Legal Communication & Rhetoric: JALWD

Fall 2014 / Volume 11

ARTICLES & ESSAYS

The Reader's Limited Capacity: A Working-Memory Theory for Legal Writers

Andrew M. Carter

The Reader's Limited Capacity

A Working-Memory Theory for Legal Writers

Andrew M. Carter*

Introduction

Early in my teaching career, to discourage students from burdening their readers with long complicated sentences, I fashioned a not-so-clever "the reader is a juggler" lecture.[1] It goes something like this: A reader is a juggler, and the ideas in a sentence are balls to be juggled. Most competent jugglers can juggle two and three balls with little effort. Likewise, most legal readers processing a sentence can keep two or three ideas aloft in their minds before the period cues that the sentence has ended and the ideas presented can finally be integrated. But things get risky from there.

Good jugglers can keep four and five balls aloft, but it requires taxing effort, and at some point a ball will surely drop. Good legal readers, in turn, can keep four or five ideas aloft before reaching the end of a sentence. But this too requires taxing effort and eventually the reader will drop an idea before final integration. In sum, long sentences jam-packed with ideas fail because the reader can juggle only so many ideas at once. The risk for the legal writer, the lecture concludes, is twofold. First, the

***** Associate Clinical Professor, Sandra Day O'Connor College of Law, Arizona State University. I thank Professors Jessica Clark and Amy Langenfeld for their labors editing this article. Their guidance was invaluable, and this article is surely better for their efforts. That said, any errors are mine alone. Finally, I owe a debt of gratitude both to the Sandra Day O'Connor College of Law for its generous support of this scholarship and to my colleagues within the Legal Method and Writing Program for their patience and good humor.

1 For learning legal writers, at least, the impulse toward long sentences makes some sense. For many students, law school is the first time they have been asked to express logical proofs in written form. And from the new legal writer's perspective, the easiest way to express a proof is to put all its elements in a single sentence. *See infra* text accompanying notes 21–22. Another, more insidious, culprit might be the "model" cases that students are asked to read early in law school. Consider *Hadley v. Baxendale*, the 1854 case that generations of law students have examined in their first-year contracts class. (1854) 156 Eng. Rep. 145 (Ct. Exch.), 9 Exch. 341. In *Hadley*, the Court of Exchequer Chamber begins its opinion with a sentence that is, amazingly, 384 words long. *Id.*

idea the reader drops might be the idea that matters most. Second, the reader will hate you for making her work so hard.[2]

I have always felt comfortable that the reader-is-a-juggler lecture is well founded. My own experience as a reader of legal prose instructs that processing long, complicated sentences demands conscious "work" and negatively affects my "reading experience." And this was hardly my own unique hang-up as a legal reader. Indeed, nearly every legal writing textbook counsels against the long complicated sentence.[3] And many offer reasoning that tracks the "too many ideas to juggle" lecture.[4] Moreover, eradicating long, bloated sentences is a plank of the influential "plain language" platform.[5] But these admonitions against long, complex sentences leave a central question unanswered: Why? That is, why do long sentences challenge a reader? One can certainly infer that long sentences exhaust some limited cognitive resource that the reader necessarily engages when reading. But what exactly was that limited resource? In section I of this article, I attempt to answer that question by exploring the corner of cognitive science dedicated to "working-memory theory." In its broadest terms, working-memory theory posits that humans process new information through a working-memory system that both stores new information and manipulates it to accomplish mental tasks. In turn, a central tenet of working-memory theory is the "limited capacity assumption," which holds that the working-memory system has a finite (and surprisingly paltry) capacity.

2 I'm not necessarily being flippant. There is research indicating that reading "disfluent" writing cues negative judgments about the writer. *See* Daniel M. Oppenheimer, *The Secret Life of Fluency*, 12 Trends Cognitive Sci. 237, 237–40 (2008); *see generally* Julie A. Baker, *And the Winner Is: How Principles of Cognitive Science Resolve the Plain Language Debate*, 80 UMKC L. Rev. 287, 287–305 (2011) (marshaling Oppenheimer's research to support the Plain Language (or anti-legalese) movement).

3 *See, e.g.* Veda R. Charrow, Myra K. Erhardt & Robert P. Charrow, *Clear & Effective Writing* 163–64 (4th ed. 2007) ("Probably no other single characteristic does more to needlessly complicate legal writing than these long sentences."); Charles R. Calleros, *Legal Method and Writing* 271 (6th ed. 2011) ("[L]egal writing often suffers from unnecessarily long and complex sentences that are poorly constructed."); Richard K. Neumann Jr., *Legal Reasoning and Legal Writing* 229 (6th ed. 2009) (noting that unnecessarily long sentences "blur[] meaning" and should be streamlined); *see also* Terrill Pollman et al., *Examples and Explanations: Legal Writing* 278 (2011) (encouraging legal writers to avoid drafting sentences that exceed four typed lines); Laurel Currie Oates & Anne Enquist, *The Legal Writing Handbook* 523 (5th ed. 2010) ("[L]egal writers can usually write sentences for their readers that average about twenty-two words per sentence with only a rare sentence exceeding a thirty-five-word limit.").

4 *See* Calleros, *supra* n. 3, at 271 ("The readers of such [long] sentences must assimilate too much information before pausing, and they often lose track of the proper relationships of the ideas expressed."); Charrow et al., *supra* n. 3, at 164 (stating that issue with long sentences is that reader can hold only a few ideas at a time).

5 In *Plain English For Lawyers* 33–39 (5th ed. 2005), Richard Wydick dedicates a separate chapter to drafting short sentences as a tonic to the lawyer's tendency "to deliver to the reader in one fat lump all their main themes, supporting reasons, details, qualifications, exceptions, and conclusions." *Id.* at 33; *see also* Joseph Kimble, *Lifting the Fog of Legalese* xi, 96 (2006) (lamenting that legal sentences tend to be "long and flabby" and encouraging writers to use short and medium-length sentences).

In section II, I review the scholarship applying working-memory theory to reading processes. Because reading necessarily consumes the working-memory resource's finite capacity, this scholarship ultimately instructs that the writing lawyer must actively manage the working-memory demands of his or her sentences and paragraphs. Practical applications of this broad insight, however, prove elusive in the absence of a model that can be easily commanded by the legal writer. Accordingly, in section III, I plumb the precepts and vocabulary of "cognitive load theory," a theory of instructional design premised on working-memory theory, to present a practical framework for legal writers to assess the cognitive demands of their sentences and paragraphs. Application of this framework, in turn, presents the legal writer with a number of guidelines for managing the reader's limited working-memory resources.

