June 2022

Improving First- and Second-Year Student Writing Using a Metacognitive and Integrated Assessment Approach

Leanne Havis
Neumann University

Improving First- and Second-Year Student Writing Using a Metacognitive and Integrated Assessment Approach

Cover Page Footnote

RW

JOURNAL OF RESPONSE TO WRITING

Improving First- and Second-Year Student Writing Using a Metacognitive and Integrated-Assessment Approach

Leanne Havis

Neumann University

Abstract: Metacognition emphasizes an awareness and understanding of one's thoughts and cognitive processes, along with management of cognition through multiple strategies, including organizing, monitoring, and adapting. Before students can truly become effective writers, they must develop an appreciation for the amount of planning, organization, and revision that a writing assignment requires. In order to improve student writing, the *exam autopsy* approach—an integrated postexam assessment model that draws upon self-assessment, peer review, and instructor feedback—was modified to include metacognitive components for use with essay exams and writing assignments. The current study employed a mixed-methods design with a quasi-experimental, nonequivalent group component across four institutions over 2 semesters, with the fall semester classes (T1) functioning as the control group and the spring semester classes (T2) functioning as the experimental group. During the spring semester of each class, the modified version of the exam autopsy process (EA 2.0) was used between two submissions of student writing (either essay exams or drafts of papers). Students who used the process in lower division classes had significantly higher scores than the control group, but not in upper division classes. Qualitative data analysis reveals some of the reasons behind the observable improvements (or lack thereof) in student writing. These reasons, as well as possible future implications for both teaching and research, are discussed in this article.

Keywords: self-assessment, peer review, metacognition, writing, writing improvement, first-year students

Author Note

The author wishes to thank the faculty members who participated in this research study.

Motivated by a desire to support undergraduate students as they work to plan, monitor, and control their own learning, Owen (2019) developed an integrated postexam self-assessment methodology known as the *exam autopsy* approach. This approach utilizes metacognitive reflective practices and incorporates feedback from peers and faculty members, as well as from the student learners themselves, to facilitate students' self-regulated learning. The underlying aim of the original exam autopsy model was to have students reflect critically on their performance and study skills and, with some direction and prompting from their peers and instructor, modify their learning strategies as needed. After Owen shared the approach with colleagues, several dozen faculty members began implementing it in various classes, and the response has been overwhelmingly positive. Student grades on tests seem to improve when this approach is incorporated into a class. Yet faculty members have consistently expressed an interest in adapting this approach to writing assignments or writing-based exams (e.g., essay tests) in a bid to improve the quality of their students' writing. This interest provided the impetus for the next iteration of the exam autopsy process (EA 2.0) to be tried and tested.

EA 2.0 is comprised of four steps aimed at providing multiple sources of insight and feedback about a sample of student writing, as well as the habits and strategies that went into producing it. The first step involves an initial self-reflection on the part of the student writer, the second seeks input from a peer, the third incorporates faculty comments and suggestions, and the concluding step provides an opportunity for the student writer to integrate all of the aforementioned observations and consider how to approach writing tasks differently in the future. Given that the literature consistently

demonstrates a clear connection between metacognitive practices and writing, two research questions guided this project. The first addressed quantitatively whether the use of EA 2.0 would result in improved student writing. The second research question involved a qualitative analysis of student comments to explore possible reasons as to why student writing did (or did not) improve as a result of EA 2.0.

Literature Review

Metacognition and Writing

Encouraging students to use metacognitive practices to monitor, control, and reflect on their own learning can be an invaluable step in promoting both academic achievement and the acquisition of transferable skills (Zimmerman, 2001). Initially coined by Flavell (1979), metacognition emphasizes an awareness and understanding of one's thought and cognitive processes, along with management of cognition through multiple strategies, including organizing, monitoring, and adapting. Given that writing as a process also involves those selfsame strategies, it is hardly surprising that metacognition and writing are often discussed within the same research frameworks. Writing is occasionally defined as a problem-solving task that requires metacognitive control of planning, text generation, and reviewing (Hayes, 2012). Before students can truly become effective writers, they must develop an appreciation for the amount of planning, organization, and revision that needs to go into a writing assignment. Planning alone extends to making decisions about the content and purpose of the work, the way in which the work should be organized in order to maximize flow and clarity, and necessary details to attend to in editing and revising (Hayes, 2012). Consequently, writing processes involve both cognitive and metacognitive mechanisms.

According to Schraw and Dennison (1994), knowledge of metacognition includes three types of awareness: declarative knowledge, which has to do with people's awareness of their own capabilities and the variables that

may influence their learning performance; procedural knowledge, which focuses on the awareness of how to execute and perform a task; and conditional knowledge, which involves an awareness of those situations in which declarative or procedural knowledge should be applied. Strengthening a sense of self-awareness, including an awareness of one's own strengths and weaknesses as well as an awareness of various appropriate strategies that may be employed if a particular approach is not achieving its desired objective, is a necessary metacognitive step that can ultimately improve student writing.

Numerous studies have investigated whether the use of metacognitive strategies results in better student writing. Wischgoll (2016) found that learners who received an additional self-monitoring-strategy intervention benefited significantly more in terms of acquiring academic writing skills and producing better quality texts than their counterparts who did not receive this intervention. These results seem to underscore the value of introducing students to self-monitoring strategies that they can apply to academic writing. Pacello (2014) suggested this value may lie in students' ability to view their learning as meaningful and relevant, arguing that pedagogical methods stressing a metacognitive-strategy approach to college reading, writing, and learning may help students to understand that learning is a process and, consequently, to consider the course as being connected to their academic, personal, and professional pursuits. Riddell (2015) added that a necessary dimension of the reflective process involves frequent opportunities for practice and feedback, inasmuch as a metacognitive approach to essay writing can provide tremendous benefits for students' writing skills if the approach includes carefully scaffolded assignments that afford opportunities to practice writing and receive feedback.

Student Self-Awareness and Self-Assessment

Promoting self-reflection and self-evaluation among students is a critical first step in the metacognitive process, and encouraging students to reflect on their own written work, as well as on the effectiveness of the steps

they took to produce that work, is a vital piece of the assessment puzzle. Feedback from outside sources (such as peers or faculty members) will carry less weight if the students themselves have not had the opportunity to self-assess. Yet most students (including upper division undergraduates) struggle with monitoring their own efforts (Warkentin & Bol, 1997). This may be attributable to the fact that many students find it challenging to judge their own knowledge and skills accurately (Ambrose et al., 2010), a phenomenon that is especially prevalent among students with weaker knowledge and skills (Dunning, 2007). Falchikov and Boud (1989) conducted a meta-analysis of 48 quantitative self-assessment studies and analyzed differences in student and faculty scores for the same assignment. They found that self-assessment was more accurate (i.e., corresponded more closely to faculty assessment of student performance) among students in upper division courses (as opposed to those in introductory or lower division courses), among students in the sciences (as opposed to those in other disciplinary areas), and in studies with a well-thought-out research design. The implications in the literature are clear: even though it is insufficient as the sole source of data upon which to draw in formulating judgments and evaluations of effective study strategies, self-assessment has definite potential for producing deeper learning (Bercher, 2012) and improvements in student writing (Fung & Mei, 2015; Mazloomi & Khabiri, 2018).

