

Mnemonic aids during tests: Worthless frivolity or effective tool in statistics education?

Karen H. Larwin
Youngstown State University

David A. Larwin
Kent State University, Salem

Jennifer Gorman.
University of Kutztown

ABSTRACT

Researchers have explored many pedagogical approaches in an effort to assist students in finding understanding and comfort in required statistics courses. This study investigates the impact of mnemonic aids used during tests on students' statistics course performance in particular. In addition, the present study explores several hypotheses that have been proposed to explain the potential benefit of using these aids during examinations. These include the student engagement hypothesis, the perception of control hypothesis, the dependency hypothesis, and a placebo effect hypothesis. The results indicate that student-generated testing aids are clearly superior to the other forms of in-test aids examined, however they were not superior to conditions where mnemonic aids are not allowed. Results are discussed in terms of their implications for the four hypotheses regarding the beneficial effects of mnemonic aids used during examinations.

Keywords: statistics education, testing aids, post-secondary students, engagement

INTRODUCTION

Eighty percent of statistics students surveyed across the United States report that they experience problems learning statistics (Onwuegbuzie, 2004). Yet it is important that students understand core statistical concepts in order to be able to effectively transfer their classroom-acquired training. Pedagogical tools that have the potential to boost student confidence and understanding, have been the focus of a good deal of recent research in statistics education. Constructive approaches, such as cooperative learning, activity-based learning, and technology-assisted learning have been found to enhance the level of student understanding; and, enhance the student's ability to effectively apply statistical concepts (Aronson, Stephan, Blaney, Skikes, & Snapp, 1978; Fernandez and Liu, 1999; Garfield, et al., 2002; Ware & Chastin, 1991; Webb, 1997; Yesilcay, 2000).

One pedagogical tool that surprisingly has not been researched in the area of statistics education is the use of what are commonly known as "cheat sheets." These cheat sheets, or as they are sometimes labeled, "crib sheets," are blank pieces of paper, or sometimes 4x6 inch note cards, on which students are permitted by their instructor to write information that they may consult during a testing situation in order to assist recall. There does not seem to be any consensus as to what is allowed as far as the size of the cheat sheet or the types of information on the paper. Many instructors will require that students turn in the prepared cheat sheet at the end of the testing situation. However, both the terms "cheat" and "crib" have negative connotations. Specifically, they suggest some dishonesty or impropriety, which makes these terms misnomers if the instructors are explicitly permitting the use of these aids. Thus, for the purpose of this investigation, cheat sheets and crib sheets will instead be referred to as mnemonic aids which are used during testing. This vocabulary better reflect the intended benefit and instructor sanctioned nature of these test-taking tools.

Although many statistics instructors may allow the use of these mnemonic testing aids, there is no available research indicating how widespread this practice is, and no research is presently available on the impact of these aids on students' performance in statistics courses in particular. This deficit is peculiar, as researchers have investigated the use of such aids in testing situations in other disciplines, such as economics (Wachsman, 2002), chemistry (Cherim, 1981), and psychology (Dickson & Bauer, 2008; Dorsel & Cundiff, 1979; Hindman, 1980; Skidmore & Aagaard, 2004). Thus, the present study appears to be the first to examine the efficacy and the functional role of these aids in statistics courses.

Competing Hypotheses

Some of the studies in other disciplines suggest that allowing the use of mnemonic aids in testing situations has a significant effect on improving student attitudes and lowering anxiety (Cherim, 1981; Skidmore & Aagaard, 2004). More importantly, researchers have argue that the preparation of these testing aids helps students to review, organize, and clarify the main points of material that will most likely appear on a test (Cherim, 1981; Hindman, 1980, Skidmore & Aagaard, 2004; Wachsman, 2002). Since active student engagement with course materials and information, such as note taking, participation in cooperative learning groups, and activities-based learning, is associated with greater performance (Skidmore & Aagaard, 2004), creating and using mnemonic aids for testing might act as another opportunity to engage students.