Before continuing, a broad caveat is necessary. I am a lawyer and clinical professor of law, not a cognitive scientist. Though I do believe my conclusions about the working-memory demands in reading legal prose are supported by mainstream cognitive science, my conceptualizations of the abstract theories at play are surely open to critique.[6] Indeed, there is surely parsimony to the analysis that betrays the complexity of the underlying concepts. But the goal of this article is not to build an airtight scientific proof. Rather, mine is the perspective of the teacher, and what the writing classroom demands is a cognitive model of the reading process that is scientifically grounded *and* useful to practicing legal writers. In that sense, this article aims only to offer what the noted cognitive scientist Daniel Kahneman has described as a "useful fiction"[7]: a model that, while not fully empirically tested, nonetheless helps us describe the complexity of the cognitive processes engaged by the legal reader.

I. Working-Memory Theory

In the spring of 1955, a young psychologist named George Miller presented a paper at a meeting of the Eastern Psychological Association in

6 There is a growing backlash against the proliferation of unfounded "pop" science that claims a grounding in cognitive research. *See* Alissa Quart, *Neuroscience: Under Attack*, N.Y. Times, Nov. 23, 2012 (reviewing backlash against "reductionist, sloppy thinking and our willingness to accept seemingly neuroscientific explanations, for, well, nearly everything"); *see also* Sally Satel & Scott O. Lillienfeld, *Brainwashed: The Seductive Appeal of Mindless Neuroscience* 14–24 (2013) (criticizing facile extrapolations of fMRI studies). Literacy scholarship is not immune to the allure of narrative models of reading based on unfounded applications of cognitive science. *See* Diane H. Tracey & Lesley Mandel Morrow, *Lenses on Reading* 10 (2012) ("The absence of the constraints to produce computational simulations or neural evidence to support hypotheses frees theorists to contemplate the cognitive processing of literacy in an almost unbounded manner, a highly appealing position from which to work.") (quoting Diane H. Tracey, Alex W. Storer & Sohrob Kazerounian, *Cognitive Processing Perspectives on the New Literacies*, in Elizabeth A. Baker ed., *The New Literacies: Multiple Perspectives on Research and Practice* 110 (2010)).

7 Daniel Kahneman, *Thinking, Fast and Slow* 28–30 (2011).

Philadelphia.[8] Initially unsure of how to fill his one-hour time allotment, Miller decided to review for the audience the results of a series of studies that measured the ability of human subjects to remember and process new information.[9] In search of a unifying theme for the talk, Miller observed that in each unique experimental setting, subjects could manage only about seven items of information before performance lagged: subjects could correctly recall only about seven numbers read to them aloud; could accurately categorize only about seven identified tones on a scale; and could accurately count up to only about seven squares blinked on a screen.[10] This apparent convergence, Miller suggested in the end, was evidence that the human mind has finite capacity to "receive, process, and remember" new information.[11]

Miller was not exactly resolute about the measure of the mind's capacity to process new information. He inelegantly titled his subsequent paper "*The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information.*"[12] And Miller struck a playful tone throughout: the paper made a subtle reference to McCarthyism and concluded with references to "the seven wonders of the world" and "the seven deadly sins."[13]

When George Miller died in 2012 at the age of 92, his obituary in the *New York Times* drew a direct line between his inauspicious Philadelphia talk and what we call today the "cognitive revolution," the ever-burgeoning interdisciplinary study of the "thinking processes" of the brain.[14] *The Magical Number Seven's* enduring impact is not rooted in its pop psychology sensibilities (think of the cocktail-party query: why do phone numbers have seven digits?). Rather, by observing that there are limits to the human brain's capacity to process new information, Miller necessarily postulated a model of the mind as an information processor with its own architectural structure.[15] Today, when even laypersons conceptualize the mind as a computer, they harken to the legacy of George Miller.[16]

8 George A. Miller, *The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information*, 63 Psychol. Rev. 81 n.1 (1956).

9 Nelson Cowan et al., *The Legend of the Magical Number Seven,* in Sergio Della Sala ed., *Tall Tales about the Mind and Brain* 45–46 (2007).

10 *Id*. at 45–49; Miller, *supra* n. 8, at 83–90.

11 *See* Miller, *supra* n. 8, at 95.

12 *Id.* at 81.

13 *Id.* at 96.

14 Paul Vitello, *George A. Miller, a Pioneer in Cognitive Psychology, Is Dead at 92*, N.Y. Times, Aug. 2, 2012, at A19. Indeed, *The Magical Number Seven* stands as one of the most oft-cited papers in modern psychology. Nelson Cowan, *Working Memory Capacity* 23 (2005).

15 *See* Vitello, *supra* n. 14, at A19.

16 In 1956 the information-processor model of human cognition was a radical notion. This was especially so within the academy, then controlled by the Behaviorist school of psychology, which insisted that the mind's operations were effectively unknowable. George A. Miller, *The Cognitive Revolution: A Historical Perspective*, 7 Trends Cognitive Sci. 141, 142 (2003). *The Magical Number Seven* decidedly broke the Behaviorist stranglehold; in that sense, Miller's article can be reasonably considered the beachhead of the cognitive revolution. *See* Vitello, *supra* n. 14, at A19.

Once one views the mind as an information processor, the familiar construction of human memory consisting of a short-term and long-term memory proves to be lacking.[17] The problem is that this "multi-store model" usually presents the short-term memory as little more than a passive storage facility.[18] But that construction fails to describe that we do more than just temporarily store new information in our memory system; rather, we also manipulate and interpret that new information in order to complete mental tasks.[19] To reflect this more-dynamic information-processing model, cognitive scientists eventually struck upon the term "working memory."[20]

To better grasp this active processor principle of working-memory theory, it can be helpful to complete a simple mental mathematics exercise. Attempt to solve the addition problem below without the use of a pen and paper; that is, complete the necessary math entirely "in your head":

$$\begin{array}{r} \mathbf{753} \\ \mathbf{+\ 118} \\ \hline \end{array}$$

Engaging their working memory, many folks can solve this problem in their heads using one of the available long-addition strategies. When I first encountered a similar mental mathematics problem, I employed a strategy learned in grade school where, beginning with the right hand column, I

17 The multi-store model has enduring appeal. In the 1890's, William James, the father of American psychology, identified a "primary memory" that could keep conscious very little information at once and a vast "secondary memory" where memories were stored long-term. Cowan, *supra* n. 14, at 13. More recently, Richard Atkinson and Richard Shiffrin followed a multi-store paradigm in their influential information-processing model of human memory. Under the "Atkinson–Shiffrin Model," new information is collected and processed through a series of memory stores. First, a sensory memory operating at the subconscious level collects vast amounts of information from the environment. Next, a short-term memory holds information selected from the sensory memory long enough to utilize it to complete a task. And finally a narrower selection of information reaches a long-term memory store. Jonathan K. Foster, *Memory: A Very Short Introduction* 27–31 (2009).