Instructor Feedback

A second source of feedback for students to consider in developing an overall sense of declarative, procedural, and conditional metacognitive knowledge is the faculty member scoring their assignments. Ross (2006) researched concurrence rates between students' self-assessments and teacher (and peer) assessments and determined that students typically rate themselves higher than their instructors rate them (with some exceptions). He contended that this is due to the differential interpretation of evaluation or assessment criteria by students and teachers (Ross,

2006). In other words, students presume they do better than they actually do because they fail to grasp fully the requirements of an assignment or the expectations of a faculty member.

Nonetheless, feedback from faculty members is tremendously valuable. In order for students' writing to improve (and, indeed, for other areas of their study strategies to improve as well), the students need to grasp fully not only what grade they earned on an assignment but also why they earned it, and along with that, what decisions they made that resulted in that grade. Research suggests that students highly value feedback from faculty members insofar as it clarifies why the final grade was awarded (Tehrani, 2018). Instructor feedback is certainly valuable, given that instructors are the ones that create the expectations and grading criteria for an assignment. However, feedback is often frustrating for both the instructor and the student. The former may feel as though the task is laborious (Bean, 2011) and end up spending hours engaging in a form of copyediting, correcting every spelling or grammatical mistake on a student's paper (consequently depriving the student of an opportunity to learn to identify and correct mistakes independently), while the latter may feel either dejected at the sheer volume of corrections that need to be made (Jonsson, 2013) or baffled because the comments noted on the paper (e.g., "avoid overgeneralizing," "don't editorialize") may seem too abstract and vague. Indeed, Bowden (2018) pointed out that students rarely make the changes suggested by faculty members on earlier drafts and noted that much of the literature aimed at uncovering why that is focuses on faculty, rather than student, perspectives on the matter. Poulos and Mahony (2008) sought to provide insight into students' perceptions of effective feedback and noted that effective feedback, from the students' point of view, was the kind that "provides emotional support and facilitates integration into the university" (p. 152). Faculty comments viewed as confusing rather than concrete, then, while offered with the best of intentions, are likely to be dismissed. Moreover, Sommers' (2012) interviews with community college students about the perceived value of instructor comments found

that students appreciated feedback that began with something positive and that had a conversational, rather than normative, tone. Interestingly, students appreciate the same characteristics in peer feedback as well, as discussed in the following section.

It may be that students fail to respond to instructor feedback because it is typically provided in writing, a format that requires students to read and process the comments. Students may not fully comprehend what instructor comments mean. Future course correcting, then, becomes somewhat impossible in this context. Students cannot be expected to change their approach to writing if they fail to understand meaningfully what it is about their approach that needs to be changed. Stannard (2008) suggests that too much written feedback can be biased toward a single learning style, essentially disadvantaging certain students. Further research suggests that providing instructor feedback using multiple modalities, such as audiotaped recordings (Rawle et al., 2018) or face-to-face conferences (Mahmoudi & Bugra, 2020), along with written comments (Bitchener et al., 2005), better meets student needs and produces greater improvements in student writing.

Peer Evaluation and Feedback

Peers constitute a third source of evaluative insight, and peer assessment or evaluation has been found to have significant cognitive and metacognitive benefits (Topping, 1998). In fact, if its chief objective is improving student writing, peer evaluation constitutes an even more integral part of the writing process because feedback and student gains in writing quality have long been strongly and positively correlated with one another (Kuyyogsuy, 2019; T. T. L. Nguyen, 2018). Furthermore, there is extensive evidence in the literature that suggests student writing improves most significantly when feedback is provided early, often, and with a subsequent opportunity for revision, redrafting, or some other form of a "next attempt" (Bean, 2011; Kolb et al., 2013; H. T. Nguyen & Filipi, 2018).

Moreover, the peer evaluation process is often as beneficial for the student providing the feedback as it is for the student receiving it (Li et al., 2010; Liu & Carless, 2006; Topping, 1998), regardless of the quality of the feedback itself. The very act of articulating and identifying a problem, and proposing an appropriate solution to it, forces students to think critically and to engage with the assignment in unprecedented ways (Boud, 1990). They learn to view their own work from the perspective of others (Nicol & Macfarlane-Dick, 2006). Students frequently comment that when they read another student's paper, they develop a better sense of what they need to work on in their own writing (Liu & Carless, 2006; Nicol & Macfarlane-Dick, 2006). Consequently, engaging in peer review promotes the development of students' self-assessment skills (Boud, 1990; DeGrez et al., 2012; Liu & Carless, 2006; Ruggiero & harbor, 2013).

As with instructor feedback, the peer evaluation or assessment process works best when feedback is not provided exclusively in writing but rather is constructed as an ongoing dialogue between teachers, learners, and peers that is individualized and aimed at providing concrete suggestions for improvement (Espasa et al., 2018; Nicol, 2010). To have students participate in this dialogue, peer reviewers must be trained, not only so that feedback can be shared with peers in a constructive, beneficial manner (Van Merrienboer, 1997) but also so that the instructor can ensure all of the students have a clear, uniform understanding of the criteria and expectations associated with scoring the assignment (Leydon et al., 2014). Hadzhikoleva et al. (2019) cautioned that students may be reluctant to comment on their classmates' work for personal reasons and that they may be unable to function effectively as peer reviewers if they fail to understand how to apply the criteria (i.e., a scoring rubric) or what the benchmark is (which reflects the instructor's expectations).

Taken as a whole, the literature establishes that student self-reflection may produce improvements in writing quality but cautions that student self-assessment as a sole means of judging performance may be inadequate

or insufficient. Peer and instructor feedback provide additional useful sources for input and evaluation, but these need to be dialogic in nature and presented using multiple modalities (i.e., not just in writing) in order to be truly meaningful. However, no study to date has triangulated these sources of insight, namely, from self, peer, and faculty, and examined the extent to which doing so might produce improvements in student writing.

Methods

The current study employed a mixed-methods design. The quantitative portion utilized a quasi-experimental, nonequivalent group component across four institutions over 2 semesters, with the fall semester classes (T1) functioning as the control group and the spring semester classes (T2) functioning as the experimental group. Since the assignment of participants to groups (i.e., either to T1 or to T2) was neither controlled by the investigator nor random, it must be acknowledged that there is a possibility the fall and spring sections of each class were not equivalent. However, the investigator sought to account for any confounding variables by choosing groups that were as similar as possible. The student populations in the control group and experimental group for each of the courses chosen are comparable, as evidenced by the lack of statistically significant differences between the means of the first scores students earned on essay exams or writing assignments in those courses. Specifically, there was no significant difference in the grade students in T1 ($M = 77.09$, $SD = 11.49$) and T2 ($M = 73.50$, $SD = 15.12$) earned on the first assignment in the sociology course; $t(46) = 0.94$, $p = 0.35$. Nor was there a significant difference in the grade students in T1 ($M = 76.95$, $SD = 9.10$) and T2 ($M = 77.39$, $SD = 9.91$) earned on their first exam in the criminal justice course; $t(35) = -0.14$, $p = 0.89$. Likewise, there was no appreciable difference in the grade students in T1 ($M = 87.0$, $SD = 6.0$) and T2 ($M = 85.11$, $SD = 7.18$) earned on their first exam in the English course; $t(15) = 0.65$, $p = 0.53$. Finally, there was no significant difference in the grade students in T1 ($M = 87.27$, $SD = 5.27$)

and T2 (M = 85.50, SD = 6.43) earned on the first writing assignment of the psychology course; $t(21)$ = 0.73, p = 0.48. These preliminary analyses suggest that writing levels in each pair of courses were similar at the outset of the semesters under study, and therefore, the student populations in the control group and the experimental group were comparable.

