9 Writing is Applied Metacognition

Douglas J. Hacker, Matt C. Keener, and John C. Kircher
University of Utah

Writing is the production of thought for oneself or others under the direction of one's goal-directed metacognitive monitoring and control, and the translation of that thought into an external symbolic representation.

Rather surprisingly, in the extensive research on writing, a definition of the writing process is rarely provided. Numerous investigations have been conducted concerning the various components of writing, the influences on writing, and what experts do that novices do not do when writing, but a definition of the writing process is seldom mentioned and often simply implied. Considering that a goal of science is to obtain understanding of the phenomena under investigation, the omission of an explicit definition in the literature is almost disturbing. Stephen Witte (1992) suggested that questions about the nature of writing are not usually discussed because no theory bridges gaps among textual, cognitive, and social dimensions of writing, and because of the uncertainties that enshroud writing as a process or product. Roy Harris (2000) admits that the “bare idea of writing” is not a popular topic “because there are so many vested interests concerned to promote ideas of writing that are far from bare” (p. 8). However, he goes on to say that the intellectual pursuit to reconstruct the “bare idea of writing” might be worth pursuing.

In this chapter, we have taken up that pursuit by starting with an explicit definition of the writing process. We divide that definition into parts and then proceed to describe and explain what those parts mean. Next, we argue that not only are metacognitive monitoring and control essential components of writing, but that writing is applied metacognition. From this argument, we propose a metacognitve theory of writing from which our definition of writing is derived. Finally, we end with a brief description of a new methodology for investigating writing, which has provided data that have contributed to our definition and metacognitive theory of writing.

Effective writing requires audience awareness. We are aware of different camps that investigate writing and the theoretical differences that separate those camps. Researchers in the cognitive camp, such as John Hayes (1996, 2006) or Deborah McCutchen (2006), when reading our definition may resonate with the word “metacognitive” and read with interest. Researchers in the sociocultural camp, such as Paul Prior (2006), who believe that “sociocultural theories represent the dominant paradigm for writing research today” (p. 54) will read the word “metacognitive,” see no mention of context, and might decide to read no further. Researchers from a semiotic camp, such as the late Stephen Witte, may see the “Witte” name cited numerous times and read with curiosity. Researchers from a social-interactive camp, such as Martin Nystrand (1989, 2006), may read “production of thought for oneself or others” and hesitate, questioning whether the “or” should be “with.” As Harris (2000) said, we all have our vested interests to promote ideas of writing,
and we admit that we have our own vested interests. But a goal in our pursuit of that “bare idea of writing” is to propose a new theory of writing that incorporates critical elements from all the various writing camps.

Thomas Kuhn (1970) identified two conditions that new paradigms must meet: The new paradigm must resolve some generally recognized problem that has not been resolved by earlier paradigms, and the new paradigm must preserve a relatively large part of the problem-solving ability of its predecessors. Although hardly a new paradigm, we believe that our theory of writing meets both conditions: Our theory resolves problems in existing writing theories by re-conceptualizing writing as primarily a metacognitive process in which the production of text is the result of a person’s goal-directed monitoring and control of their cognitive and affective states, and we preserve a great deal of the critical elements from other theories of writing.

A Definition

In this section, we divide our definition into four meaningful parts: (a) production of thought, (b) oneself or others, (c) goal-directed metacognitive monitoring and control, and (d) translation of thought into an external symbolic representation. We describe and explain what those parts mean with the goal of constructing the “bare idea of writing” in a way that bridges gaps among textual, cognitive, and social dimensions of writing and addresses whether writing should be defined as process or product. We end with our argument that writing is applied metacognition.

Production of Thought

Production of thought is the core of writing. Regardless of what symbolic form it takes, writing is imbued with thought. Whether it is scribble writing accompanied by a drawing, a grocery list, pictographs, an essay, a novel, or a scientific report, the goal of the author is to produce meaningful thought either by translating his or her thoughts into a tangible written form or by discovering meaning through this translation process (e.g., Flower & Hayes, 1984; Kellogg, 1994; Nystrand, 1989; Prior, 2006; Witte, 1992). Disagreements do arise as to whether the production of thought is a solitary cognitive act on the part of the writer or a joint social act on the part of the writer in negotiation with a reader. Those disagreements will be addressed more fully in the next section. The point being made here is only to argue for the first part of our definition of writing, which seems to be common to all modern theories of writing, “Writing is the production of thought . . .”

Oneself or Others

The issue of writing being a solitary cognitive process or a joint social process has been at issue, sometimes hotly contested, for at least 30 years. The dominance of the cognitive perspective of the 1970s gave way rather quickly to social perspectives with the works of Shaugnessy (1977), Nystrand (1982), Heath (1983), and Prior (1991), to name just a few (see Nystrand, 2006, for an excellent review of the social and historical context of writing research). More recently, some of the original proponents of the cognitive perspective on writing have acknowledged a more salient role for social influences (e.g., Flower 1994; Hayes, 1996).

On the one hand, one cannot deny that writing usually is a social artifact. The form and content of our writing is shaped by the conventions of our society and culture at a particular time in history (Bakhtin, 1981; Vygotsky, 1991). We write within a world bound by
those conventions, and even if an author breaks the writing conventions of a particular segment of society (e.g., James Joyce’s stream of consciousness in *Ulysses*), the departure from conventions is still understood and judged by that society’s established writing conventions. We use a socially shared language, a socially shared writing system, socially shared artifacts, mores, values, and history, and a socially shared culture. And, to write effectively to another, one must always keep in mind the intended generalized social audience. In all these ways, we can say that writing is social.