18 *See* Foster, *supra* n. 17, at 33 ("Short-term memory was previously conceptualized (either explicitly or implicitly) as a relatively *passive* process.") (emphasis in original); Cowan, *supra* n. 14, at 16 ("The term short-term memory also became associated, in time, with a simplified notion of performance in which items that were to be remembered in the short term were placed in a simple location in the mind to be retrieved, intact, later."); Marilyn L. Turner & Randall W. Engle, *Is Working Memory Capacity Task Dependent?*, 28 J. Memory and Lang. 127, 127 (1989) (stating that early theories of memory "viewed short-term memory as a fixed number of slots or bins").

19 *See* Foster, *supra* n. 17, at 33 (describing how information held in short-term memory is subject to mental operations); Marcel Adam Just & Patricia A. Carpenter, *A Capacity Theory of Comprehension: Individual Differences in Working Memory*, 99 Psychol. Rev. 122, 122–23 (1992) ("Most recent conceptions of working memory extend its function beyond storage to encompass the actual computations [or] symbolic manipulations that are at the heart of human thinking . . . .").

20 The term was first used in a 1960 book by George Miller and two colleagues, Eugene Galanter and Karl Pribram. *See* George Miller et al., *Plans and the Structure of Behavior* 65 (1960). Note that a precise, universally accepted definition of "working memory" is elusive. *See* Priti Shah & Akira Miyake, *An Introduction* in Akira Miyake & Priti Shah eds., *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* 1 (1999) (stating that "the term *working memory* is used in quite different senses by different communities of researchers"). Here, I use the term in its broadest sense to describe the mind's "concurrent storage and manipulation of [new] information" in service of completing a mental task. Alan Baddeley, *Working Memory*, 255 Sci. 556, 556 (1992).

completed a series of smaller addition problems moving right to left. Critically, this strategy requires not only that my mind complete three separate sub-problems (the series of simpler addition problems) but also that my mind simultaneously recall the digit information representing the solution to the preceding sub-problems. That is, to solve the problem, my mind (read: my working memory) needs to both *store and actively process* the new information. Thus, the definition of working memory as a "short-duration system in which small amounts of information are simultaneously stored and manipulated in the service of accomplishing a task."[21]

Once you ponder upon it, you can become a bit breathless in describing the role that the working-memory system plays in human cognition. It is our working-memory system that provides temporary storage and manipulation of new information necessary for such fundamental cognitive tasks as language comprehension, learning, and reasoning.[22] Indeed, if "human thought processes are underpinned by an integrated system for temporarily storing and manipulating information,"[23] then it can be reasonably posited that our working memory is the gateway to knowledge.[24] But if working memory serves as a gateway to knowledge, it is a pretty small gateway.

A foundational precept of working-memory theory is that working memory is a limited capacity resource.[25] As Miller first deduced in *The Magical Number Seven*, the mind can effectively process only so much new information.[26] Indeed, our ability to hold and process new information can strike one as surprisingly paltry. To understand this critical notion of the limited capacity of working memory, it is helpful to attempt a second mental mathematics problem. Again, try to solve the following long addition problem entirely in your head:

21 David Caplan & Gloria S. Waters, *Verbal Working Memory and Sentence Comprehension*, 22 Behavioral & Brain Sci. 77, 77 (1999).

22 *See* Baddeley, *supra* n. 20, at 559 (stating that studies "have linked working memory to performance on a range of important tasks, including language comprehension and reasoning").

23 Alan Baddeley, *Working Memory: Looking Back and Looking Forward*, 4 Nat. Revs./Neurosci. 829, 837 (2003).

24 *See* Baddeley, *supra* n. 20, at 559 ("Working memory stands at the crossroads between memory, attention, and perception.").

25 *See* Shah & Miyake, *supra* n. 20, at 10 (describing "wide consensus" of the existence of capacity limits in working memory, despite a lack of consensus on "underlying mechanisms responsible for the limitations"); Baddeley, *supra* n. 23, at 829 ("The theoretical concept of working memory assumes that a limited capacity system, which temporarily maintains and stores information, supports human thought processes.").

26 *See* Miller, *supra* n. 8, at 96; Nelson Cowan, *Visual and Auditory Working Memory Capacity*, 2 Trends Cognitive Sci. 77, 77 (1998) ("It has been clear that the capacity of working memory is limited ever since George Miller described various research studies suggesting that people can recall at most about seven independent, meaningful items or 'chunks' at a time.").

89789
+ 42112

There are surely some readers who were able to solve this problem in their heads, even a few, perhaps, who did so with little conscious difficulty.[27] But when I encountered a similar exercise, I quit about half-way through the problem. The problem demanded that my working memory complete too many sub-problems while simultaneously retaining recall of too many digits. In sum, the exercise's working-memory demand—"the number of procedures required to reach a sub-goal and the number of elements maintained in each procedure"[28]—exceeded my working-memory capacity. And so I quit on it. Just as a reader would quit an equally demanding sentence of legal prose.

Much of the modern working-memory scholarship builds upon a model of the functional architecture of the working-memory system proposed in 1986 by Alan Baddeley and Graham Hitch.[29] The Baddeley–Hitch model identifies three separate functional components of the working-memory system: two memory stores and a so-called "central executive."[30]

The first memory store, the "phonological loop" (also referred to as the articulatory or verbal–acoustic loop) is charged with storing new speech-based information; i.e., spoken and written language.[31] Under the Baddeley–Hitch model, the phonological loop itself has two subcomponents: a phonological store for new speech-based information and a "loop" where new articulatory information is "rehearsed" to avoid decay (i.e., forgetting).[32] If, while solving the mental math problems above, you remembered certain digit information by repeating (rehearsing) the digits in your head, you employed this phonological loop.[33] The second memory store, the descriptively named "visuospatial sketchpad" is where visual information is temporarily stored.[34] As one might expect, a person's

27 For those who finished the problem and want to confirm their solution, the sum of the two numbers is 131,901.

28 Caplan & Waters, *supra* n. 21, at 78.

29 Cowan, *supra* n. 14, at 21 ("[T]he working memory model of Baddeley and colleagues has stood as the industry standard for many years, and for a reason. It is easy to grasp and serves to explain many important phenomena.").

30 *See generally* Baddeley, *supra* n. 23, at 830–37; Baddeley, *supra* n. 20, at 556–57.

31 Baddeley, *supra* n. 20, at 558–59; Foster, *supra* n. 17, at 35.

32 Baddeley, *supra* n. 23, at 830.

33 *See generally* Alan Baddeley, *Working Memory and Language: An Overview*, 36 J. Commun. Disorders 189, 191 (2003) (reviewing subvocal rehearsal system).

34 Baddeley, *supra* n. 23, at 833–35; Foster, *supra* n. 17, at 36.

capacity to hold and manipulate visuospatial representations "predicts success in fields such as architecture and engineering."[35]

The final component of the Baddeley–Hitch model is "the central executive," which controls the focus of the working-memory resource and coordinates the information stored in the phonological loop and the visuospatial sketchpad to complete various sub-goals.[36] In this sense, "[t]he central executive is the workhorse and mastermind of human cognition. It allocates attention to a task and performs informational storage and computational functions within a given task."[37] Returning to the mental math problems above, you may have engaged the phonological loop and visuospatial sketchpad subsystems to store the information to solve the problems. But it was the central executive that focused attention upon and manipulated the stored information in order to complete the sub-problems necessary to solve the exercise.

