Sperling, G. (1988). The magical number seven: Information processing then and now. In W. Hirst (Ed.), *The making of cognitive science: Essays in honor of George A. Miller* (pp. 71-80). Cambridge, UK: Cambridge University Press.

5 The magical number seven: information processing then and now

---

George Sperling

The magical hypotheses

George Miller read an invited address to the Eastern Psychological Association on April 15, 1955: "The Magical Number Seven, Plus or Minus Two: Some Limits in Our Capacity for Processing Information." He complained: "I have been persecuted by an integer." The article was published the following year in the *Psychological Review*. Both the spoken version and the written version (here referred to as 7 ± 2) were immediate successes. A survey conducted twenty years later found that 7 ± 2 was the single most often cited paper in cognitive psychology (Garfield, 1975). Every student of cognitive psychology has been exposed to it, and many cognitive psychologists, including myself, have been profoundly influenced by it.

The central thesis of 7 ± 2 is that the number 7 occurs in two contexts. The first context is absolute judgments of brightness, loudness, pitch, extent, and so on. In absolute judgments – that is, classification of sensory stimuli into categories – 7 ± 2 *is* the effective number of categories that the subject can maintain. This number is derived from what was then a novel statistical computation: the 2 to 3 bits of *information transmitted* by an observer in these tasks.

To transmit more than 3 bits of information in an absolute judgment, a subject requires sensory stimuli that vary in more than one dimension. The transmitted information in each component dimension suffers somewhat as new dimensions are added; nevertheless, subjects can transmit enormously greater amounts of information about multidimensional stimuli than about one-dimensional stimuli.

The second context in which 7 ± 2 occurs is memory. In short-term recall (the classical immediate memory test), 7 ± 2 describes the number of items that a subject can recall. In contrast to the absolute judgments, the number of recallable items does not depend on their information content, so the information-transmitted statistic does not predict performance. Although binary digits contain less than one-third the information of decimal digits, a subject can recall only very slightly more binary digits than decimal digits. However, by recoding several separate elements (e.g., binary digits) into a unitary "chunk" (e.g., an octal digit), the subject can enormously increase his or her recall capacity for binary digits.

George Miller classified many of the experimental procedures that had been used to study information processing into two categories (absolute judgments and immediate recall) and invented twin hypotheses, one to characterize the human performance limit in each category. One-dimensional absolute judgments could transmit up to 3 bits of information; immediate recall was limited to 7 ± 2 chunks. All of the experiments that George Miller advanced in support of these twin hypotheses had been performed by others. The contribution of 7 ± 2 is entirely theoretical – its succinct classification of a great deal of data and its clear formulation of two hypotheses that demanded theoretical explanations.

Professor George Miller

In the spring of 1958, Professor Miller offered a seminar in information processing for the graduate students in the Psychology Department at Harvard. I enrolled in that class and volunteered to give the seminar presentation on 7 ± 2. How I got to this point is a story in itself.

The previous spring, I had been a student in the neighboring Department of Social Relations. Professor Miller conducted the last sessions of the required proseminar (offered jointly with Jerome Bruner, Richard Solomon, and George Mandler). I presented a class report on a paper by Lawrence and Laberge (1956). After describing their experiments, I proposed a partial report experiment that could better address the same issues. Professor Miller was sufficiently interested in this proposed experiment to offer to support it. Thereby began our association.

In the summer of 1957, Professor Miller obtained permission for me to use Jerry Bruner's tachistoscope during Bruner's absence, and he supported my application for a transfer to the Psychology Department. In the fall, just as I entered the Psychology Department, my draft board officially notified me that they now expected me to turn my attention to their needs. This was obviously to be my last chance in graduate school. Fortunately, during the summer, with the encouragement of Roger Shepard (Professor Miller's postdoctoral fellow, whom he designated to oversee the research), I had completed the experiments for my Ph.D. thesis on what Ulric Neisser (1967) later dubbed "iconic memory."

In those days, from the students' point of view, Harvard's Psychology Department was clearly divided on ideological grounds that were reflected in its geographical layout. Fred Skinner was located at the north end of the basement of Memorial Hall, Smitty Stevens was entrenched at the opposite end, and George Miller, Phil Teitelbaum, Edwin Newman, and everyone else who had not chosen sides was thrown together in the middle. At the opposite ends of Memorial Hall, students attended competing discussion groups offered weekly by Stevens and Skinner. Cognitive science was centered in the middle.

