Students' Use of Note Sheets on Accounting Exams

Terry Baker
Wake Forest University

Carol Cain
Winston-Salem State University

George Violette
University of Maine

Students' use of note sheets (or, "cheat sheets") during exams is an instructional design choice for accounting instructors. We find that accounting students invest significant effort in the preparation of note sheets before exams and that they increase this investment for a subsequent exam. We conclude that students perceive value in and increase the quality of note sheets with succeeding exams. Student gender and prior exam-performance appear to be significant factors. Our results could raise a question of equity for instructors who choose to include note sheets in their instructional design.

INTRODUCTION

An important choice in the design of exams and other assessments is whether to allow students access to books or notes during the assessment ("closed-book" versus "open-book," Durning et.al., 2015). Within that design choice, if the instructor allows access to books or notes, the next question is how broad that access may be. One narrow approach is to allow the use of "cheat sheets" or "crib sheets" as exam-taking aids. Hamouda and Shaffer (2016) define such aids as a limited set of notes that are (1) created by the student and (2) explicitly allowed by the instructor as a reference during an exam. For simplicity, we offer the term "note sheets" as a way of encompassing the Hamouda-Shaffer definition, but avoiding potentially pejorative connotations of the term cheat sheets or crib sheets. Although commonly used, the phrases "cheat sheet" or "crib sheet" could connote the use of a prohibited aid. For clarity, we adopt the term "note sheet" throughout the manuscript. The impact of note sheets on student performance has been studied in disciplines such as economics, psychology, computer science, and the health sciences. However, little if any research has been conducted on their use in business or accounting courses. As such, the extent to which accounting students invest in preparing note sheets, or rely on them during exams, is unknown. Note sheets appear to be used as an approved exam reference in certain types of accounting courses. An informal survey of the authors' colleagues indicates that about 30% of the accounting faculty at our institutions routinely allow and encourage the use of note sheets. The use of note sheets seems to be concentrated in courses in taxation, financial statement analysis, and MBA-level financial and managerial accounting.

We believe research on the use of note sheets is important for several reasons. First, although note sheets appear to be common in certain types of courses, little is known about how students prepare or use them. For example, one debate documented in the literature from other disciplines is whether the benefit of note sheets is primarily a learning effect from pre-exam preparation or an exam-taking effect, such as improved recall or reduction of anxiety. Second, our anecdotal evidence suggests that an instructor's choice to allow note sheets is autonomous and course-specific rather than guided by a department- or school-level policy. As such, issues of equity across instructors, courses, and semesters could arise. For example, if multiple sections of a course are taught by different instructors in the same semester, and only one instructor allows the use of note sheets, then are the resulting exam scores and course grades comparable? If those resulting scores and grades are used for decisions such as admission, continuation, and financial aid, then the use of note sheets could have implications beyond the instructor's course design choice. Third, if note sheets can be linked to improvements in learning and performance, should their use be explored further as part of a standard course design? And, can students be coached on how to prepare and use them for best effect? Fourth, we note a potential concern about allowing note sheets given that the professional certification exams do not allow such exam-taking aids. Is the use of note sheets in courses inconsistent with the goal of preparing students for certification exams? To begin to address these questions, we investigate students' preparation and use of exam note sheets in a graduate-level accounting course. Our inquiry focuses on the quantity and quality of the information in the note sheets and what role note sheets play as the student studies for and takes an exam. In addition, by investigating across two exams, we explore how students change their preparation and use of note sheets with experience. To investigate this issue, we triangulate data from three sources: 1) ratings of note sheets in terms of organization, clarity, coverage, and other dimensions, (2) student surveys, and (3) exam performance. Our main finding is that students appear to perceive value in the preparation and use of note sheets for exams and that, with experience, students invest more in the overall quality of their note sheets. In addition, we find strong evidence that gender and prior exam performance are factors in this increased investment with experience. Given the scarcity of prior research on note sheets in business or accounting courses, and the mixed results from prior research in other disciplines, our research is exploratory in nature. Our intent is to establish a baseline for future inquiry and to bring attention to what we see as an important instructional design choice. We argue that this topic deserves further study and hope that our work spurs accounting academics to consider the implications of note sheets as a learning tool.

