Speech Perception and Reading: Two Parallel Modes of Understanding Language and Implications for Acquiring Literacy Naturally

DOMINIC W. MASSARO
University of California, Santa Cruz

I review 2 seminal research reports published in this journal during its second decade more than a century ago. Given psychology's subdisciplines, they would not normally be reviewed together because one involves reading and the other speech perception. The small amount of interaction between these domains might have limited research and theoretical progress. In fact, the 2 early research reports revealed common processes involved in these 2 forms of language processing. Their illustration of the role of Wundt's apperceptive process in reading and speech perception anticipated descriptions of contemporary theories of pattern recognition, such as the fuzzy logical model of perception. Based on the commonalities between reading and listening, one can question why they have been viewed so differently. It is commonly believed that learning to read requires formal instruction and schooling, whereas spoken language is acquired from birth onward through natural interactions with people who talk. Most researchers and educators believe that spoken language is acquired naturally from birth onward and even prenatally. Learning to read, on the other hand, is not possible until the child has acquired spoken language, reaches school age, and receives formal instruction. If an appropriate form of written text is made available early in a child's life, however, the current hypothesis is that reading will also be learned inductively and emerge naturally, with no significant negative consequences. If this proposal is true, it should soon be possible to create an interactive system, Technology Assisted Reading Acquisition, to allow children to acquire literacy naturally.

We are celebrating 125 years of *The American Journal of Psychology* (*AJP*), which is synonymous with the duration of the discipline of psychological inquiry. My charge was to find and discuss a few seminal articles that set the stage for progress in our field. A sobering thought is whether we have actually made progress, and indeed we have. One of the first research paradigms was centered on the introspective method, whereas research since that time has repeatedly revealed how fallible our introspections and memories actually are. Much remains to be accomplished, however. My favorite illustration, one that is relevant to this article, is education. I believe that psychology has yet to inform most of education practice. We know that a medical doctor returning from a century ago would be lost in current medical practice. Not so for a schoolteacher! Admittedly, they would have to learn to use some technology, but

American Journal of Psychology
Fall 2012, Vol. 125, No. 3 pp. 307–320 •

the instruction procedures and classroom dynamics would be very familiar.

When we search back for meaningful events such as influential articles, we find that our extant mindset creates an imposing filter on what we discover. I have a story to tell, and I use a few historical articles from the archives of the *AJP* to embellish this tale. Given the rapid technological, academic, and political changes since the first days of *AJP,* it might be surprising and enlightening to find relevant articles that are still valuable today. I will focus on just two articles to begin my story. For those familiar with my empirical and theoretical work, this use of history will offer another bully pulpit to lobby for language learning and understanding as a process of pattern recognition.

At issue is how we come to perceive and experience meaningful events, in our case via listening and reading. If you asked this question at the beginning of psychological inquiry, you would turn to philosophy. As we know only too well, psychology evolved mainly from philosophy and somewhat from biology. Wilhelm Wundt, who founded the first psychology laboratory, was grounded in the German philosophical tradition in which the concept *apperception* was prominent. Apperception was described by Leibniz and Kant and referred to the mind's experience of its inner states. As in most scholarship, development of one idea usually emerged to contrast with another. In the current scenario, the combatant was the school of British associationism and its concept of the mind being composed of many elementary sensations. Apperception imposed a unity on one's experience and involved a constructive activity of the mind rather than a passive reaction to environmental events. Initiating our empirical science, Wundt proposed experimental investigations of the process of apperception. Well before today's popularity of brain measures, Wundt even speculated on an "apperception centre" in the forebrain that coordinated the activity of lower sensory and motor centers.

A student of Wundt, Edward Bradford Titchener, brought the empirical laboratory to Cornell University, located in the wilds of western New York State. My two targeted articles were dissertations from his laboratory before the turn of the 20th century; the latter authors naturally cited the earlier author's work, and their experimental subjects were existing or future esteemed scientists of mind and behavior.

Walter Bowers Pillsbury devised an early and direct experimental test of Wundt's assumption that perception depends on a preexisting structure of knowledge. After a prolonged discussion of apperception and its relationship to other extant concepts, Pillsbury proposed,

> The purpose of our investigation was to obtain definite and, if possible, quantitative results as regards one feature of the phenomena of association and apperception—to determine the amount of change which might be made in an object ordinarily perceived or assimilated in a certain way without change in the character of the resultant perception or assimilation. (Pillsbury, 1897, p. 339)

It should be noted that Pillsbury did not begin his research with an interest in reading, but his goal was simply to find stimulus objects in which the apperceptive process could be measured. After preliminary experiments with pictures, colors, and geometric objects, he chose typewritten words. A word can be named more reliably than these other possible stimuli (a picture of a robin can be named *robin, bird,* or *animal,* and we often disagree on color names). (This reminds me of a walk with a naturalist when a tweet came from a nearby tree. Adapting the Socratic method, he asked, "What was that?" A preschooler expressed his amazement at the leader's innocence and reprimanded him with, "It's a bird.")

Thus, Pillsbury's study of apperception evolved into a study of word recognition in reading. The apparatus was "a lantern so arranged as to project an image upon a ground glass screen, with a delicate photographic shutter fixed in front of the lens to control the length of exposure" (Pillsbury, 1897, p. 343). Pillsbury made three types of changes in a letter making up the test word: He omitted a letter, substituted another letter for it, or printed an *X* over it to give it a shapeless blur instead of the original letter. These displays were presented very briefly to members of the psychology department (including Titchener and other faculty members), who were then asked to report what they had seen. Not surprisingly, the subjects were able to identify the distorted words correctly, and sometimes they even failed to perceive anything unusual in the display. For example, the presentation of *fashxon* was read as *fashion, foyever* as *forever, disal* as *deal, uver-*

more as *evermore,* and *danxe* as *danger.* According to the prevailing theory, the constructive apperceptive process, using the reader's existing word knowledge, enabled a successful and perhaps an uncompromised perception of the presented words.

Pillsbury's research along with other direct studies of reading (e.g., Cattell, 1886; Erdmann & Dodge, 1898) revealed that the act of reading is only partially limited by what is being read. This generated a great deal of excitement in psychological and educational circles, and for some reason it convinced scientists and educators alike that word recognition did not depend on the recognition of individual letters, a notion that still raises its ugly head periodically in various disguises (Elingsh, 2011). Some educators, impressed by these early results, began to advocate the whole-word method of teaching reading, more marketable today as literature-based reading instruction. Thus began the controversies (dubbed the reading wars by Jeanne Chall, 1967) that have raged ever since around the proper method for teaching children to read. If skilled readers perceive entire words as a whole, it seemed reasonable that the method of teaching children letters, spelling patterns, and spelling-to-sound correspondences (phonics) could only compromise their developing the optimal skill for deriving meaning from the text.