Although the Baddeley–Hitch model enjoys broad acceptance, things get tricky when scholars attempt to ground the limited-capacity assumption within its framework. When completing the mental math problems above, one certainly feels as though the storage demands of the exercise (recalling the necessary digit information) and the processing demands of the exercise (completing the necessary sub-problems) competed for some unitary working-memory resource. That notion finds purchase in the research of Deborah Daneman, which suggests that the storage functions of the phonological and visuospatial stores compete with the processing function of the central executive for some finite pool of working-memory "energy."[38] But other scholars dispute that the working-memory resource is, in the parlance, "domain general" (i.e., shared by the storage and processing components); instead, these scholars argue that the working-memory resource is domain specific (i.e., each component has its own dedicated "energy" resources).[39]

35 Baddeley, *supra* n. 23, at 834. A chess player processing available moves (and their consequences by way of the opponent's likely moves in response) is thought to rely on the visuospatial sketchpad. *See* Foster, *supra* n. 17, at 36. Another example of this memory store at work might be when you stop at an intersection and look right to see an oncoming car. When you then look to your left, you temporarily retain the visual information about the oncoming car on the right in your visuospatial sketchpad. This allows you to consider all the relevant visual information before making a subsequent driving decision.

36 *See* Baddeley, *supra* n. 20, at 557.

37 Caplan & Waters, *supra* n. 21, at 77; *see also* Baddeley, *supra* n. 23, at 835 (suggesting the central executive be conceptualized as "a homunculus, the little man taking all the important decisions").

38 Meredyth Daneman & Patricia A. Carpenter, *Individual Differences in Working Memory and Reading*, 19 J. Verbal Learning & Verbal Behav. 450, 451 (1980); *see also* James W. Montgomery, *Verbal Working Memory and Sentence Comprehension in Children with Specific Language Impairment*, 43 J. Speech, Lang. & Hearing Research 293, 294 (2000) ("The Daneman model is a computational model in which both storage and processing functions must share a limited pool of resources during comprehension."); Just & Carpenter, *supra* n. 19, at 134.

39 *See generally* Klaus Oberauer, *Access to Information in Working Memory: Exploring the Focus of Attention*, 28 J. Experimental Psychol. 411, 411 (2002) (reviewing research that "casts doubt on the resource-sharing hypothesis"); Cowan, Fortunately, we do not have to take sides in that debate. Rather, a narrower and broadly accepted explanation is available for one's sense that a singular shared working-memory resource is consumed during complex mental tasks: The working memory's true finite resource is the central executive's focus of attention.[40] Roughly, under this view, the central executive has a finite span of attention that can be focused to process and integrate information actively stored in the subsystem stores.[41] Thus, I failed in my attempt to solve the second math problem above because the exercise exhausted my central executive's limited focus (or span) of attention.

II. Working-Memory Theory and Reading Processes

Psycholinguistics—i.e., the study of how we acquire, use, and comprehend language, including written language—has long been a foundational discipline within cognitive science. And the role of working memory in reading tasks generally, and sentence processing specifically, has been well researched.[42] There is, therefore, a bounty of research establishing that reading written text engages the working memory, specifically the memory store of the phonological loop and the limited attention resource of the central executive.[43] All of which provides certain purchase to the-reader-is-a-juggler lecture: long, jam-packed sentences fail because they exhaust the reader's working-memory capacity. The number of ideas in the sentence (and their interactions) overwhelm the central executive's limited focus of attention. But working-memory theory can explain more than *why* the reader labors under long, jam-packed sentences. It also illus-

supra n. 14, at 52–66 (reviewing studies inconsistent with conclusion that storage and processing functions draw from a shared resource); *see also* Laurie Stowe et al., *Encoding and Storage in Working Memory during Sentence Comprehension* in Paolo Merlo & Suzanne Stevenson eds., *The Lexical Basis of Sentence Processing: Formal, Computational and Experimental Issues* 181, 202 (2002) (reviewing data from neuroimaging studies which suggest "disassociation of processing and storage" and arguing that processing and storage "must be distinguished from each other as separate functions").

40 *See* Cowan, *supra* n. 14, at 7–10 (reviewing value of "capacity of the focus of attention" as a measure of working-memory capacity); Randall W. Engle, *Working Memory Capacity as Executive Attention*, 11 Current Directions Psychol. Sci. 19, 20 (2002); Oberauer, *supra* n. 39, at 420–21 (postulating that the central executive within the Baddeley model has a separate and limited focus of attention resource).

41 *See* Cowan, *supra* n. 14, at 7–9.

42 *See infra* n. 43.

43 *See, e.g.* Daneman & Carpenter, *supra* n. 38, at 450–66 (studying the role of working memory in reading comprehension); Caplan & Waters, *supra* n. 21, at 78–79 (reviewing evidence "indicating that comprehension of sentences requires verbal working memory resources"); *see generally* Evelina Fedorenko, Edward Gibson & Douglas Rhode, *The Nature of Working Memory Capacity in Sentence Comprehension: Evidence Against Domain-Specific Working Memory Resources*, 54 J. Memory & Lang. 541, 541–53 (2006) (investigating consumption of working-memory resources during sentence comprehension).

trates the consequences for the writer when a reader must invest substantial executive attention to simply decipher a sentence.

To assess working-memory demands in reading, it is helpful to bifurcate sentence processing into two stages: an interpretive stage and a post-interpretive stage.[44] In the interpretive stage, the reader dedicates working-memory resources to discerning the "meaning" of a sentence.[45] Toward that end, the working memory must decode linguistic signals; letters must be decoded into phonemes, phonemes into words, words into a syntactic structure, and all of this, finally, into a full meaning (or a semantic whole).[46]

In the post-interpretive stage, the working memory must first store the full "meaning" of the sentence obtained during the interpretive stage. Next it must mediate an interaction with the long-term memory where the extracted meaning is used "for storing information in long-term semantic memory, reasoning, planning actions, and other functions."[47] Unfortunately, more-precise conceptualizations of the metacognitive tasks completed during post-interpretive stage are elusive. It is telling, for instance, that in his early influential information-processing model of reading, the reading scholar Philip Gough was relegated to describing the post-interpretive stage as "The Place Where Sentences Go When They Are Understood."[48]

The interpretive and post-interpretive stages of sentence processing draw from the same limited-attention resource of the central executive.[49] They are, therefore, in a trading relationship: executive attention dedicated to decoding the meaning of a sentence in the interpretive stage is unavailable for the higher-end cognitive tasks to be completed in the post-interpretive stage.[50] This means, of course, that if a reader expends an

44 This approach is associated with David Caplan and Gloria S. Waters. *See* Caplan & Waters, *supra* n. 21, at 78. Other researchers employ a similar bifurcation of the sentence processes but prefer different labels. *See e.g.* Fedorenko et al., *supra* n. 43, at 542 (bifurcating sentence processing into "on-line" sentence processing and subsequent "non-linguistic verbally-mediated tasks").