During the spring semester of each class (T2), EA 2.0 was used between two submissions of student work (either two essay exams or successive drafts of the same research paper) in order to determine whether the process resulted in improved student writing. The hypothesis guiding the study predicted that students who completed EA 2.0 would have statistically significant improvements in their writing from one submission to the next. The qualitative portion of the study involved an analysis of written student comments submitted as part of EA 2.0 in order to explore why significant improvement in writing was (or was not) observed.

Participants

Faculty at four institutions were approached and asked to participate in the current study. Given the research focus on improving student writing and developing students' self-regulated learning skills, interested faculty members were asked to select classes that met two criteria for inclusion in the study. First, the class under consideration had to utilize writing-based assignments; that is, either essay exams (at least two, for the purposes of the study) or a research paper (involving the submission of at least two separate successive drafts for feedback). Second, the class under consideration had to be taught by the same faculty member using the same delivery method (i.e., hybrid, face-to-face, or online) two semesters in a row (fall and spring). The fall sections (T1) functioned as the control group; no interventions were introduced, and the classes were taught using the pedagogical approach typically adopted by the particular faculty member. Students in these sections received instructor feedback on their writing when the papers or exams were returned to them but did not spend any time practicing

self-assessment or peer review. Instead, the faculty member continued delivering course content in preparation for the next writing task. The spring sections (T2) functioned as the experimental or treatment group; faculty members teaching the classes used EA 2.0, either between essay exams or between two consecutive drafts of a writing assignment. Ultimately, four courses were selected to participate in the study, representing various disciplines and assorted class levels: a 100-level sociology course, predominantly taken by first-year students to fulfill a general-education core requirement; a 200-level criminal justice course, mainly taken by second-year students in the major; a 300-level English course, comprised of students in their third or fourth year who were taking the course either as an elective or as majors; and a 400-level psychology capstone course, exclusively taken by seniors in the major (see Table 1). In the interests of universality, 100- and 200-level courses will be referred to as "lower division" classes or courses, and 300- and 400-level courses will be referred to as "upper division" classes or courses.

Table 1

Course Characteristics by Class Level, Course Type, and Delivery Format

Course characteristics	SOC T1	SOC T2	CJ T1	CJ T2	ENG T1	ENG T2	PSYCH T1	PSYCH T2
Participants' class levels								
First-year student	19	20	1	1	0	0	0	0
Second-year student	2	3	15	9	1	0	0	0
Third-year student	1	2	2	4	5	6	0	0
Fourth-year student	1	1	3	4	7	3	11	12
Total	23	26	21	18	13	9	11	12

Havis, L. (2022). Improving First- and Second-Year Student Writing Using a Metacognitive and Integrated-Assessment Approach. *Journal of Response to Writing*, 8(1), 4–42.

Course type for participants

Major	1	0	16	10	4	5	11	12
Core	22	25	0	0	1	1	0	0
Elective	0	1	5	8	8	3	0	0
Total	23	26	21	18	13	9	11	12
Delivery format	F2F	F2F	F2F	F2F	H	H	F2F	F2F

Note. "SOC" refers to sociology. "CJ" refers to criminal justice. "ENG" refers to English. "PSYCH" refers to psychology. "F2F" refers to face-to-face courses. "H" refers to hybrid courses.

Materials

In order to test both of the research questions, the four-step EA 2.0 process was used as an intervention. The prompts and questions themselves are presented in Appendix A, and the procedure by which the steps were introduced is described in the following section.

Procedure

IRB approval was secured at each of the participating faculty members' home institutions and a 2-hour training session was held to orient faculty to EA 2.0. Details were provided about the rationale for the approach and about possible challenges faculty could expect to encounter. Particular attention was paid to Step 3 of the process (i.e., providing faculty feedback to the students) and to the peer reviewer training presentation described later. The investigator scheduled time for faculty to participate in active role-playing from a student perspective, engaging with the worksheets and student artifacts from previous semesters.

At the start of each semester under study (both T1 and T2), faculty members distributed informed consent forms to the students in their courses. During T1, students were asked to indicate whether they would consent to share their grades on two essay exams, or two subsequent drafts of the writing assignment selected (with their names redacted to

ensure anonymity). All students in all classes consented to participate in the study. For the protection of their privacy and anonymity, students were assigned random numbers for data entry purposes.

During T2, students were likewise asked to indicate whether they would consent to share their grades on two essay exams or two subsequent drafts of the writing assignment selected (with names redacted to ensure anonymity); however, they were also told that they would be expected to participate in EA 2.0 and that any of the contents of the related worksheets could also be shared with the investigator, albeit anonymously. Once again, all students in all classes consented to participate in the study, and students were assigned random numbers for data entry purposes.

The fundamental difference between T1 and T2 was that, between the first and second essay exams (in the lower division criminal justice class and the upper division English class) or between the first and second drafts of the writing assignment (in the lower division sociology class and the upper division psychology class), the students in T2 participated in EA 2.0. At the start of the class periods in which the first two steps of EA 2.0 took place, faculty members reminded students that they could withdraw from the study at any time. No student across all four of the T2 classes declined to participate at this point. The classes then commenced EA 2.0 using the steps represented in Figure 1 and described in detail in Appendix A.

During the class period when the selected first paper or exam was returned to students, faculty members explained that a postexam or postassignment metacognitive self-assessment would be taking place and that the objective of the self-assessment (Step 1) was for students to think critically about what they had done well and areas in which they could improve; this step was intended to raise awareness around what Schraw and Dennison (1994) described as declarative knowledge. Students were instructed to focus on process over content: They were to concern

Figure 1

The Modified Exam Autopsy Process (EA 2.0)
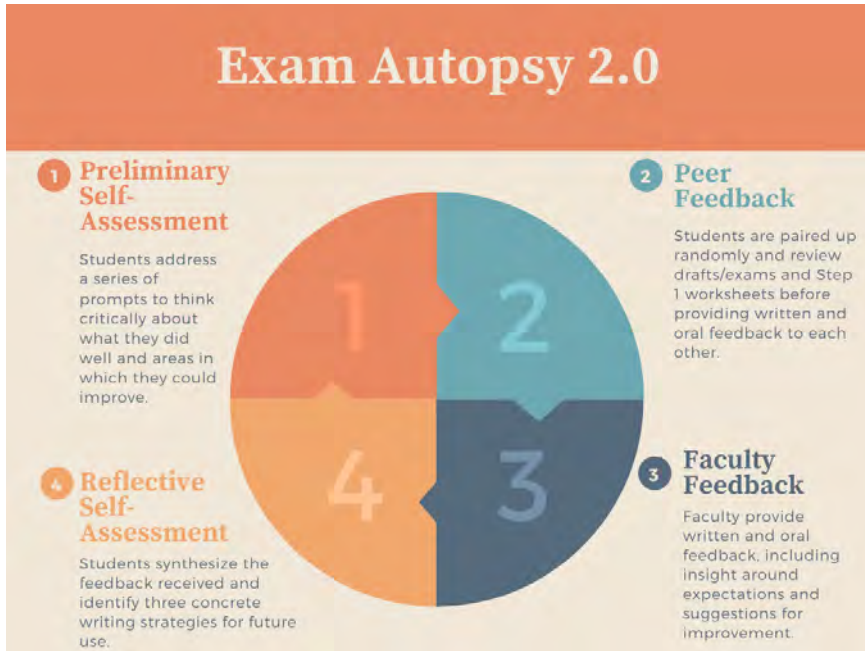
themselves not with whether they got any of the material wrong but rather with how they made sense of, planned out, and executed the assignment (or the exam).