Many researchers argue in favor of such a student engagment explanation of the beneficial effects of testing aids; the notion that the preparation encourages students to review, organize, and clarify the main points of the course material that will most likely appear on a test (Cherim, 1981; Hindman, 1980, Skidmore & Aagaard, 2004; Wachsman, 2002). Loftman (1975) and Zimmerman & Pons (1986) investigated the impact of having students review course materials prior to exams. Both

researchers asked students to rewrite, revise, and reorganize the course notes in an effort to increase overall engagement with course content, increase familiarity with materials (and thus potential exam questions), and increase study time. They found that students who participated in the reviews of the exam materials significantly outperformed students who did not. Thus, having students prepare aids should engage students with course material in similar ways.

Another possible means by which testing aids may have their beneficial effect is through influencing perceptions of control. Research demonstrates a strong connection between perceived control and a number of positive psychological effects (Burger, 2005; Lefcourt, 1992; Rotter, 1966, Thompson, 1991). In particular, several studies indicate that perceptions of control are associated with greater academic performance and achievement (Burger 1987; 1992; 2005; & Findley & Cooper, 1983) suggesting that feeling a sense of control can be important for students' learning. For example, Endler, Speer, Johnson, and Flett (1999) maintain that during high anxiety events, such as testing, perceptions of control can be predictive of increased performance. These authors refer to perceptions of control, or perceived control, as "...an individual's appraisal of the extent to which the situation itself maybe brought under the control of the individual" (p. 37). The relationship may be explained, in part, by the effect of perceived control on anxiety. According to Thompson (1991), when an individual perceives control over a stressful situation, the level of anxiety is decreased, as the availability of options, "...increases perceptions of control, although that is not usually measured as a part of the research." (p. 608). Endler, Speer, Johnson, and Flett found that when they could successfully manipulate students' perceptions of control during a testing situation, such that the students perceived themselves as having increased control, resulting in reducing anxiety and improving performance.

However, the foregoing pattern of results about the effectiveness of mnemonic aids for use during testing and the explanations for why they might be effective has not been uniformly supported. Dickson and Bauer (2008) examined a sample of students in psychology courses in a repeated-measures study, finding that there was indeed significantly better performance when students prepared and used these aids on a multiple-choice exam, relative to when those same students prepared but did not use the aids on a pre-test composed of identical questions. As all the students had prepared the mnemonic aids before the pretest then and exam, the better performance on the exam when the aids were allowed suggests that constructing the testing aid did not actually enhance student learning; instead, the use of testing aid simply enhanced student performance. Dickson and Bauer reasoned that if the act of constructing aids actually enhanced learning, all students should have performed similarly on the pre-test and the exam,

However, there were a number of potential extraneous variables operating within the Dickson and Bauer (2008) study that qualifies this evidentiary challenge. First, students were aware that the pre-test would not count toward their grade and thus their performance quality on the pre-test was of little significance to them personally. This may have undermined some of their motivation to perform well. Secondly, students were under the impression that they would be able to use their aids on their examinations. Thus, even though the pre-test didn't "count," having an unexpected pretest to take without the potential benefit of their aids may have generated some test anxiety that could have worked to attenuate pre-test performance. Finally, as this study was a repeated-measures study where the pre-test items and exam items on which performance was assessed were identical, it is possible the improved scores on the exam, the second assessment, were influenced by a practice effect. Indeed, Dickson and Bauer presented evidence that the students performed significantly better on the specific exam questions that were identical to those in the pre-test relative to the exam questions that were not a part of the pre-test.

Another possible explanation for the beneficial effect of mnemonic aids on student performance may be the result of a placebo effect. Mnemonic aids used during testing, may not provide cognitive

benefit for test performance and mastery of course material but instead, emotional comfort and associated expectancy of greater success—and increased self-efficacy. This is similar to the perception of control hypothesis as mentioned above however, there is a subtle difference. The perception of control hypothesis assumes that students' sense of confidence, emotional comfort, and control is a reflection learning, knowledge, and mastery of course material. The placebo effect hypothesis, assumes that the belief in the value of the aid is responsible for the beneficial effect on performance.

Dorsel and Cundiff's (1979) findings can be interpreted in terms of a placebo effect. Students who prepared and were unexpectedly unable to use their mnemonic aid performed worse on exams relative to students who prepared and were able to use their aids. If preparing the mnemonic aid was responsible for the benefits of greater student engagement with course material, then both groups should have benefited from the preparation. However, the group that was unable to use their prepared aids performed worse on the course exams is consistent with a placebo effect account of the impact of the testing aid. These students were suddenly without the tool that they had expected to help them. If these aids function as placebos, not having the aid available should indeed attenuate the placebo effect.