The remainder of this paper is organized as follows. In section 2, we review prior research on this issue. In sections 3 and 4, we describe our sample and analysis of data. Our conclusions are presented in section 5.

REVIEW OF PRIOR RESEARCH

The majority of prior studies have explored a direct association between note sheets and student performance, as measured by scores on midterm or final exams. The research design typically takes two forms: (1) manipulating the presence or absence of note sheets (for example, Funk & Dickson, 2011), or (2) manipulating the presence of note sheets versus alternative exam aids such as an "open book" format (for example, Gharib et.al., 2012). Hamouda and Shaffer (2016) provide a thorough review of this line of research and conclude that relatively little evidence has been found linking note sheets and exam performance. Based on our independent review of the literature, we concur with Hamouda and Shaffer on this point. The results are, at best, mixed on a link between note sheets and performance.

Hamouda and Shaffer identify another line of research on the potential moderating effect of note sheets on performance through reduction of anxiety before and during exams. Their conclusion from a review of the research is that students typically report reduced levels of anxiety from having note sheets. However, the evidence does not yet support the notion of a link to improved performance.

More recently, research has focused on cross-student variation in the content of note sheets and exam performance (Edwards & Loch, 2015; Ludorf & Clark, 2014). Given our focus on students' preparation of note sheets, we discuss this line of research in more detail. Ludorf and Clark (2014) examined student

performance on exams in an undergraduate psychology course in statistical methods. Relevant to our research, Ludorf and Clark develop several measures of the content of students' note sheets including density, organization, and a subjective assessment of overall quality. Their results showed a positive association of the overall quality of note sheets with exam performance, but a negative association for density. They speculate that density is driven by the student's command of the material aside from the note sheet. Therefore, students with a less robust understanding of the material tend to load more content into their note sheets. A possible limitation of the Ludorf-Clark results, relative to our study, is the small sample size (n = 21) and the undergraduate course setting. (Undergraduates could have less motivation to invest in note sheets than graduate students.) Nonetheless, the constructs of organization and density appear to be useful dimensions for our design.

Edwards and Loch (2015) also measured density and organization to test note sheet content and exam performance in an undergraduate math course (n = 85). We note two insights from the Edwards-Loch study. First, given that 14% of the student-sample declined to prepare a note sheet for the exam, the authors were able to test for, and find evidence of, improved performance for those students who prepared a note sheet. Second, like the Ludorf-Clark study, Edwards and Loch found a negative correlation between performance and density of content. As discussed later, we use the dimension constructs from prior research in our design.

Research Questions

We explore two main questions in our research. First, if given the option to prepare a note sheet for an exam, how do students respond? How many students invest in a note sheet, and what levels of organization, clarity, and other dimensions are exhibited in the sheets?

Second, do students change their investment in and use of note sheets with experience? Do more or fewer students choose to prepare a note sheet for the next exam, and do the levels of organization, clarity, etc. change between exams? Do the responses differ across exam performance level or student demographic such as gender? We explore these questions with sample data, described below.

SAMPLE

Description of Course

We collected data from a semester-long graduate course in taxation. The course is required for all students, is conducted in a traditional on-campus format, and covers topics that are typical in a master's-level curriculum including corporate, partnership, and estate and gift taxation. Total enrollment was 81 students across three sections of about equal size, with all sections taught by the same instructor.