Once students completed Step 1, faculty members spent approximately 30 minutes providing peer reviewer training in accordance with the suggestions made by Van Merrienboer (1997). Faculty members clarified expectations and benchmarks (further capitalizing on an opportunity to provide students insight into how the scoring and grading process takes place), students engaged with the scoring rubric in a norming exercise, and faculty members explicitly and repeatedly reminded students to ensure that

their comments as reviewers were respectful and, above all, constructive; the process of providing feedback was framed as an opportunity not to denigrate or humiliate one another but rather to provide support in a very concrete way. The faculty member also clarified that the peer review/feedback process was not intended to position students (when providing feedback) in the role of an editor; that is, they were not responsible for pointing out sentence-level errors or correcting faulty citations. Instead, they were supposed to consider their classmates' writing holistically and prioritize suggestions for improving their classmates' writing.

Once faculty members' expectations were clarified, Step 2 began in earnest. Students were paired up randomly to exchange exams or drafts (something that was not in the original version of exam autopsy), as well as their Step 1 worksheets, before being asked to address (in written format first and then oral format, in case expansion or elaboration were needed) the Step 2 worksheet prompts. Students shared their feedback with one another in peer conferences that took approximately 20 minutes of class time. This collaboration with peers offered student writers access to what Schraw and Dennison (1994) called procedural and conditional metacognitive knowledge. Before leaving for the day, students turned in both their Step 1 and Step 2 worksheets, attached to one another.

Step 3 afforded faculty members the opportunity to share their unique perspective on students' procedural and conditional knowledge while simultaneously providing insight into their thoughts about the best way to approach the assignment in a low-stakes, nonthreatening way. Additionally, the final question on the worksheet for Step 3 allowed faculty members to model their own conditional knowledge by reflecting on specific steps that they would take if they were in their students' shoes (see Appendix A). Faculty members shared these thoughts with individual students in a 5- to 10-minute face-to-face conference, at which time students were given back all three worksheets (including the written version of their instructor's feedback). Students were then asked to complete Step 4 (see Appendix A).

Two things are noteworthy about Step 4 of EA 2.0. First, unlike the metacognitive self-assessment in Step 1, in which students were asked to engage exclusively in an exploration of their declarative knowledge, this final step required the students to synthesize procedural and declarative knowledge in a way that would allow them to integrate what they had learned into future decision-making processes. Second, in requiring them to identify three concrete strategies to consider implementing in the future (faculty members had only provided them with two), students had to flex their metacognitive skills further as they worked to identify appropriate solutions for some of their own personal challenges. That required a great deal of self-awareness, as well as an appreciation of the assignment's requirements.

EA 2.0 took as little as 1 week for the upper division English class and as long as 2 weeks for the lower division sociology class (partly because of the time commitments involved in meeting individually with each student in the larger class). In the upper division psychology class, the second draft of the paper was due approximately 2 weeks after the first draft had been submitted. In the sociology class, the second draft of the paper was due approximately 3 weeks after the first draft. In the lower division criminal justice class, the second exam was administered approximately 5 weeks after the date of the first exam. In the English class, the second exam was administered approximately 6 weeks after the date of the first exam. Details about the exams and writing assignments are provided in Appendix B.

The same criteria pertaining to writing quality (i.e., context of and purpose for writing, content development, genre and disciplinary conventions, sources of evidence, and control of syntax and mechanics) were applied to all assignments, regardless of whether they were essay exams or drafts of papers. Faculty members participating in the current study scored student papers using the Association of American Colleges and Universities' (AAC&U) written communication VALUE rubric in the

interest of consistency (AAC&U, 2009). Students' scores on these criteria for the first and second exams/drafts, along with the mean scores and the standard deviation, are presented in the following section. This study only evaluated scores associated with the quality of students' writing, so students who wrote well may still have scored poorly overall on an exam if they misunderstood an essay question or failed to answer it altogether.

Percent agreement was used to calculate interrater reliability. Faculty members randomly drew one quarter of the artifacts of student writing from each course (12 for sociology, 10 for criminal justice, 6 for English, and 6 for psychology) and forwarded these to the investigator, who used the AAC&U written communication VALUE rubric to score the assignments and exams as a second independent grader before collecting the faculty members' scores. The investigator and the faculty member teaching the sociology course agreed in 10 out of 12 cases, or 83% of the time. The investigator and the faculty member teaching the criminal justice course agreed in nine out of 10 cases, or 90% of the time. The investigator and the faculty member teaching the English course agreed in four out of six cases, or 67% of the time. Finally, the investigator and the faculty member teaching the psychology course agreed in five out of six cases, or 83% of the time. The mean percent agreement across all four courses was 80.75%.

Results

In order to test the hypothesis related to the first research question guiding this study, namely that students who completed EA 2.0 would have statistically significant improvements in their writing from one submission to the next, two sample (unpaired) t tests assuming unequal variances were conducted separately for each of the four classes to determine whether the mean differences in the T1 and T2 changes in student grades from the first to the second assessment (i.e., exams or drafts of written work) were statistically significant. The alpha level was set at 0.05.

For the lower division sociology class, there was a significant difference in the grade changes between T1 (M = 7.0, SD = 3.87) and T2 (M

= 10.8, SD = 3.87); $t(47)$ = −3.28, $p < 0.01$ (see Table 2). In the lower division criminal justice class, there was also a significant difference in the grade changes between T1 (M = 5.7, SD = 3.20) and T2 (M = 8.4, SD = 3.62); $t(34)$ = −2.47, p = 0.02 (see Table 3). In these two classes, the p value was less than the alpha value of 0.05, so the null hypothesis could be rejected, and EA 2.0 could be assumed to be effective in improving student writing.

Table 2

Grade Changes in SOC T1 (no EA 2.0) and SOC T2 (EA 2.0)

Statistics	T1 grade change	T2 grade change
M	7.043	10.808
Variance	14.953	17.362
Observations	23	26
Hypothesized mean difference	0	
df	47	
t	−3.2880	
p one-tail	0.001	
t critical one-tail	1.679	
p two-tail	0.002	
t critical two-tail	2.012	

Table 3

Grade Changes in CJ T1 (no EA 2.0) and CJ T2 (EA 2.0)

Statistics	T1 grade change	T2 grade change
M	5.667	8.389
Variance	10.233	13.075
Observations	21	18

Havis, L. (2022). Improving First- and Second-Year Student Writing Using a Metacognitive and Integrated-Assessment Approach. *Journal of Response to Writing, 8*(1), 4–42.

