Clause</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PS</td>
<td>The Ganges Princess was the dramatic sensation of a decade</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>It had been running for a solid year at the huge Hendrik Hudson Theatre in New York</td>
</tr>
<tr>
<td>3</td>
<td>A</td>
<td>(It had) weathered a hot summer with hardly a noticeable falling off of receipts</td>
</tr>
<tr>
<td>4</td>
<td>PS</td>
<td>It was Chester Bartlett's first venture into (X)</td>
</tr>
<tr>
<td>5</td>
<td>MS</td>
<td>(What) is technically known as the legitimate field</td>
</tr>
<tr>
<td>6</td>
<td>A</td>
<td>He had staged it with that lavish disregard for expense and with that keen sense of the artistic</td>
</tr>
<tr>
<td>7</td>
<td>PS</td>
<td>Which had given him preeminence as a producer of light musical entertainment</td>
</tr>
<tr>
<td>8</td>
<td>A</td>
<td>(It was) written by one of America's most flamboyant playwrights</td>
</tr>
<tr>
<td>9</td>
<td>A</td>
<td>It told a turgid story of Oriental passion and treachery</td>
</tr>
<tr>
<td>10</td>
<td>A</td>
<td>(It was) set against a spectacular background</td>
</tr>
<tr>
<td>11</td>
<td>A</td>
<td>(It) depicted scenes in ancient India</td>
</tr>
<tr>
<td>12</td>
<td>A</td>
<td>As sheer spectacle it quite transcended anything</td>
</tr>
<tr>
<td>12+</td>
<td>A</td>
<td>hitherto attempted in the United States</td>
</tr>
<tr>
<td>13</td>
<td>A</td>
<td>It presented a series of settings</td>
</tr>
<tr>
<td>14</td>
<td>PS</td>
<td>(They) were so flaming in their color</td>
</tr>
<tr>
<td>Number</td>
<td>Type</td>
<td>Clause</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td>--------</td>
</tr>
<tr>
<td>15</td>
<td>PS</td>
<td>(They) were so permeated with the mystery of the East</td>
</tr>
<tr>
<td>16</td>
<td>PS</td>
<td>(They) were so splendid in their suggestion of great size and vast distances</td>
</tr>
<tr>
<td>17</td>
<td>A</td>
<td>Each new revelation was invariably greeted with gasps of amazement from the audience</td>
</tr>
<tr>
<td>18</td>
<td>PS</td>
<td>A cast (bristled) with distinguished names</td>
</tr>
<tr>
<td>19</td>
<td>A</td>
<td>(The cast) gave versimilitude to the somewhat bombastic dialogue</td>
</tr>
<tr>
<td>20</td>
<td>PS</td>
<td>Purely incidental members of the company included a troup of fifty real nautch girls, six elephants, five camels and a flock of sheep</td>
</tr>
</tbody>
</table>
2. "Ghost" *(Atlantic Monthly)*

<table>
<thead>
<tr>
<th>Number</th>
<th>Type</th>
<th>Clause</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MS</td>
<td>He heard X</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>(As) he waked next morning</td>
</tr>
<tr>
<td>3</td>
<td>A</td>
<td>It was heavily raining</td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>(When) he looked out</td>
</tr>
<tr>
<td>5</td>
<td>PS</td>
<td>The trees stood still in gray sheets of straightly falling rain</td>
</tr>
<tr>
<td>6</td>
<td>PS</td>
<td>There was no wind</td>
</tr>
<tr>
<td>7</td>
<td>A</td>
<td>The mournful, obliterated scene did not appress him</td>
</tr>
<tr>
<td>8</td>
<td>PS</td>
<td>The weather was all to the good</td>
</tr>
<tr>
<td>9</td>
<td>MS</td>
<td>He thought (X)</td>
</tr>
<tr>
<td>10</td>
<td>MS</td>
<td>He had always liked a rainy day in the country</td>
</tr>
<tr>
<td>11</td>
<td>A</td>
<td>Ghosts don't walk in the rain</td>
</tr>
<tr>
<td>12</td>
<td>A</td>
<td>If Malcolm hadn't come in the moonlight</td>
</tr>
<tr>
<td>13</td>
<td>A</td>
<td>He wouldn't come now</td>
</tr>
<tr>
<td>14</td>
<td>MS</td>
<td>He felt sunken, exhausted, and rather sick</td>
</tr>
<tr>
<td>15</td>
<td>MS</td>
<td>Yet his spirits were not bad</td>
</tr>
<tr>
<td>16</td>
<td>MS</td>
<td>He was fit for the encounter with Antonia</td>
</tr>
<tr>
<td>17</td>
<td>A</td>
<td>(When) he went down to the dark dining-room</td>
</tr>
<tr>
<td>18</td>
<td>PS</td>
<td>(It was) darker than ever to-day</td>
</tr>
<tr>
<td>19</td>
<td>A</td>
<td>He found (X)</td>
</tr>
<tr>
<td>20</td>
<td>A</td>
<td>Only one place (was) laid</td>
</tr>
<tr>
<td>21</td>
<td>A</td>
<td>The maid told him (X)</td>
</tr>
<tr>
<td>22</td>
<td>A</td>
<td>Both the ladies were breakfasting in their rooms</td>
</tr>
<tr>
<td>23</td>
<td>MS</td>
<td>This was unexpected and disconcerting</td>
</tr>
</tbody>
</table>
But he made the best of it

(He) drank his coffee and ate kedgeree and toast with not too bad an appetite

A little coal-fire had been lighted in the library

He went in there after breakfast and read the papers and wrote some letters

The morning passed not too heavily

But at luncheon-time his heart sank almost to the qualm of the night before

(When) he found (X)

(There was) still only one place laid

After half an hour of indecision over his cigarette, he wrote a note and sent it up to Antonia