On the other hand, we cannot ignore that writing also is a cognitive activity that occurs within an individual. Writing reflects the unique phenomenology of an individual. According to Charles S. Peirce, the writer links an object (i.e., concept, entity, idea, feeling, or emotion) to a sign or series of signs (Moore, 1972). Linking an object to a sign requires that the writer first brings personal understanding to the object. What is the object that I hope to describe in words? What are its properties? What does it mean to me and others? What does it do? What does it not do? The personal understanding must then be constructed through a system of signs such that a link is established between the object and sign. Peirce contends that the sign must activate the object and the object must activate the sign (Moore, 1972). Kellogg (1994) refers to this cognitive process as translating one’s personal symbols into consensual symbols. Further, the signs selected by the writer (i.e., the consensual symbols) must transmit to the reader a meaning that at least approximates the original object of the writer.

The approximation of the writer’s thoughts by the reader requires a negotiation of the “intellectual and emotional space between the ‘self’ and the ‘other,’ between the individual and the social” (Witte, 1992). Therefore, the chances that the reader’s configuration of meaning will match exactly that of the writer are fairly minimal. Witte goes on to say:

Hence what I write to mean will likely be read by someone else—in some way—to mean something else. The downside, of course, is that we—writer and reader—can only ever communicate approximately, which is generally the case between a “self” and an “other.” (p. 288)

A similar sentiment was expressed about 2500 years earlier by Plato (1961) in the *Phaedrus*:

Socrates: You know, Phaedrus, that’s the strange thing about writing, which makes it truly analogous to painting. The painter’s products stand before us as though they were alive, but if you question them, they maintain a most majestic silence. It is the same with written words; they seem to talk to you as though they were intelligent, but if you ask them anything about what they say, from a desire to be instructed, they go on telling you just the same thing forever. And once a thing is put in writing, the composition, whatever it may be, drifts all over the place, getting into the hands not only of those who understand it, but equally of those who have no business with it.

(Phaedrus, 275c)

If chances are minimal that we can fully negotiate a meaning in our writing with another, and an “intellectual and emotional space” will likely always exist, there is some part of writing that is not social. That is, even the best of writers with the best of intentions to communicate meaning can produce words that are meaningful to him or her but will fail to be meaningful to another. This failure to communicate occurs, in part, because not all thought is dependent on language. We have spatial, visual, tactile, auditory, emotional, and other non-linguistic thoughts that do not translate easily or at all into a verbal form. In
these cases, not only must the writer bring interpretation into play to understand his or her own thoughts (verbal and non-verbal), a reader must bring interpretation into play to understand the writer's interpretation (Gadamer, 1993; Nystrand, 1982). Wittgenstein comes to the conclusion that the arguments between cognitive and social approaches “may ultimately point as much to the inadequacy of a purely ‘social perspective’ as it does to the inadequacy of the Flower and Hayes cognitive process theory” (p. 262). This seems to be a reasonable conclusion. To address the inadequacies in each theoretical perspective, a viable theory of writing perhaps must incorporate both perspectives.

Writing can be a solitary cognitive act of producing meaning for oneself. For example, Leonardo da Vinci’s personalized system of writing was known only to him for the purposes of expressing only to his “self” his verbal and non-verbal thoughts. Or, writing can be a social act of producing meaning through negotiation with others. When writer and reader are face-to-face, the negotiation of meaning can be a powerful tool to enhance understanding for both writer and reader. But, when writer and reader are not in spatial, chronological, cultural, or historical contact, the negotiation of meaning occurs in more of an abstract sense, but it nonetheless carries with it a “potential for interaction” (Nystrand & Himley, 2001, p. 199). Thus, “Writing is the production of thought for oneself or others . . .”

Goal-Directed Metacognitive Monitoring and Control

Although there are declarative and procedural knowledge components to metacognition, we will focus mainly on the procedural components. These cognitive processes “monitor the selection and application as well as the effects of solution processes and regulate the stream of solution activity” (Kluwe, 1982, p. 204). Metacognitive monitoring is often described as an awareness of one’s current thoughts or behavior, and metacognitive regulation or control is described as the modification of one’s current thoughts or behavior. The importance of both components of metacognition has been recognized in the area of writing for at least 30 years. In the original Hayes and Flower (1980) model, the actions of three working memory processes, planning, translating, and reviewing, are controlled and coordinated by a monitor. Bruecher (1993) stated that a well-developed monitor is necessary to keep the complex, recursive processes of writing manageable. Kellogg (1994) proposed that a writer must monitor and evaluate the progress of his or her thinking and writing using metacognition. Deliberate control of writing subskill was identified by Bracewell (1983) as important for the acquisition of new language subskills that are specific to writing and for solving problems that occur during writing. In an update of the Hayes and Flower model, Hayes (1996) maintained the importance of metacognitive control in writing, which is accomplished through a task schema that guides and controls the writing process; and Flower (1994) continued to acknowledge the importance of metacognition in the tacit or explicit control of cognitive acts that contribute to writing. Finally, Nystrand (1989) acknowledged that a monitor is necessary to coordinate the components of writing, but that “a true theory of writing requires more than an inventory of components; a theory of writing needs a principled explanation of the monitor itself” (p. 69). We agree with Nystrand on that point, and our goal, in part, is to provide that “principled explanation.”

Every act of writing is an act of meaning production. Reading, re-reading, reflecting, and reviewing—processes traditionally associated with writing—serve as monitoring strategies to ensure that the production of meaning is in conformance with the author’s goals for writing, and as long as the author’s monitoring is accurate, this conformance is more likely to occur. Editing, drafting, idea generation, word production, translation, and revision serve as control strategies that are responsible for the actual production of meaning. In most cases of writing, by the time the first words have been written, the author has already
monitored and controlled the production of numerous ideas that are potentially appropriate to his or her goals for writing. Of these numerous ideas, only a select few have probably been translated into written text, although this varies with writing experience, with less experienced authors translating more of the ideas produced, and more experienced writers translating less. Bereiter and Scardamalia (1987) referred to the former as knowledge-telling writers and the latter as knowledge-transforming writers.