Hypothesized mean difference	0
df	34
t	−2.471
p one-tail	0.009
t critical one-tail	1.691
p two-tail	0.020
t critical two-tail	2.032

However, in the upper division English class, there was no significant difference in the grade changes between T1 ($M = 3.0$, $SD = 1.63$) and T2 ($M = 3.67$, $SD = 0.71$); $t(17) = -1.31$, $p = 0.21$ (see Table 4). Nor was there a significant difference in the grade changes between T1 ($M = 2.3$, $SD = 1.42$) and T2 ($M = 3.0$, $SD = 1.54$) in the upper division psychology class; $t(21) = -1.18$, $p = 0.25$ (see Table 5). In these two classes, since the p value was not less than the alpha value of 0.05, the null hypothesis could not be rejected.

Table 4

Grade Changes in ENG T1 (no EA 2.0) and ENG T2 (EA 2.0)

Statistics	T1 grade change	T2 grade change
M	3	3.667
Variance	2.667	0.5
Observations	13	9
Hypothesized mean difference	0	
df	17	
t	−1.306	
p one-tail	0.105	
t critical one-tail	1.740	

Havis, L. (2022). Improving First- and Second-Year Student Writing Using a Metacognitive and Integrated-Assessment Approach. *Journal of Response to Writing*, 8(1), 4–42.

p two-tail	0.209
t critical two-tail	2.110

Table 5

Grade Changes in PSYCH T1 (no EA 2.0) and PSYCH T2 (EA 2.0)

Statistics	T1 grade change	T2 grade change
M	2.273	3
Variance	2.018	2.364
Observations	11	12
Hypothesized mean difference	0	
df	21	
t	−1.179	
p one-tail	0.126	
t critical one-tail	1.721	
p two-tail	0.252	
t critical two-tail	2.080	

Given that EA 2.0 appeared to have a differential impact on lower division and upper division classes (i.e., mean differences in grade changes for the lower division courses were significant, but those for the upper division courses were not), it became necessary to test whether a correlation existed between grade changes and class level (i.e., first-year, second-year, etc.).

For the purposes of calculating Spearman's rho, *x* was an individual's grade change and *y* was the individual's class level (i.e., a first-year student was scored as *1*, second-year scored as *2*, third-year as *3*, and fourth-year as *4*). In the sociology T1 and T2 classes, the association between the two variables was statistically significant, $r_s = 0.638$, *p* (two-tailed) = 0.001 and $r_s = 0.731$, *p* (two-tailed) < 0.001, respectively. In the criminal justice T1

Havis, L. (2022). Improving First- and Second-Year Student Writing Using a Metacognitive and Integrated-Assessment Approach. *Journal of Response to Writing, 8*(1), 4–42.

and T2 classes, the association between the two variables was statistically significant, $r_s = -0.770$, p (two-tailed) < 0.001 and $r_s = -0.904$, p (two-tailed) $= 0$, respectively. In the English T1 and T2 classes, the association between the two variables also was statistically significant, $r_s = -0.807$, p (two-tailed) < 0.001 and $r_s = -0.750$, p (two-tailed) $= 0.020$, respectively. Given that all of the students in the psychology T1 and T2 classes were fourth-year students, the Spearman correlation coefficient could not be calculated because there would have been a division by zero; thus, the correlation coefficient r for these data is undefined.

In order to investigate the second research question guiding this study, namely, why EA 2.0 did (or did not) result in improved student writing, written comments from among the four steps were analyzed to provide some valuable insight into students' thinking. For example, all nine students in the T2 English class (100.0%) and 11 out of 12 students in the T2 psychology class (91.7%) reported in Step 1 that the grade they earned on the exam/assignment was what they were expecting to earn. Five out of nine students in the T2 English class (55.6%) and 10 out of 12 students in the T2 psychology class (83.3%) commented in Step 4 that they "knew what [they] needed to do but just didn't do it this time" (with minor alterations in the exact wording). These trends seem to reinforce the notion that upper division students may already possess the necessary metacognitive and cognitive skills for effective writing but simply lack the motivation (or time) to implement appropriate strategies. In contrast, only seven out of 26 students in the T2 sociology class (26.9%) and four out of 18 students in the T2 criminal justice class (22.2%) reported in Step 1 that the grade they earned was what they were expecting to earn. None of the student comments in Step 4 in either the T2 sociology or the T2 criminal justice class mentioned already knowing what needed to be done and simply choosing not to act on it. While it is impossible to intuit accurately why students may or may not have made certain comments on their worksheets, the patterns are interesting and seemingly suggest

that first- or second-year students are more in need of exposure to (and opportunities to practice) metacognitive strategies.

This need is perhaps illustrated most concretely in the area of student study habits. When asked how they spent their time preparing for the essay exam or working on the assignment, the students in the lower division classes typically did not mention active learning strategies, such as practicing writing out essay responses or organizing their thoughts with an outline before beginning to write. Step 1 comments around this topic from the T2 sociology group included such statements as "I made sure I read all of the assigned readings before I started," "I reviewed the PowerPoints the teacher gave us," and "I looked at the textbook again to make sure I had all the information I needed to know." Interestingly, many of the same students who referred to these types of study strategies noted that they spent "a lot" of time studying or preparing; time estimates ranged from 2 to 6 hours, and several other students used descriptors such as "lots of time" or "more time than I did for other classes" rather than providing specific amounts. This pattern was also true for the T2 criminal justice group, whose Step 1 comments included such insights as "I studied for a really long time" and "I spent ages preparing for this" while describing such strategies as "I looked at the review questions at the end of the chapter" and "I printed out my PowerPoint slides so I had them to go over."

The fact that lower division students' strategies focused on expending time rather than energy and utilizing passive, not active, techniques was something that both peers and faculty picked up on. In Step 2, several students in both the T2 sociology group and the T2 criminal justice group wrote to their partners some variation of the following: "I think you studied for a really long time but I don't know that what you did was enough" (example student comments are quoted exactly as written). Unfortunately, since their peers lacked the same prior exposure to reflective practice as the student authors themselves, they were often unlikely to suggest an appropriate way to pivot. That is, they could identify that it appeared their

classmates were not doing "enough," but they did not know what "enough" preparation might look like or what form it might take. In Step 3, when faculty were asked to express an opinion of the methods students used to prepare or study, suggestions were more concrete, as in: "It sounds like you spent a lot of time reading, which is great, but how did you test your knowledge to make sure you understood the material? Did you prepare an outline? Did you have a thesis statement in mind? How did you plan to support your argument?"

In contrast, students from the upper division courses were less likely to mention in Step 1 that they studied for a long time and more likely to mention that they used active, rather than passive, learning strategies. For example, in the T2 English group, comments included, "I knew the topics we would have to pick from and I made sure I had at least a thesis statement worked out for the questions I planned to answer" and "I found it helpful to outline some thoughts ahead of time, not so I would have to memorize anything but so I could remember where I wanted the essay to go." Likewise, in the T2 psychology group, one student wrote, "[I prepared by] thinking through ahead of time what I wanted to write in response to the prompts and then getting organized so I could back up my answers with evidence from the case."

