Title

What Do We Reward in Reflection? Assessing Reflective Writing with the Index for Metacognitive Knowledge

Author

Ratto Parks, Amy

Peer reviewed

What Do We Reward in Reflection? Assessing Reflective Writing with the Index for Metacognitive Knowledge

Amy Ratto Parks, The University of Montana, US, Amy.Ratto-Parks@mso.umt.edu

Abstract: Reflection is a staple of contemporary writing pedagogy and writing assessment. Although the power of reflective writing has long been understood in writing studies, the field has not made progress on articulating how to assess the reflective work. Developed at the crossroads of research in reflection and metacognition, the Index for Metacognitive Knowledge (IMK) is designed to help writing researchers, teachers, and students articulate what is being rewarded in the assessment of reflection and to articulate the role of metacognitive knowledge in critical reflective writing. The IMK was used to code final portfolio introductions from first-year writing courses in order to analyze the distribution of the three kinds of metacognitive knowledge (declarative, procedural, and conditional) and to explore the quality and complexity of students' metacognitive knowledge. Inter-rater reliability testing on the IMK showed that it is highly reliable; the Fleiss' kappa was 83% (K=.834). The IMK offers researchers, teachers, and students language with which to explore the unique work of reflective writing in order to develop more metacognitively rich observations. It provides a framework to explain the evolving complexity of students' reflective writing and to assess and describe the impacts of other pedagogical interventions.

Reflection is a staple of contemporary writing pedagogy and writing assessment (Huot & Williamson, 2007; White, 2007; Yancey, 1999/2007). Variously framed as an academic behavior (Serra & Metcalfe, 2009; Tobias & Everson, 2009), a part of the writing process (Camp, 2009; Taczak & Robertson, 2017; Yancey, 1999/2007), and a genre of writing (Reynolds & Rice, 2006), reflection has permeated pedagogical literature since the third wave (1986-present) of writing assessment theory (Yancey, 1999/2007) when the portfolio solidified its centrality in the writing classroom. Research about reflection shows persuasive, subtle attention to its inextricable connection to the writing process (Camp, 1992; Pianko, 1979; Yancey, 1998), its fundamental role in portfolio assessment (White, 2007; Yancey, 1998), and, most recently, its connection to learning transfer and metacognition (Beaufort, 2016; Gorzelsky et al., 2016; Taczak & Robertson, 2017).

Although writing studies researchers are now articulating the relationship between reflection and metacognition (Beaufort, 2016; Gorzelsky et al., 2016; Taczak & Robertson, 2017), reflection's relationship to metacognition has historically been more difficult to characterize. In *A Rhetoric of Reflection* (2016), Kathleen Blake Yancey writes, "in writing studies, *reflection* has been the key term, while in higher education contexts, *reflection* and *metacognition* are often used interchangeably" (p. 6). Both terms have been variously defined in the fields of writing studies and educational research. In writing studies, reflection is broadly understood as "the process by which we know what we have accomplished and by which we articulate accomplishment" (Yancey, 1998, p. 6). In educational research, metacognition is broadly defined by "two separate but related aspects: (1) knowledge/awareness of cognitive processes, and (2) control of cognitive processes" (Williams & Atkins, 2009, p. 27).

As Yancey (2016) suggests above, the relationship between reflection and metacognition sits at the intersection of writing assessment conversations happening in two distinct fields: writing studies and educational research. Educational researchers have been assessing metacognition since 1979 (Flavell, 1979) and while writing studies began defining reflection and its role in the writing process around the same time (Pianko, 1979), the field has made little progress in articulating how it is assessed. Even Yancey (1998), after outlining the many benefits and uses of critical reflective work, and despite her clear and informed belief in the importance of reflection, goes on to say that "the function of reflection in an assessment context isn't entirely clear" (p. 145) and concludes, "no one really knows what we reward in reflection" (p. 147).

Developed at the crossroads of research in reflection and metacognition, the Index for Metacognitive Knowledge (IMK) is designed to help writing researchers, teachers, and students do two things: 1) articulate what is being rewarded in the assessment of reflection (Yancey, 1998), and 2) articulate the role of metacognitive knowledge in critical reflective writing (see Appendix A). The IMK's demonstrated reliability positions it for use by researchers interested in large-scale explorations into the assessment of reflective writing. However, the IMK also functions at the classroom and individual levels as an "instructive assessment" (Huot, 2002, p. 18) that informs the progress of student learning. As an instructive assessment, the IMK meets Huot's (2002) demand that "all procedures used to assess writing . . . also contain properties that work toward the improvement of the teaching and learning of writing" (p. 18).

Reflection and Metacognition

Reflection and metacognition are clearly valued by national governing bodies in writing studies. The Council of Writing Program Administrators, the Conference on College Composition

and Communication, and the National Council of the Teachers of English all include *reflection* and/or *metacognition* as important pedagogical outcomes or directives.[1] In the historical literature of writing assessment, the terms reflection and metacognition are often used interchangeably to describe a student's behavior of reviewing the thinking and behavioral choices in the writing process and/or the development of (or change in) writing assignments. Writing studies researchers have long assumed, but not necessarily clarified, a connection between these concepts. For example, in his article "The Scoring of Writing Portfolios: Phase 2," White (2005) describes the role reflection plays in portfolio assessment. He writes, "it is a powerful metacognitive act—thinking-about-thinking—that no other assessment device includes" (p. 583). In a discussion of the contents of a particular portfolio assignment, Hamp-Lyons and Condon (2007) include "a reflective piece . . . that deals in some way with the writing in the portfolio (metacognition)" (p. 317). Reynolds and Rice (2006) imply (but do not directly say) that metacognition is a result of reflective work. In *Portfolio Teaching: A Guide for Instructors*, they begin a paragraph by defining reflection and then, almost without transition, move to defining metacognition (p. 8). The juxtaposition of the definitions of the two terms implies a relationship between them, but does not further elaborate on how or why they are related.

More recent works though, particularly essays in *A Rhetoric of Reflection* (Yancey, 2016), have dug into defining and exploring the relationship between cognition, metacognition, and reflection. Beaufort's (2016) article "Reflection: The Metacognitive Move Towards the Transfer of Learning" explores reflection as a site of learning transfer. Taczak and Robertson's (2017) article, "Metacognition and the Reflective Writing Practitioner: An Integrated Knowledge Approach," argues that students need "integrated knowledge" which includes (among other concepts) "cognition, metacognition, and reflection" (p. 212). They argue that the relationship between reflection and metacognition is "similar, yet distinct, and separate but interrelated" (p. 216) and they directly articulate that cognition and metacognition "are accessed together through reflection" (p. 212).