Online monitoring and control of writing continues uninterrupted until the author experiences a breakdown in meaning, which indicates that something in the text no longer conforms to his or her goals for writing, or additional meaning is simply not forthcoming. To re-establish meaning production, the author must exert control to ensure that any inappropriate mapping is remediated and further production continues. At a minimum, this requires an orchestration of monitoring and control, involving a variety of strategies, such as accurately diagnosing the breakdown in meaning, reviewing what has been written, generating new ideas, and rewriting to produce new text that is in better conformance with the writer's purposes. Once the breakdown is resolved, the author again continues with writing until another breakdown in the production of meaning occurs. As larger units of text are completed, each larger unit can be monitored through a review process for meaning at a macro-level, and if the review indicates non-conformance to the author's goals, further control must be exerted. If the review indicates conformance to one's goals, meaning production continues until the author decides to stop.

Very likely, some portion of the monitoring and control of writing is tacit or implicit (Flower, 1994; Newell & Simon, 1973; Reder & Schunn, 1996). Within just the language production portion of writing, several monitoring functions occur, many requiring less than 100 ms (Postma, 2000). Moreover, control strategies such as planning, detecting, and diagnosing problems can be highly automated, implicit processes (Flower, 1994). The extent to which monitoring and control of writing is implicit is associated with the degree of writing expertise, with higher degrees of expertise associated with greater levels of implicit processing. For younger or inexperienced writers, language production may be the only implicit process available to them. This makes every other process involved in writing explicit, demanding of cognitive resources, and potentially an impediment to writing. However, it is important to keep in mind that although automatic processing can be a valuable asset to effective writing, explicit monitoring and control are essential components in the production of meaningful text. Only through explicit reading and re-reading, reviewing, revising, editing, and deliberate production and reproduction of text can the writer gain confidence that the written text is a good representation of his or her thoughts.

Granted that writing involves metacognitive monitoring and control of the production of meaning, we must ask the question, what meaning? We do not intend to delve into a philosophical discussion of the "meaning of meaning," but we do need to think about the meaning an author is trying to produce. The meaning produced during writing is directed by the writer's goals. The idea that human thought and behavior are best analyzed and understood in terms of people's goals has become a commonly accepted concept across disciplines (e.g., Cantor & Kihlstrom, 1987; Elliott & Dweck, 1988; Markus & Nurius, 1986). Carver and Scheier (1991) are perhaps two of the strongest proponents of this view: "In our view, virtually all of human behavior has this goal-oriented quality" (p. 168).

Goals are structured in a hierarchy, with higher-level goals being translated into lower-level goals (Conway, 2005). In the case of writing, an example of a higher-level goal would be to write a paragraph that describes how symbols differ from signs. The writer must then monitor and control a production of meaning that conforms to this goal. As the writer monitors and controls the production of text, he or she will adopt lower-level goals. For example, the writer may monitor for appropriate word choice, maintain coherent meaning
within and between sentences, or ensure proper grammatical structure. The knowledge necessary for successful goal-directed monitoring and control of meaning production is drawn from the writer’s memory, but if the knowledge is insufficient or lacking, other sources must be sought (Carver & Scheier, 1991), such as other texts or people. The seeking of other sources appropriately illustrates the need identified by Witte (1992) for a theory of writing to bridge the textual, cognitive, and social dimensions of writing. The text emerges as a consequence of the knowledge of the writer and the social collaborations with others, all under the direction of the writer’s goal-directed monitoring and control of the production of meaning.

Also important to note is that goals change as the task progresses; therefore, the focus of one’s metacognitive monitoring and control changes as well. As a text grows, the task situation continually changes so that the writer is never thinking about the text in exactly the same way (Rijlaarsdam & van den Bergh, 1996). With each word written, the task environment presents a new stimulus to which the writer must generate a new response. This is not a passive behavioral relationship between an external environmental prompt and an associated response that is reinforced. This is an active within-person relationship, with the writer knowingly generating a stimulus (i.e., the written text), which is representative of his or her thoughts, and knowingly responding to that stimulus with further thoughts. As the writer’s thoughts change, different goals are evoked (Carver & Scheier, 1991), and as the goals change so does the focus of the writer’s metacognitive monitoring and control. Thus, the writer’s goals set the stage for production of meaning and for the character and quality of the monitoring and control of that production. Therefore, “Writing is the production of thought for oneself or others under the direction of one’s goal-directed metacognitive monitoring and control...”

Translation of Thought into an External Symbolic Representation

Writing is the author’s production of meaning, and that production must be translated into an external symbolic representation. But what symbolic representation? We most often think of traditional, conventionalized writing systems; however, do these conventionalized writing systems adequately represent the meaning-making that humans are capable of producing? As mentioned previously, not all thought is dependent on language. We have spatial, visual, tactile, auditory, emotional, and other non-linguistic thoughts that do not translate easily or at all into a verbal form. Therefore, we need to consider the instantiation of meaning in text in ways that do not privilege traditional, conventionalized forms of language. Witte (1992) argues that if we restrict ourselves only to traditional language we will not be able to account for many different kinds of “written text,” written text that is culturally determined and able to account for the “meaning-constructive and social-constructive dimensions of writing” (p. 249).