Description of Performance Assessment

Assessment included two midterm exams of equal length and of equal weight in determining the course grade. Each exam covered four to six broad topics corresponding to assigned chapters in the required textbook. Students were told that in preparing for these exams, they were allowed, but not required, to create a sheet of notes to use during the exam. The instructor did not provide guidance or coaching on how to prepare or use these notes. Other than a space limit – one side of an 8-1/2 x 11-inch sheet of paper – students were free to determine the style, content, and organization of these notes. Also, there was no limit on how far in advance of the exam date the student could begin preparing a note sheet.

Data Collection

Of the 81 enrolled students, 80 prepared note sheets for both exams. One student prepared a note sheet only for the first exam. As a result, our sample includes 80 pairs of student-sheets within 161 sheets overall. One student declined to prepare a note sheet for the second exam because note sheets are not permitted on professional examinations (e.g., the CPA and CMA exams) and the student did not want to use them as exam aids in this course. As such, this student seems to have recognized a potential inconsistency with preparing for exam conditions for professional certification. We draw two preliminary

insights from this result: (1) 100% of students perceived at least some level of value in preparing and using a note sheet for the first exam, and (2) 98% of students continued to perceive value for the second exam.

We collected data from the exercise in two ways. First, between the first and second exam, the instructor asked students to complete a survey on how they used their note sheets. Second, after the course ended, the instructor and one co-author independently rated the note sheets on dimensions such as organization, clarity, and density, as discussed in detail below. Initial agreement from these independent ratings was over 85%. Differences between the initial ratings were resolved by judgment of the third co-author. The resulting data set is a single, majority/consensus rating on each dimension within the note sheets.

In our ratings, we assessed six dimensions of the students' note sheets as follows:

- Presentation – The portion of the notes typed versus hand-written (1=primarily typed; 2=mix of typed and hand-written; 3=primarily hand-written).
- Organization – The extent to which the notes are built hierarchically on topic within chapter (1=topic and chapter; 2=topic or chapter, but not both; 3=moderately by topic or chapter; 4=little apparent organization).
- Coverage – the extent to which the notes are thorough and detailed in coverage of topics (1=extensive; 2=general; 3=somewhat; 4=little).
- Clarity – The extent to which the notes are readable by and useful to an informed reader (1=very; 2=somewhat; 3=little).
- Neatness – The extent to which the notes exhibited neatness and orderliness (1=very; 2=somewhat; 3=little (messy)).
- Density – The portion of the sheet filled with words or tables (1=high density; 2=moderate density; 3=low density (sparse)).

As constructed, lower rating numbers correspond to more effort invested in the preparation of a note sheet. For example, high density (rating = 1) reflects more preparation effort by the student than low density (rating = 3).

Descriptive Statistics

Descriptive statistics on these dimension ratings are reported in Table 1. Table 1 includes the full sample of note sheets from both exams (n=161). For each dimension, we report the range of the rating scale, the mean and standard deviation of the rating, and the distribution across the range. For example, on presentation (scale: 1-3), the mean (std. dev.) rating is 1.75 (0.96), and 99 (61.5%) of the note sheets were rated as "primarily typed."

TABLE 1
DESCRIPTIVE STATISTICS: RATINGS OF DIMENSIONS ON STUDENTS' NOTE SHEETS

Overall sample (n=161)							
Dimension[1]	**Rating scale**	**Mean**	**Std. dev.**	**Distribution of rating-level[2]**			
				N(1)	**N(2)**	**N(3)**	**N(4)**
Presentation	1-3	1.75	0.96	99 (61.5%)	3 (1.9%)	59 (36.6%)	n/a --
Organization	1-4	1.91	0.97	69 (42.9%)	51 (31.7%)	27 (16.8%)	14 (8.6%)
Coverage	1-4	1.77	0.98	88 (54.7%)	33 (20.5%)	29 (18.0%)	11 (6.8%)
Clarity	1-3	1.67	0.77	83 (51.6%)	48 (29.8%)	30 (18.6%)	n/a --
Neatness	1-3	1.24	0.55	131 (81.4%)	21 (13.0%)	9 (5.6%)	n/a --
Density	1-3	1.64	0.69	78 (48.4%)	63 (39.1%)	20 (12.5%)	n/a --