45 Caplan & Waters, *supra* n. 21, at 78.

46 *See id.* (describing interpretive stage as "the processes of recognizing words and appreciating their meanings and syntactic features; constructing syntactic and prosodic representations; and assigning thematic roles, focus, and other aspects of propositional and discourse-level semantics").

47 *Id.*

48 Philip B. Gough, *One Second of Reading*, in James F. Kavanaugh & Ignatius G. Mattingly eds., *Language by Ear and by Eye* 340 (1972).

49 *See* Fedorenko et al., *supra* n. 43, at 542–43 (reviewing evidence favoring the view that sentence processing draws upon a singular working-memory resource); Just & Carpenter, *supra* n. 19, at 141, 145 (offering the theory that the central executive has limited "available activation" during language comprehension); Daneman & Carpenter, *supra* n. 38, at 450–51 (suggesting that sentence processing draws from single "shared[,] limited capacity" resource). *But see* Caplan & Waters, *supra* n. 21, at 92–93 (arguing in favor of a domain-specific resource for the interpretive stage that is unrelated to the post-interpretive stage).

50 Daneman & Carpenter, *supra* n. 38, at 450–51.

inordinate amount of working-memory energy making meaning out of a sentence, metacognition will lag because there are insufficient working-memory resources for the critical post-interpretive processes.[51]

And this brings us back to the reader-is-a-juggler lecture. Long, jam-packed sentences are a pox upon the page not only because they make the reader expend working-memory resources to simply extract the intended meaning of the sentence; the real danger of long complicated sentences is that they leave the reader with few, if any, executive-attention resources for the post-interpretive stage where critical reasoning takes place. In sum, by making a reader endure long arduous sentences, the legal writer assures that the reader will not have the working-memory resources to complete the very reasoning that the sentences are presumably designed to foster.

III. Toward a Practical Framework: Cognitive-load theory

The reader-is-a-juggler lecture, then, finds support in cognitive science. But is the underlying working-memory theory of any real use to legal writers? Surely, the classroom lecture that references "the central executive's limited focus of attention" is an invitation for students to check their e-mail. What is required is a practical framework that can guide legal writers toward drafting sentences that effectively manage the reader's working-memory resources. That brings us to cognitive load theory.

A. Cognitive-Load Theory

In the 1980s, John Sweller and his colleagues in the instructional design department at the University of New South Wales began studying the dynamic between student learning and the working-memory demands of instructional models.[52] Observing that learning failed when a learning exercise exceeded a student's working-memory capacity, the scholars devised "cognitive[-] load theory," a framework to describe and assess the working-memory demands of a particular mental task.[53] Under cognitive-load theory (or CLT), a mental task's "total cognitive load"—the aggregate of working-memory resources required to successfully complete a mental

51 Just & Carpenter, *supra* n. 19, at 123–24 (offering a theory of language comprehension and observing that when "task demands are high (either because of storage or computational needs), then processing will slow down and some partial results may be forgotten. . . . When the task demands exceed the available resources, both storage and computational functions are degraded.").

52 Fred Paas, Alexander Renkil & John Sweller, *Cognitive Load Theory and Instructional Design: Recent Developments*, 38 Educ. Psychol. 1, 1–4 (2003); Stephen N. Elliott et al., *Cognitive Load Theory: Instruction-Based Research with Applications for Designing Tests* 3 (Feb. 24, 2009) (unpublished manuscript).

53 Paas et al., *supra* n. 52, at 1; Elliott et al., *supra* n. 52, at 2–3.

task—is made up of two subcomponents, the "intrinsic load" and the "extraneous load."[54]

The intrinsic load represents the inherent (and inescapable) working memory demands of a mental task.[55] A core CLT doctrine is that intrinsic load does not wholly turn on the number of bits of information that the working memory must hold. Rather, the intrinsic load is also heavily weighted by the interactions between and among the information bits that the working memory must process.[56] This "elemental interactivity"—"the extent to which multiple content components must be processed simultaneously in working memory to allow for problem solving"[57]—is a core draw upon working-memory resources.[58] CLT would explain the challenge of the mental math problems above as a function of the intrinsic load being freighted with elemental interactivity. The challenge of the exercises was not so much storing the specific digit information in one's working memory; rather, it was the simultaneous processing of the relationships between and among the digit information that exhausted the working-memory resource.[59]

Extraneous load, in turn, is any component of the instructional setting that consumes working-memory resources with no attendant positive impact on learning.[60] Cognitive-load theory offers a number of well-known examples of extraneous load in the instructional setting. One example is the "redundancy effect," which underlies the familiar admonition against reading aloud text presented in a PowerPoint slide.[61] The redundancy effect explains that the dual presentation of the same information in two separate instructional modalities (in written text and in spoken word) actually manages to result in less learning than if one chose a singular modality of presentation.[62] This is because "one source will be redundant and having to process both will lead to an extraneous cognitive load."[63]

54 John Sweller, Paul Ayres & Slava Kalyuga, *Cognitive Load Theory* 57–58 (2011). Though not included in the classic model, some expressions of cognitive-load theory include a third component of total cognitive load labeled the "germane cognitive load." *See id.* at 57 (discussing the term "germane cognitive load" and suggesting the concept is better understood as a component of intrinsic load). Roughly, in these models, the germane load equals the working-memory resource dedicated to mediating interactions with the long-term memory and the formation of schema. *See generally* John Sweller, Jeroen J.G. van Merriënboer & Fred G.W.C. Paas, *Cognitive Architecture and Instructional Design*, 10 Educ. Psychol. Rev. 251, 259 (1998). Though beyond the scope of the article, application of "germane load" scholarship to legal-drafting techniques, especially those based on organizational principles, warrants further examination.

55 Sweller et al., *supra* n. 54, at 57–58.

56 *Id.* at 58–60.

57 Elliott et al., *supra* n. 52, at 2–3.

58 *See* Paas et al., *supra* n. 52, at 3.

59 Sweller et al., *supra* n. 54, at 59 ("The elements interact and so they must be processed simultaneously rather than as single, unrelated elements because they cannot be understood as single elements. Such material is high in element interactivity and high in intrinsic cognitive load.").

60 *Id.* at 57.

61 *See id.* at 144–45.

62 *Id.*

63 *Id.* at 144.