In Professor Miller's seminar, each student was to present an extended two-hour seminar on an important paper. I chose 7 ± 2 because it was supposed to be a very important paper. On first reading, 7 ± 2 seemed to offer the student an opportunity for a dazzling display of critical acumen. Its assertion that the span of absolute judgment could be so low was patently absurd. Simply searching out the references would uncover their procedural artifacts. For visual judgments of spatial extent, this nitpicking paid off. The displays were tiny (subtending less than ½ degree) and, even so, 3.25 bits of information were transmitted (corresponding to 9.5 alternatives). I imagined that larger displays would undoubtedly produce still better performance and thereby significant violations of 7 ± 2. In all of the other modalities, however, the data withstood scrutiny, and they firmly contradicted my intuition. I almost began to feel persecuted, too.

Miller had noted that, with multidimensional stimuli in absolute judgment experiments, the additional dimensions enable the subject to surpass the 7 ± 2 restriction on the number of effective categories. I intended to propose using this property in reverse to discover what the underlying dimensions of judgments were. Since this seminar was often attended by postdoctoral fellows (such as Roger Shepard) and students from Social Relations (such as Saul Sternberg), as well as by the Psychology Department students, it was an occasion for lively exchanges. However, on this occasion, another psychology student, Jerry Shickman, objected so vehemently and persistently to my use of the concept of dimension that I was unable to proceed. By the time Professor Miller intervened to get things going again, so much tension had built up that subsequent discussion was totally inhibited.

The seminar meeting was a failure. However, the intellectual seed had been sown. The particular set of problems and the issues surrounding them have remained with me ever since. And although I have been persecuted by editors and critics for many more than seven years, my experiments and the data they generated have been a source of much comfort.

The challenge of a theory

What was it about 7 ± 2 that made it such a milestone in cognitive psychology? As usual, it was not just one thing but a propitious combination of many factors. The framework of the presentation was masterful: the provocative challenge of the theoretical approach. Like Sherlock Holmes, the theoretician demonstrates that the evidence is already at hand. One need only be clever enough to perceive it. By implication, of course, the data-bound experimenters and the rest of us in the audience were not up to the task. It was unnecessary for George Miller to proclaim ponderously that there must be a role for theoreticians in a discipline dominated by experimentalists; the demonstration was a better proof. A cognitive psychologist could be a theoretical psychologist proposing simple hypotheses that organized large amounts of data.

The classical period and the dark ages of information processing

To understand why George Miller's twin hypotheses – his theory – had such an impact, we must examine the field as it was when 7 ± 2 burst forth upon it. The three main textbooks of the time were Woodworth and Schlosberg's *Experimental Psychology* (1954), Osgood's *Method and Theory of Experimental Psychology* (1953), and Stevens's *Handbook of Experimental Psychology* (1951).

Osgood offers no treatment whatsoever of any of the paradigms, data, or issues raised in 7 ± 2. In Stevens's *Handbook,* C. H. Graham's chapter on visual perception has a subsection suggestively entitled "Span of Perceptions," but its four pages are devoted entirely to Hunter and Sigler's (1940) study of estimated dot numerosity as a function of luminance and exposure duration in brief displays. The chapter on cognitive processes by Leeper is utterly astounding in terms of how the field is defined today. The beginning is bogged down in a discussion of consciousness, and the remainder is devoted to concept formation in monkeys.

Not only George Miller's interests in 7 ± 2 (information transmitted in perceptual judgment and in short-term recall) were given short shrift in Osgood's and Stevens's compendia of the 1950s. Attention, in the context of human performance, is entirely absent. Yet, in classical psychology, attention is the heading under which the paradigms of 7 ± 2 would be treated. Related subjects were also ignored. For example, even a subject with sensory as well as cognitive implications – saccadic eye movements as our means of acquiring visual information – is not mentioned in either text. Anyone who thinks that 7 ± 2 did not represent a leap forward in our conceptualization of the important issues of psychology need only look at the primordial ooze from which it sprang.

Woodworth and Schlosberg's views of the important issues in psychology fare better by today's standards. Most of the topics previously mentioned are considered. The treatment is dust bowl empirical; many experimental procedures are summarized and the results stated matter-of-factly. If one already knows what the interesting questions are, the data speak for themselves. But there is not a trace of theory.