[1] Definition of dimensions and rating scale:
Presentation: Portion of the notes typed versus hand-written (1=primarily typed; 2= mix of typed and hand-written; 3=primarily hand-written).
Organization: Extent to which the notes are built hierarchically on topic within chapter (1=by topic and chapter; 2=by topic or chapter, but not both; 3=moderately by topic or chapter; 4=little apparent organization).
Coverage: Extent to which the notes are thorough and detailed in coverage of topics (1=extensive; 2=generally; 3=moderate; 4=little).
Clarity: Extent to which the notes are readable and useful to an informed reader (1=very; 2=somewhat; 3=little).
Neatness: Extent to which the notes exhibit neatness and orderliness (1=very; 2=moderate; 3=little).
Density: Portion of the sheet filled with words or tables (1=high; 2=moderate; 3=low (sparse)).

[2] N(x) = number of times each rating-level (x) was assessed for that dimension (and relative percentage).

Our general observation from the results in Table 1 is that for the most part students invested significant effort in preparing their note sheets. On all dimensions, the proportion on N(1) is the largest rating for that dimension. Predominantly, note sheets were rated as typed, organized by topic within chapter, extensive in coverage, very clear, very neat, and highly dense.

A second source of data is from student surveys on their use of note sheets during the exam. After the first exam, we asked students to report on two categorical questions: (1) how many times they referred to their note sheets during the exam (Q1), and (2) how their actual usage compared to their expected usage during the exam (Q2). Results are compiled in Table 2. For Q1 (actual usage), 58% of respondents reported using their note sheets between one and five times during the exam, and 28% reported using them between six and ten times. In contrast, the reported usage was zero for only three students (4% of respondents). Recall that n=81 for the number of students taking the first exam. Of these, 74 responded to the survey on Q1 and 72 on Q2. The percentages in each category are based on the number of responses received for each question. The data indicates that a large majority of students referred to their note sheets during the exam, and did so numerous times.

TABLE 2
DESCRIPTIVE STATISTICS: STUDENT SURVEYS

Students' use of note sheets, actual versus expected, during the exam	
Q1: About how many times did you refer to your help sheet during your exam?	
Response	**Frequency (%)**
1. None	3 (4.05%)
2. 1-5	43 (58.11%)
3. 6-10	21 (28.38%)
4. >10	7 (9.46%)
Total Observations	74 (100%)
Mean (Median)	2.4324 (2)
Standard Deviation	0.7230
Q2: Was this usage more, less, or about how frequently you expected to refer to your help sheet coming into the exam?	
Response	**Frequency (%)**
1. Less	5 (6.94%)
2. Same	25 (34.72%)
3. More	42 (58.33%)
Total Observations	72 (100%)
Mean (Median)	2.5139 (3)
Standard Deviation	0.6278

The result for Q2 (expected use compared to actual) is also shown in Table 2. We note that 58% of respondents reported that they used their note sheets more than expected, and only 7% reported using them less than expected. Overall, it appears that students consistently relied on their note sheets during the exam, and that most used them more than expected. This insight suggests that students might learn from their use of note sheets and change how they prepare them for the second exam. We explore this possibility with further analysis.

ANALYSIS

Recall that our second research question focuses on how students change their investment in note sheets between exams. In Table 3, we compare dimension ratings partitioned by exam. Table 3 shows the mean rating on each dimension for exam 1 versus exam 2. As reported in the table, n=81 for exam 1, and n=80 for exam 2. Recall that one student chose not to prepare a note sheet for exam 2. We compute the difference in means (exam 2 score minus exam 1 score) and test the difference. A negative difference indicates, for example, an increase in typed presentation, in organization detail, or in density. The test results suggest that on all dimensions students invested more effort in the preparation of note sheets between exams. The difference in means is statistically significant at $p < .05$ or better for all dimensions except coverage, which is marginally significant ($t = 1.89$). We also note that the mean exam score increased by 2.6 points (out of 100 possible points) on the second exam, but the difference is not statistically reliable.