Dickson and Bauer (2008) results also provide evidence supporting a placebo effect interpretation. Students who prepared the aids and then used them on an exam performed better than they did on a pre-test where they were not allowed to use them. Both performance on the pre-test and the actual exam should have reflected benefits from preparing the aids. The resulting difference between the pretest and exam scores was interpreted by the authors as evidence that the use of these aids results in simple performance benefits, but not genuine learning. However, if the benefits are a placebo effect, it would be expected that when students were able to use them they would perform better than when they couldn't. So the pattern of results obtained by Dickson and Bauer is consistent with a placebo effect. The possibility of a placebo effect is an idea that has not yet been specifically examined, and needs to be more fully explored.

While both the Dorsel & Cundiff (1979) and Dickson & Bauer (2008) studies may not offer a clear test of the placebo hypothesis, the authors themselves interpreted their results as supporting a dependency hypothesis interpretation. According to this explanation, these aids benefit students in that they serve as a source of information that students need for an exam, and students then rely on or depend on the aid, rather than learning the information and committing it to memory. According to this explanation, the act of creating a testing aid does not enhance student learning and memory; it amounts instead to a clerical exercise in creating a reference tool that is depended upon to provide the information the student needs during an exam. Whatever the reason for this pattern of results, a consideration of these findings, along with those of Dickson and Bauer (2008) discussed above reveals that the precise impact of these aids on student learning and performance is varied and still far from clear.

Current Study

The purpose of the present study is threefold. First, despite the common perception that the statistics course is a likely situation where instructors might allow the use of mnemonic aids on exams, there is no research examining the use or effectiveness of mnemonic aids in the statistics course. In many of the studies examined in the foregoing review, the exams that constituted the dependent measures were multiple choice tests. However, examinations in statistics courses are more likely to include computation problems and other constructed-response questions. The manner in which the construction and use of testing aids assist student learning and exam performance, given these testing characteristics, may be different than that observed in other disciplines with other types of exams.

Second, the present study seeks to not only determine what effects the use of mnemonic aids have with a distinct population of statistics students, but to also provide evidence to help explain why aids have the effects that they do; evidence that should be relevant to the use of aids, and an explanation of their effects, across disciplines. Despite the several studies that have examined the use of these aids thus far, a number of important questions remain to be answered. Are there indeed any benefits associated with the creation and use of mnemonic aids during testing at all? If there are benefits, are these simply the result of a placebo effect and student beliefs in what aids might be able to do for them? And if there are benefits beyond a possible placebo effect, what is the basis for such benefits? Is it that the act of creating an aids for use during an exam enhances student engagement, cognitive encoding of information, and thus student learning? The present study was designed to provide some evidence indicating which of these explanations best accounts for the pattern of results observed in the context of students creating and using these aids in a statistics course. In this process, the present study may help explain some of the discrepancies that exist in the results of studies conducted on aids to date, and in doing so it may also help determine if there is a general pattern of effects surrounding these aids use that is common across disciplines, or if these effects will inevitably vary from one discipline area to the next.

Lastly, the present investigation examined the effects of four different testing conditions on student exam performance in introductory-level statistics courses. Specifically, the present study examined whether students perform differently when they use aids provided for them by the instructor, when they use these aids they prepare themselves, when they are permitted to use their text book during an exam, and when they are not permitted to use any testing aids during an examination. If the idea that creation and use of mnemonic aids forces students to engage course material, enhance their learning, and produce a sense of control is indeed correct, then it is expected that students who prepare testing aids will demonstrate significantly greater levels of performance on their class exams relative to the students in the three other testing conditions. In accordance with the Engagement/Perceived control hypothesis, students taking exams without study aids of any sort should perform better than students with the instructor provided aids and the open textbook. The inclusion of the open textbook condition allows for an additional test of the distinction between the student engagement hypothesis and the placebo hypothesis. There is research to indicate that, like a presumed placebo effect, open textbooks available during an exam serve to reduce anxiety; but they also reduce student motivation to study and students' self-reported time spent studying (Ioannidou, 1997). The beneficial effect they would have, if any, seems less consistent with an engagement hypothesis and more consistent with a placebo effect. In addition, the dependency hypothesis argues that students rely too heavily on testing aids and this undermines student learning. Presumably, the same would be true of students relying on an open textbook during an exam. Thus, along with the inclusion of the instructor-provided aids condition, the open textbook condition provides additional tests of both the placebo effect and the dependency hypothesis.