The treatment of spans in Woodworth and Schlosberg derives directly from the earlier edition of Woodworth (1938). And this book, in turn, owes much to Woodworth's long internship in Wundt's laboratory.

The debt to Wundt

Wundt was concerned with precisely the issues that 7 ± 2 raises and devoted considerable attention to them immediately in Chapter 1 of his popular *Introduction to Psychology* (1912), a condensation of his longer *Outlines*. Wundt asks, how many elements can be maintained simultaneously in consciousness? When the elements are ticks of a metronome, the answer depends on how the subject groups them. Subjectively grouping ticks by twos (imagining 2/8 time) yields sixteen conscious clicks grouped into eight "chunks" (to use Miller's term). Grouping by eights (imagining 4/4 time) yields only five chunks, although chunking increases the total number of ticks from sixteen to forty. Similar advantages of chunking appear in the visual domain when viewing clearly visible tachistoscopic flashes of unrelated letters or words. Viewers report they can perceive as many unrelated words as unrelated letters (about four). In the auditory modality, observers can report back six spoken nonsense syllables. Braille characters are designed to use only six tactile dot positions. Basically, Wundt proposes the magical number 6 ± 2, and indeed, for immediate memory, this describes the situation.

Unlike his mentor, Helmholtz, whose procedures are as timely today as they were one hundred years ago, Wundt's usual methods are entirely subjective. This was by Wundt's design. When he thought the occasion required it, Wundt could be quantitative and rigorous by today's standards. For example, his *Physiological Psychology* is full of examples of the comparative structure of sense organs in various species and of mathematical formulations of significant relations. Wundt was versed in biology and mathematics. In trying to forge a distinctive science of psychology, different from physiology, Wundt wanted to use distinctively different procedures. Thereby, he chose the wrong path. The new psychology was not then, and still is not, ready for Wundt's introspective methods. Cognitive psychology still succeeds best with experimental procedures that place the complexity in the stimulus and leave the response simple and constrained. Wundt's reverse procedures, such as presenting the subject with a simple red patch and asking him to introspect at length about what he sees, are still beyond our reach.

Helmholtz and Wundt raised many questions that we regard as core issues of cognitive psychology. Helmholtz concentrated primarily on perception, and his methods were universal. Wundt went further, to information processing, thinking, and beyond, but his methods too often were introspective. Many of Wundt's followers were less well versed in scientific protocol than he. In their hands, the introspective methods ultimately stimulated psychologists to create the behavioral revolution. Unfortunately, as in many revolutions, the good was overthrown with the bad. For almost half a century, not only the methods but even the questions were discarded by most American psychologists.

Traces of interest in cognition survived in a new empirical garb, as noted, for example, by Woodworth and Schlosberg. Although casting cognitive questions in empirical, behavioral terms was an improvement, in the absence of theory the spark was lost. Aside from observing empirical relationships, as in the span of apprehension experiments, there was no inkling of how to cast theories.

The renaissance

When Shannon's (Shannon & Weaver, 1949) information theory came on the scene, it was quickly adapted to a variety of paradigms because it was the only systematic framework psychologists had for dealing with information. [At the same time, Wald's (1950) statistical decision theory blossomed in psychology as signal detection theory.] But the routine application of information theory to psychological paradigms was unfruitful. In the morass of information-oriented experiments, 7 ± 2's contribution was the clear delineation of a domain in which information theory was useful (absolute judgments) and a complementary domain (short-term list recall) in which it was not. This was an important and necessary step in moving forward. In a larger scale, in relation to its time, 7 ± 2 redefined the classical subject of span experiments in terms of information processing, an area of cognitive psychology with clearly phrased problems and the possibility of significant theoretical approaches.

The baroque era and the age of computers

The classical past offered the fascinating cognitive issues posed by Wundt and, in the United States by James but completely lacked an adequate methodology. In the post–World War I period of unmitigated empiricism, the intellectual thread was lost. A post–World War II flurry of information theoretic studies culminated in 7 ± 2. What was and is yet to come?

In outline, the path ahead looks straightforward. The 7 ± 2 theory is a descriptive macro theory. That is, it offers mathematical descriptions of stimulus–response relations at a very global level. The descriptive formulations of 7 ± 2 will eventually be supplemented with process theories – models that embody the step-by-step computations carried out in the cognitive microprocesses that underlie performance. Eventually, the process models will be fleshed out with neural components that represent the biological structures that carry out the cognitive microprocesses. The early stages of this process of scientific evolution can already be discerned.