TABLE 3
DIMENSION RATINGS OF NOTE SHEETS BY EXAM

	Exam 1 (n=81)	Exam 2 (n=80)	Difference in means (exam2 - exam1)	
Dimension [1]	**Mean**	**Mean**	**Difference**	**t-statistic** [2]
Presentation	1.97	1.53	-0.44	3.05***
Organization	2.07	1.75	-0.32	2.14**
Coverage	1.91	1.63	-0.28	1.89*
Clarity	1.81	1.53	-0.28	2.42**
Neatness	1.36	1.13	-0.23	2.77***
Density	1.77	1.51	-0.26	2.34**
Exam score (range)	81.9 (56-99)	84.5 (52-100)	2.6	1.54

[1] For definitions of dimensions, see Table 1.
[2] Two-tailed significance: * <.10, ** <.05, *** <.01.

We extend our exploration of the second research question by investigating differences in dimension ratings across exam performance and student gender. We investigate gender because a sizable stream of research finds some evidence of gender differences in learning styles, pre-exam preparation, and exam performance (Richardson & King, 1991; Severiens & ten Dam, 1994). Gammie et.al. (2003) also finds that female undergraduate accounting students have a tendency to be more apprehensive about examination performance relative to their peers. In more recent work examining gender performance differences, Fogarty and Goldwater (2010) find that female accounting students may not significantly outperform males, but did work significantly more practice questions and thus expended more effort in preparing for exams.

We speculate that gender differences could therefore impact how much students invest in the preparation of note sheets before an exam. Panel A of Table 4 shows dimension ratings partitioned by gender for the full sample of note sheets (n = 161). We compute the difference in mean ratings (female – male) and test for significance. Note that for all dimensions, the difference is negative, suggesting more investment of effort by female students, consistent with the results in Fogarty and Goldwater (2010). However, only the differences for coverage and density are significant at better than .05. Our tentative conclusion from these results is that gender could impact the preparation of note sheets, at least on some dimensions.

In Panel B of Table 4, we also report mean exam scores by gender. The overall mean score on both exams combined is 2.4 points higher for female students, but not at a reliable level of significance. However, examining the scores separately by exam, we find a significant difference in mean score by gender on exam 1. For exam 1, female students achieved a score that was 4.6 points higher on average, and the difference is significant at better than .05. But, we note that for the second exam, the difference in mean score is not reliably different from zero. One possible explanation is that as students learn about the efficacy of note sheets, they increase their investment in them and that this investment correlates to higher exam performance. But, the effect is more pronounced for students who invested less for the first exam. One possible explanation for the disappearance of this gender effect for the second exam is that, as students learn about the efficacy of note sheets, they increase their investment in them and this increased effort correlates to higher exam performance. We expect this effort would be more pronounced for students who invested less for the first exam.

TABLE 4
DIMENSION RATINGS AND EXAM SCORES BY GENDER

	Male	Female	Difference in means (female – male)	
Panel A: dimension ratings				
Dimension [1]	**Mean**	**Mean**	**Difference**	**t-statistic** [2]
Presentation	1.75	1.76	0.01	0.07
Organization	2.03	1.74	-0.29	1.87*
Coverage	1.97	1.49	-0.48	3.18***
Clarity	1.68	1.65	-0.03	0.40
Neatness	1.26	1.21	-0.05	0.58
Density	1.79	1.42	-0.37	3.39***
Panel B: exam scores				
Exam score - all (range)	82.2 (52-100) n=95	84.6 (65-99) n=66	2.4	1.42
Exam 1 score (range)	80.0 (56-97) n=48	84.7 (65-99) n=33	4.6	1.98**
Exam 2 score (range)	84.4 (52-100) n=47	84.6 (67-99) n=33	0.2	0.08

[1] For definitions of dimensions, see Table 1.
[2] Two-tailed significance: * <.10, ** <.05, *** <.01.