It would be remiss of me not to be at least partially sidetracked by this debate. As will be described later, readers use several sources of information or constraints that contribute to perception and understanding. This statement could be interpreted as a contemporary version of Wundt's apperceptive process, as exemplified by Pillsbury's findings. Of course, readers can recognize both written and spoken words without complete resolution of their components (letters and phonemes). This interpretation does not mean that all the components are not processed; it simply means that word recognition can be achieved before the processing of all the components is finished. It would be a very foolish strategy indeed to instruct the child learning to read to ignore the middle letters just because accomplished readers can read a text with just the first and last letters and scrambled internal letters (Elingsh, 2011; see Massaro & Jesse, 2005, for a comprehensive analysis).

It is important for our story to note that psychologists intensified their study of reading, convinced that it held the key to a great many crucial psychological issues. Every reading scholar is all too familiar with Edmund Huey's seminal book in 1908 and his popular quote,

> To completely analyze what we do when we read would almost be the acme of a psychologist's achievements, for it would be to describe very many of the most intricate workings of the human mind, as well as to unravel the tangled story of the most remarkable specific performance that civilization has learned in all its history. (Huey, 1908/1968, p. 6)

It is important for our story to note that this type of enthusiasm was not generated for spoken language, even though an analogous finding was reported in our second targeted article. William Chandler Bagley (1900) published his dissertation a few years later in this same journal. We can speculate what was in Bagley's mind when he recruited Pillsbury's paradigm from reading for speech. He must have believed in analogous processes in these two different domains. In both cases, language perceivers are requested to make sense of a degraded input. One comes in by eye and one by ear. I can just envision the early discussion between mentor and student: "Professor Titchener, Walter found these results in reading words; shouldn't we also determine whether the same result occurs in listening to words?" Herr Professor evidently did not chide his student for proposing only "incremental" questions. To quote from Bagley's publication,

> Above all else, this (Pillsbury's) work upon visual perception bears overwhelming testimony to the significance which "context" has for the perception of symbols which appeal to the eye. It was the primary object of the present study to determine whether a similar condition obtains in the case of symbols appealing to the ear. (Bagley, 1900, p. 86)

In Bagley's experiment, naturally spoken words were recorded and played back to eight members of the Cornell University Psychology Department on Edison phonograph cylinders. The words were pronounced by deleting one of the consonant sounds in the word. Bagley expanded on Pillsbury's study by evaluating the effects of context beyond that given by the word itself: The word was presented alone, with one or two related words, and at the beginning, middle, and end of complete sentences. This experiment may have been the first systematic distortion of the spoken message, although people doubtless have been playing with the words they utter since they began uttering them (witness Pig Latin, for example). The results demonstrated that subjects were often able to correctly recognize the distorted words, and the word's context influenced performance. Correct word recognition was improved if the word was placed in the middle of a sentence, for example, relative to being presented alone. (See the Appendix for a profile of Bagley and Pillsbury.)

Although this journal was the leading psychological journal of the time, this intuitive result generated little interest and was quickly forgotten. It was not used to advocate how we should speak to our children, and it did not initiate "speech perception" wars of any kind. (This observation is meant to be only partially frivolous because it anticipates my concern with the lack of synergy between reading and speech research.) In contrast to the huge impact of early reading research, speech research fell outside the domain of experimental psychology. Reading warranted its own chapter in Woodworth's *Experimental Psychology* (1938), whereas Bagley's seminal study was not even cited in his handbook, and a 20th-century survey of psychology in America omitted any reference to speech perception (Hilgard, 1987). It also remained somewhat foreign during the "cognitive revolution," at the end of 20th century, although the technical goal of automated speech recognition by machine brought speech perception some attention from experimental psychologists, linguists, and other explorers of the talking mind.

To bridge the apparent century-old gap between reading and speech perception, I now turn to research that supports analogous processes in these two domains. In both experiments, bottom-up (sensory-driven) and top-down (context-driven) sources of information are independently varied in a factorial design. The results illustrate how these two sources of information influence language perception in both reading and speech. To illustrate these analogous processes, I will simply describe one experiment in each of these domains, using the fuzzy logical model of perception (FLMP) as an explanatory framework (Massaro, 1998; Massaro et al., 2011). Several other explanatory frameworks would serve a similar purpose, but the FLMP has been quantified to provide exact predictions of the experimental results of individual subjects. These features are still uncommon today in research and theory, for reasons that elude me.

In the FLMP (Figure 1), language processing is best described as a form of pattern recognition, which involves influences from multiple sources of information. Understanding both spoken and written language is constrained by a variety of auditory, visual, and gestural cues, as well as lexical, semantic, syntactic, and pragmatic constraints. Research in a variety of domains and tasks supports the conclusions that perceivers have continuous rather than categorical information from each of these sources; each source is evaluated with respect to the degree of support for each alternative; each source is treated independently of other sources; the sources are integrated to give an overall degree of support for each alternative; decisions are made with respect to the relative goodness of match among the viable alternatives; evaluation, integration, and decision are necessarily successive but overlapping stages of processing; and crosstalk among the sources of information is minimal, as expected from an optimal Bayesian-type process.