If we agree that writing is the production of meaning, then we must be open to other external symbolic representations of that meaning. Children’s emerging writing, which is often a cross between pictures and conventionalized letters or letter strings, is (a) a translation of meaning that is certainly not entirely conventional but is nonetheless culturally-constructive as determined by the type and content of the pictures and letters drawn, (b) is meaning-constructive as determined by the ideas the child wishes to convey, and (c) is social-constructive as determined by the intended message that is sent to another. Guidebooks to birds make use of conventionalized writing systems, but without the pictures of each bird, the guidebooks would be useless. The usefulness of an operator’s manual on how to fix a mechanical device depends as much on having the actual mechanical device to make sense of the manual as it does to have the manual to make sense of the mechanical
device (Kellogg, 1994). The manual/mechanical devices are cultural contrivances that are bound by meaning- and social-constructive processes. Finally, text messaging is a production of meaning that uses bits and pieces of a conventionalized writing system but goes beyond that system in a way that is bound to a specific culture, context, and is both meaning- and social-constructive. Consider the following example: “omg its ez 2 bffs so y ru not going 2 da part t 8tr 2 nite ppl r saying ura baby 4 nt going so u betr go K tyl.” Although bits of conventionalized writing can be identified, the translation of the author’s meaning and the conveyance of that meaning are highly dependent on the social and cultural milieu in which it was written.

The translation of meaning in an external symbolic representation entails more than a simple choice of what writing system to use. If writing is the production of meaning, and meaning can be more broadly defined than linguistic meanings, then we need to think about a more broadly defined semiotic system (i.e., theory of signs and symbols) than the traditional, conventionalized systems that are commonly used (Witte, 1992). We must acknowledge the fact that with any narrowly defined semiotic, such as a conventionalized writing system, it is unlikely that we can perfectly represent the world through language, and it is also unlikely that we can communicate perfectly our personal representation to another (Sapir, 1961; Witte, 1992). But we should be willing to accept these shortcomings if we accept writing as both a conveyance of meaning and a discovery of meaning (Witte, 1992). The discovery occurs for both the writer and reader: For the writer, when he or she searches for just the right symbol or sign to represent his or her ideas; for the reader, when he or she attempts to translate that sign or symbol. Whether conveyance or discovery, “Writing is the production of thought for oneself or others under the direction of one’s goal-directed metacognitive monitoring and control, and the translation of that thought into an external symbolic representation.”

**Writing is Applied Metacognition**

In the foregoing discussion, we argued that metacognitive monitoring and control are essential components of writing. We take this argument one step further by proposing that writing is applied metacognitive monitoring and control. We will pursue this argument by using Stephen Witte’s (1992) question of whether writing should be defined as process or product. The distinction between process and product makes sense when we are readers of other authors; however, when we are readers of our own writing, the distinction between process and product is not so clear. In this case, the process of writing is a reflection of our thinking, and the product of writing is a reflection of our thinking. How can we look at our own writing and not also look at our own thoughts?

We will start our argument by drawing on the multitude of definitions of writing provided in the Oxford English Dictionary (2008). There are a variety of definitions of writing as a noun (i.e., the product of writing) and as a verb (i.e., the process of writing). As a noun, we have selected . . .

The using of written characters for purposes of record, transmission of ideas, etc. Expression of thoughts or ideas in written words; literary composition or production. Words, letters, etc., embodied in written (or typewritten) form.

* We would like to thank Emily Hacker for sharing her text messaging with us. The message reads: “Oh my gosh, it’s easy to be best friends so why are you not going to the party later tonight? People are saying you are a baby for not going, so you better go, OK? Talk to you later.”
As a verb...

To give expression to (one's feelings, thoughts, etc.) by means of writing.
To express or present (words, etc.) in written form.
To compose and set down on paper (a literary composition, narrative, verse, etc.).

As readers of another person's writing there is a clear distinction between product and process. The product lies before us. We can access the author's written letters, characters, and words, read them, and create meaning from them. However, the process of writing, during which the author translated his or her thoughts or ideas in written form, is hidden in the product. We can infer the process from the product, but with any inference made during reading, it is difficult to know how much our own knowledge and beliefs are influencing the intended text. As readers, we must reconstruct how the author used the processes of writing to express the meaning that he or she had in mind, and we must do this in a way that resembles as closely as possible the author's meaning and purpose for writing (i.e., engage in hermeneutical interpretation). Indeed, the goal of literary criticism is to expose the writer's thinking through an analysis of the writer's finished text (Kellogg, 1994). Sometimes our interpretations may not be accurate representations of the author's process of writing, and as was mentioned previously in the quotes from Witte and Plato, what we "write to mean will likely be read by someone else—in some way—to mean something else." Good enough interpretations may be the best we can hope for, although good enough at times may be acceptable and even desired.

As readers of our own writing, the distinction between product and process becomes blurred. Writers are in a privileged position. Not only do they generate the thoughts that they wish to write, and monitor and control that generation of thoughts, but they translate those thoughts into writing, and they monitor and control that translation. The only difference between monitoring and controlling the thoughts being generated and the monitoring and controlling of the thoughts being written is that the thoughts being written likely represent only a subset of the thoughts being generated. Not all thoughts make it to print, but as writers we know what thoughts do and what thoughts do not. We have no need to infer the intended meaning. We know the meaning, and even if that is in question, we are privy to the interpretations that may be used to find that meaning. Reading, re-reading, reflecting, and reviewing are used as monitoring strategies of our own thoughts. Editing, drafting, idea generation, word production, translation, diagnosing, and revision are used as control strategies of our own thoughts. The monitoring and control of our own thinking is metacognition. Writing is applied metacognition. Joan Didion (1980) put it much more eloquently, "I write entirely to find out what I'm thinking, what I'm looking at, what I see and what it means" (p. 335).

A Metacognitive Theory of Writing

The Nelson and Narens (1990) model of metacognition has served as a versatile theoretical framework for the conceptualization of metacognition and as a heuristic for further theorizing and empirical research. Three principles underlie their model: (a) Mental processes are split into two or more specifically inter-related levels, a cognitive and a metacognitive level; (b) the metacognitive level contains a dynamic model of the cognitive level; and (c) there are two dominance relations called control and monitoring, which are defined in terms of the direction of flow of information between the meta-level and the object-level. Figure 9.1 illustrates a modification of the Nelson and Narens model as applied to writing. Based on the state of the dynamic model of the object-level at the
meta-level, control information flows from the meta-level and modifies the state of the object-level process or knowledge or changes the object-level altogether. Modifications or changes at the object-level are monitored at the meta-level so that modifications to the model of the object-level can be made. A change in state of the model of the object-level can lead to additional control information flowing to the object-level.