To explore this possible explanation further, we partition dimension ratings by both gender and exam. Results are reported in Table 5. The point of this analysis is to investigate whether gender correlates to a difference in students' investment in their note sheets from exam 1 to exam 2. Panel A shows the change in ratings from exam 1 to exam 2 for male students. We note that on all dimensions, the difference in mean ratings for males is negative (indicating more investment) and highly significant. This result suggests that males – who on average invested less in their note sheets for the course overall (Table 4, Panel A) – significantly increased their investment from exam 1 to exam 2. Panel B shows a similar analysis for female students. While the change in means for females is also negative for all dimensions, the differences are not statistically significant. The overall increase in investment between exams indicated in Table 3 appears to be driven by gender, as suggested in Tables 4 and 5. More specifically, male students appear to have significantly increased their investment in note sheets following exam 1. This result suggests a differential learning effect by gender following students' experience with note sheets from the first exam.

TABLE 5
DIMENSION RATINGS BY EXAM AND GENDER

Panel A: Male students (exam 1 n=48; exam 2 n=47)				
	Exam 1	**Exam 2**	**Difference in means (exam 2 – exam 1)**	
Dimension [1]	**Mean**	**Mean**	**Difference**	**t-statistic** [2]
Presentation	2.06	1.43	-0.63	3.43***
Organization	2.25	1.81	-0.44	2.23**
Coverage	2.15	1.79	-0.36	1.70*
Clarity	1.85	1.51	-0.34	2.12**
Neatness	1.44	1.09	-0.35	3.05***
Density	1.96	1.62	-0.34	2.34**
Exam score (range)	80.0 (56-97)	84.4 (52-100)	4.4	1.84*
Panel B: Female students (exam 1 and exam 2 n=33)				
Presentation	1.85	1.67	-0.18	0.75
Organization	1.82	1.67	-0.15	0.66
Coverage	1.58	1.39	-0.19	0.91
Clarity	1.76	1.55	-0.21	1.18
Neatness	1.24	1.18	-0.06	0.51
Density	1.48	1.36	-0.12	0.84
Exam score (range)	84.7 (65-99)	84.6 (67-99)	-0.1	0.03

[1] For definitions of dimensions, see Table 1.
[2] Two-tailed significance: * <.10, ** <.05, *** <.01.

Next, we investigate changes in students' preparation of note sheets based on their exam performance. In Table 6, we compare changes in mean ratings partitioned by the median score on exam 1. For high performers on exam 1 (above median score), we compute the mean change in note sheet ratings from exam 1 to exam 2. We do the same for low performers on exam 1 (below median) and test the difference in means. On all dimensions, low performers exhibited significantly more investment in the preparation of their note sheets for the second exam. In addition, we note that the mean score on exam 2 was 10.2 points higher for students who scored below the median on exam 1 (p-value <.01).

TABLE 6
CHANGE IN DIMENSION RATINGS FROM EXAM 1 TO EXAM 2, BY EXAM 1 PERFORMANCE[1]

	Exam 1 - high (n=43)	Exam 1 - low (n=66)	Difference in mean change (low – high)	
Change in dimension [1]	**Mean change**	**Mean change**	**Difference**	**t-statistic [2]**
Presentation	-0.26	-0.65	-0.41	1.77*
Organization	-0.05	-0.68	-0.63	3.44***
Coverage	-0.19	-0.43	-0.24	1.67*
Clarity	-0.09	-0.49	-0.40	2.55**
Neatness	-0.05	-0.43	-0.38	2.82***
Density	-0.16	-0.38	-0.22	1.65
Exam score – 2 vs. 1 (range)	-2.2 (-26 to +11)	8.0 (-13 to +29)	10.2	4.51***

[1] Exam 1 - high (low) = exam 1 score above (below) median.
[2] For definitions of dimensions, see Table 1.
[3] Two-tailed significance: * <.10, ** <.05, *** <.01.