METHOD

Participants

Participants included $n = 94$ second and third year undergraduate students that ranged from 18 to 52 years of age ($M = 20.82$, $SD = 5.08$), including $n = 40$ male (42.6%) and $n = 54$ female (57.4%) students who were enrolled in one of four sections of applied statistics at a small Eastern U.S. university. Two different instructors each taught two of the four sections. Students in these applied statistics courses were expected to develop the statistical tools used in research decision making, including but not limited to determination and interpretation of measures of central tendency, variance, probability,

regression and correlation analysis, hypothesis testing, frequency and probability distributions, and sampling methods. Students also were introduced to graphical, tabular, and mathematical depictions of statistical information. All sections incorporated activity-based cooperative learning groups and supported learning with the use of statistical computer packages (MiniTab and SPSS).

Measures

Class exams were used to measure student performance. There were four exams administered for each of the four course sections. In each section, the exams were of similar length and difficulty level. All students completed the exams within the allotted period. These exams included selected-response items, constructed response items, and application problems appropriate for the material covered by each exam. The application portion of these exams accounted for approximately 75% of the examination grade. Although the exams were not intended to be cumulative, the nature of the course material required that students understand foundational course material presented earlier in the semester in order to understand material presented later in the semester.

Procedures

Students in the present study experienced four different testing conditions in a randomly determined order. These were: (1) NO aid allowed on an exam; (2) Instructor-prepared aid, where aid prepared by the instructor were provided for student use during an exam; (3) Student-prepared aid, where students prepared and used their own aid; and (4) an Open-textbook condition where students were allowed to consult their course textbook during the exam.

Students in the NO AID testing condition were given their exams in class, and were expected to complete exams by recalling information without the assistance of any study aids. The students in the Instructor-prepared AID condition received an 11 X 8.5 inch sheet of paper from the instructor. This paper contained any formulas and/or definitions that had been discussed and used in class activities that might be useful for answering questions and solving problems on the particular exam. Students were instructed to write their names on the provided AID, and they were required to return them to the instructor at the end of the testing period. The students in the Student-prepared AID condition were given an 11 X 8.5 inch piece of paper on which they were permitted to write any information that they deemed essential. Students were instructed to write their names on these AIDs and were required to turn them into the instructor at the end of the testing period. Students in the open-textbook condition were allowed to consult the required textbook for their section during the examination period. Students were instructed to empty their textbooks of any note papers prior to the start of the exam.

Each course section experienced each one of the four testing conditions on one of their four exams. Prior to the administration of the first exam, the order of testing conditions was randomly determined for each of the four sections. The order of the testing conditions was randomly determined so that order effects would not impact the results of the investigation. Students were informed about the experiment and the different testing conditions, and asked to indicate their willingness to participate by signing the IRB approved consent form. Students not giving consent still participated in the different testing conditions as a part of their course experience; however their data was not used in the study.

Analysis

A 2 (gender) by 2 (instructor) by 4 (testing condition) Factorial Analysis of Variance was conducted on the student exam scores in which, student gender and testing condition were examined as

between-subjects factors. In addition, as part of the course evaluation process, students were provided with a brief questionnaire and asked to indicate which testing condition they preferred, and why they preferred the condition they indicated over and above the others.

RESULTS

The factorial analysis indicated that there were no significant differences in the exam scores between the different gender groups or classes taught by different instructors, $p > 0.05$. The specific mean differences for gender are presented in Table 1 (Appendix).

Because the class sections differed in terms of their randomly determined order of presentation of the four testing sections, there were no potentially confounding testing condition order effects at work in the present data set. The descriptive data for each of the four testing conditions is presented in Table 2 (Appendix).