Research Rooted in Cognitivism

The study of metacognition grew out of the study of cognition. Cognitive models of the writing process have sought to establish a roadmap for the "set of distinctive thinking processes" (Flower & Hayes, 1981/2011, p. 254) involved in writing (Fleckenstein, 2012; Flower & Hayes, 1981/2011). Many writing researchers have shown that writing is an inherently recursive process and that all writers, no matter how successful they are, engage in a series of many recursive loops between thinking, planning, writing, and revising (Berthoff, 1984/2011; Emig, 1971/2009; Murray, 1972/2011; Sommers, 1980/2011).

Cognitive models have offered important insight into the writing process. However, many scholars have argued that these models are still limited (Berthoff, 1984/2011; Fleckenstein, 2012). Fleckenstein (2012) writes that the critics of the cognitive models had two major objections. They "questioned the foundational premises of the model: Is writing a species of problem solving?" (p. 89). Fleckenstein goes on to show that Berthoff also argued "a problem solving approach strips

---

1 Statements include the Council of Writing Program Administrators' Outcomes Statement for First-Year Composition (3.0) (2014) and Framework for Success in Postsecondary Writing (2011); the Conference on College Composition and Communication's "Writing Assessment: A Position Statement" (2022), "Principles and Practices in Electronic Portfolios" (2015b), and "Principles of the Postsecondary Teaching of Writing" (2015a); and the National Council of Teachers of English's "Resolution on Grading Student Writing" (1993) and "Statement on Professional Knowledge for the Teaching of Writing" (2016).

away from the writing the formative power of imagination, reducing writing to manipulation, rather than creation" (pp. 89–90). Even John H. Flavell (1976), considered to be the father of metacognition, turned away from cognition and toward metacognition because, he wrote, "I suspect that a lot of human thought, even in problem-solving situations, may be erratic and inconsistent in direction, subject to multiple embedded interruptions and detours, and generally replete with vague, difficult-to-model ideas" (p. 234).

The limitations in the theories of cognitive models of writing have led to a call for a more complex model that incorporates these internal authorial questions. Fleckenstein (2012) suggests an eco-cognitive paradigm. Berthoff (1984/2011), recalling Paulo Freire's pedagogy of knowing, argues, "until the mind of the learner is engaged, no meaning will be made, no knowledge can be won" (p. 310). A metacognitive model helps expose the thinking behind the actions, or as Bruner (1991) says, it would show the "Self *and* [emphasis added] its doings" (p. 66).
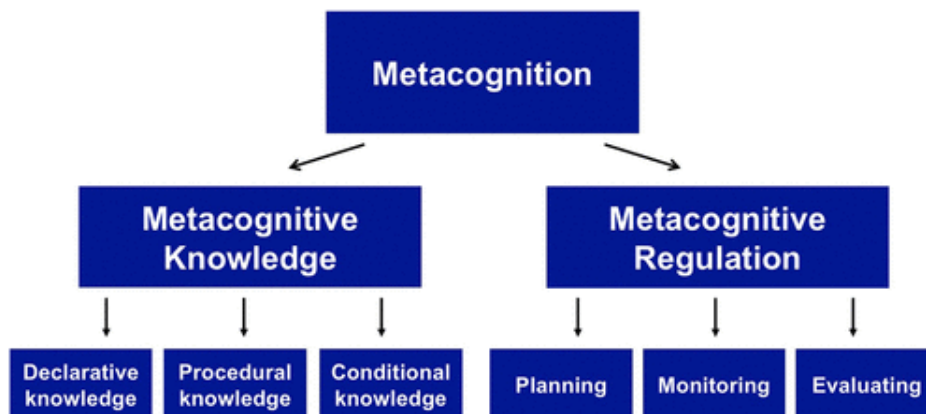
Metacognition and Writing

The awareness of and the ability to observe and control the cognitive process interests researchers who study metacognition and writing studies. Thinking and writing are closely tied (Bean, 2011). Metacognition is so closely tied to writing that Hacker et al. (2009) argue "writing *is* [emphasis added] applied metacognition" (p. 154); they define *writing* as "the production of thought for oneself or others under the direction of one's goal-directed metacognitive monitoring and control, and translation of that thought into an external symbolic representation" (p. 154).

Although metacognition is often used as a singular concept in writing studies, educational researchers understand it as a category of cognitive work that includes metacognitive knowledge and metacognitive regulation (see Figure 1). Metacognitive knowledge "includes what you know about your own thinking and what you know about strategies for learning" (Stanton et al., 2021). Metacognitive regulation "involves the actions you take in order to learn" (Stanton et al., 2021).

Figure 1
Metacognition Framework (Stanton et al., 2021)

Note. This Figure appears in Stanton et al. (2021) "Fostering Metacognition to Support Student Learning and Performance"

Only one major study, "Cultivating Constructive Metacognition: A New Taxonomy for Writing Studies" (Gorzelsky et al., 2016), has explored the relationship between metacognition and transfer in the writing classroom by engaging "the specific subcomponents of metacognition and their relationships" in order to "promote metacognitive development that supports the transfer of writing-related knowledge across courses and contexts" (p. 215). The large study (N=398) included analysis of a data set of reflections, writing produced for classes, follow-up interviews, and interviews that included a think-aloud protocol. They organized their code development around five major metacognitive concepts—"knowledge of cognition, planning, monitoring, regulation/ self-control, and evaluation" (p. 224) —in order to "reveal how specific metacognitive components and subcomponents operate and intersect in writing" (p. 238).[2]

The resulting taxonomy offers a broad, deep articulation of how these five metacognitive concepts appear in student reflections, and has the potential to provide a foundation for research and curricular development. However, as their article concludes, they also call for more research into reflective work. They write, "To help teachers to foreground and build more effectively on students' existing metacognitive capacities, further research should investigate fully what metacognitive moves students are already making" (Gorzelsky et al., 2016, p.241). Like Taczak and Robertson (2017), they find that that reflection is the site of learning about metacognition and draw attention to the need for a renewed attention to research on the assessment of reflective writing.

Reflection and Writing Assessment

The idea of reflection so permeates the theory and practice of contemporary writing studies that it is difficult to believe that so little attention has been paid to how to actually assess it in a classroom setting. For 30 years, the field of writing studies has been exploring the uses and impacts of reflective writing. However, as Taczak and Robertson (2017) write, "definitions and, perhaps, more importantly, perceptions of what reflection means with regards to writing, have little consensus among us" (p. 217).

Researchers have carefully and productively characterized different types of reflective work (Sommers, 2006; Yancey, 1998) and theorized its role in portfolio assessment (Huot & O'Neil, 2009; Reynolds & Rice, 2006; White, 2007), and the field seems to be relative agreement about the "educational force of reflection" in the writing classroom (White, 2007, p. 168). In his book, *Assigning, Responding, Evaluating*, White (2007) argues that in some situations, the portfolio's reflective letter is so central the goals and outcomes of the course that in reviewing portfolios for final grades, instructors "can focus almost entirely on the reflective letters" (p. 171). Reynolds and Rice (2006) describe the reflective introduction as "the equivalent of a final exam" (p. 41).