As was mentioned previously, some portion of the monitoring and controlling of writing is tacit or implicit (Flower, 1994; Newell & Simon, 1973; Recher & Schunn, 1996). For example, detection and repair of an error at the phonemic level can occur about 100 ms before there is overt realization of the error (Postma, 2000). Moreover, highly practiced strategies, such as planning, diagnosing, or reviewing can be automatized (Flower, 1994). Therefore, the monitoring and control of cognitive processes and knowledge at the object level can be either explicit or implicit. This is illustrated in Figure 9.1 by the solid and dashed lines indicating the explicit and implicit flow of information in monitoring and control. Deliberate and intentional writing relies more on explicit than implicit processing, although writing processes that are put into action at the object-level might always include...
some implicit processing. This is illustrated in Figure 9.1 by the dashed curved lines within the object-level.

Writing is an act of meaning production, and during the goal-directed monitoring and control of meaning production, the author uses reading, re-reading, and reviewing as monitoring strategies to ensure that the production of meaning is in conformance with the author’s goals for writing. Word production, revision of text, planning, and production of text serve as control strategies that are responsible for the actual production of meaning. Ideas produced as a consequence of the control strategies are monitored and compared against the goals the author has established at the meta-level. Some of the ideas are written and end up in the text but others do not. The author must continually monitor not only the ideas produced but the selected ideas that end up in the text so that intended goals are being met.

Hayes and Flower (1980) originally posited that writing processes are hierarchical in nature. Three working memory processes—planning, translating, and reviewing—work in concert during writing, and within each of these three processes are embedded sub-processes: Planning includes the generation and organization of content and goal setting; reviewing includes reading and editing. Furthermore, each of these mental acts may occur at any time in the writing process and are considered recursive (Hayes & Flower, 1980).

We agree with the idea of recursion, but we contend that rather than a hierarchy of writing processes as proposed by Hayes and Flower, which suggests that processes higher in the hierarchy have precedence over processes lower in the hierarchy, the writer either implicitly or explicitly engages whatever control or monitoring process is needed to maintain meaning production at the moment of writing. Moreover, although control processes are distinct from monitoring processes, the line that separates the two is often tenuous. This distinction is represented in Figure 9.1 by the dashed line between control and monitoring at the object-level. Consider these examples. As writers plan, they simultaneously monitor their plans for meaning and conformance to their goals, which can result in immediate alteration of the plans, further review of their plans, or writing of a few words or entire strings of words. As writers translate their thoughts into words, they simultaneously monitor those words for meaning and conformance to their goals, which can lead to further translation, major or minor revision of what was written, or regressing back to earlier sections for review and possible revision. As writers re-read what they have written, the re-reading can serve as a stimulus for the production of more words. A reorganization of a sentence for better meaning can lead imperceptibly to re-reading the prior sentence, which can lead imperceptibly into altering goals for writing. Using either pen and paper, keyboard and monitor, or crayon and the bedroom wall, the writer’s goal is to produce meaning in an external symbolic form, and whatever process will help in that production will be used whenever it is needed. Each monitoring or control process in use contains within it the potential for every other process, and what determines the selection of process is whether the writer’s intended meaning is being produced.

Moreover, as the text develops the task situation continually changes. The writer never attends to the same meaning in the same way anytime during writing (Rijlaarsdam & van den Bergh, 1996). Therefore, the quality and character of any writing process, whether it involves control or monitoring of meaning production, could change considerably depending on how the production of meaning proceeds for the writer. Reviewing occurring at one time may be much different from reviewing that occurs at another time. Revision of the text at the earlier stages of writing will be much different from revision that occurs at later stages (Rijlaarsdam & van den Bergh, 1996). And this can be said of every writing process. As was mentioned earlier, each monitoring or control process contains within it the
potential for every other process, and this potential changes as the writing task situation changes.

The complexities and difficulties of writing stem in large part because of what seems to be a completely chaotic process. Writing may appear to be much like the “Blooming, buzzing, confusion” that Williams James (1890) described of our world without the benefit of perceptual categories. But humans are meaning-making creatures, and as long as the purpose of writing is to produce meaning, the blooming, buzzing, confusion of writing will result in meaning for the author and hopefully for the reader.

New Methodology to Investigate Writing

With the complexities presented by even the simplest writing task, the investigation of writing is a difficult task. Hayes and Flower (1980) introduced the think-aloud methodology to writing, which had been used extensively to gain access to people's cognitions as they problem solve (e.g., Bereiter & Scardamalia, 1987; Chi, Lewis, Reimann, & Glaser, 1989; Montgomery & Svenson, 1989; Pressley & Afflerbach, 1995). Hayes and Flower's research was seminal and led to innumerable studies of writing.

Ronald Kellogg (1987) provided critical understanding of writing by investigating processing time and effort people invested in the various cognitive processes that had been identified by Hayes and Flower. He made use of the directed retrospection technique (Ericsson & Simon, 1980) in which writers were interrupted at various times during writing and asked to retrospectively identify the writing process in which they were engaged prior to the interruption. The retrospective reports indicated what processes were involved in writing, and they allowed investigators to calculate the proportion of time devoted to each process. Cognitive effort was assessed by secondary response times (Kahneman, 1973). This was measured by recording the amount of time writers required to identify the writing process in which they were engaged. Cognitive effort was defined in terms of the increase over a baseline response time calculated when writers were not engaged in writing.