Finally, we address both factors – gender and exam performance – with an analysis of ratings in Table 7. Given the evidence from Table 6 indicating that the biggest increase in investment between exams is for low performers, we focus on that portion of the sample. In Table 7, we compare the mean change in ratings by gender, for students who scored below the median on exam 1. We note that for all dimensions, the increase in investment is greater for male students. However, only the differences for coverage and density are significant at conventional levels.

TABLE 7
CHANGE IN DIMENSION RATINGS FROM EXAM 1 TO EXAM 2 FOR BELOW-MEDIAN PERFORMERS ON EXAM 1, BY GENDER

Below median exam 1 scores	**Exam 1 - male (n=25)**	**Exam 1 - female (n=12)**	**Difference in mean change (male – female)**	
Change in dimension [1]	**Mean change**	**Mean change**	**Difference**	**t-statistic [2]**
Presentation	-0.72	-0.50	-0.22	0.57
Organization	-0.80	-0.42	-0.38	1.16
Coverage	-0.60	-0.08	-0.52	2.25**
Clarity	-0.52	-0.42	-0.10	0.35
Neatness	-0.52	-0.25	-0.27	1.01
Density	-0.52	-0.08	-0.44	2.20**
Exam score – 2 vs. 1	10.1	3.5	6.6	1.67

[1] For definitions of dimensions, see Table 1.
[2] Two-tailed significance: * <.10, ** <.05, *** <.01.

We also compute the change in exam scores from exam 1 to 2 and note two insights. One, for both genders, low-performing students achieved a significant increase in their scores for exam 2. And, two, the increase for male students (10.1 points) is nearly three times that for female students (3.5 points). The difference is marginally significant (t-statistic of 1.67 and p-value of .103). Our overall conclusion from Table 7 is that students' investment in the preparation of note sheets is linked to both gender and performance on the previous exam.

CONCLUSION

This research explores how students prepare and use note sheets as an exam-taking aid in a graduate accounting course. We explore two questions: (1) if an instructor gives the option to use a note sheet for an exam, how many students invest in a note sheet, and what levels of organization, clarity, and other dimensions are exhibited in the sheets, and (2) how do students change their investment in and use of note sheets with experience, across exams?

We address these questions with a sample of 161 note sheets prepared by students in a required course in taxation. To conduct our analysis, we rated students' note sheets on six dimensions, including organization, clarity, and density, as suggested by prior research. We combine these dimension ratings with feedback from students on their use of note sheets during an exam, and with exam performance data.

Our main findings are as follows. The evidence strongly suggests that students perceive value in the preparation of note sheets for an exam. In our sample, 161 note sheets were prepared out of 162 opportunities across two exams, a response rate of 99.4%. In addition, most students indicated that they used their note sheets during the exam more often than they expected to. Next, we examined how students changed their approach to note sheets for the second exam. Our focus was on how students' experiences from the first exam affected the preparation of their note sheets for exam 2, and which students invested more or less compared to the first exam. Again, the evidence strongly suggests that students perceived value from the preparation and use of note sheets on the first exam. On all six rating dimensions, students on average increased their investment in note sheets for the second exam. In addition, we found that gender and exam performance appear to be linked to this increase. Male students and students with below-median scores on the first exam exhibited the biggest increases in their note sheets for the second exam.

We offer the following implications from our findings. First, research in other disciplines provides some evidence of a link between note sheets and performance. However, there appears to be little if any research in business or accounting on this aspect of course design, or on how students prepare or use note sheets during assessments. Our research strongly suggests that students perceive note sheets to have value for learning or exam performance (or both) and that with experience, students increase their investment in note sheets for later exams. Yet, little is known about the empirical link between note sheets and those outcomes. In addition, our research strongly suggests that students invest differently in the preparation of their note sheets. Specifically, gender and exam performance appear to be linked to an increase in the overall quality of note sheets for later exams. As such, if they are indeed a factor in performance, note sheets could benefit some students more than others.