However, there is very little specific guidance offered in how to respond to or assess the reflective letter in particular. White (2007) views the reflective letter as an argument (p. 169) for what was learned in the class and he is very careful to clarify that portfolio assessment only succeeds pedagogically if the program sets clear learning goals for the course (p. 169). In his view, the strength of the reflective writing can be gauged by how well the writing meets the course goals.

---

2 The field of educational research continues to discuss and refine the subcomponents of metacognition. Gorzelsky et al. (2016) developed their codes based on concepts from Scott and Levy (2013) in "Metacognition: Examining the Components of a Fuzzy Concept." Figure 1 cites a different map of the concepts developed by Stanton et al. (2021).

This makes broad pedagogical sense, but does not necessarily specify how reflective writing might function as its own genre or what one teacher or another might reward in reflective work.

The book *Portfolio Teaching: A Guide for Instructors* (Reynolds & Rice, 2006) also works toward a direct discussion of assessing reflective writing with a section called "Assessing the Reflective Introduction." However, the section includes only "passages from several students' reflective introductions" offered not as "models so much as illustrat[ions] of different kinds of passages" (pp. 52–53). It offers a detailed discussion of pitfalls of reflective work such as "glow and schmooze" (p. 59) and "stock narratives" (p. 61), but does not offer a reader direct guidance on what might make one reflective letter stronger than another.

White (2007) and Reynolds and Rice (2006) offer an illustration of a deep tension at the heart of assessing reflective writing: is reflective writing a genre of argument and should it be assessed as one? Or do we also hope to see something else? And if so, *what is the something else*? When White (2005) clarifies that reflective letters need to show "genuine reflection" (p. 181), not a "hasty overview of the portfolio contents" (p. 180), he communicates a value judgement about the perception of a student's personal investment. Perhaps Reynolds and Rice (2006) are hinting at the same mystery when they acknowledge that "even the most impressive passages of writing may lack something that you know you expect to see when you make this assignment" (p. 59).

As the field moves toward understanding reflection less as a final exam (Reynolds & Rice, 2006) and more as the site of metacognitive work (Gorzelsky et al., 2016; Taczak & Robertson, 2017), our understanding of what to reward in this reflective letter becomes even less clear. Take the sample student writing below, for example. A generic assignment for a portfolio introduction or reflective letter might ask students to reflect on their writing and revision over the course of the semester; it might also ask them to demonstrate an understanding of course concepts and/or an understanding of the rhetorical situation, to explain their portfolio selections, and/or demonstrate their awareness of their own writing choices.[3] The following examples of student writing offer common and contrasting traits of reflective writing:

A) This semester I wrote three papers. I did research for one of them, but one was a personal essay so it was really more about stories than quoting. The other one I didn't like as much. The argument one. I have never been good at writing arguments and I guess I'm still not! But I do think that I understand pathos, ethos, and logos better now. So thank you for that! The personal essay was my favorite because I liked talking about my grandmother's baking.

B) Of all of our assignments this semester, the research-based argument was my favorite. I was excited to learn so much about how technology and happiness relate to each other. Because my little sister was my audience, I spent a lot of time thinking about how to paraphrase my research so that it would be in my own voice. I think it's important to find ways to connect with a reader on their own terms.

It seems like it might be immediately obvious that one of these samples is stronger than the other. But why? Both essays offer reflections on writing completed during the course, both offer insight into how the student understood and interacted with the assignment(s), and both engage course concepts to explain something they learned about writing. Of course, the samples differ in tone, in

---

3 Assignments for portfolio introductions vary widely. This description is intentionally vague, but attempts to capture the core elements addressed in the assignment.

style, and in a demonstrated ability to organize sentences in a paragraph. But is that what we value here? Or is it something else?

Does student A's writing seem like "genuine reflection" (White, 2005, p. 181), not a "hasty overview of the portfolio contents" (p. 180)? If the goal of assigning critical reflective writing is to help students understand themselves and their processes more clearly, will it help student A further understand themselves and their processes if they are told to improve the organization? Will it help them at all if a teacher is placing judgements about their perceived personal investment in "genuine reflection" (White, 2005, p. 181)? Isn't it possible that we don't know what defines "genuine reflection" and that we are mostly being guided by an unexamined bias toward students that successfully navigate the expectations of Standard Edited Academic English (SEAE) (Inoue, 2015)? Or can there be decent metacognitive work happening inside the student's varying ability to meet the expectations of SEAE?

There are, of course, countless ways to discuss possible analyses of these two samples. These samples are meant as an illustration of our general lack of articulation about what teachers in writing classrooms reward in this genre of writing. One sample demonstrates stronger SEAE *writing* than the other, but is it also stronger *reflection*? And if reflective writing offers a window into a student's metacognitive development—if it is potentially a site in which instructors might measure progress toward national best practices and outcomes—then how can we articulate the metacognitive work happening in these two samples?

The Index for Metacognitive Knowledge

The IMK was designed to help teachers, researchers, writing program administrators— and students—begin to see reflective writing through a metacognitive lens and articulate how their metacognitive knowledge works on the sentence level (see Appendix A). Gorzelsky et al. (2016) established a broad articulation of how five concepts of metacognition present themselves in the writing classroom. The IMK narrows the scope of assessment in two ways: it focuses on only one metacognitive concept, metacognitive knowledge (called *knowledge of cognition* in Gorzelsky et al., 2016), and it explores how metacognitive knowledge appears in only one kind of reflective work, the reflective letter or portfolio introduction.

Stanton et al. (2021) show that metacognition is often understood as a combination of metacognitive knowledge (or "what you know about your own thinking and . . . about strategies for learning") and metacognitive regulation ("the actions you take in order to learn"). Metacognitive theorists describe learning situations as "problem-solving" situations; they say that in order for a student to negotiate the situation well, they need to understand the relationship between themselves, the task, and the strategies available for the task. The three kinds of problem-solving knowledge—self, task, and strategy knowledge—form an interdependent, triangular relationship (Flavell, 1979) (see Figure 2). All three elements are present in any problem-solving situation and a change to one of the three requires an adjustment of the other two (i.e., if the task is an assignment given to the whole class, then the task will remain the same; however, since each student is different, each student will need to figure out which strategies will help him or her best accomplish the task).

A combination of a learner's skills or understanding in person, task, and strategy variables helps define the kind of metacognitive knowledge they are able to employ. There are three kinds of metacognitive knowledge: declarative knowledge, procedural knowledge, and conditional knowledge (see Figure 3). While all three kinds of metacognitive knowledge should be in place

for effective task performance (Harris, et al., 2009), students at various skill levels often present dominant competence in one area over the others (Waters & Schneider, 2010).