Reading

In this illustrative experiment, readers were asked to read a letter string and to identify one of its letters as *c* or *e* (Massaro, 1979). It is possible to gradually transform a lowercase *e* into a *c* by decreasing the length of the horizontal bar. To the extent the bar is long, there is good visual information for an *e* and poor visual information for a *c*. Now consider the letter presented as the first letter in the context *-oin* and in the context *-dit*. Only *c* is admissible in the first context *-oin* because the sequence *eoin* is not a recognizable word, and the three consecutive vowels *eoi* are in general violation of English orthographic patterning. Only *e* is admissible in the second context *-dit* since *cdit* does not form a word, and the initial cluster *cd* is not an English pattern. For a skilled reader of English, the context *-oin* favors *c,* whereas the context *-dit* favors *e*. The contexts *-tsa* and *-ast* can be considered to favor neither *e* nor *c*. The first remains an inadmissible context whether *e* or *c* is present, and the second is orthographically admissible for both *e* and *c.*

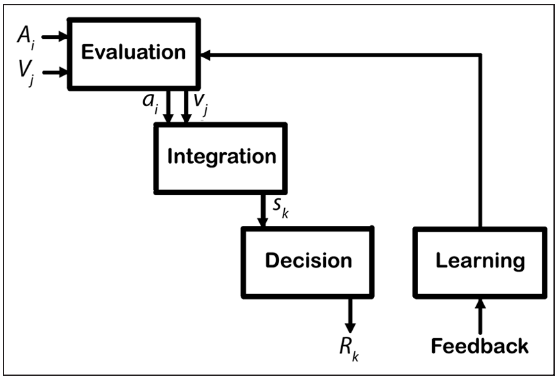

FIGURE 1. Schematic representation of three processes involved in pattern recognition, shown to precede left to right in time to illustrate their necessarily successive but overlapping processing. These processes use knowledge stored in long-term memory. Two information sources are represented by uppercase letters (A_i and V_j). The evaluation process transforms these sources of information into psychological values (indicated by lowercase letters a_i and v_j) These sources are then integrated to give an overall degree of support, s_k, for each alternative k. The decision operation maps the outputs of integration into some response alternative, R_k. The response can take the form of a discrete decision or a rating of the degree to which the alternative is likely. After recognition is completed, feedback about the response allows the learning process to adjust the information value for each source

The experiment combined six levels of visual information about the critical letter with these four levels of orthographic context in a factorial design, giving a total of 24 experimental conditions. For the six levels of visual information, the bar length of the letter took on six values going from a prototypical *c* to a prototypical *e*. The four orthographic contexts were NA (*neither* c *nor* e *admissible*), BA (*both admissible*), CA = (c *admissible*), and EA (e *admissible*). The test letter was also presented at each of the four letter positions in each of the four different contexts, giving 96 unique displays. One of these displays occurred randomly on each trial and was presented briefly for 30 ms, followed by a blank interval and then a masking stimulus after one of four intervals, which influenced the difficulty of the task. Subjects were told that there was a test letter in each display and instructed to identify it as *e* or *c* on the basis of what they saw.

The data points in Figure 2 present average performance as a function of the test letter and the context (pooled across letter position and masking interval). Both variables influenced performance in the expected direction, and the nature of their interaction is important. Although a first impression of this interaction might be interpreted as additive, the effect of context is significantly larger for the more ambiguous test letters in the middle of the stimulus continuum between *c* and *e*.

The lines in Figure 2 give the predictions of the FLMP. In producing predictions for the FLMP, it is necessary to estimate parameter values for the six levels of bottom-up information and the four orthographic contexts. Thus, 10 free parameters are used to predict the 24 independent probabilities of an *e* response. The good description of the results is apparent in Figure 2 and the small root mean square deviation (RMSD) between predicted and observed values.

The time between the onset of the test stimulus and masking stimulus was also systematically varied at processing intervals of 35, 70, 125, and 270 ms. This manipulation can illuminate the time course of processing the test letter and the context. Figure 3 shows the FLMP's predictions of the amount of support from the test letter and the context for the test letter *e* with increases in processing time before the onset of the masking stimulus. As predicted, the support from these two sources grew at the same rate in a negatively accelerated manner, reaching an asymptote at around 250 ms. In addition, the total support for the

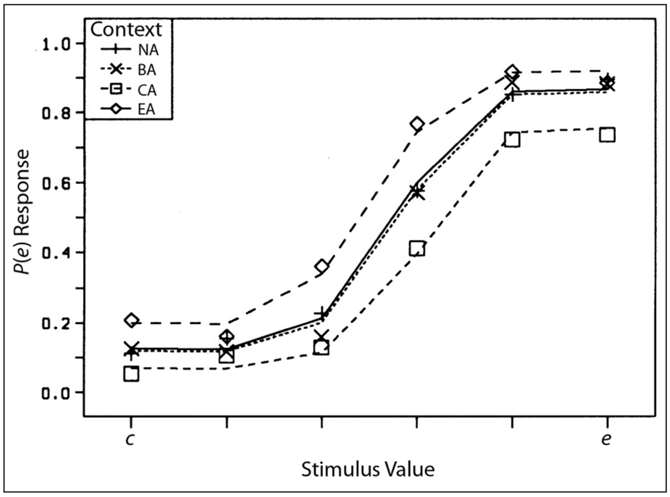

FIGURE 2. Observed (points) and fuzzy logical model of perception predicted (lines) probability of an *e* identification, *P*(*e*), as a function of the bar length of the test letter and the orthographic context. BA = both admissible; CA = *c* admissible; EA = *e* admissible; NA = neither *c* nor *e* admissible. (Results and predictions after Massaro, 1979)

test letter asymptoted at a significantly higher level than the support for its orthographic context. This result is consistent with our experience of reading more slowly when we are proofreading a text.

Figure 4 gives the predicted *e* judgments as a function of the letter information and orthographic context at each of the four processing times. The change in performance can be understood by the growth curves in Figure 3, which show more influence with more processing time. These predictions matched the results very closely, with an RMSD of .050, an impressive result given that the 96 independent data points were predicted with just 11 free parameters. These results were replicated in a second study using the letters *n* and *h* (Massaro, 1979).

We now turn to an experiment on speech perception, which also varied bottom-up and top-down influences on perception.

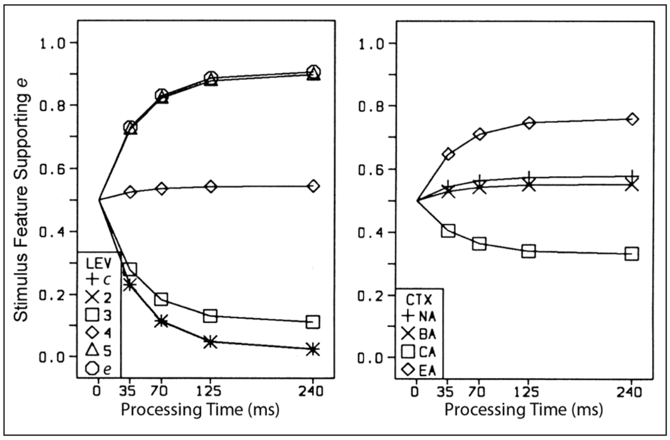

FIGURE 3. The fuzzy logical model of perception's predictions of the amount of stimulus feature supporting *e* from the test letter level (lev) and the context (ctx) with increases in processing time before the onset of the masking stimulus. BA = both admissible; CA = *c* admissible; EA = *e* admissible; NA = neither *c* nor *e* admissible. (Results and predictions after Massaro, 1979)