Second, based on our informal survey of colleagues, allowing note sheets or other exam-taking aids appears to be an autonomous choice by instructors, made without department- or school-level guidance. If students benefit differentially, as our research indicates, then a possible question of equity can be raised. Our findings suggest the possibility that note sheets "level the playing field" across student demographics and performance. We argue that accounting faculty should be mindful of this possibility and how it could impact decisions such as admission and financial aid.

We acknowledge a potential limitation in our research. As the data was drawn from three sections of a single graduate course taught by one instructor, the generalizability of the results could be limited. However, as an initial baseline, focusing on one instructor and one course has the benefit of more control. We have the same instructor, same students, and same assessments for our treatment. In doing so, we avoid potentially confounding variables such as differences in the courses students took concurrently or

before the subject course, or differences in peer dynamics. To address this potential limitation, we extended the level of our analysis. In Tables 3 through 7, we partition the data into successively smaller subsets based on gender and prior exam performance. The deeper we delve into the data, the more evidence we seem to find of a link between these factors and student investment in note sheets.

Finally, we offer our research findings as a baseline for further inquiry and suggest the following as possible directions for future research. The first focuses on student perception and motivation: Why do students invest differently in the preparation and use of note sheets for an exam and how does this difference relate to learning in general and to performance on assessments? The second poses questions for instructors: When an instructor allows note sheets, does this choice affect other course design factors such as exam style or difficulty or overall course rigor? If yes, then additional issues of equity could arise when, for example, instructors have different policies across sections of the same course.

REFERENCES

Durning, S.J., Dong, T., Ratcliffe, T., Schuwirth, L., Artino, A., Boulet, J.R., & Eva, K. (2015). Comparing open-book and closed-book examinations: A systematic review. *Academic Medicine*, 91, 583-599.

Edwards, A. & Loch, B. (2015). Are crib sheets worthwhile? An exploratory data analysis. https://fyimaths.files.wordpress.com/2015/07/2015-edwards-fyi-maths.pdf

Fogarty, T.J., & Goldwater, P.M. (2010). Beyond just desserts: The gendered nature of the connection between effort and achievement for accounting students. *Journal of Accounting Education*, 28, 1-12.

Funk, S.C., & Dickson, K.L. (2011). Crib card use during tests: Helpful or a crutch? *Teaching of Psychology,* 38, 114-117.

Gammie, E., Paver, B., Gammie, R., & Duncan, F. (2003). Gender differences in accounting education: An undergraduate exploration. *Accounting Education,* 12, 177-196.

Gharib, A., Phillips, W., & Mathew, N. (2012). Cheat Sheet or Open-Book? A Comparison of the Effects of Exam Types on Performance, Retention, and Anxiety. *Psychology Research,* 2, 469-478.

Gharib, A., & Phillips, W. (2013). Test anxiety, student preferences and performance on different exam types in introductory psychology. *International Journal of e-Education, e-Business, e-Management and e-Learning*, *3*(February). doi: 10.7763/IJEEEE.2013.V3.183

Hamouda, S., & Shaffer, C.A. (2016). Crib sheets and exam performance in a data structures course. *Computer Science Education*, 26, 1-26.

Ludorf, M., & Clark, S. (2014). Help sheet content predicts test performance. Published in *Essays from Excellence in Teaching*, ed: W. Altman, Stein, L., & Stowell, J. 60-63.

Richardson, J.T.E., & King, E. (1991). Gender differences in the experience of higher education: quantitative and qualitative approaches. *Educational Psychology,* 11, 363-382.

Severiens, S., & ten Dam, G. (1994). Gender differences in learning styles: a narrative review and quantitative meta-analysis. *Higher Education,* 27, 487-501.