The factorial analysis of variance also indicated that there were significant differences in student exam scores across the different testing conditions, $F(3,87) = 22.92$, $p < .001$, partial $\eta^2 = .441$. Specifically, Scheffe post hoc analyses revealed that significant differences existed between the Student-prepared aids condition and both the Instructor-prepared aids condition, M difference = 5.73, $p < 0.001$ and the Open-textbook condition, M difference = 13.19, $p < 0.001$. Students in the Student-prepared aids condition performed significantly better on the course exams than did students in both the Instructor-prepared aids and the Open-textbook conditions. However, although students in the Student-prepared aids testing condition performed better than students in the No aids condition, these differences were not statistically significant.

In addition, significant differences in mean exam scores were found between the Open-textbook condition and both the No aids testing condition, M Difference = 11.57, $p < 0.001$, and the Instructor-prepared aids condition, M Difference = 7.45, $p < 0.001$. In other words, students in the Open-textbook condition performed significantly worse on the course exams than students in both the No aids and Instructor-prepared conditions. And while the students in the No aids condition outperformed students in the Instructor-prepared aids condition on the course exams, the post hoc differences revealed only a marginal difference, M difference = 4.12, $p = 0.057$.

Finally, students were asked at the end of the semester to indicate which testing condition they preferred, and why. Most students indicated they preferred the Student-prepared aids condition ($n = 36$, 38.3%). Students who indicated that the Student-prepared aids condition was their favorite testing condition indicated that they preferred to use the notes they prepared themselves for reasons such as:

It makes it easier for us to reinforce what we don't know during the test. It allows me to put what I think is necessary to know on a paper for the exam.

I prefer the student-prepared notes condition because when I write notes down it also helps me to study as well for the exam and therefore helps to reinforce the material covered in the exam double as opposed to the other options.

Because my notes are written in a way that I can understand and the notes I make on my notes may not be understood by someone else but help me make connections.

I feel that sometimes the book is hard to understand so my own notes simplify these concepts into my own words.

Some students indicated a preference for the Instructor-prepared aids condition ($n = 23$, 24.5%), closely followed by the Open-textbook condition ($n = 20$, 21.3%). Students indicating that the Instructor-prepared aids condition was their favorite indicated such things as:

I prefer having professor provided notes because that guarantees that you have the necessary formulas available. If a student makes his/her own study guide he/she may leave out an equation or formula that is vital to completing a test answer.

The professors notes cover all of the materials that they think we should know ...my notes might not be as clear and I might of missed something during the discussion.

Because I know the information I am getting is correct and will help me in some way.

Students who stated their preference was for the open-textbook condition indicated such things as:

Using your book makes you read and understand the information.

Easier to prepare for an open book test.

When you are able to use your book you are not forced to memorize a bunch of definitions for essays and you can be much more relaxed going into the exam.

The condition that was the least favorite of the students in the present study was the No aid testing condition, ($n = 15$, 16%). The students who stated a preference for this condition offered reasons such as:

It takes less time to recall information once it is learned than to look it up in a book. Also I think it is more beneficial to memorize the equations and information.

I find that I do better when I need to learn the material than reading out of a book during the test. Also I take too much time and second guess my answers when I use a testing aid.

I feel that I would learn the most and make myself study more if I had to prepare the night before and not have anything to look at during the test. This gives me a real idea of how well I know the information.

When I knew I could rely on the book or notes I didn't study as much or pay attention as much in class I should have. As a result the test is not an accurate indication of what I actually know.

DISCUSSION

The present investigation sought to examine the usefulness and efficacy of mnemonic aid use during testing on student exam performance in college-level statistics coursework. The pattern of results observed provides evidence supporting the instructional use of mnemonic aids during student exams. However, the potential benefit of student use of aids observed in the present study was not as clear and as robust as that observed in other research (e.g. Wachsman, 2002). Specifically, while the students in the Student-prepared aids condition did exhibit superior exam performance relative to students in all of the remaining testing conditions, the differences were statistically significant only with respect to the Instructor-prepared and Open-textbook conditions. In contrast, the mean difference in exam performance between the Student-prepared aids condition ($M = 79.38$) and the No aids condition ($M = 77.78$) did not reach statistical significance. That the students who prepared and used their own aids outperformed students in the other two aids conditions indicates that the self-generated aids is likely the mnemonic aid of choice for the applied statistics course. But the present study does not clearly support the notion that self-generated aids result in superior performance relative to standard examination conditions where no aids of any sort are permitted.