Figure 2
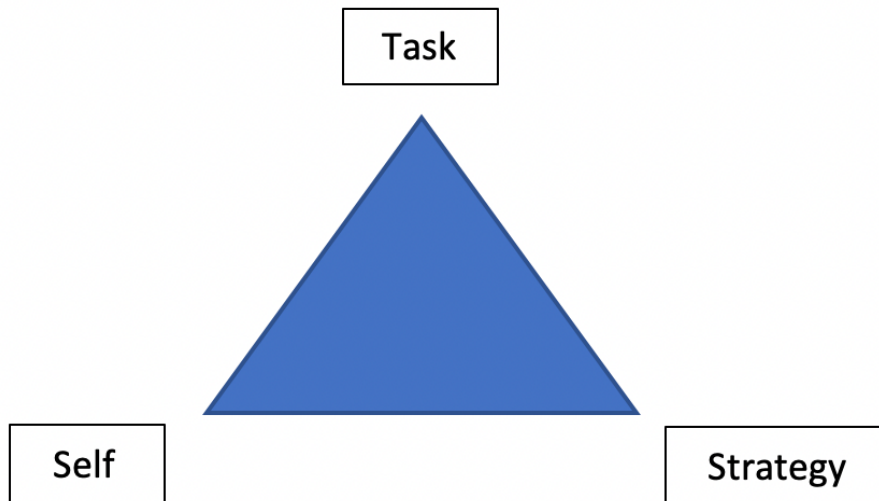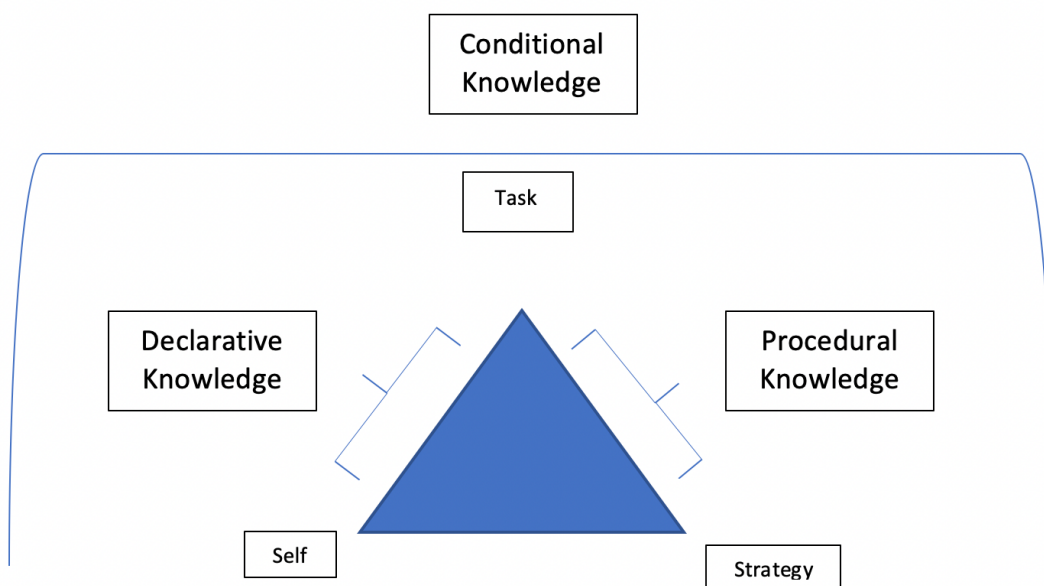
Problem-Solving Knowledge (Flavell, 1979)

Figure 3

The Relationship Between Problem-Solving Knowledge and Metacognitive Knowledge (Flavell, 1979

Three Kinds of Metacognitive Knowledge

Tarricone (2011) defines declarative knowledge as "stable, familiar, constant, established long-term knowledge which involves self-knowledge, self-awareness and a sensitivity to and evaluation of this knowledge" (p. 156). Tarricone (2011) adds, "essentially, declarative knowledge is knowing when and what you know and what you do not know" (p. 157). The definition of declarative knowledge shifts some when considered in the writing classroom. Harris et al. (2009) state that declarative knowledge "refers to one's knowledge about oneself as a learner, including knowledge about one's own abilities (strengths and weaknesses). Declarative knowledge also refers to knowledge regarding the task, including knowledge, skills, and strategies needed for effectively completing the task under one or more conditions" (p. 133).

Tarricone (2011) defines procedural metacognitive knowledge as "knowledge of processes and actions or essentially knowing how . . . to meet task demands or task objectives" (p. 160). Harris et al. (2009) define procedural knowledge in the writing classroom similarly, as "the knowledge needed to carry out procedures, including strategies, in order to apply declarative knowledge and reach goals. This knowledge is about 'how to do it'" (p. 133). Procedural knowledge includes understanding and martialing "general and genre-specific strategies" (Tarricone, 2011, p. 135). Tarricone (2011) adds that procedural knowledge is "developed through application and experience" (p. 160) and that "beliefs about task difficulty can positively or negatively influence the identification and successful application of strategies" (p. 163). If students do not accurately assess their writing situation, they could easily apply unproductive strategies and it would present as weak procedural knowledge.

Conditional knowledge is the third category of metacognitive knowledge, but its relationship to the other two kinds is its most important feature. Tarricone (2011) says, "conditional knowledge involves knowing when and why to use declarative and procedural knowledge . . . it is discussed mainly in terms of declarative and procedural strategy knowledge, application and effectiveness in task situations" (p. 165). In Tarricone's analysis, conditional knowledge reins over the other two forms of metacognitive knowledge, but those other two forms operate on the same plane (one is not preferred over the other). Harris et al. (2009) say that conditional knowledge "refers to knowing when, where, and why to use declarative knowledge as well as particular procedures or strategies, and is critical to effective use of strategies" (p. 133). Conditional metacognitive knowledge helps 1) evaluate the writing task in order to 2) determine the strategies needed and, 3) to select from alternative strategies. Conditional knowledge is the kind of metacognitive knowledge that helps the student see the big-picture context of the task situation; it leads a student to transfer knowledge and strategies into new situations (Tarricone, 2011).

Methodology

In order to test the reliability of the IMK, relational analysis was used to analyze end-of-semester portfolio introductions collected from College Writing I courses through an IRB-approved (#171-14) process. All sections of College Writing I used the same portfolio curriculum. The course was designed based on the principles detailed in Reynolds and Rice's (2006) book *Portfolio Teaching: A Guide for Instructors*, which emphasizes "choice, variety, and reflection" (p. v) throughout the semester. As guided by Reynolds and Rice (2006), informal reflective writing was integrated throughout the portfolio course alongside formal essay assignments. The final portfolio

was assigned as a curated collection of writing samples selected by the student. The portfolio introduction (sometimes also referred to as a reflective letter, reflective memo, or cover letter) was assigned as the final major writing assignment of the semester in order for students to demonstrate "their ability to think, to analyze, and to respond to the rhetorical situation" (p. 38).