Analysis of the comments provided in Step 2 suggests that students in lower division classes differ from their peers in upper division classes not only in study strategies but also in the nature of the constructive criticism they provided. Peer comments in the lower division classes focused more on sentence-level, local editing issues (e.g., two students in the T2 sociology group said, "You have a lot of typos" and "You should spell check more carefully") while comments in the upper division classes bypassed those types of issues and centered more around providing constructive criticism relating to flow, clarity, and global editing (e.g., two students in the T2 English group commented, "This could move down here, it would flow better" and "Help the reader see where you're coming from with more examples or explanations." Students from the T2 psychology

group provided similar feedback: "Provide more support through examples" and "You seem to contradict yourself"). Based on the peer review comments, students in lower division classes seem to believe that good writing is writing that is mechanically correct and free of sentence-level errors (despite having received peer reviewer training to the contrary), while students in upper division classes recognize that good writing benefits from coherence, cohesion, and the construction of a sound argument.

Discussion

The fact that the EA 2.0 intervention resulted in statistically significant improvement in student writing in the lower division courses and not in the upper division courses merits further exploration. The distribution of the mean grade changes is far more spread out in the lower division courses and more closely clustered around the mean in the upper division courses (as shown in Tables 2–5). This may have to do with the level of the courses but may also be attributable to the nature of the courses, the types of writing tasks assigned, and the faculty members' pedagogical approaches. However, given the greater likelihood that students taking upper division courses (a) take them as a major requirement (as opposed to an elective) and (b) are more likely to be upper division students (who would have had previous exposure to some of the disciplinary content or, at the very least, to expectations surrounding writing in the discipline), the level and nature of a course are presumed to be key factors associated with the effectiveness of EA 2.0. One upper division class was an English seminar with many English majors who, according to the faculty member teaching the course, were accustomed to reading and critiquing one another's work, revising, and redrafting. Likewise, the upper division psychology class was a capstone seminar, which presupposes that students have already taken most (if not all) of their required courses and are well positioned to write a final integrative paper. Students in upper division classes did not have as much meaningful improvement in their writing because the incoming students were (in both semesters)

already fairly familiar with the writing tasks expected of them and with the faculty member's scoring criteria. There was little EA 2.0 could add to what they had already learned about the writing process. These results regarding the skill levels of upper level students confirm findings by Falchikov and Boud (1989), who determined that upper division students were better able to self-assess than their lower division peers.

The finding about the differential skill levels of upper and lower level students is further borne out by the fact that the students in the upper division courses started off (with their first draft or first essay exam) at a higher grade than their counterparts in the lower division courses, so there was less room for them to improve on their performance. In contrast, there was a great deal more variability among the grades earned on first drafts (or first exams) by students in the lower division classes, which may account for the fact that the actual grade change from one draft (or exam) to the next was, in some cases, far more dramatic. What is encouraging is that students in those lower division classes that had more room for improvement appeared to benefit significantly more from the intervention, a finding that has critical implications for faculty members teaching lower division classes that are predominantly populated by first- and second-year students.

Embedding a metacognitive approach to self- and peer-assessment early in the university experience could be highly effective in promoting better writing and better study habits in younger students, as indicated by Wischgoll (2016). There are two fundamental reasons for this. First, as previously stated, many students are poor judges of their own knowledge and skills (Ambrose et al., 2010), especially if those knowledge and skills are somewhat lacking (Dunning, 2007). Qualitative analysis of the data in this study supports these findings when it comes to students' perceptions of their own study habits. Generally speaking, students in the lower division classes believed studying for longer was a key to success, while students in the upper division classes believed it was less about the amount of time spent and more about the organization and thoughtfulness that went

into the process of studying. This difference clearly indicates a need for students in introductory or lower division classes to understand that not all study strategies are equally effective and that what may have worked for them in high school may not be helpful in higher education. Forcing them to reflect and self-monitor is a critical component of prompting them to self-evaluate and recognize what changes they may need to make moving forward.

Second, providing an opportunity to engage in reflective practice could also help students in lower division classes gain greater insight into faculty expectations pertaining to the writing process, a factor identified by Ross (2006). To improve student understanding of writing tasks and reinforce the organizational and evaluative aspects of metacognitive practices, instructors could not only ask students to self-reflect individually but also work through the trends in the peer comments from Step 2 as a group to unpack the importance of certain elements in a writing sample. As noted by Poulos and Mahony (2008) and Pacello (2014), students appreciate and benefit from feedback that is practical, meaningful, and concrete rather than vague, ambiguous, or subject to interpretation. EA 2.0 provides a useful opportunity for discussing collaboratively what constitutes good writing so that students in lower division classes might recognize, as their peers in upper division classes do, that good writing is less about local, sentence-level mechanics and syntax and more about coherence, cohesion, and the construction of a sound argument, which all stem from a holistic, global approach to planning and editing.

Conclusion

The findings of this study have meaningful implications both for teaching and for future research. Utilizing an approach like EA 2.0 in class between two essay exams or two drafts of a writing assignment could serve to clarify a number of issues, including the role of college students, an instructor's expectations relating to scoring and grading, and the comparative effectiveness of certain study strategies over others. Moreover, by

collaborating with students to promote self-monitoring, self-evaluation, and self-reinforcement in a safe, low-stakes environment, faculty have a unique opportunity to model for students what the practice of providing feedback looks like. In doing so, they can highlight that a critical component of lifelong learning is not that mistakes never happen but rather that it is possible to identify them, learn from them, and pivot accordingly.

This study has two key limitations. First, although courses were deliberately chosen in four separate disciplines and across four different institutions, with every effort taken to ensure equivalence and comparability among lower division students and likewise among upper division students, it is possible that the student populations were not as alike as initially believed. Perhaps certain incoming student populations were underprepared, while others had received more extensive writing instruction in high school. Another researcher in the future might wish to compare apples to apples, as it were, by employing the same methodological design across courses within the same institution (or even within the same program). Second, faculty members self-identified and volunteered to participate in this study, a fact which could potentially have compromised the study's internal validity. The willingness to employ a new pedagogical approach in the classroom may have stemmed from some underlying factor (i.e., either a commitment to innovation that suggests enthusiasm for focusing on certain skills or, conversely, a desire to revamp an entire course because the status quo was deemed to be ineffective), and it may be that this unknown factor, rather than EA 2.0 itself, ultimately contributed to the observed results.

Several opportunities exist for further inquiry into the effectiveness of EA 2.0. It would be interesting to note whether there are differences along gender lines when it comes to the effectiveness of EA 2.0. Alternatively, it might be useful to design a more long-term study that could track students who participated in EA 2.0 during a lower division class, ideally during the first semester of their college experience, and those who did not, in order to determine whether there are significant differences in

terms of either their overall academic performance (relating to study habits that may or may not have improved) or their subsequent writing assignments in other courses (relating to writing skills that may or may not have improved). Another longitudinal project could attempt to introduce EA 2.0 to students before they enter higher education, possibly during middle school or high school (with some modifications to make the process developmentally appropriate), in order to track whether students with greater exposure to metacognitive reflective practice of this type are more likely to succeed in college. Embedding opportunities for students to engage in metacognitive reflective practice can have significant positive implications for both their writing specifically and for their academic performance overall. Such a practice promotes the development of critical skills for students to master while in college so they can use them in the professional world when they graduate.