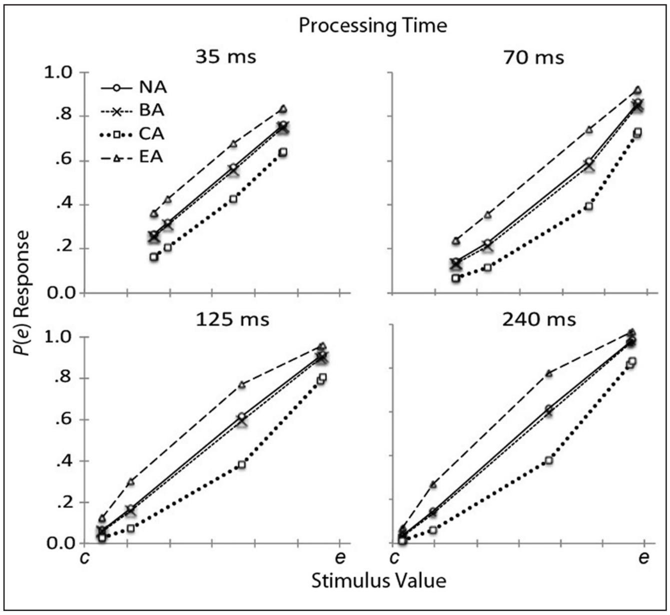

FIGURE 4. The predicted probability of *e* judgments as a function of the letter information and orthographic context at each of the four processing times. BA = both admissible; CA = *c* admissible; EA = *e* admissible; NA = neither *c* nor *e* admissible. (Results and predictions after Massaro, 1979)

Speech Perception

Exactly analogous to the reading experiment, a speech segment and its lexical context were manipulated in a factorial design (Pitt, 1995). Subjects were asked to identify a speech segment that was varied between two test alternatives and embedded in one of two different contexts. The initial consonant segment was varied along six steps between /g/ and /k/, and the following context was either the phonemic string /Ift/ or /Is/. In this case, the context /Ift/ would bias subjects to perceive /g/ as in *gift,* and the context /Is/ would bias subjects to perceive /k/ as in *kiss*.

Given that a large number of observations were recorded, we were able to analyze the exact outcome for each of the 12 subjects. The points in Figure 5 give the observed results, which resemble those in the reading experiment shown in Figure 2. Ten of the 12 subjects were influenced by lexical context in this expected direction. Subject 1 gave an inverse context effect, and Subject 7 was not influenced by context.

The lines in Figure 5 give the predictions of the FLMP (Massaro & Oden, 1995). It is not surprising that each subject was differently influenced by the two independent variables. We each have unique genetic structure and life experiences, so this variability is to be expected. The FLMP handles this variability by estimating subject-specific parameter values for the six levels of bottom-up information and the two lexical contexts. The good description of the results is apparent in Figure 5 and in the small RMSD between predicted and observed values.

Subject 1's results were the exception to the rule, which can highlight how top-down and bottom-up sources are typically combined in language perception. As can be seen in Figure 5, the other subjects were more likely to perceive the speech segment that was consistent with the word context (e.g., they were

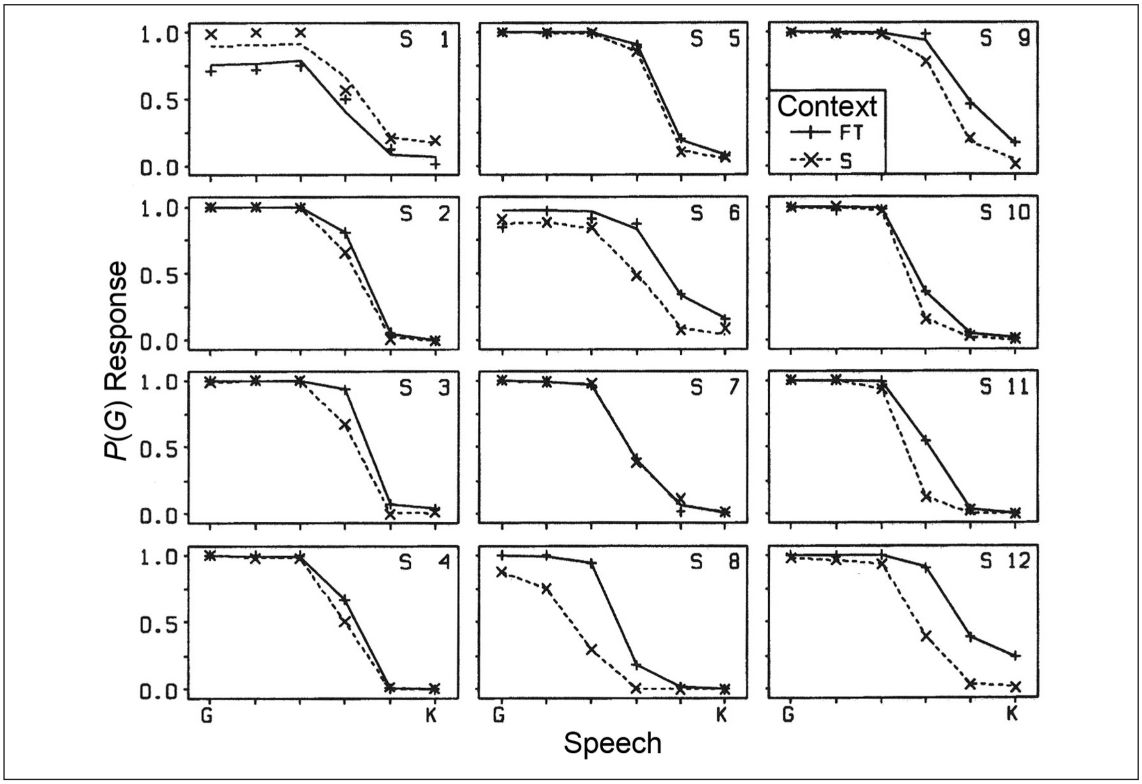

FIGURE 5. Observed (points) and fuzzy logical model of perception predicted (lines) probability of a /g/ identification, *P*(*G*), as a function of the initial speech segment along the /g/–/k/ continuum and the lexical context /Ift/ or /Is/. The 12 panels correspond to 12 different subjects. (Results and predictions after Massaro & Oden, 1995)

more likely to claim that /g/ was presented when the word ended in /Ift/ rather than /Is/). This context effect was larger to the extent that the bottom-up information was ambiguous, that is, in the middle range of the /g/–/k/ speech continuum. Subject 1 showed a reverse context effect, and the context effect was larger at the endpoints rather than in the middle of the /g/–/k/ continuum. The FLMP gave a very poor description of this subject's results, yielding an RMSD of .066. This RMSD is five times larger than the average RMSD of .013 across the independent fits of the other 11 subjects. This poor fit of the deviant results of Subject 1 is impressive because it demonstrates in another way that the FLMP is not so powerful that it can fit any possible result, a claim that has surfaced several times but not gone unanswered (Massaro, Cohen, Campbell, & Rodriguez, 2001).