Perhaps the most reasonable conclusion with respect to the potential benefits of self-generated aids is that indeed there is reason to believe there may be such benefits, but they were not demonstrated conclusively in the present data. While past research studies have yielded mixed results as to the beneficial effects of aids, the studies that argue against the potential beneficial effects (e.g. Dickson & Bauer, 2008; Dorsal & Cundiff, 1979) seem to be plagued with greater methodological problems than the studies which support them (e.g. Wachsman, 2002). And the Self-prepared aids condition was the top performing condition in the present study. These considerations are suggestive of some potential benefits of self-generated AIDs, even in the statistics course. But indeed the results of the present study are also consistent with the possibility that self-generated aids in a statistics course—regardless of effects that may be present in other courses—do not in fact offer any benefits to student exam performance relative to standard testing conditions in which aids are disallowed. The benefits of aids use may be different for the statistics course than they are for other disciplines. Without an interpretive context provided by other investigations of aids use in statistics courses, the answers to these questions await further study and the patterns of results that emerge therein.

The results of the present study also offer a number of implications pertinent to the various theoretical accounts of why aids may be effective in enhancing student learning. First, the Placebo Effect account of why aids might benefit student exam performance seems to have received little support in the present investigation. Simply providing students with instructor-generated aids and allowing them to use open textbooks for consultation during exams resulted in significantly inferior exam performance relative to the Student-generated aids and No aids conditions. The placebo effect hypothesis would predict the reverse. Thus, whatever benefits aids might have for student exam performance, these benefits do not seem to be due to a placebo effect.

However, it might be argued that there is something unique about one's own self-generated aids that causes students to have greater faith and confidence in them, and in that way a placebo effect may be operating within the present data set and account for the superior performance of the Student-generated aids condition. However, there is both logic and evidence that argues against this interpretation. The placebo effect is about belief and expectations in the power of things outside of and separate from oneself. This characterizes the Open-textbook and Instructor-provided aids conditions to a greater extent than it does the Self-generated aids condition. The aids created and provided by the instructor and the course textbook represent authoritative sources of expertise in the context of a college-level course; sources that presumably generate more in the way of expectations of worth and value.

Indeed, the qualitative responses students provided to the open-ended questionnaire items about the various testing conditions at the end of the semester indicated that students perceived considerable value in these external authoritative sources as in-class testing resources. They indeed stated that they perceived the instructor provided aids and the textbook as trusted sources of information to rely on. For example, one student remarked that the aids prepared by the instructor were valuable “Because I know the information I am getting is correct and will help me in some way.” Students did not indicate these sorts of beliefs and expectations with respect to the self-generated aids. In light of this pattern of evidence, it seems unlikely that the Self-generated aids condition created some sort of exaggerated placebo effect relative to the potential placebo effects generated by the Instructor-provided aids and the Open textbook testing conditions. Thus, the pattern of evidence obtained in the present study argues against a placebo effect account of the benefits of aids use in college level statistics examinations.

Another rival interpretation for the potential beneficial effects of aids use is similar in some ways to the placebo effect hypothesis. That is the dependency hypothesis. This account argues that aids may partially benefit students in that they come to depend on them for enhanced exam performance, but do not in fact fully learn the course material. In other words, the aids function as a something of a crutch to rely on during an exam. If that crutch were unavailable, presumably student exam performance would suffer. But as long as it remains available, student exam performance should not negatively be impacted. Dorsel & Cundiff (1979), for example, found that students who prepared but were unexpectedly not allowed to use their aids performed worse than students who prepared and used them and students who neither prepared nor used aids. The authors argued that these students must have been depending too heavily on their aids and were unable to perform without them. However, one potentially confounding factor in this interpretation is that the unexpected denial of the opportunity to use the prepared aids may have generated considerable anxiety, and this may have been what attenuated exam performance for this group.