Havis, L. (2022). Improving First- and Second-Year Student Writing Using a Metacognitive and Integrated-Assessment Approach. *Journal of Response to Writing, 8*(1), 4–42.

References

Ambrose, S. A., Bridges, M. W., DiPietro, M., Lovett, M. C., & Norman, M. K. (2010). *How learning works: Seven research-based principles for smart teaching*. Jossey-Bass.

Association of American Colleges and Universities. (2009). *Written communication VALUE rubric*. https://www.aacu.org/value/rubrics/inquiry-analysis

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

Bercher, D. A. (2012). Self-monitoring tools and student academic success: When perception matches reality. *Journal of College Science Teaching*, *41*(5), 26–32.

Bitchener, J., Young, S., & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing*, *14*(3), 191–205. https://doi.org/10.1016/j.jslw.2005.08.001

Boud, D. (1990). Assessment and promotion of academic values. *Studies in Higher Education*, *15*(1), 101–111. https://doi.org/10.1080/03075079012331377621

Bowden, D. (2018). Comments on student papers: Student perspectives. *Journal of Writing Assessment*, *11*(1), 1–17. Retrieved from http://journalofwritingassessment.org/article.php?article=121

DeGrez, L., Valcke, M., & Roozen, I. (2012). How effective are self- and peer assessment of oral presentation skills with teachers' assessments? *Active Learning in Higher Education*, *13*(2), 129–142. https://doi.org/10.1177%2F1469787412441284

Dunning, D. (2007). *Self-Insight: Roadblocks and detours on the path to knowing thyself*. Taylor & Francis.

Espasa, A., Guasch, T., Mayordomo, R. M., Martínez-Melo, M., & Carless, D. (2018). A dialogic feedback index measuring key aspects of feedback processes in online learning environments. *Higher Education Research & Development*, *37*(3), 499–513. https://doi.org/10.1080/07294360.2018.1430125

Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research*, *59*(4), 395–430. https://doi.org/10.2307/1170205

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

Fung, Y. M., & Mei, H. C. (2015). Improving undergraduates' argumentative group essay writing through self-assessment. *Advances in Language and Literacy Studies*, *6*(5), 214–224.

Hadzhikoleva, S., Hadzhikolev, E., & Kasakliev, N. (2019). Using peer assessment to enhance higher order thinking skills. *TEM Journal*, *8*(1), 242–247. https://doi.org/10.18421/TEM81-34

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication*, *29*(3), 369–388. https://doi.org/10.1177%2F0741088312451260

Jonsson, A. (2013). Facilitating productive use of feedback in higher education. *Active Learning in Higher Education*, *14*(1), 63–76. https://doi.org/10.1177%2F1469787412467125

Kolb, K. H., Longest, K. C., & Jensen, M. J. (2013). Assessing the writing process: Do writing-intensive first-year seminars change how students write? *Teaching Sociology*, *41*(1), 20–31. https://doi.org/10.1177%2F0092055X12448777

Kuyyogsuy, S. (2019). Promoting peer feedback in developing students' English writing ability in L2 writing class. *International Education Studies*, *12*(9), 76–90.

Leydon, J., Wilson, K., & Boyd, C. (2014). Improving student writing abilities in geography: Examining the benefits of criterion-based assessment and detailed feedback. *Journal of Geography*, *113*(4), 151–159. https://doi.org/10.1080/00221341.2013.869245

Li, L., Liu, X., & Steckelberg, A. L. (2010). Assessor or assessee: How student learning improves giving and receiving feedback. *British Journal of Educational Technology*, *41*(3), 525–536. https://doi.org/10.1111/j .1467-8535.2009.00968.x

Liu, N. F., & Carless, D. (2006). Peer feedback: The learning element of peer assessment. *Teaching in Higher Education*, *11*(3), 279–290. https://doi .org/10.1080/13562510600680582

Mahmoudi, F., & Bugra, C. (2020). The effects of using rubrics and face to face feedback in teaching writing skill in higher education. *International Online Journal of Education and Teaching*, *7*(1), 150–158.

Mazloomi, S., & Khabiri, M. (2018). The impact of self-assessment on language learners' writing skill. *Innovations in Education and Teaching International*, *55*(1), 91–100. https://doi.org/10.1080/14703297.2016 .1214078

Nguyen, H. T., & Filipi, A. (2018). Multiple-draft/multiple-party feedback practices in an EFL tertiary writing course: Teachers' and students' perspectives. *International Education Studies*, *11*(8), 1–26.

Nguyen, T. T. L. (2018). The effect of combined peer-teacher feedback on Thai students' writing accuracy. *Iranian Journal of Language Teaching Research*, *6*(2), 117–132.

Nicol, D. (2010). From monologue to dialogue: Improving written feedback processes in mass higher education. *Assessment and Evaluation in Higher Education*, *35*(5), 501–517. https://doi.org/10.1080/026029 31003786559

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback.

Studies in Higher Education, *31*(2), 199–218. https://doi.org/10.1080/03075070600572090

Owen, L. (2019). The exam autopsy: An integrated post-exam assessment model. *International*

Journal for the Scholarship of Teaching and Learning, *13*(1), Article 4. https://doi.org/10.20429/ijsotl.2019.130104

Pacello, J. (2014). Integrating metacognition into a developmental reading and writing course to promote skill transfer: An examination of student perceptions and experiences. *Journal of College Reading and Learning*, *44*(2), 119–140. https://doi.org/10.1080/10790195.2014.906240

Poulos, A., & Mahony, M. J. (2008). Effectiveness of feedback: The students' perspective. *Assessment & Evaluation in Higher Education*, *33*(2), 143–154. https://doi.org/10.1080/02602930601127869

Rawle, F., Thuna, M., Zhao, T., & Kaler, M. (2018). Audio feedback: Student and teaching assistant perspectives on an alternative mode of feedback for written assignments. *The Canadian Journal for the Scholarship of Teaching and Learning*, *9*(2), Article 2. https://doi.org/10.5206/cjsotl-rcacea.2018.2.2

Riddell, J. (2015). Performance, feedback, and revision: Metacognitive approaches to undergraduate essay writing. *Collected Essays on Learning and Teaching*, *8*, 79–96. https://doi.org/10.22329/celt.v8i0.4256

Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research, & Evaluation*, *11*, Article 10. https://doi.org/10.7275/9wph-vv65

Ruggiero, D., & harbor, j. (2013). Using writing assignments with calibrated peer review to increase engagement and improve learning in an undergraduate environmental science course. *International Journal for the Scholarship of Teaching and Learning*, *7*(2), Article 21. https://doi.org/10.20429/ijsotl.2013.070221

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, *19*(4), 460–475. https://doi.org/10.1006/ceps.1994.1033

Sommers, N. (Director). (2012). *Beyond the red ink: Students talk about teachers' comments* [DVD]. Bedford St. Martin's.

Stannard, R. (2008). A new direction in feedback. *Humanising Language Teaching*, *10*(6). http://old.hltmag.co.uk/dec08/mart04.htm

Tehrani, F. A. (2018). Feedback for writing or writing for feedback? *Journal of Language and Linguistic Studies*, *14*(4), 162–178.

Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, *68*(3), 249–276. https://doi.org/10.3102%2F00346543068003249

Van Merrienboer, J. J. G. (1997). *Training complex cognitive skills.* Educational Technology Publications.