A New Proposal: Acquiring Literacy Naturally

Although we have presented only two prototypical results, dozens of empirical and theoretical studies in speech science, reading, and psycholinguistics support the FLMP's critical assumption of analogous processes in language processing regardless of its input modality (Massaro, 1975, 1987, 1994, 1998; Massaro & Jesse, 2005; McClelland, 2009; Movellan & McClelland, 2001).

These same processes appear to occur in many other languages, not just English (Bates et al., 1984; Massaro, 1987). The actual sources of information can differ dramatically in different languages, but the underlying processes appear to be the same. In English sentence processing, for example, word order is more important than animacy of the constituents, whereas the opposite holds in Italian.

The processes that have been uncovered also occur in language acquisition, not just in accomplished language users (Fennell & Waxman, 2010; Hollich et al., 2000; Massaro, 1987, chapter 8). Similar to the FLMP, an emergentist coalition model assumes that children rely on multiple cues in mapping words onto referents (Golinkoff & Hirsh-Pasek, 2008; Hirsh-Pasek, Golinkoff, & Hollich, 2000). The use of and the weight given to these cues change systematically across development. For example, infants initially rely mostly on perceptual cues and gradually begin to use a speaker's intent and linguistic cues to determine word reference.

Although most of the supporting research for analogous processes involves people at least somewhat experienced in speech perception and reading, the findings also have implications for their initial acquisition. Currently, speech perception and reading are acquired in very different ways. Speech perception is learned implicitly and unintentionally with spoken language experience, whereas reading is systematically taught well after much of spoken language is acquired. The audacious hypothesis I am proposing is that reading can be learned in the same manner as speech if appropriate print input is constantly available from an early age. This proposition challenges the commonly held belief that written language necessarily comes after speech and requires formal instruction, whereas spoken language does not.

The broader impacts of this proposal are far reaching. The inability to read (illiteracy) is prevalent around the world and discouragingly present even in much of civilized society. The cost of illiteracy and the huge cost of formal literacy instruction are a major social and financial burden on societies. In addition to the high cost of public schooling, educational applications, toys, and other devices to teach reading are a huge commercial market. If this proposal is correct and it can be implemented successfully, it would have a vast impact on society because there would be no need for the many resources currently devoted to reading instruction.

This new perspective would also help redirect financial resources where they will have the most impact. Although 90% of public educational spending is on children between the ages of 6 and 19 (Karoly et al., 1997), 90% of brain growth occurs before age 6 (Gale, O'Callaghan, Godfrey, Law, & Martyn, 2004; Sakai, 2005), putatively corresponding to the stage most optimal for learning. Nurturing children for literacy before age 6 has the potential to improve the quality of children's lives, especially children who currently reside on the wrong side of the digital divide. As just one example, it is reasonable (but apparently not well documented) that reading vocabulary and grammar is much richer than spoken language. I compared the occurrence of words in a corpus of infant-directed speech to a corpus of 32 popular picture books for children. There were about 2.4 unique words in the reading corpus for every unique word in the speech corpus. The earlier children are reading, the greater is the opportunity for acquiring complex language.

Developmental, behavioral, and brain sciences have documented critical periods in audition, vision, and language. These critical periods are important for development because the brains of young children are especially plastic, or malleable. Deprivation of sensory or linguistic input during these critical periods can diminish neural cell growth, produce cell loss, and reduce the number of dendritic connections between neural cells. This can result in a substantial deficit in the functions of sensory and language systems of the child (Huttenlocher, 2002; Mayberry, 2010). Although it has not yet been studied, it is possible that limited written input during early development makes learning to read more difficult than necessary.

My proposal implies some independence between speech perception and reading. Belanger, Baum, and Mayberry (in press) measured the use of phonological and orthographic codes in reading by deaf and hearing readers with different skill levels. They found that reading skill in deaf readers was not predicted by skill in phonological processing. Mayberry, del Giudice, and Lieberman (2011) carried out a meta-analysis of the relationship between reading ability and phonological coding and awareness skills in severely and profoundly deaf people. These skills predicted only 11% of the variance in reading proficiency in the deaf subjects. In another meta-analysis of beginning hearing readers, Scarborough (2005) found that early phonological skill is not a strong predictor of later reading ability. Although this is not a direct test, evidence is accumulating for the possibility of learning to read without speech as a necessary component.

We might envision reading as an additional language learned in parallel with speech. Early reading would be of particular value for deaf children because they could learn written language in parallel with the learning of sign language or spoken language. Most children born deaf have hearing parents who do not know sign language. Early learning of a first language is the major determinant of learning a second language (Mayberry, 2007). It follows that early reading should facilitate the acquisition not only of spoken language but also of sign language. In the oral deaf community, deaf children are often bootstrapped into language via written language rather than spoken language (Mirrielees, 1947). Anne Sullivan "talked" to deaf and mute Helen Keller by drawing letters on the palm of her hand, long before she learned spoken language. Early availability of written language could be a boon for hard of hearing and deaf children.

Infants' Reading Ability

Because spoken language is present continuously from birth, speech is learned inductively. Written language is not usually pervasive enough in the growing child's world in a salient form allow inductive learning. When it is presented, usually at the onset of schooling, it is in the context of directed instruction in learning how to read. Instruction usually takes the form of building phonological awareness, learning the alphabet, learning sight words, and sounding out words (National Reading Panel, 1999). However, if an appropriate form of written text were made available early in life, and continuing through the preschool years, reading might also be learned inductively without direct instruction.