Relational analysis is a type of content analysis that "examines the relationship between a number of concepts" (Huckin, 2009, p. 14). Specifically, the concepts that were examined via relational analysis were the three kinds of metacognitive knowledge presented in the texts. The distribution of the three kinds of metacognitive knowledge—declarative, procedural, and conditional—communicated the quality and complexity of a students' metacognitive knowledge. The relational content analysis of the three kinds of metacognitive knowledge were coded and analyzed using the Index of Metacognitive Knowledge in Critical Reflective Writing (IMK).

The IMK was used to code each sentence of the students' portfolio introductions. The IMK designates sub-traits for each of the three kinds of metacognitive knowledge and also includes two categories for sentences that do not convey metacognitive knowledge. The IMK's descriptions, traits, and sub-traits are as follows:

> **Declarative Knowledge**: Knowledge of oneself and one's abilities (strengths and weaknesses). Also includes task knowledge; knowledge, skills and strategies for completing the task. In reflective writing, statements that convey declarative metacognitive knowledge are focused on the relationship between self and task. These student writers make statements that:
>
> **D1**: assess general personal strengths and challenges or recount personal experiences and preferences.
> **D2**: simplistically recount or explain what happened (or didn't happen) or what was completed (or was not completed).
> **D3**: demonstrate task awareness by naming a discreet task and beginning to explain his/her strategies for completion (but does not name strategy).
>
> *Annotation*: Commonly coded declarative sentences focus on the connection between self and task (Flavell, 1979) and express personal preferences or explanations of "what happened" during an assignment. The following are a few examples of declarative quality writing: "I was very overwhelmed by that paper." "I did not understand what rhetoric was before this class." "The first assignment we did was the blind draft." "I spent more time on my second paper."
>
> **Procedural Knowledge**: Knowledge needed to carry out procedures in order to apply declarative knowledge; tells the learner how to complete a task. In reflective writing, statements that convey procedural metacognitive knowledge are focused on the relationship between task and strategy. These student writers make statements that:
>
> **P1**: name a strategy or behavior the student did or did not engage in a single assignment.
> **P2**: explain why a particular strategy was engaged in a single assignment.
> **P3**: observe more than one strategy or writing behavior in a single assignment.

P4: show that the student observes connection between strategies and tasks in a single assignment.

Annotation: Procedural writing focuses on the relationship between task and strategy (Flavell, 1979) and focuses on writing strategies, behaviors, or connections between assignments. The following are a few examples of procedural quality writing: "The arrangement could still use a little revision." "It outlined how to set up my subject, how to identify my audience and different techniques to steer my paper." "It proved very helpful for my next essay, the research-based persuasion." "When reviewing my own materials . . . I was surprised by the differences I saw."

<u>**Conditional Knowledge**</u>: Knowing when, where, and why to use declarative and procedural knowledge. In reflective writing, statements that convey conditional metacognitive knowledge are focused on the relationship between person, task, and strategy. These student writers make statements that:

C1: observe behaviors across writing assignments or throughout the course.
C2: explain when, where, or why a certain writing strategy was or would be of use across assignments or throughout the course.
C3: demonstrate an understanding of the class as a collection of interrelated writing tasks; may state or imply future use.
C4: demonstrate beginning theories about writing, rhetoric, learning, school, or their own practices, processes, or behaviors.

Annotation: Conditional quality writing is highly contextualized and considers self, task, and strategy (Flavell, 1979). Sentences indicating conditional knowledge attempt to or succeed in communicating why certain writing strategies work in certain contexts; the learners can observe relationships between writing tasks and can offer beginning theories about writing, rhetoric, or their own practices, processes, or behaviors. The following are a few examples of conditional quality writing: "I learned just how much gearing a work toward a particular group changes the rest of the evidence used in writing." "One of the biggest obstacles I found was simply choosing an audience, as this choice determines the voice and content of a work." "[In the future] I will consider what is relevant to my audience, what my purpose in speaking to that audience is, and the message I intend to deliver." "The essay is short (college students are known for their small attention spans) [sic] uses colloquial language, and leaves out scientific data and citations."

Transitional/Miscellaneous: Sentences were coded as T/M if they were
  •a quote from a peer or another source and were included without context or analysis.
  •comments directly posed to the teacher ("you").

Fragmented: Sentences were coded F if they were
  •a fragmentary interjection such as "so fun!" or "wow."

•a sentence fragment or impartial thought that did not offer enough information for analyzing metacognitive knowledge.

Inter-rater Reliability of the IMK

Before establishing final codes for the writing samples from the current study, the researcher conducted an inter-rater reliability study of the IMK. In order to prepare for the inter-rater reliability study, the researcher scored all portfolio introductions (a total of 1,108 sentences) twice. For each coding round, the researcher used a random number generator to establish a coding order. For this study, each sentence was coded using one of the sub-trait codes from the IMK. For example, rather than simply assigning D to mark a sentence as declarative, the researcher assigned the more specific sub-trait codes of D1, D2, or D3. During these first two rounds of coding, the researcher took careful notes about potential confusions or inconsistencies in applying the traits of the IMK to these samples of student writing.

After the first two rounds of preliminary coding were complete, the researcher examined the codes side by side, reviewed her process notes, and made small clarifying revisions to the IMK so that it would communicate more clearly to the coders. The researcher also calculated her own inter-rater reliability as a coder using Cohen's kappa calculation.

A third round of preliminary coding was conducted in order to pilot a potential method of coding for the coding training; the researcher sought to discover whether or not sentences could be reliably coded when removed from the context of the larger paper. Therefore, during the third round of preliminary coding, a sample of student sentences was selected randomly for coding. The sample size, which would offer only a 5% margin of error, was determined by a sample size calculator at raosoft.com to be 286 of the 1,108 sentences. In order to randomize the sentences chosen for coding, all 1,108 sentences were uploaded in a flashcard application on the educational website Quizlet.com. Each sentence was one side of the flashcard while its code was on the other side. The researcher coded the first 286 sentences presented by Quizlet.com's flashcard application. However, any sentences that were inconsistently coded in the first two rounds were skipped. A Cohen's kappa calculation of reliability was run between the second and third round codes and it was determined that coding sentences removed from their original written context was not nearly as reliable as coding them in the context of their original writing.

Coder Training

The researcher identified four potential coders who met certain qualifying parameters. Potential coders were considered to be qualifying if they had at least a master's degree in English or English Education; an extensive background teaching writing specifically to students in the approximate developmental range of traditional first-year writing students (approximately 18 years old); familiarity with portfolio instruction, assessment and introductions; and a background in writing assessment. Two of the identified and invited coders were able to attend the coding training and both of them completed coding the materials.