In light of this concern, the present study was designed to avoid the extraneous factor of test anxiety that might confound the interpretation of the potential effects of aids use. There was no condition where students prepared but were not allowed to use their aids. However, there were a number of conditions used and results observed in the present study that shed some light on the utility of the dependency hypothesis. While there were no unexpected prohibitions of aids use in the present study, there were two additional conditions that should have also promoted student aids dependency effects of the sort Dorsel & Cundiff (1979) indicate, similar to what would presumably be in effect for a self generated aids condition. These conditions were the Instructor provided aids condition and the Open textbook condition. If the benefit of using an mnemonic aids during testing is simply the result of depending upon a memory aid for an exam, then students in these two conditions and students in the Student-generated aids condition should have been subject to such a dependency effect and its associated benefits. While the Dorsel and Cundiff study used a strategy of trying to compare students who were either allowed or disallowed from using their prepared aids, the present study compared groups that might all be expected to exhibit dependency effects, if that general explanation for the benefits of aids were correct. Variations in exam performance between these groups would thus be unexpected, and indicate something other than a dependency effect operating in the data.

Such a pattern of variation is exactly what the present analysis revealed. Students who generated their own aids clearly outperformed students who relied on their course textbook or instructor provided aids during the exams. That all these groups might benefit from an aid to depend upon, but the Student-generated condition clearly benefited more, suggests perhaps something more than a simple dependency effect for the Student-generated condition. In addition, the logic that an aids only works to improve exam scores by a simple dependency or over reliance on the aids would suggest students in the No- aids condition should perform better on exams than all of the other groups who have some sort of in-test

memory aid. But that was not the case in the present study. The No- aids condition was not the top performing condition on the student exams; the Student-generated aids condition was. That the No- aids condition—in which no dependency effect would be operating---performed similarly to the Student-generated aids condition and also better than the Instructor-generated aids and Open textbook conditions suggests that if a mere dependency effect is indeed at work, it is more evident in the latter two conditions, and not in the Self-generated aids condition. Thus, while the possibility of a dependency effect might not be something that can be completely ruled out by the results of the present study, the pattern of results obtained is not what would be expected by such a dependency effect. With this in mind, the dependency effect as an explanation for the beneficial effects of student aids use seems rather tenuous, at best.

The final substantive theoretical interpretation for why aids might work to enhance student learning and exam performance that was examined in the present study is the engagement hypothesis; the notion that aids serve to foster greater student engagement with course material as students review that material in preparing an aids for use during an exam. From this perspective, it might also be expected that students in the Instructor-provided aids condition and students in the Open textbook condition would rely more on the power of the expert-provided aids and experience less motivation to study the course material. Such a perspective would predict that students in the Self-generated condition would perform superior to all of the three remaining conditions, and that the No aids condition would also perform better on exams than students in the Instructor-provided and Open textbook conditions. In fact, this expected pattern of resulting mean differences among the four testing conditions is exactly what was observed in the present study. The only deviation from this expected pattern is that not all the expected mean differences were statistically significant. While the Student-generated aids group did perform the best on the exams, the difference was statistically significant only for the comparisons with the Instructor-provided and Open textbook conditions, but not for the No aids condition. Also, while students in the No aids condition performed better on exam scores than students in the Instructor-provided and Open textbook conditions, the difference was only clearly significant for the comparison with the Open textbook group; it was only marginally statistically significant for the comparison with the Instructor-provided aids group. Why these two mean differences were not also statistically significant is not clear. Perhaps a larger sample might have found them to be so, but the present sample was not overly small to result in an insufficient test of the research hypotheses. In the end, while the argument that the present study offers support for the student engagement hypothesis would have been stronger had these two observed mean difference among testing conditions been clearly significant like the others predicted by the student engagement hypothesis were, it is still the case that the pattern of results obtained in the present study best fits the student engagement explanation.

Strengths of the present study

The present study examined the potential benefits of using aids in a college-level statistics course, and compared several theoretical explanations attempting to account for the beneficial effects of aids use. What's more, the present study examined these questions in a context characterized by greater realism than what was true of many previous studies. For example, some studies have examined measures of student performance on pre-tests and other measures that did not count toward their course grade (Dickson & Bauer, 2008; Dorsel & Cundiff, 1979). Students in these conditions may not have been fully motivated and engaged to perform at their best, as they presumably would be in an actual testing situation in one of their college courses. The present study examined student performance under various conditions with aids both available and unavailable in actual college course testing situations. The test performance of each student counted toward their final course grade.