It is possible that children cannot learn to read if they cannot simultaneously write. But infants understand speech long before they produce it, and children would start writing earlier if they are learning to read earlier. James (2009) found evidence that preliterate preschool children showed greater blood oxygen level–dependent activation in the visual association cortex during letter perception only after sensorimotor learning (printing the letters). Given the synergy between reading and writing, early reading can produce a positive feedback situation in which early reading can encourage early writing, which in turn would facilitate reading acquisition.

VISION CAPABILITIES OF INFANTS AND TODDLERS

The behavioral and neuroscience literature is a rich source to assess the vision capabilities of infants and toddlers and their perceptual development in the first years of life. I have documented that there is a substantial body of measurements and research reports that provide an unambiguous conclusion that the visual system of infants and toddlers is capable of processing written language (Massaro, 2011c). Infants and toddlers quickly develop the fundamental capacities (e.g., eye movements and tracking, visual acuity, discriminating letter-like forms, and categorizing objects and events) necessary for reading. Some of the vision milestones for infants are the perception of color by 1 month, focusing ability at 2 months, eye coordination and tracking at 3 months, depth perception at 4 months, and object and face recognition at 5 months (Beddinghaus, 2010). Infants' visual acuity also improves dramatically from birth onward, reaching close to adult acuity by 8 months of age (Smith-Kettlewell, 2011). The well-documented research and measurement of vision development show that infants have the capacity to perceive appropriately structured text.

CAPABILITY OF INFANTS TO LEARN TO READ

There is no doubt that infants are sophisticated information processors and quick learners of the varied situations they experience. A typical experiment habituates infants to a sequence of inputs with specific statistical constraints and then changes it to determine whether the infants notice the change. As an example, 9-month-old infants learned about scenes that differed in the spatial juxtaposition of several colored geometric objects (Fiser & Aslin, 2010). A recent study by Téglás et al. (2011) revealed an impressive ability of 12-month-old infants to form time-varying expectations of complex displays of multiple moving objects. This behavior was described as being consistent with a Bayesian ideal observer embodying abstract principles of object motion, which is consistent with the FLMP theoretical framework we have described.

Infant perception studies have used speech, music, and two- and three-dimensional objects, but none of these recent studies have used letters as test stimuli. In addition to my intuition, there is one finding that infants would be equally competent with letters. It appears that some of the stimuli that have already been used have the same visual and optical characteristics as letters. Changizi (2009) analyzed the topological properties of the shapes of a wide range of nonpictorial signs and alphabets. These topological properties describe the manner in which the separate lines of a shape intersect or join with other lines. The outcome indicated that only a subset of possible topographical shapes are used, and, most impressively, these shapes conformed to the contours of objects found in our natural environments across many geographic settings. This analysis reveals that there is nothing special about letters. Our visual system evolved to efficiently perceive our surroundings, and it appears that alphabet forms evolved to replicate those same properties. Thus, we expect that infants would be able to show off their perceptual processing and memory skills with letters also.

In the speech domain, research has shown that infants can recognize the sound patterns of their native language within just a few days after birth and their own names by 4 months and perceive individual words by 6 months (Metcalf Infant Research Lab, 2011). Thus, infants have the capacity to perceive, process, and abstract semantic components that occur in language (Parish, Ma, Hirsh-Pasek, & Golinkoff, 2010). If an appropriate form of written text is continuously available during infant development, the hypothesis is that reading acquisition should approximately follow the same time course as speech perception.

Most children are first instructed in learning to read when they begin formal schooling. In the last decades, however, there have been advocates of teaching reading well before schooling begins (Sanger, 2011). What is important to note, however, is that all these proposals accept current reading pedagogy and simply implement it for much younger children. Thus, even advocates of early reading accept, at least implicitly, that speech occurs naturally and reading must be taught. The video demonstrations posted on the Web have many examples of preschool children being taught to read and succeeding in recognizing, remembering, sounding out individual words and phrases. Scrutiny of these demonstrations reveals that the reading behavior of these young children matches that of children in school. For example, deliberate word-by-word reading occurs early in the trajectory of school children when they are learning to read.

Language proficiency is critical for learning to read, and I am not proposing that simply having written language in the child's world preempts its importance. In a meta-analysis, Mayberry et al. (2011) found that language ability predicted 35% of the variance in reading proficiency. As concluded by the authors, "These meta-analytic results indicate that PCA (phonological coding and awareness) skills are a low to moderate predictor of reading achievement in deaf individuals and that other factors, most notably language competence such as vocabulary size, have a greater influence on reading development, as has been found to be the case in the hearing population" (p. 164). Very early language development appears to be especially critical because competence measures at ages 3 and 4 predict later reading skill as well as these measures taken at age 5 (Scarborough, 2001, 2005). In addition, early language skills predict phonological awareness when school begins as well as direct measures of phonological awareness (Scarborough, 2005).

Some professionals might question whether a toddler can negotiate written language simultaneously with his or her typical social interactions. Spoken language is easier in this regard because the child can attend to an object or event while simultaneously hearing language. Written language requires the child to pay consecutive attention to the event and to the written language. Written language can be learned like sign language in which caregivers will either attract the child's attention to the sign or simply sign in front of the object or event that the child is attending to (Lieberman, Hatrak, & Mayberry, 2011). The caregiver illustrates how toast is buttered and then attracts the child's attention to her depiction of the event in sign language. If the child is learning to read naturally, the caregiver–child interaction will involve consecutive attention to what is meaningful in the exchange and the written language.

It is possible that early introduction of written language will reduce the occurrence of reading disabilities to the level of spoken language disabilities. Given the assumed comparability between speech perception and reading processes, early literacy in reading would make it comparable to competence in speech understanding.

We reviewed fairly convincing evidence indicating that young preliterate children have a visual system that is mature enough to accurately track, discriminate, and categorize written language, and therefore they should be able to learn to read. Even if the child is capable, providing the appropriate written input is more challenging. Ideally, we need an environment that enables a child to learn to read naturally, without intention, and with no negative consequences. I have proposed a method and system that would elucidate the meaning of a child's experience and display this meaning in written language appropriate to the child's reading skill (Massaro, 2011a).

One possible instantiation of this system is shown in Figure 6. It uses automated speech recognition of the caregiver's speech to infer the child's experience, which can then be described in written form that the child sees (Massaro, 2011b). A second possibility is to implement object, scene, and action recognition to determine what is most meaningful to the child at a given instant and to present it in written form. Admittedly, these technologies are not quite ready for primetime, but it is my firm belief that they are nearly good enough to make written language pervasive during a child's early exposure to language. My vision is an interactive system, technology assisted reading acquisition (TARA), to allow children to acquire literacy naturally.