Eliminating the extraneous cognitive load of an instructional task makes good sense, of course. But what if after the instructor eliminates extraneous load, the intrinsic load of instructional material still threatens to exceed a student's working-memory capacity? Then one must "chunk" the material. The properly descriptive term "chunking" was coined by George Miller;[64] in *The Magical Number Seven*, Miller observed that in order to remember digits in excess of seven, we necessarily "chunk" the numbers into smaller, more-memorable components.[65] (Thus, we do not remember our social-security number as a series of nine digits; instead we chunk the numbers into three shorter, more recallable, sets of numbers). Cognitive-load theory instructs that learning tasks with excessive intrinsic load should be segmented into component "chunks," each of which can be efficiently processed into long-term memory schemas by the limited-capacity working memory.[66]

Finally, under CLT, the extraneous and intrinsic demands on working memory are additive; the difficulty of an instructional task—its total cognitive load—is measured by the combined weight of the intrinsic and extraneous loads.[67] In the end, CLT offers a broad imperative to instructional designers: Learning tasks must be designed so that their intrinsic and extraneous loads do not exceed the student's working-memory capacity.[68] Stated another way, instructional designs that manage the intrinsic load and eliminate extraneous load will best promote learning through wise management of the learner's working-memory resources.[69]

B. A Cognitive-Load Theory of Legal Writing

As a theory of instructional design, cognitive-load theory has obvious application to the law-school classroom.[70] But here I propose a novel

64 *See* Miller, *supra* n. 8, at 92–93.

65 *Id.*

66 Fred Paas et al., *Cognitive Load Measurement as a Means to Advance Cognitive Load Theory*, 38 Educ. Psychol. 63, 63–64 (2003) (reviewing role of chunking in schema formation); *see also* Glenda Andrews & Graeme S. Halford, *Recent Advances in Relational Complexity Theory and Its Application to Cognitive Developments* in *Cognitive Development and Working Memory: A Dialogue Between Neo-Piagetian Theories and Cognitive Approaches* 50 (Pierre Borrouillet & Vinciane Gaillard eds. 2011) ("Complexity is reduced when tasks are decomposed or when variables are chunked . . . .").

67 Sweller et al., *supra* n. 54, at 58.

68 *Id.* at 59.

69 *Id.;* Paas et al., *supra* n. 52, at 3–4.

70 A number of scholars have specifically discussed how CLT might influence modes of legal education. *E.g.* Larry O. Natt Gantt, *The Pedagogy of Problem Solving: Applying Cognitive Science to Teaching Legal Problem Solving*, 45 Creighton L. Rev. 699, 738–48 (2012); Hillary Burgess, *Deepening the Discourse Using the Legal Mind's Eye: Lessons from Neuroscience and Psychology that Optimize Law School Learning*, 29 Quinnipiac L. Rev. 1, 25–29 (2011); Stefan H. Krieger & Serge A. Martinez, *A Tale of Election Day 2088: Teaching Storytelling Through Repeated Experiences*, 16 Leg. Writing 117, 125–30 (2010); Deborah J. Merritt, *Legal Education in the Age of Cognitive Science and Advanced Classroom Technology*, 14 B.U. J. Sci. & Tech. L. 39, 44–48 (2008); Stefan H. Krieger, *Domain Knowledge and the Teaching of Creative Legal Problem Solving*, 11 Clinical L. Rev. 149, 164–65, 184–85 (2004). This is scholarship that specifically references CLT. As for scholarship that discusses lessons for the law-school classroom drawn from other corners of the exceedingly broad field of cognitive science, the citations are too numerous for inclusion here.

application where a CLT-*like* framework is used by the legal writer to manage the reader's working memory loads in order to maximize learning and reasoning. Application of CLT to legal writing finds certain support in the reading and literacy scholarship; a number of scholars submit that management of a reader's working-memory resources (generally through automated reading processes) underlies full reading comprehension.[71] Moreover, applying CLT within the writer-reader context makes intuitive sense. CLT is intended to inform instructional settings where instructional designers (teachers) seek to maximize learning by the audience (students).[72] In a very real sense, legal writing is an instructional setting; the legal writer is a teacher who aims to maximize learning by the reader. Legal writers, it follows, should care deeply about how their instructional design—the structure of their sentences and paragraphs—manages their readers' working-memory resources.

In applying cognitive-load theory to the legal writing process, two guiding principles emerge that firmly ground the concision and economy lesson of my reader-is-a-juggler lecture. First, the legal writer must reduce extraneous load in the interpretive stage by drafting sentences and paragraphs that can be processed automatically, i.e., without the need of dedicated working-memory resources. This allows full executive-attention resources to focus on completing the critical metacognitive functions of the post-interpretive stage. Second, the legal writer must manage the intrinsic load for the reader by chunking sentences and paragraphs that, after elimination of extraneous load, still threaten to exceed the reader's working-memory capacity.

1. Reduce the extraneous load by promoting automaticity.

It makes sense here to pause and consider the mind-bending complexity of the interpretive stage of the reading process. Consider the sentence, "The defendant failed the test." To make meaning out of the sentence, the reader must decode a long series of linguistic signals. Let's begin with the letter "d" in "defendant." The shape of the letter (its orthographic structure) communicates what letter it is; the right-sided vertical line differentiating the "d" from an "o." We know that "d" has a unique phoneme (the "d" sound). By combining the "d" phoneme with the phonemes of other letters and letter couplings (the phonological structure), we can identify the word "defendant." Next, by looking at the

---

71 *See generally* Jeffrey J. Walczyk, *The Interplay Between Automatic and Control Processes in Reading,* 35 Reading Research Q. 554, 554–66 (2000) (presenting selection of reading comprehension theories that explore dynamic between "automatic processes," which generally do not consume working-memory resources, and "control processes," which do consume working-memory resources).

72 *See generally* Paas et al., *supra* n. 52, at 1.

order of the words (the syntactic structure), we can discern that "defendant" is the subject of the sentence, that "failed" is the verb, and that "test" is the verb's object. Finally, we process all of this into the meaning (the semantic structure), and the sentence ultimately communicates new information to the reader, i.e., that the defendant failed the test.[73] At this level of magnification, as others have noted, reading a single sentence matches in complexity "conducting a symphony" or performing brain surgery.[74]

This does raise a question. Considering the number and complexity of the tasks that must be completed simply to decode the linguistic signals, how is it that the interpretive stage does not entirely exhaust the limited executive-attention resource of working memory? Or stated another way, how is it that a reader ever has working-memory resources left to complete the metacognitive tasks of the post-interpretive stage? It was effectively this same question that puzzled researchers David LaBerge and S. Jay Samuel in the 1970s. Laberge and Samuel assessed the demands of the interpretive stage of reading and observed, "If each component process requires attention, performance of the complex skill will be impossible, because the capacity of attention will be exceeded."[75]

But, of course, readers perform the complex skill of reading all the time without exceeding the capacity of executive attention. Laberge and Samuels famously explained that reading does not exceed the reader's attention capacity because most of the interpretive tasks involve automatic processing—i.e., the processing can be completed "while attention is directed elsewhere."[76] Thus, with sentence processing, "if enough of the components and their coordinations can be processed automatically, then the load on attention will be within tolerable limits and the skill can be successfully performed."[77]

Recall the sentence I discussed earlier: "The defendant failed the test." Most readers will expend few, if any, working-memory resources to decode the linguistic cues necessary to extract the meaning of that sentence. Rather, through learning, repetition and practice, skilled readers can decode the letters into phonemes, phonemes into words, and words and syntactic structure into meaning without conscious effort and with no investment of working-memory resources. Reading the sentence is

73 *See generally* Maryanne Wolf, *Proust and the Squid: The Story and Science of the Reading Brain* 81 (2007). For a fascinating general discussion of how those learning to read move from letter decoding to comprehension, see pages 90–133.