The coding training was carefully planned and organized based upon holistic scoring methods described by Myers (1980) and White (1985). The coder training for this research project took place in stages. First, the researcher gave a small lecture about metacognition and its role in problem solving, education, and writing. The lecture also offered the coders a brief introduction to

the concepts built into the IMK. The second stage of the coder training included a study of anchor sentences and the actual IMK. The anchor sentences were carefully chosen by the researcher because they were clear representations of the way the traits and sub-traits of the IMK typically appeared in students' portfolio introductions. The researcher and the coders read through each trait and sub-trait of the IMK and then reviewed the corresponding anchor sentence, making sure to verbally process how the sentence expressed the kind of metacognitive knowledge described in the sub-trait. Any questions about the IMK were carefully explored.

During the next stage of the coder training, the researcher gave each coder a supplemental coding guide (Appendices B-G) that included visual depictions of declarative, procedural, and conditional knowledge, as well as a guide to some commonly confronted questions, potential inconsistencies, and examples of difficult sentences to code. The coders and the researcher reviewed the anchor sentences while referencing the new supporting materials, and then moved on to coding two practice papers together. The practice papers were actual student portfolio introductions from this research study; they were chosen because they were short and the majority of their sentences were relatively straightforward. With the first practice paper, the coders and the researcher talked through the first 10 sentences aloud and came to consensus on the appropriate code. After that, they each coded the next same 10 sentences quietly and then everyone shared their codes, discussed their reasoning behind assigning their codes, and worked toward a consensus code. By the end of the second paper, the researcher and the coders believed that the coders were ready to begin coding their sample papers independently.

Each coder was assigned seven student samples, which was a total of 289 sentences. The sample size, which would offer a 5% margin of error and therefore, 95% confidence level, was determined by raosoft.com to be 286 sentences. Both coders completed two of their student papers while still at the coding training; the rest were completed independently within the following week. The researcher emailed each coder an Excel file in which to record each code they completed after the training. When the researcher received the data from the coders, a Cohen's kappa was calculated between the researcher's codes and each coder's codes. A Cohen's kappa was then calculated between the codes from each coder.

Coding with the IMK

After the IMK was shown to be a reliable coding index, the researcher reviewed all available sub-trait codes for all 1,108 sentences and determined the appropriate dominant trait code for each sentence. For example, if sentence 17.139 had sub-trait codes of D1, D2, and D3, the sentence's dominant trait was D, or declarative. Because the sub-trait designations often overlapped considerably in order to accommodate a coder's interpretation of the sentence, the sub-traits were important for training and cross reference. However, because of the overlap between the possible characteristics leading to sub-trait designations, and because one sentence could easily be characterized by more than one sub-trait, the sub-traits themselves did not communicate anything specifically about the work. The important characteristic was the dominant trait. If a sentence had sub-trait codes that were inconsistent across dominant traits (for example, one coder coded it a D, while the other two coded it a C), the researcher marked the sentence and created a list of sentences that still had inconsistencies. After collecting all inconsistencies, the researcher blindly coded those sentences again and in each case was able to establish a dominant code.

Data Analysis

Cohen's kappa measures inter-rater reliability by measuring the consistency between codes and removing the chance of a coder randomly guessing at a code (Landis & Koch, 1977). According to Landis and Koch (1977) a Cohen's kappa of 80% or higher demonstrates high reliability. Between the first two preliminary rounds of coding, which included 2,216 sentences, the Cohen's kappa was 98% (K=.98) which means that the researcher consistently coded the sentences with the same type of metacognitive knowledge 98% of the time.

The third round of preliminary coding was conducted differently. The goal of the third round of coding was to identify the best method of presenting trained coders with sentences to code. The researcher was trying to determine whether sentences would be more reliably coded in the context of their original essays or removed from that context and presented as individual sentences. In order to remove each sentence from its original context, all 1,108 sentences were uploaded to the educational website Quizlet.com. Quizlet.com generated flashcards with one sentence per card. In "flashcard" mode, the website then presented each sentence in a scattered, random order. After removing any sentences that were inconsistently coded in the first two rounds, the researcher coded the first 286 sentences presented by Quizlet.com. According to raosoft.com, 286 sentences created a statistical sample of the sentences that would generalize results with only a 5% margin of error. The Cohen's kappa for the third round of coding was 60% (K = .5997). Landis and Koch (1977) consider this to be moderately strong reliability; however, 60% reliability is not nearly high enough to be acceptable and, therefore, the coder training for the inter-rater reliability study was conducted with complete student writing samples rather than individual sentences (see Table 1).

After the coder training, the coders independently scored seven complete portfolio introductions; collectively, the portfolio introductions included 289 sentences, which was a statistical sample that would ensure generalizability of the results with only a 5% margin of error. The Fleiss kappa statistic is used to measure inter-rater reliability with more than two raters (Landis & Koch, 1977). According to Landis and Koch (1977) a Fleiss kappa score in the range of .81-1.00 is considered to be the highest degree of agreement and is labeled, "almost perfect" (p.165). The Fleiss kappa score for all three scorers of this sample was .834 or 83.4% agreement, which demonstrates high reliability.

Limitations

Possible limitations to the coding methodology should be acknowledged. Although strategic efforts were made throughout the process to neutralize the impacts of the researcher's perspective, it is impossible for it to be completely removed. Because the researcher conducted the coder training, she was directly involved in the coding process. It is also always possible for coder

Table 1

Comparison of Cohen's kappa Values for Two Coding Methods

Coding Method	Cohen's kappa (*K*)
Complete Papers	.980
Individual Sentences	.599

bias to influence the results. Replication of the study's methodology and coder training would offer insight into the broader scope of the index's reliability.

Findings and Implications

The results of the inter-rater reliability study show that the IMK is a highly reliable index that could be applied by teachers and administrators. Its high reliability score means that if the sentences being analyzed are considered in their original context (i.e., not coded individually or removed from their original context) and the coders are properly trained, the IMK could potentially offer a vital tool for educators who need to assess the metacognitive complexity of their students.

Earlier in this article, student writing was presented as an example of how difficult it can be to describe the reflective work happening in portfolio introductions. Tables 2 and 3 below offer an illustration of the application of the IMK.

Table 2

IMK Codes Applied to Student A's Writing

Sentence	IMK Code
This semester I wrote three papers.	D2
I did research for one of them, but one was a personal essay so it was really more about stories than quoting.	P4
The other one I didn't like as much.	D1
The argument one.	F
I have never been good at writing arguments and I guess I'm still not!	D1
But I do think that I understand pathos, ethos, and logos better now.	P1
So thank you for that!	T/M
The personal essay was my favorite because I liked talking about my grandmother's baking.	D1

Table 3

IMK Codes Applied to Student B's Writing

Sentence	IMK Code
Of all of our assignments this semester, the research-based argument was my favorite.	D1
I was excited to learn so much about how technology and happiness relate to each other.	D1
Because my little sister was my audience, I spent a lot of time thinking about how to paraphrase my research so that it would be in my own voice.	P4
I think it's important to find ways to connect with a reader on their own terms.	C4

The IMK shows that Student A is working with declarative and procedural knowledge. They demonstrate declarative knowledge in their communication of strengths, weaknesses, and personal preferences and they demonstrate procedural knowledge in their communication of strategies used in writing. Student A also includes a fragmented sentence (F) and a direct address to the reader which is considered transitional/miscellaneous (T/M) because it does not communicate metacognitive knowledge. Student A isn't communicating any conditional knowledge in this text.