Retrospective

We began this journey with a discussion of two innovative research articles, which envisioned the apperceptive process as common to both speech perception and reading. We modernized this process within the framework of the FLMP by conceptualizing both speech perception and reading as pattern recognition. However, the study of speech and reading split off into different subdisciplines that prolonged behavioral science's study of their similarity and the resulting potential for applications. Thus research and education have not considered the proposition advocated here. If indeed it is correct, we should be able to create an environment for the developing child whereby written text is constantly available so that learning to read occurs automatically, without intention, without direct instruction, and with no negative consequences. Implementing this goal challenges and seeks to shift current research and clinical practice paradigms and would

FIGURE 6. Four scenes of a caregiver interacting with a 19-month-old boy who is playing with a Lego man. The caregiver might say, "You have a toy man" or "This is a man." This speech is recognized by a speech recognition system implemented on an iPad, and an edited written form "MAN" is shown on its screen. Like sign language, shared attention has to be negotiated between looking at the written word and processing the signified event

solve a previously believed unsolvable problem of a guaranteed method of learning to read.

NOTE

I was not able to adequately address many of the issues mentioned in this article, and the reader should refer to my research papers at http://mambo.ucsc.edu/people/dominic-massaro.html. Applications related to learning to read naturally are described at http://www.earlyread.org/ and http://psyentificmind.com/read-with-me-2. Correspondence should be addressed to massaro@ucsc.edu.

REFERENCES

Bagley, W. C. (1900). The apperception of the spoken sentence: A study in the psychology of language. *American Journal of Psychology, 12,* 80–130.

Bates, E., MacWhinney, B., Caselli, C., Devescovi, A., Natale, F., & Venza, V. (1984). A crosslinguistic study of the development of sentence interpretation strategies. *Child Development, 55,* 341–354.

Beddinghaus, T. (2010). *Top 5 milestones in vision development.* Retrieved from http://vision.about.com/od/childrensvision/tp/vision_develop.htm

Belanger, N. N., Baum, S. R., & Mayberry, R. I. (in press). Reading difficulties in adult deaf readers of French: Phonological codes, not guilty! *Scientific Studies of Reading*.

Cattell, J. M. (1886). The time it takes to see and name objects. *Mind, 11,* 63–65.

Chall, J. (1967). *Learning to read: The great debate.* New York, NY: McGraw-Hill.

Changizi, M. (2009). *The vision revolution: How the latest research overturns everything we thought we knew about human vision.* Dallas, TX: Benbella.

Elingsh. (2011). Retrieved from http://www.snopes.com/language/apocryph/cambridge.asp

Erdmann, B., & Dodge, R. (1898). *Psychologische Untersuchung über das Lesen auf experimenteller Grundlage* [Psychological study of reading on an experimental basis]. Halle, Germany: Niemeyer.

Fennell, C., & Waxman, S. R. (2010). What paradox? Referential cues allow for infant use of phonetic detail in word learning. *Child Development, 81*(5), 1376–1383.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences, 99,* 15822–15826.

Gale, C. R., O'Callaghan, F. J., Godfrey, K. M., Law, C. M., & Martyn, C. N. (2004). Critical periods of brain growth and cognitive function in children. *Brain: A Journal of Neurology, 127*(2), 321–329.

Golinkoff, R., & Hirsh-Pasek, K. (2008). How toddlers learn verbs. *Trends in Cognitive Science, 12,* 397–403.

Hilgard, E. R. (1987). *Psychology in America.* San Diego, CA: Harcourt Brace Jovanovich.

Hirsh-Pasek, K., Golinkoff, R., & Hollich, G. (2000). An emergentist coalition model for word learning: Mapping words to objects is a product of the interaction of multiple cues. In R. M. Golinkoff, K. Hirsh-Pasek, N. Akhtar, L. Bloom, G. Hollich, K. L. Smith, . . . A. Woodward (Eds.), *Breaking the word learning barrier: What does it take?* New York, NY: Oxford University Press.

Hollich, G., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., . . . Rocroi, C. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Society for Research in Child Development Monograph Series, 65*(3), i–vi, 1–123. Chicago, IL: University of Chicago Press.

Huey, E. B. (1968). *The psychology and pedagogy of reading*. Cambridge, MA: MIT Press. (Original work published 1908)

Huttenlocher, P. R. (2002). *Neural plasticity: The effects of environment on the development of the cerebral cortex.* Cambridge, MA: Harvard University Press.

James, K. H. (2009). Sensori-motor experience leads to changes in visual processing in the developing brain. *Developmental Science, 13,* 279–288.

James, W. (1890). *The principles of psychology*. Mineola, NY: Dover.

Kandel, I. (1961). *William Chandler Bagley: Stalwart educator.* New York, NY: Columbia University Press.

Karoly, L., et al. (1997). *Rand Corporation, April 1997. Public expenditures (in constant 1992 dollars) on children aged 0 to 5 are for early childhood development, welfare programs and health care; expenditures on children aged 6 to 18 are for education, criminal justice, youth employment and other programs.* Retrieved from http://results.techriver.net/website/article.asp?id=1073

Lieberman, A. M., Hatrak, M., & Mayberry, R. I. (2011). The development of eye gaze control for linguistic input in deaf children. In N. Danis, K. Mesh, & H. Sung (Eds.), *BUCLD 35: Proceedings of the 35th annual Boston University Conference on Language Development* (pp. 391–403). Somerville, MA: Cascadilla.

Massaro, D. W. (1975). *Understanding language: An information processing analysis of speech perception, reading and psycholinguistics.* New York, NY: Academic Press.

Massaro, D. W. (1979). Letter information and orthographic context in word perception. *Journal of Experimental Psychology: Human Perception and Performance, 5*(4), 595–609.

Massaro, D. W. (1987). *Speech perception by eye and by ear: A paradigm for psychological inquiry*. Hillsdale, NJ: Erlbaum.

Massaro, D. W. (1994). Psychological aspects of speech perception: Implications for research and theory. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 219–263). New York, NY: Academic Press.

Massaro, D. W. (1998). *Perceiving talking faces: From speech* *perception to a behavioral principle*. Cambridge, MA: MIT Press.