74 *See* Elizabeth S. Norton & Maryanne Wolf, *Rapid Automatized Naming (RAN) and Reading Fluency: Implications for Understanding and Treatment of Reading Disabilities*, 63 Annual Rev. Psychol. 427, 428 (2012).

75 David LaBerge & S. Jay Samuels, *Toward a Theory of Automatic Information Processing in Reading*, 6 Cognitive Psychol. 293, 293 (1974).

76 *Id.* at 295.

77 *Id.* at 293.

automated; in the vocabulary of working-memory theory, it can be completed without making demands on the central executive's limited focus of attention.[78]

In the CLT framework, extraneous load is comprised of any facet of the learning task that consumes working memory without a positive impact on learning. In legal writing, I propose, the extraneous load of a sentence is largely represented by disruptions in automaticity during the interpretive stage. These disruptions require that the reader dedicate working-memory resources to make meaning out of a sentence without any positive impact on learning. Indeed, sentences that undermine automatic sentence processing will often *preclude* the reader from grasping the sentence's intended lesson. As discussed above, if a reader expends an inordinate amount of executive attention making meaning out of a sentence during the interpretive stage, there will be insufficient working-memory resources for the critical post-interpretive processes where learning and reasoning take place.[79]

It follows that the skilled legal writer will eliminate extraneous load by drafting sentences that promote automatic reading and conserve working-memory resources for the post-interpretive stage. As anticipated by the reader-is-a-juggler lecture, this means that the legal writer must avoid long sentences in which the sheer number of ideas to be processed disrupts automaticity.[80] But it also means that the legal writer must zealously edit for the more-mundane disrupters of automaticity—missing punctuation, grammar oversights, awkward syntax, and semantic inconsistencies.[81]

Finally, legal writing often suffers from its own unique form of extraneous load, namely, legalese—the ornate syntax, obscure Latin phrases, and chummy argot that despite our best efforts, still infects legal

78 *See* Walczyk, *supra* n. 71, at 555 (explaining that automatic processes in reading "occur effortlessly . . . and make minimal demands on attention and working [memory]"); Norton & Wolf, *supra* n. 74, at 432 (reviewing the proposition that "successful reading depends on not only accuracy but automaticity of multiple cognitive and linguistic processes, requiring minimal conscious effort").

79 Norton & Wolf, *supra* n. 74, at 429 ("The development of automaticity at all the lower levels of reading represents the great apex of development that provides us with the bridge to true reading with its capacity to direct cognitive resources to the deepest levels of thought and comprehension.").

80 Just and Carpenter explain this phenomenon with some precision: "[I]f the total amount of activation that is available to the system is less than the amount required to perform a comprehension task, then some of the activation that is maintaining old elements will be deallocated, producing a kind of forgetting by displacement. Thus representations constructed early in a sentence may be forgotten by the time they are needed later on in the sentence, if enough computational activity has intervened." Just & Carpenter, *supra* n. 19, at 123.

81 A good example of a sentence that consumes high levels of working memory in the interpretive stage is the "garden path" sentence, where syntactic structure invites the reader to parse the sentence inaccurately. Consider this common example: "The horse raced past the barn fell." Because of the imprecise syntactic structure, the reader must focus considerable executive attention to extract the proper meaning of the sentence.

writing.[82] Legalese disrupts automated reading, demanding that readers dedicate substantial working-memory resources during the interpretive stage just to extract the writer's intended meaning. This will, of course, leave less executive attention that can be focused on the higher comprehension tasks of the post-interpretive stage.[83] There is, then, a firm cognitive basis for the Plain Language Movement. The effective goal of the Plain Language Movement is to eradicate words, phrases, and syntactic structures that disrupt automated decoding during the interpretive stage of reading.[84] It is, in the end, a movement to reduce the extraneous cognitive loads imposed on legal readers.

2. Manage the intrinsic load by chunking sentences and paragraphs.

Consider the inherent demands on working memory presented by the following sentence from a memorandum reviewing a Department of Justice employment issue:

> If the Justice Department denies an employee a security clearance; that is, if the Department determines that the employee is not eligible for access to classified information—or if the Department revokes such eligibility, the Department must provide the employee with a comprehensive and detailed written explanation of the basis of the decision, to the extent that the national security interests of the United States and other applicable law permit.

To be sure, the sheer number of ideas in the sentence produces a high intrinsic cognitive load. But what more heavily weights the intrinsic load is the sentence's "elemental interactivity," the interactions among the ideas that the reader must manage. A logical proof of the sentence is "if A or if B, then C, but not if D." The sentence imposes on the reader's working memory not just the chore of temporarily storing each of the ideas in the sentence but also the chore of tracking the interactions (the stated relationships) between and among the ideas.[85]

Legal analysis, and concomitantly legal writing, trades in building relationships among ideas. Thus, when drafting, the legal writer is constantly faced with high elemental interactivity. Indeed, even where the extraneous load is low, because of elemental interactivity, the intrinsic load

82 *See* Wydick, *supra* n. 5, at 4 (noting that although the legal profession has made progress in adopting plain language style, "the victory is not yet won").

83 *See supra* n. 78–79.

84 For a separate argument of how cognitive science supports the Plain Language Movement, see Baker, *supra* n. 2.

85 *See generally* Jeanne S. Chall, *Varying Approaches to Readability Measurement,* 25 Revue Québécoise de Linguistique 23, 28 (1996) (reviewing work by Kintsch & Vipond suggesting that "propositional density of a text and the number of new concepts per proposition" heavily impacts readability).

of a sentence by itself might exceed the reader's working-memory capacity. It is, therefore, incumbent on the legal writer to ensure that the number of ideas in a sentence and the number of interactions between and among those ideas by itself does not exceed the reader's working-memory capacity. But this just prompts the question, What is the legal reader's working-memory capacity? That is, how many ideas and interactions can the skilled legal reader handle in one sentence?