The IMK shows that student B is moving more smoothly between all three kinds of metacognitive knowledge. They begin with declarative knowledge, move to procedural details, and end with a theorizing comment that communicates conditional knowledge. Student B demonstrated a pattern that combines specificity and abstraction in a way that characterizes what many teachers consider strong reflective work.

In both cases, the IMK offers language with which to talk to a student about the metacognitive work happening in their writing without necessarily talking about other basic writing elements (style, organization, etc.) at the same time. If a teacher wanted to respond to student A in a way that guides metacognitive growth, they could 1.) show them where their language is communicating declarative and procedural knowledge and encourage them to add detail in those places, 2.) ask questions that might guide the student to make connections required by conditional knowledge, (for example, "why do you think we quote in research papers, but not personal essays?"), and 3.) show the student that the fragmented and transitional/miscellaneous sentences don't add new information in this genre. Similarly, with student B, a teacher could show the student how their observations move well from specific to abstract, then offer a suggestion of either moving to larger abstraction ("how else do writers connect with readers on their own terms") or by moving back down into specifics ("what other strategies do writers use to connect with readers?"). In this way, the IMK not only helps teachers and writers understand the metacognitive moves, but it also becomes a way to leverage and expand the metacognitive growth itself. A student could, in theory, look at their own work and think "I only have declarative work. What could I say about strategies?" Although this study was conducted on portfolio introductions, the IMK offers teachers and students a common language with which to discuss any kind of reflective work throughout the course of the semester.

Applied in this way, the IMK can become a tool for pedagogical transparency and authentic conversation about the thinking work of reflective writing. It makes it possible to respond to reflective work without an intervening and unproductive value judgement about whether a student has made a "hasty" or "genuine" (White, 2007, pp. 180–181) reflection. In a time of rich research and conversation about embodiment (Fleckenstein, 2012) and the impact of the supremacy of white literacies in writing classrooms and college campuses (Inoue, 2015), the IMK could bolster vital questions about language, cultural power, and academic expectations. Because analysis with the IMK focuses solely on the content of student sentences (and not on stylistic conventions), it lends itself well to all students in conversation with the academic standards that center Standard Edited Academic English. For example, in the context of the IMK, we can argue that Standard Edited Academic English is not the default set of writing choices, but instead a set of writing strategies that can be marshaled. Across the scope of a whole class, it can also help teachers see the metacognitive work happening alongside the writing choices in order to make observations about what they are rewarding in this genre of writing. Over time or at a larger scale, it also makes it possible to observe change over time or to gauge the impact of specific pedagogical interventions.

In a more specific sense, the IMK can offer an avenue for writing assessment in the age of AI and generative writing tools. Teachers across many disciplines have begun to reconsider and adapt their use of reflective writing as a measure of accountability by asking students to reflect on their writing choices and strategies in a given assignment. Assessment using the IMK could bolster this use of reflective writing by drawing the students' attention to the powerful and irreplaceable role their own minds play in the creation of scholarly writing.

Conclusions

The field of writing studies has long understood the power of reflective writing in the classroom, but it has not made much concrete progress on articulating how to assess the work assigned in writing classrooms. Yancey (1998) wrote that teachers are really assessing "two (related) performances: the writing performance and the reflecting/self-assessing performance" (p. 147) and concluded "no one really knows what we reward in reflection" (p. 147).

The IMK is presented here as a reliable index from which we can begin to articulate the work of metacognitive knowledge in portfolio introductions. The index offers reliable language that help teachers articulate the "reflecting/self-assessing performance" separate from the "writing performance" (Yancey, 1998, p. 147). Applied to drafts of portfolio introductions, it can offer teachers and students language with which to talk about the content work of reflective writing in order to overtly work toward more metacognitively rich reflective observations. As an important site of learning transfer (Beaufort, 2016; Taczak & Robertson, 2017), it is vital for teachers and researchers to not only be able to explain the work of reflective writing, but to assess and describe the developing complexity of their students' thinking in response to other pedagogical interventions.

References

Bean, J. (2011). *Engaging ideas: The professor's guide to integrating writing, critical thinking, and active learning in the classroom* (2nd ed.). Jossey-Bass.

Beaufort, A. (2016). Reflection: The metacognitive move towards transfer of learning. In K. B. Yancey (Ed.), *A rhetoric of reflection* (pp. 23–41). University Press of Colorado.

Berthoff, A. (2011). Is teaching still possible? Writing, meaning, and higher order reasoning. In V. Villianueva & K. Arola (Eds.), *Cross-talk in comp theory: A reader* (pp. 743–755). NCTE. (Reprinted from "Is teaching still possible? Writing, meaning, and higher order reasoning," 1984, *College English*, 46[8], 743–755).

Bruner, J. (1991). Self-making and world making. *Journal of Aesthetic Education*, 25(1), 66–78.

Camp, R. (2009). Changing the model for the direct assessment of writing. In B. Huot & P. O'Neill (Eds.), *Assessing writing: A critical sourcebook* (pp. 102–130). Bedford/St.Martin's.

Conference on College Composition & Communication. (2015a). Principles for the postsecondary teaching of writing. *Conference on College Composition & Communication*. Retrieved November 6, 2022, from https://cccc.ncte.org/cccc/resources/positions/postsecondarywriting

Conference on College Composition & Communication. (2015b). Principles and practices in electronic portfolios. *Conference on College Composition & Communication*. Retrieved November 6, 2022, from https://cccc.ncte.org/cccc/resources/positions/electronicportfolios

Conference on College Composition & Communication. (2022). Writing assessment: A position statement. *Conference on College Composition & Communication.* https://cccc.ncte.org/cccc/resources/positions/writingassessment

Council of Writing Program Administrators. (2014). WPA Outcomes Statement for First-Year Composition. *Council of Writing Program Administrators.* https://wpacouncil.org/aws/CWPA/pt/sd/news_article/243055/_PARENT/layout_details/false

Council of Writing Program Administrators. (2011). Framework for success in postsecondary writing. *Council of Writing Program Administrators.* https://wpacouncil.org/aws/CWPA/pt/sd/news_article/242845/_PARENT/layout_details/false

Emig, J. (2009). The composing processes of twelfth graders. In S. Miller (Ed.), *The Norton book of composition studies* (pp. 228–251). NCTE. (Reprinted from "The composing processes of twelfth graders," 1971, National Council of Teachers of English).