Acoustic confusibility

The first important progress following 7 ± 2 came with the observation that not all items were equivalent, even when they conveyed precisely the same information. It was discovered that the number of items recalled from visual or auditory presentations depends on the acoustic structure of the items. Items that are acoustically confusible (such as the letters *b, c, d, g, p, t,* and *v*) are not recalled as well as items that are less confusible (Conrad & Hull, 1964; Sperling, 1963; Sperling and Speelman, 1970). Thus, all chunks are not equivalent; how well a chunk is remembered depends on its sound.

That acoustic structure is critical suggests an acoustic basis for short-term memory (Sperling, 1968). Articulatory coding is an alternative possibility (Hintzman, 1967). There are severe difficulties with any purely structural theory – acoustic or articulatory – since familiarity, which is not easily embodied in any of these theories, has an enormous role in short-term recall (Sperling, Parish, Pavel, & Desaulniers, 1984). Any contemporary theory of short-term recall must deal in much more detail with much more detailed cognitive processes than Miller was forced to do. For example, we know that the phonemic and acoustic structure of the to-be-remembered chunks matters; that recoding, rehearsal strategies, and grouping have specifiable mnemonic effects; and that prior learning experiences with the items are critical. These are some of the presumed components of process theories.

Neural models for immediate memory have been proposed by Grossberg (1980), bypassing the functional process description. However, in the absence of a functional model to explain recoding, rehearsal, grouping, and other strategic options available to the subject, neural specification is probably premature.

Memory noise in absolute judgments

The limiting factor in absolute judgments seems to be that subjects cannot remember the precise boundaries of their categories. There is some uncertainty (noise) in coding stimulus intensity, but the main bottleneck seems to be limited capacity memory (Durlach and Braida, 1969).

One particular formulation of the memory bottleneck (Heinemann, 1984) has been extensively tested on pigeons as well as humans. Chase (1984) and Heinemann find that pigeons make absolute judgments of auditory intensity that are qualitatively quite similar to human judgments. They model the limited capacity memory by assuming that, in memory, the outcome of a trial is represented by a record containing the stimulus, the response, and the feedback on that trial. Memory contains about 1,000 locations. Each new trial is stored independently at a random location of memory, overwriting the previous contents of the location. When it comes time to make a judgment of a new stimulus, some records, typically estimated to be about seven, are extracted from the 1,000-location long-term memory and placed into a short-term working memory. The judgment is made by comparing the unknown stimulus to the contents of working memory.

Heinemann's model accounts nicely for a number of second-order effects relating to absolute judgments, such as how discimination in various parts of the stimulus dimension depends on the spacing of stimuli and how a relabeling of the stimuli is gradually learned. Although Heinemann avoids the inference, for me the attractive aspect of this kind of process model is that the working memory that holds the records of previous trials of the judgment experiment may be the same memory that holds the chunks in the immediate memory experiment. This kind of theory is representative of the exciting kinds of models of mental microprocesses that cognitive psychology offers. It also illustrates the incompleteness of theories in which the complex control processes needed to utilize limited-capacity memories are axiomatically assumed rather than explicitly modeled.

The coming era

Physiologically, it is unlikely that formatted records of the sort described in Heinemann's model are written in a memory with a fixed number of locations. This description of a process is best regarded as a convenient, workable conceptualization of a neural network. Indeed, it is easy to design neural networks that behave like the stack memory previously described. Whether the network that actually performs the stacklike memory function in the brain is describable in simple terms is not known. It may or may not be simply organized. However, a complete functional description that can be applied to any particular experimental situation will undoubtedly be very complex.

One of the remarkable emerging properties of the many recently proposed neural networks is that they are quite similar in their overall learning properties, even though their structures are profoundly different. These networks can be used interchangeably to fill the black boxes of the functional process models in much the same way that computer memory chips of different manufacturers can be interchanged on a central processing unit board. At least, that is how one aspect of the future of cognitive models appears to me in the 1980s.

The future

What is the future of 7 ± 2? It offers a descriptive theory of two classes of phenomena. We look forward to better process theories and eventually, to physiologically plausible neural theories. Does this mean that 7 ± 2 will become outmoded and replaced? Not necessarily. Predicting the fate of any psychological theory requires a careful consideration of the more general role of the theory.