George Miller, of course, suggested the magic number was around seven, that humans could generally store about seven "bits" of information in working memory before the information began to decay.[86] Others argue that working-memory capacity is smaller, at four bits of information in healthy young adults (working-memory capacity appears to decline with age).[87] And still others hold that the entire enterprise of expressing working-memory capacity in the form of a number is a fool's errand.[88] But whatever their validity, numerical measurements of working memory capacity are not particularly helpful to the legal writer.

The problem is that numeric measures of working-memory capacity focus largely on storage capacity; they generally elide the processing challenges posed by elemental interactivity.[89] Thus, the number four (or seven) might be helpful to establish an upper limit for the *number* of ideas in a sentence that a reader can reasonably store in working memory. But this number does not capture whether the working memory can also *process* interactions among those ideas. Indeed, with elemental interactivity between and among the ideas, four ideas in a sentence will often prove too much for the reader's working memory.

Recall the sentence offered above reviewing a Department of Justice employment policy regarding security clearances. The sentence presents four ideas and the three interactions among those ideas. When I first encountered the sentence, it certainly pushed the limits of my working-memory capacity. (As a reader, a good sign that you are reaching the capacity of executive attention is when automaticity is disrupted, and you become consciously aware of the mental effort involved in processing the sentence.)[90] Presuming that I have the working-memory capacity of the average legal reader, I propose this general guideline: In legal writing, a

86 Miller, *supra* n. 8, at 96.

87 *See* Cowan, *supra* n. 14, at 110–37. Cowan argues that study results suggesting broader capacity, like the seven items identified by Miller, are a consequence of mental strategies such as rehearsal and chunking. *Id*.

88 *See* Klaus Oberauer, *In Search of the Magic Number*, 54 Experimental Psychol. 245, 245 (2007) (describing "mostly silent consensus among researchers" that precisely delineating working-memory capacity is impossible because of extreme variability of human subjects' performance of cognitive tasks).

89 *See* Cowan, *supra* n. 14, at 149 (noting there is reason to doubt "whether the idea of holding a few separate chunks is really adequate to understand the range of processing situations").

sentence with four ideas and three interactions marks the upper limit of what is reasonable to impose upon the reader's working-memory system (and even then such sentences must be employed only sparingly). Sentences beyond that limit will disrupt automaticity and hamstring the reader's completion of the higher metacognitive tasks associated with the post-interpretive stage.

What is the remedy if the intrinsic load of a sentence approaches or exceeds the reader's anticipated working-memory capacity? In the CLT framework, the remedy is to "chunk" the ideas in the sentence into smaller groupings of ideas that the working memory can effectively process. In application, this means breaking long sentences down into shorter sentences that the reader's working memory can efficiently process.[91] Note, however, that for new legal writers, chunking ideas into shorter sentences can be challenging because students often do not command techniques for cuing the reader to relationships that exist across sentences. It makes good sense, therefore, to combine instruction regarding shorter sentences with instruction on the use of generic and substantive transitions. For instance, using generic and substantive transitions, our sample sentence from the Department of Justice memorandum can be chunked as follows:

> The Justice Department may take an adverse action regarding an employee's security clearance in two ways. First, the Department may determine that the employee is not eligible for access to classified information in the first place. Second, the Department may revoke a previously granted eligibility. In either circumstance, the Department must provide the employee with a comprehensive and detailed written explanation of the basis of the decision. This requirement, however, may be limited by the national security interests of the United States and other applicable law.[92]

The "chunked" rewrite, of course, breaks down what was once a single sentence pressing the limits of the reader's working-memory capacity into

90 *See* Fred G.W.C. Paas & Jeroen J.G. van MerriënBoer, *The Efficiency of Instructional Conditions: An Approach to Combine Mental Effort and Performance Measures,* 35 Human Factors 737, 738 (1993) ("The intensity of mental effort can be considered as an index of mental workload."). *Cf.* Oppenheimer, *supra* n. 2, at 237 (describing "cognitive fluency" as "the subjective experience of ease or difficulty with which we are able to process information").

91 This is, of course, exactly what the legal writing textbooks instruct. *See, e.g.* Charrow et al., *supra* n. 4, at 165 (instructing legal writers to break long passages into shorter sentences); Neumann, *supra* n. 3, at 229 (instructing legal writers to break long sentences "in two"). Naturally, there are limits to the chunking principle. A legal writer can "over-chunk" by breaking sentences into short sentences with such low intrinsic loads that working-memory capacity is underutilized, rendering the reading process inefficient (what the reader might experience as "skimpy" writing). Thus, working-memory theory and CLT support an ancillary principle that short sentences with just two ideas or less and no elemental interactivity should not be stacked upon one another.

a set of five sentences, each with an intrinsic load well within the average legal reader's capacity. And because of the generic and substantive transitions, the essential relationships between and among the ideas in the original sentence remain clearly delineated. To be sure, the overall presentation is slightly longer, but concision is not a virtue when it translates to more, not less, work for the reader.

IV. Conclusion

As with any relationship, there are mutual obligations that underlie the relationship between legal writers and their readers. The reader carries three obligations: (1) to actually read the legal writer's work, (2) to make a good-faith effort to comprehend its sentences and paragraphs, and (3) to consider the embedded ideas. For all this, the legal writer must do only one thing—reasonably manage the reader's workload.

The legal reader is a working reader. The judge does not pick up a lawyer's brief because she likes the writer's Faulknerian manner or because she has a Saturday afternoon to while away in the hammock. The judge reads the lawyer's brief because that is her job. From this perspective, asking the reading lawyer to endure long working-memory-challenging sentences is akin to asking the tired mason to carry a few more bricks when his wheelbarrow is already spilling.

The legal writer is obliged to manage the working-memory loads that his or her writing imposes on the reader. It shows respect for the reader's efforts and time. And yet managing the reader's working-memory loads is more than a professional courtesy. Working-memory theory teaches that the real danger of difficult-to-decipher sentences and paragraphs is that they leave the reader with few, if any, cognitive resources for the higher-level reading processes where critical reasoning takes place. In sum, by making a reader endure overloaded sentences and paragraphs, the legal writer ensures not only the reader's resentment. The legal writer also ensures that the reader will not have the working-memory resources to complete the very reasoning that the writer's sentences are presumably designed to invite. The reader's limited working-memory capacity matters. It is the poor legal writer who ignores it.

92 In the "chunked" model paragraph above, the writer used the generic transitions "First" and "Second" to communicate to the reader that the ideas presented in the second and third sentences are related and sequential. The writer also used the generic transition "however" in the final sentence to communicate to the reader that the idea presented in the sentence contrasts with (or presents an exception to) the idea in the preceding sentence. There are three substantive transitions in the chunked model. By repeating the phrase "the Department may" in the second and third sentences, the writer communicated to the reader that the ideas in these sentences connected back to the idea in the first sentence that was originally introduced with that phrasing. And by starting the fourth sentence with the summarizing phrase "in either circumstance," the writer communicates that the idea to be presented relates back to the ideas presented in the two preceding sentences.