Although these measures have increased our understanding of writing, they do not entirely reveal the moment-to-moment processes involved while writing. Problems with retrospective reports and secondary response times include their coarse granularity, lack of specificity, unnatural writing experiences, invasiveness, and disjointedness in providing a cohesive representation of writing. As an alternative to these approaches, we have developed a new methodology to investigate the online writing behaviors that are essential for understanding the role that metacognitive monitoring and control play in writing. Our new methodology makes use of eye-tracking technology, which, although not problem-free, is considered the state of the art for online investigations of language comprehension (Boland, 2004; Carreiras & Clifton, 2004; McConkie, Hogaboam, Wolverton, Zola, & Lucas, 1979; Rayner, 1998). Eye tracking provides continuous measures of processing time, attention, and effort, does not disrupt the reader or writer from the main task, produces data reflecting attentional shifts in periods of time as short as a few ms, can pinpoint comprehension problems at a word and even intra-word level, and provides a more natural way of examining word and sentence processing (Boland, 2004; Carreiras & Clifton, 2004).

Traktext

We developed an extension of eye-tracking technology to track eye movements during writing. Our computer program, called Traktext, displays several windows on a computer.
monitor: some windows present static texts used in traditional eye-tracking analyses of reading behaviors, and one presents a word processing program used to record participants’ writing behaviors. We continuously record participants’ eye movements as they freely navigate among the reading and writing windows. The dynamic nature of writing is captured offline at the word level by generating a bitmap of the writing window for each word that the participant types, edits, or deletes. The derived spatial coordinates for each word are used in conjunction with the eye positions recorded 60 times per second by the eyetracker to measure the number and duration of eye fixations on the word.

Our Traktext program provides numerous measures of the writing process. A record of each and every word written, deleted, or edited is compiled along with the amount of time spent writing each word (measured in ms/character) and the amount of time spent reading and re-reading each word (also measured in ms/character). Cognitive effort allocated to the writing and reading of each word is measured using pupil diameter. Numerous studies have shown that the greater the demand on cognitive capacity, the more the pupil dilates (e.g., Ahern & Beatty, 1979; Granholm, Asarnow, Sarkin, & Dykes, 1996). Kahneman (1973) has shown that pupillometric response is a reliable and sensitive psychophysiological index of the momentary processing load during performance of a wide variety of cognitive activities, including attention, memory, language, and complex reasoning. Moreover, pupillometric response shows within-task, between-task, and between-individual variations in processing load and cognitive resource capacity (Kahneman, 1973).

In our current research paradigm, we are interested in both reading and writing behaviors during problem solving; however, for the purposes of this chapter, we will focus exclusively on writing behaviors. Traktext first presents participants with a simple pre-problem exercise that involves both reading and writing. Participants read a short prompt about the Jack and Jill nursery rhyme and then write as much as they can remember about the nursery rhyme. This allows us to obtain baseline rates for both keyboarding speed and pupil diameter for reading, language production, and keyboarding behaviors. These baseline rates are used as a standard from which writing and reviewing measures on the actual problem are calculated.

Following the baseline task, participants are presented with a text that describes a problem about gravity. It is a classic problem often used in refutational-text research about two balls falling from the same height but with different trajectories (Guzzetti, Snyder, Glass, & Gamar, 1993). Participants are asked which of the two balls would hit the ground first and why. After writing their answers, participants are presented with another text that contains information that can be used to help solve the problem. Once finished with the informational text, participants can then revise their answers in any way they wish and navigate freely between the text and the problem as they revise. Participants are allowed as much time as they wish to complete the task.

**Analysis of Writing**

Using these methods to analyze writing, we have identified six distinct processes: (a) area of monitoring and revision (AMR); (b) area of monitoring and production (AMP), which includes word (language) production and graphomotor production; (c) planning, which includes generating ideas, organizing, and goal setting (Hayes & Flower, 1980); (d) area of review—formative (ARF); (e) area of review—summative (ARS); and (f) reading at the point of production. Each of these is summarized in Table 9.1 and illustrated as a monitoring or control process in Figure 9.1.
Table 9.1 Control and Monitoring Processes, Acronyms, and Definitions

<table>
<thead>
<tr>
<th>Control strategies</th>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area of monitoring and production</td>
<td>AMP</td>
<td>Contiguous area of text produced prior to a writer making a saccade beyond 12 characters or spaces from his or her last fixation on a currently produced word to another part of the text or prior to a navigation off the text.</td>
</tr>
<tr>
<td>Planning</td>
<td></td>
<td>Generating ideas, organizing ideas, goal setting, and re-setting goals.</td>
</tr>
<tr>
<td>Area of monitoring and revision</td>
<td>AMR</td>
<td>Area of the text in which a writer makes a revision.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Monitoring strategies</th>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area of review—formative</td>
<td>ARF</td>
<td>Area of text in which the writer re-reads beyond 12 characters or spaces from the last point of production and then continues with writing.</td>
</tr>
<tr>
<td>Area of review—summative</td>
<td>ARS</td>
<td>Area of text re-read just before the writer finishes with writing or before additional information is encountered that could lead to a change in the answer.</td>
</tr>
<tr>
<td>Reading at point of production</td>
<td></td>
<td>Reading the current word that is written and one or two words prior to the current word.</td>
</tr>
</tbody>
</table>

An AMR is an area of the text in which a revision occurs. Revision is broadly defined according to Fitzgerald (1987, p. 484):

Revision means making any changes at any point in the writing process. It involves identifying discrepancies between intended and instantiated text, deciding what could or should be changed in the text and how to make the desired changes, and ... making the desired changes. Changes may or may not affect the meaning of the text, and they may be major or minor.

From this definition and from our description of our metacognitive model of writing, it is clear that revision involves both monitoring (“identifying discrepancies between intended and instantiated text”) and control processes (“making the desired changes”). The line that separates control from monitoring is not always distinct. As we suggested earlier, each monitoring or control process contains within it the potential for every other process, and what determines the selection of process is whether the writer’s intentions for meaning are being met.