Warkentin, R. W., & Bol, L. (1997, April). *Assessing college students' self-directed studying using self-reports of test preparation* [Paper presentation]. Annual Meeting of the American Educational Research Association, Chicago, IL.

Wischgoll, A. (2016). Combined training of one cognitive and one metacognitive strategy improves academic writing skills. *Frontiers in Psychology*, *7*, Article 187. https://doi.org/10.3389/fpsyg.2016.00187

Zimmerman, B. J. (2001). Theories of self-regulated learning and academic achievement: An overview and analysis. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-Regulated Learning and Academic Achievement* (2nd ed., pp. 1–38). Erlbaum.

Appendix A

Exam Autopsy 2.0 Description

Step 1

Students were given a self-reflection worksheet with the following questions:

- How did your actual grade on this exam/assignment compare with the grade you expected? How do you explain the difference, if there is any?
- How do you feel about your exam/assignment grade? Are you surprised, pleased, relieved, disappointed, or what?
- How many hours did you spend preparing for this exam/working on this assignment? Was this enough time to get the grade you wanted, or should you have spent more time preparing/working?
- How did you spend your time preparing for the essay exam/working on the assignment? (For instance, did you practice writing out essay responses? Did you organize your thoughts with an outline before beginning to write? Did you ensure that you had a clear thesis statement and that you were prepared to make all supporting points persuasively?)*
- Examine the items on which you lost points and look for patterns. Did you include an introduction and conclusion? Did you follow a clear and logical structure? Did you take the time to edit your work and proofread carefully for grammatical and spelling errors? Were all the necessary citations included and properly formatted according to the style guidelines indicated by your professor?*
- Set a goal to get a certain percentage or score in the next exam/draft. What study strategies and schedule will enable you to earn that score?

Step 2

Students were given a worksheet during peer review with the following questions:

- Do you agree with your partner's assessment of how and why s/he earned a different grade than expected? Why or why not?
- Any and all feelings your partner may express about his/her exam grade are valid. What words of wisdom or comfort could you share in light of how s/he feels?
- What can you highlight that s/he did well?*
- What is your opinion of the time your partner spent studying for this test/preparing for this assignment?
- What is your opinion of the methods your partner used in studying for this test/preparing for this assignment?
- What is your opinion of your partner's assessment of the items s/he got wrong, or where s/he lost points? Do you have another interpretation of or explanation for what might have happened?
- What do you think of the goals that your partner has set out for him/herself? Are they realistic? What are two additional ideas you could suggest to help him/her achieve those goals?

Step 3

Faculty members were given a worksheet with the following questions:
- Do I agree with your assessment of why you got a different grade than expected? Why or why not?
- Any and all feelings you may express about your exam grade are valid. What words of wisdom or comfort could I share in light of how you feel?
- What is my opinion of the time you spent studying for this test/preparing for this assignment?
- What is my opinion of the methods you used in studying for this test/preparing for this assignment?
- What is my opinion of your assessment of the items you got wrong, or where you lost points? Do I have another interpretation of or explanation for what might have happened?

Havis, L. (2022). Improving First- and Second-Year Student Writing Using a Metacognitive and Integrated-Assessment Approach. *Journal of Response to Writing*, 8(1), 4–42.

- What do I think of the goals that you have set for yourself? Are they realistic? What are <u>two</u> additional ideas I could suggest to help you achieve those goals?

Step 4

Students were given the following instructions and prompts:

> Think about your original answers to the self-assessment questions, as well as the feedback that you received from your partner and from me [the faculty member]. In a brief paragraph, write down what, if anything, has changed in terms of how you prepared for the first test or assignment and how you plan to prepare for the next one. Be concrete and specific in describing <u>at least three strategies</u> that you plan to use to study for (or take) the next test (or to organize and plan for your next paper draft). Why do you think those strategies are the most promising for you? What can I do to help support your learning and your preparation for the next exam/assignment?

Note. An asterisk [*] denotes a question that was modified substantively from the original exam autopsy approach.

Appendix B

Description of Writing Tasks

For the lower division sociology course, students were assigned a paper. The paper prompt was as follows:

> In a 3- to 5-page paper, analyze either a movie poster or a movie trailer through the lens of race, gender, and/or social class introduced in this course. You are encouraged to be creative and have some fun with this, but I want to see you apply (and not just mention) specific readings we have covered in class this semester. Make sure you cite your sources appropriately using APA format. You are encouraged to discuss your topic with me sooner rather than later to ensure that you have chosen something appropriate.

For the lower division criminal justice course, students were instructed to choose two essay questions to answer on a timed, 90-minute exam.

Students chose from the following three questions:

- Discuss psychological theories of crime. You should make sure you address the focus of this group of theories and describe, in detail, each of the theories covered in class (including psychoanalytic theory, maternal attachment theory, and moral development theory) and their provenance. Your descriptions should explicitly state how each of these theories explain criminal behavior.
- Describe the changes brought on by industrialization, urbanization, and immigration in the United States during the late 19th and early 20th centuries and explain how these related to the development of the Chicago School and its approach to criminology. Your answer should begin with a historical overview and end with a discussion of social disorganization, concentric zone theory, and the differences between Durkheim's concept of anomie and Merton's concept of anomie-strain.

- Explain how social process perspectives view criminal behavior as the product of interpersonal dynamics. Your answer should include a comparison between social process and social structural perspectives, as well as a detailed discussion of such theories as differential association, differential identification, differential reinforcement, and labeling.

For the upper division English course, students were instructed to answer five short essays on a timed, 3-hour exam.

The faculty member selected questions randomly from a pool in order to minimize the likelihood that two students would receive the same test with the same questions. The course introduced students to contemporary nature writing and environmental literature through three genres: poetry, fiction, and creative nonfiction. All assigned texts focused on the natural world and humans' relationships with it. Exam questions asked students to grapple with such topics and questions as the following:

- The role of literature in how we perceive and conceptualize nature;
- Earth as a literary setting and stage but also as a habitat;
- Utopias versus dystopias;
- Sense of place;
- Pollution, climate change, the fossil and postfossil fuel economies, and other environmental problems and potential catastrophes;
- Notions of private versus communal property;
- Mammals, birds, and concepts of the more-than-human world;
- Extirpation and extinction of animal and plant species;
- Notions of wilderness and wilder places;
- Settlers and nomads;
- Observation and contemplation of nature;
- Nature and silence;
- The relationship between landscape and story; and
- Ethics, environmental activism, and questions of responsibility to the earth

Havis, L. (2022). Improving First- and Second-Year Student Writing Using a Metacognitive and Integrated-Assessment Approach. *Journal of Response to Writing, 8*(1), 4–42.

For the upper division psychology course, students were assigned a case study.

After reading the case study carefully, students were given the following instructions:

> Using your knowledge of risk assessment, mental disorders and offending behavior, and interview and treatment strategies, answer the following questions:

- Describe the type(s) of mental disorder Mr. D may be suffering from.
- Consider whether those disorders are likely to contribute to the risk he poses of future violence.
- Identify those risks that Mr. D poses to himself and others.
- Consider whether you would discharge Mr. D from the hospital at this time and give your reasons why.
- Highlight what challenges Mr. D may pose in treatment and how you might overcome them.