Massaro, D. W. (2011a). *Method and system for acquisition of literacy.* Patent Application No. 13/253,335, October 5, 2011.

Massaro, D. W. (2011b). *A reading window for visually challenged individuals.* Paper presented at the 26th Annual International Technology & Persons with Disabilities Conference, San Diego, CA.

Massaro, D. W. (2011c). *Universal literacy: The digital age engages the learning brain*. Paper presented at "iGeneration: How the Digital Age Is Altering Student Brains, Learning & Teaching," 28th International Conference for Pre-K Through University Educators, Parents and Clinicians, San Francisco, CA. Retrieved from http://www.earlyread.org/wp-content/uploads/2011/07/massaroLiteracy2.pdf

Massaro, D. W., Cohen, M. M., Campbell, C. S., & Rodriguez, T. (2001). Bayes factor of model selection validates FLMP. *Psychonomic Bulletin & Review, 8,* 1–17.

Massaro, D. W., Cohen, M. M., Meyer, H., Stribling, T., Sterling, C., & Vanderhyden, S. (2011). Integration of facial and newly learned visual cues in speech perception. *American Journal of Psychology, 124,* 341–354.

Massaro, D. W., & Jesse, A. (2005). The magic of reading: Too many influences for quick and easy explanations. In T. Trabasso, J. Sabatini, D. W. Massaro, & R. C. Calfee (Eds.), *From orthography to pedagogy: Essays in honor of Richard L. Venezky* (pp. 37–61). Mahwah, NJ: Erlbaum.

Massaro, D. W., & Oden, G. C. (1995). Independence of lexical context and phonological information in speech perception. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1053–1064.

Mayberry, R. I. (2007). When timing is everything: Age of first-language acquisition effects on second-language learning. *Applied Psycholinguistics, 28*(3), 537–549.

Mayberry, R. I. (2010). Early language acquisition and adult language ability: What sign language reveals about the critical period for language. In M. Marschark & P. Spencer (Eds.), *Oxford handbook of deaf studies, language, and education* (Vol. 2, pp. 281–291). Oxford, England: Oxford University Press.

Mayberry, R. I., del Giudice, A. A., & Lieberman, A. M. (2011). Reading achievement in relation to phonological coding and awareness in deaf readers: A meta-analysis. *Journal of Deaf Studies and Deaf Education, 16*(2), 164–188.

Metcalf Infant Research Lab. (2011). *The Metcalf Infant Research Lab at Brown University.* Retrieved from http://www.cog.brown.edu/research/infant_lab/

McClelland, J. L. (2009). The place of modeling in cognitive science. *Topics in Cognitive Science, 1,* 11–38.

Miles, W. R. (1964). *A biographical memoir. Walter Bowers Pillsbury.* Washington, DC: National Academy of Sciences.

Mirrielees, D. I. (1947). *Education of the young deaf child: Special subjects and methods*. Chicago, IL: University of Chicago Home Study Department.

Movellan, J., & McClelland, J. L. (2001). The Morton–Massaro law of information integration: Implications for models of perception. *Psychological Review, 108,* 113–148.

National Reading Panel. (1999). *Teaching children to read. Summary report.* Retrieved from http://www.nationalreadingpanel.org/publications/summary.htm

New World Encyclopedia. (2011). *William Chandler Bagley (March 15, 1874-July 1, 1946).* Retrieved from http://www.newworldencyclopedia.org/entry/William_Chandler_Bagley

Parish, J., Ma, W., Hirsh-Pasek, K., & Golinkoff, R. M. (2010). A world of relations: Relational words. In B. Malt & P. Wolf (Eds.), *Words and the world: How words capture the human experience* (pp. 219–242). New York, NY: Oxford University Press.

Pillsbury, W. B. (1897). A study in apperception. *American Journal of Psychology, 8,* 315–393.

Pitt, M. A. (1995). The locus of the lexical shift in phoneme identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1037–1052.

Sakai, K. L. (2005). Language acquisition and brain development. *Science, 310*(4), 815–819.

Sanger, L. (2011). *Reading Bear.* Retrieved from http://larrysanger.org/category/education/reading-bear/

Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. Neuman & D. Dickinson (Eds.), *Handbook for research in early literacy* (pp. 97–110). New York, NY: Guilford.

Scarborough, H. S. (2005). Developmental relationships between language and reading: Reconciling a beautiful hypothesis with some ugly facts. In H. W. Catts & A. G. Kamhi (Eds.), *The connections between language and reading disabilities* (pp. 3–24). Mahwah, NJ: Erlbaum.

Smith-Kettlewell. (2011). *What can my baby see?* Retrieved from http://www.ski.org/Vision/babyvision.html

Téglás, E., Vul, E., Girotto, V., Gonzalez, M., Tenenbaum, J. B., & Bonatti, L. L. (2011). Pure reasoning in 12-month-old infants as probabilistic inference. *Science, 27*(332), 1054–1059.

Woodworth, R. S. (1938). *Experimental psychology*. New York, NY: Holt.

Walter Bowers Pillsbury

William Chandler Bagley

APPENDIX. CAREER PATHS OF PILLSBURY AND BAGLEY

Pillsbury and Bagley did not initiate careers centered on their PhD research. By today's standards, either of their dissertations would have set the stage for many years of follow-up questions and experimental investigations. However, Pillsbury published his dissertation research in a French book called *L'Attention,* with the English version following 2 years later. He was able to take the elusive concept *apperception* and describe it in terms of a familiar concept ("everyone knows what attention is," as described by William James, 1890). Pillsbury took a faculty position at the University of Michigan, being an active and productive member for 63 years. During this same impressive period, he was also active as a member of the editorial board of this journal. At Michigan, he established an experimental psychology laboratory (although he was in the philosophy department for many years) and nurtured many advanced students) (Miles, 1964).

Bagley took a somewhat different path. He trained teachers, was a school superintendent, and settled in at Teachers College, Columbia University, founding a teacher education department. He lectured and wrote many books for teacher training, and indeed it was *training,* given his authoritarian view of education. His model was a principal issuing orders to the teachers, who enforced tasks for students, who would obtain skills and knowledge. Although conservative, he advocated educational foundations based on scientific principles. He cofounded the *Journal of Educational Psychology* in 1910 and remained active in teaching students and collaborating with colleagues (Kandel, 1961; New World Encyclopedia, 2011). During this long career, Bagley authored and coauthored more than 30 books and 400 articles (all without a word processor).