We define an AMP as a contiguous area of text produced prior to a writer making a saccade beyond 12 characters or spaces from his or her last fixation on a currently produced word to another part of the text or prior to a navigation off the text. We selected 12 characters or spaces because this distance has been typically defined as the amount of text that can be in foveal view during a fixation (Rayner, 1998). Thus, moving a gaze beyond 12 characters or spaces indicates that the writer has made a saccade and likely is focusing his or her attention on parts of the text that are not currently being produced.

Figure 9.2 illustrates a participant’s writing that followed immediately after her reading of the problem on gravity. The ovals on each line indicate eye fixations. Lines a, b, and c constitute the ending of the writer’s first AMP. The writer has continuously produced the words “I think that the juggler who was standing still would have the” and in Line a has
Writing is Applied Metacognition

167

a) I think that the juggler who was standing still would have _faster_.

b) I think that the juggler who was standing still would have the _fastest_.

c) I think that the juggler who was standing still would have the fastest ball. The juggler was _standing_.

d) I think that the juggler who was standing still would have the fastest ball. The juggler was _standing_.

e) I think that the juggler who was standing still would have the fastest ball. The juggler was _standing_.

f) I think that the juggler who was standing still would have the fastest ball. The juggler was _standing_.

Figure 9.2 Writing sample illustrating eye fixations, writing processes, and descriptive statistics on processes.

produced “fastest.” As indicated by the fixations, the writer is currently fixating on the word being produced, “fastest,” and on the previous word “the.” In Line b, the writer is fixated on the word being produced, “ball,” and on the previous word, “fastest.” Fixating on the word being produced and one or two words prior to that word is commonly observed in our participants. Writers often appear to read the last word that is written as a lead-in to the word that is about to be produced. This is represented in our model of writing in Figure 9.1 as Reading (Word Level during Production). Looking at the word previous to the word produced may serve as a micro-review process that the writer uses to maintain the stream of thought contributing to continued production of meaning. In Line c, the word “The” is produced and fixated, which ends the AMP because the writer then regresses in the AMP, and as is shown in Line d, begins an ARF, which is defined as a review of text that lies beyond 12 characters or spaces from the last point of production. In other words, the writer made a saccade from the point of production to another portion of the text. This is a formative area of review because the review leads to further production of their current answer. In contrast, an ARS is the final review of text before finishing the
task or encountering additional information that could change their current understanding and answer.

The values presented with the AMP1 in Figure 9.2 indicate that the production of AMP1 took 19,467 ms, which is 2,992 ms faster than the writer’s writing base rate, and the 1.09 is a ratio of the average pupil diameter during the AMP as compared to the baseline pupil diameter that was measured during the pre-problem exercise. The 1.09 indicates that greater cognitive effort was being expended as compared to the writer’s baseline value. As an approximate estimate of what a 9% increase in pupil diameter means, Hess and Polt (1964) conducted pupil diameter studies in which they produced a rough calibration of pupil increase and cognitive effort. A 5% increase in pupil diameter is approximately the amount of increase when a person begins to solve a fairly easy multiplication problem such as 7 x 8. A 10% increase is approximately the amount of increase when a person is solving a problem such as 8 x 13. A 20% increase is approximately the amount of increase when solving a difficult problem such as 16 x 23. Therefore, the writer produced AMP1 relatively quickly and showed a small increase in cognitive effort as compared to writing the Jack and Jill rhyme. A reasonable interpretation of AMP1 is that the writer likely had mentally produced his or her answer while reading the problem on gravity and quickly wrote her answer. This would be an illustration of a writing style identified by Bereiter and Scardamalia (1987) as knowledge telling, that is, the writer already has the knowledge formed and needs only to write it.

The values presented by the ARF1 indicate that ARF1 took 1,417 ms, which is 117 ms slower than the writer’s reviewing base rate, and that the ratio of the writer’s pupil diameter was 1.11. Thus, the writer interrupted the AMP to make a quick review of what she had written, and because of the increase in pupil diameter, was using a moderate amount of cognitive effort to conduct the review. The ARF is then interrupted to continue a second AMP in which “who,” “was,” “speeding,” etc. were produced. In Line 5, the second AMP is fully represented. Similar to AMP1, the values indicate that the production of AMP2 was fairly quick, taking 22,150 ms, which is 6,457 ms faster than the writer’s base rate, and that the ratio of the writer’s pupil diameter was 1.07, which indicates that less cognitive effort was being expended as compared to AMP1 and ARF1, but more than what was required for the Jack and Jill nursery rhyme. These values further support our interpretation that the writer was engaging in knowledge-telling writing in that she likely had already mentally produced her answer while reading the problem and quickly wrote her answer.

We are able to change our unit of analysis from the process level to the word level, and it is interesting to note that as the participant was writing the words “I don’t really understand” in Line 5, her pupil diameter ratios for the words were 1.11, 1.08, 1.12, and 1.11, respectively. These ratios indicate that she was expending a moderately high level of cognitive effort as she was attempting to understand her quandary in solving the problem.

After writing the word “understand,” the writer interrupts her production for another short ARF that lasts 1,250 ms, which is 834 ms slower than the writer’s base rate, and continues to expend a moderate amount of cognitive effort as indicated by the 1.07 pupil diameter ratio. ARF2 is followed immediately with another AMP in which “happening with his ball, not just gravity” is written. AMP3 is written quickly, taking only 8,917 ms, which is 3,914 ms faster than the base rate, and the pupil diameter ratio decreases to 1.03, indicating a slight decrease in cognitive effort. Possibly, during ARF2, the writer came to realize that her lack of understanding of the problem was due to some component other than gravity, and then used a portion of the ARF (i.e., the additional 834 ms) to plan her thoughts that were expressed in AMP3.