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. Resnick (Ed.), *The nature of intelligence*. Lawrence Erlbaum Associates.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive development inquiry. *American Psychologist*, 34(10), 906–911.

Fleckenstein, K. (2012). Reclaiming the mind: Eco-cognitive research in writing studies. In L. Nickoson, & M. P. Sheridan (Eds.), *Writing studies research in practice: Methods and methodologies* (pp. 86–97). Southern Illinois University Press.

Flower, L., & Hayes, J. (2011). A cognitive process theory of writing. In V. Villanueva & K. Arola (Eds.), *Cross-talk in comp theory: A reader* (pp. 253–277). NCTE. (Reprinted from "A cognitive process theory of writing," 1981, *College Composition and Communication*, 32[4], 365–387).

Gorzelsky, G., Driscoll, D., Paszek, J., Jones, E., & Hayes, C. (2016). Cultivating constructive metacognition: A new taxonomy for writing studies. In C. M. Anson & J. L. Moore (Eds.), *Critical transitions: Writing and the question of transfer*. WAC Clearinghouse. https://wac.colostate.edu/books/perspectives/ansonmoore/

Hacker, D. J., Keener, M. C., & Kircher, J. C. (2009). Writing is applied metacognition. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 154–172). Taylor and Francis.

Hamp-Lyons, L., & Condon, W. (2007). Questioning assumptions about portfolio-based assessment. In B. Huot & P. O'Neill (Eds.), *Assessing writing: A critical sourcebook* (pp. 315–329). Bedford/St. Martin's.

Harris, K. R., Graham, S., Brindle, M., & Sandmel, K. (2009). Metacognition and children's writing. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 131–153). Taylor and Francis.

Huckin, T. (2009). Content analysis: What texts talk about. In C. Bazerman & P. Prior (Eds.), *What writing does and how it does it: An introduction to analyzing texts and textual practices* (pp. 13–32). Lawrence Erlbaum Associates.

Huot, B. (2002). *(Re)articulating writing assessment for teaching and learning.* Utah State University Press.

Huot, B., & Williamson, M. M. (2007). Rethinking portfolios for evaluating writing: Issues of assessment and power. In B. Huot & P. O'Neill (Eds.), *Assessing writing: A critical sourcebook* (pp. 330–342). Bedford/St. Martin's.

Huot, B., & O'Neil, P. (2009). *Assessing writing: A critical sourcebook*. NCTE.

Inoue, A. B. (2015). *Antiracist writing assessment ecologies: Teaching and assessing writing for a socially just future*. The WAC Clearinghouse, Parlor Press.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. http://www.jstor.org/stable/2529310

Murray, D. (2011). Teach writing as a process not product. In V. Villanueva & K. Arola (Eds.), *Cross-talk in comp theory: A reader* (pp. 3–7). NCTE. (Reprinted from "Teach writing as a process not product," 1972, The Leaflet).

Myers, M. (1980). *A procedure for writing assessment and holistic scoring*. National Council of Teachers of English.

NCTE Executive Committee. (1993). Resolution on Grading Student Writing. *National Council for the Teaching of English*. https://ncte.org/statement/gradingstudentwrit/

NCTE Executive Committee. (2016). Professional Knowledge for the Teaching of Writing. *National Council for the Teaching of English*. Retrieved November 6, 2022, from https://ncte.org/statement/teaching-writing/

Pianko, S. (1979). Reflection: A critical component of the composing process. *College Composition and Communication*, 30(3), 275–278. https://www.jstor.org/stable/356394

Reynolds, N., & Rice, R. (2006). *Portfolio Teaching A Guide for Instructors (2nd ed.)*. Bedford/St. Martin's.

Ryan, K. (2014a). Final portfolio assignment. *Writing course curriculum, Introductory writing: WRIT 101*. The University of Montana.

Ryan, K. (2014b). Portfolio introduction assignment. *Writing course curriculum, Introductory writing: WRIT 101*. The University of Montana.

Scott, B. M., & Levy, M. G. (2013). Metacognition: Examining the components of a fuzzy concept. *Educational Research eJournal*, 2(2), 120-131.

Serra, M. J., & Metcalfe, J. (2009). Effective implementation of metacognition. In D. J. Hacker, J. Dunlosky. & A.C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 278-298). Taylor and Francis.

Sommers, J. (2006). Enlisting the writer's participation in the evaluation process. In R. Straub (Ed.), *Key works on teacher response: An anthology* (pp. 328–335). Boynton/Cook.

Sommers, N. (2011). Revision strategies of student writers and experienced adult writers. In V. Villanueva & K. L. Arola (Eds.), *Cross-talk in comp theory: A reader* (pp. 43–53). (Reprinted from "Revision strategies of student writers and experienced adult writers," 1980, *College Composition and Communication*, 31[4], 378–388).

Stanton, J. G., Sebesta, A. J., & Dunlosky, J. (2021). Fostering metacognition to support student learning and performance. *CBE Life Sciences Education*, 20(2). https://doi.org/10.1187/cbe.20-12-0289

Taczak, K., & Robertson, L. (2017). Metacognition and the reflective writing practitioner: An integrated knowledge approach. In P. Portanova, J. M. Rifenburg, & D. Roen (Eds.), *Contemporary perspectives on cognition and writing* (pp. 211–229). The WAC Clearinghouse, University Press of Colorado.

Tarricone, P. (2011). *The taxonomy of metacognition*. Psychology Press.

Tobias, S., & Everson, H. (2009). The importance of knowing what you know: A knowledge monitoring framework for studying metacognition in education. In D. J. Hacker, J. Dunlosky, & A. C. Grasesser (Eds.), *Handbook of metacognition in education* (pp. 107–127). Routledge.

Waters, H. S., & Schneider, W. (Eds.). (2010). *Metacognition, strategy use, & instruction*. The Guilford Press.

White, E. M. (1985). Holisticism. *College Composition & Communication*, 35, 400–409.

White, E. M. (2005). The Scoring of Writing Portfolios: Phase 2. *College Composition & Communication*, 56(4), 581–600.

White, E. M. (2007). *Assigning, Responding, Evaluating* (4th ed.). Bedford/St. Martin's.

Williams, J. P., & Atkins, J. G. (2009). The role of metacognition in teaching reading comprehension to primary students. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 26-43). Taylor and Francis.

Yancey, K.B. (1998). *Reflection in the writing classroom*. Utah State University Press.

Yancey, K. B. (2007). Looking back as we look forward: Historicizing writing assessment. In B. Huot & P. O'Neill (Eds.), *Assessing writing: A critical sourcebook* (pp. 131–150). Bedford/St. Martin's. (Reprinted from "Looking back as we look forward: Historicizing writing assessment," 1999, *College Composition & Communication*, 50[3], 483–503).

Yancey, K. B. (2016). *A Rhetoric of Reflection*. Utah State University Press.