The goal of theory in experimental psychology is to provide the best possible description of a class of phenomenon at a given level of complexity (Sperling, 1978). Subsequent theories may be more complex and detailed, but they will not replace earlier theories unless they can explain more with equal or less complexity. Insofar as a theory offers the best description at a given level of complexity, it is eternal and will not be replaced, though it certainly will be supplemented. There are many difficulties with this formulation of the goal of theory, not the least of which is the continuously changing nature of complexity. But the magical cognitive powers of the number seven make 7 ± 2 a probable candidate for the best theory at its chosen level of complexity. If so, 7 ± 2 will endure. That is the ultimate achievement of any theory.

References

Chase, S. 1984. Pigeons and the magical number seven. In M. L. Commons, R. J. Herrnstein, & A. R. Wagner (Eds.), *Quantitative Analyses of Behavior: Discrimination Processes* (pp. 36–47). Cambridge, Mass.: Ballinger.

Conrad, R., & Hull, J. A. (1964). Information, acoustic confusion, and memory span. *British Journal of Psychology, 55,* 429–32.

Durlach, N. I., & Braida, L. D. (1969). Intensity perception. I. Preliminary theory of intensity resolution. *Journal of the Acoustical Society of America, 46,* 372–83.

Garfield, E. (1975). Highly cited articles. 19. Human psychology and behavior. In E. Garfield, *Essays of an Information Scientist* (pp. 262–8). Philadelphia: Institute for Scientific Information.

Graham, C. H. (1951). Visual perception. In S. S. Stevens (Ed.), *Handbook of Experimental Psychology* (pp. 868–920). New York: Wiley.

Grossberg, S. (1980). How does a brain build a cognitive code? *Psychological Review, 87,* 1–51.

Heinemann, E. G. (1984). A memory model for decision processes in pigeons. In M. L. Commons, R. J. Herrnstein, & A. R. Wagner (Eds.), *Quantitative Analyses of Behavior: Discrimination Processes* (pp. 3–19). Cambridge, Mass.: Ballinger.

Hintzman, D. L. (1967). Articulatory coding in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 6,* 312–16.

Hunter, W. S., & Sigler, M. (1940). The span of visual discrimination as a function of time and intensity of stimulation. *Journal of Experimental Psychology, 26,* 160–79.

Lawrence, D. H., & Laberge, D. L. (1956). Relationship between accuracy and order of reporting stimulus dimensions. *Journal of Experimental Psychology, 51,* 12–18.

Leeper, R. (1951). Cognitive processes. In S. S. Stevens (Ed.), *Handbook of Experimental Psychology* (pp. 730–57). New York: Wiley.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97.

Neisser, U. (1967). *Cognitive Psychology.* New York: Appleton-Century-Crofts.

Osgood, C. E. (1953). *Method and Theory in Experimental Psychology.* New York: Oxford University Press.

Shannon, C. E., & Weaver, W. (1949). *The Mathematical Theory of Communication.* Urbana: University of Illinois Press.

Sperling, G. (1963). A model for visual memory tasks. *Human Factors, 5,* 19–31.

Sperling, G. (1968). Phonemic model of short-term auditory memory. *Proceedings, 76th Annual Convention of the American Psychological Association, 3,* 63–4.

Sperling, G. (1978). *The Goal of Theory in Experimental Psychology.* Bell Telephone Laboratories Technical Memorandum 78-1221-12.

Sperling, G., Parish, D. H., Pavel, M., & Desaulniers, D. H. (1984). Auditory list recall: Phonemic structure, acoustic confusibility, and familiarity. *Bulletin of the Psychonomic Society, 18,* 36.

Sperling, G., and Speelman, R. G. (1970). Acoustic similarity and auditory short-term memory: Experiments and a model. In D. A. Norman (Ed.), *Models of Human Memory* (pp. 149–202). New York: Academic Press.

Stevens, S. S. (Ed.). (1951). *Handbook of Experimental Psychology.* New York: Wiley.

Wald, A. (1950). *Statistical Decision Functions.* New York: Wiley.

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

Woodworth, R. S., & Schlosberg, H. (1954). *Experimental Psychology* (rev. ed.) (pp. 90–105). New York: Holt.

Wundt, W. *An Introduction to Psychology.* (1912). (R. Pintner, translator, from the second German edition). London: Allen & Unwin. (Reprinted 1924.)