Finally, she ends her answer with an ARS. She reviews selected portions of her text one
last time, and even though she spends 13,533 ms reviewing, which is 8,733 ms more than her base rate, her pupil diameter ratio is quite low, 0.94, which indicates that there is a sharp decrease in her cognitive effort at this point. ARS1 ends with her decision to read the informational text on gravity, so perhaps the writer resigned herself to the idea that she was unable to solve the problem and would end her answer and consult with the informational text for further input. She then reads the informational text, returns to her writing, and through a series of AMRs, AMPs, ARFs, and ARSs, completely changes her writing to reflect a correct answer.

With this particular research paradigm, it is difficult to determine how much time was spent planning, because at least a portion of her planning was done during the reading of the task and the informational texts. However, within the AMPs, ARFs, ARSs, and AMRs, 14 seconds were unaccounted for, which could have been used for planning. In the future, we intend to use a paradigm that will more clearly isolate planning time.

Our methodology offers advantages over the traditional methodologies of think-aloud (Hayes & Flower, 1980) and directed retrospection (Kellogg, 1987). By investigating writing online, we get a more naturalistic representation of the processes in which writers engage as compared to think-aloud or retrospective report paradigms that may have unintended effects on the writing process. Sometimes we tell more and sometimes less than what we know (Nisbett & Wilson, 1977). As compared to self-report, eye-tracking methodology provides measures of behavior and physiology. Such measures may offer new insights into the kinds of processes that underlie the production of text, the transitions between processes, the timing of those processes, and the effort required to engage and maintain those processes.

Earlier, we had discussed how every writing process contains within it the potential for every other process and that changes can occur in a seemingly chaotic fashion. In our illustration, as well as with other participants we have observed writing, the writer went from producing text, jumped to reviewing, went back to producing text, jumped again to reviewing, back to producing, and finally back to reviewing. All of this occurred in 66.7 seconds and resulted in the production of 38 words. Some of these processes occurred in 1.25 seconds, and some represent the lion’s share of time, such as production of text. Hayes and Flower (1980) discussed the recursiveness of writing, and our illustration shows that not only is writing recursive, it is furiously recursive at times. Moreover, we found evidence for three types of reviewing: a micro-review that occurs at the point of word production, a review that pops up seemingly unpredictably during word production, and a review that occurs at the end of word production. Likely, each review serves a different purpose, and our future research will investigate these purposes. Finally, measures of processing time and cognitive effort fluctuate considerably during writing. Both of these measures have been used in prior research to make inferences about the nature of the processes writers use. In our illustration, we identified areas of writing that went smoothly, efficiently, and seemingly without taxing the writer—areas of writing that would fit with the knowledge-telling concept of Bereiter and Scardamalia (1987). We also identified areas of writing that seemingly taxed the writer and led the writer to seek additional resources—areas that would fit with the knowledge-transforming concept of Bereiter and Scardamalia. We are only at the beginning of refining this writing methodology and disseminating it. The potential for further investigations with this technology are considerable.

Conclusion
In this chapter, we have provided a definition of writing that begins to bridge the gaps among textual, cognitive, and social dimensions of writing (Witte, 1992). Our definition
derives from our proposed theory of writing that conceptualizes writing as a metacognitive process in which meaning is produced under the direction of metacognitive monitoring and control processes. In turn, our writing theory was influenced by the new ways in which we are investigating writing. Writing requires both thinking and thinking about that thinking, and the best way to derive a better understanding of the complexities of writing is by getting as proximal to thinking as possible. Think-aloud protocols and directed retrospection have done well at this, but investigating writing at the moment of production is getting us perhaps as close to the thinking and the thinking about thinking as is possible. As in any science, observations lead to the development of theory, and theory leads to definitions. Our online observations of writing have led to changes in existing theories of writing, and the changes have led to a revised definition of writing. Our overall goal has been to get closer to that "bare idea of writing."

Writing can be a solitary cognitive act of producing meaning for oneself, and writing can be a social act of producing meaning through negotiation with others. The very symbols that are used to express ideas, the manner in which the symbols are arranged, and the ways those symbols are interpreted by the writer and reader are socially, culturally, and historically bound. These aspects of writing cannot be ignored. But we also cannot ignore that there is a mind/brain that stores, manipulates, and uses the symbols for oneself or makes them available for others to use.

Kuhn (1970) proposed that new paradigms must meet two conditions: The new paradigm must resolve some generally recognized problem that has not been resolved by earlier paradigms, and the new paradigm must preserve a relatively large part of the problem-solving ability of its predecessors. Several writing theories that are still current consist mostly of lists of component processes that contribute to writing, such as generating ideas, organizing ideas, planning, reviewing, reading, content and rhetorical problem solving. These processes are still a part of our theory—although we conceptualize them as writing strategies—and therefore, a large part of the problem-solving ability of prior theories is maintained.

The problem that remains from earlier paradigms is that writing theories have not convincingly described how these component processes are coordinated under the direction of a monitor. "A true theory of writing requires more than an inventory of components; a theory of writing needs a principled explanation of the monitor itself" (Nystrand, 1989, p. 69). We have addressed this problem by providing a principled explanation of the monitor by re-conceptualizing writing as primarily applied metacognition in which the production of text is the production of meaning that results from a person's goal-directed monitoring and control of their cognitive and affective states. Moreover, each monitoring or control process contains within it the potential for every other process, and what determines the selection of process is whether the writer's intended meaning is being produced. Online monitoring of writing behavior reveals rapid and erratic changes from one writing process to another, with variable time courses and fluctuations in cognitive effort. These observations are consistent with the idea that one writing process can lead directly to any other, and all result as a consequence of the writer's goal-directed monitoring and control.

References
Writing is Applied Metacognition

Boland, J. E. (2004). Linking eye movements to sentence comprehension in reading and listening. In M. Carreira & C. Clifton, Jr. (Eds.), The on-line study of sentence comprehension: Eyetracking, ERPs and beyond (pp. 51–76).