Some studies have also examined student aids use with mnemonic strategies and performance measures that are not likely representative of genuine testing situations in a college course. For example, Dorsel & Cundiff (1979) used what they termed “substitution coding” as the mnemonic strategy for students to use. This involved students picking letters from key words to write down and organize however they wished that they would use later as a recall aid, after studying for 10 minutes a short prose passage describing some concepts from Freud’s Psychoanalytic theory of personality. The performance measure constituted student recall of the concepts in the prose passage. This highly constrained form of mnemonic coding bears little resemblance to what is used on student generated aids prepared for a typical college course exam. And a test of recall for a brief passage of material is hardly representative of the breadth and depth of material students must review and master in preparation for a college level course exam. It is hard to imagine how well such manipulations, measures, and results would generalize to an actual circumstance where college student exam performance is being evaluated. In contrast, the present study examined student exam preparation, the use of aids, and actual exam performance precisely in the actual context to which our research questions and hypotheses pertain.

In similar fashion, the present study enhanced the realism and generalizability of the dependent variables by having them include more encompassing measures of student exam performance. For example, many studies examining the effects of aids in real or simulated testing situations employ primarily selected-response item exams, such as exams composed of multiple-choice questions (e.g. Dickson & Bauer, 2008). In contrast, the present study employed course examinations that included a few selected-response items, but also included a number of constructed-response items—short essays and computational problems—as well. The applied statistics examinations used in these sections required students to demonstrate their ability to apply the information covered in the course. These application-heavy assessments necessitate that students demonstrate higher-order thinking about the requisite course material. With this type of testing, students have to demonstrate their ability to take a plethora of formulas and terms, use them appropriately and interpret the results correctly within the given context. Thus, the present study broadened the scope of the student performance evaluations often used in the aids literature.

Another feature of the present research design is that the potential confounding factors observed in prior research studies were strategically avoided. Conditions employed by other researchers that attempt to examine theoretical suppositions but in practice also create other unintended effects were not included in the present research. Several studies have included manipulations that likely increased student anxiety—like having students prepare aids and then unexpectedly telling students they would not be allowed to use them (e.g. Dickson & Bauer, 2008; Dorsel & Cundiff, 1979)—and thus may have confounded the measures of student recall or exam performance. The present study contained no such surprises for students or violations of their expectations, or what they had been told by their instructors, which might have adversely affected measures of their exam performance.

Likewise, the present study was designed to avoid another potentially confounding factor in prior research studies on aids that likely generates considerable practice effects. Specifically, the present study did not employ a pre-test post-test design where students were repeatedly assessed on the same test items as other studies have done (e.g. Dickson & Bauer, 2008). While creating testing conditions that attempt to allow the examination of various theoretical interpretations, these other studies introduce potentially confounding effects on their dependent measures. While it is true the present study made use of a within-participants component in the research design, this was done in order to expose each of the participants to different testing conditions, and utilized completely different exams and test items from one measure to another. This within-participants component was utilized to control for the potential confound of test condition order effects. The key comparisons of note that addressed the effectiveness

of various aids, and the various theoretical explanations for the potential benefits of aids use, were all between-participant comparisons that avoided these potentially confounding practice effects.

Another strength of the present study is that it is the first to examine the possibility that the beneficial effect of aids use might be due to a simple placebo effect. Given the ubiquitous nature of placebo effects in the world of human performance and behavior, it seems surprising this possibility has not been considered and examined before. In a novel contribution to this research literature on student aids use, the present study included an instructor-provided aids comparison condition and an open textbook control condition to better examine both this potential placebo effect and the dependency hypothesis. While neither of these explanations emerged as tenable answers in light of the pattern of results obtained, the unique design of the present research allowed these ideas to be examined and dismissed as likely explanatory candidates.

Another unique contribution of the present study is that it is the first attempt to examine the use and efficacy of aids in college level statistics courses. It would seem such a context would be a likely place where student use of aids would occur, which makes the dearth of empirical research in this area that much more surprising. While the present study did not definitively answer the question as to whether or not aids might be effective in promoting student learning and performance in the statistics course, hopefully it will serve to initiate substantial research in this direction.

Finally, the present research is unique in its attempt to address multiple competing theoretical explanations for the potential benefits of aids use in a single study. While a number of other studies have examined specific hypotheses, the present design examined simultaneously the student engagement hypothesis, the dependency hypothesis, the placebo effect hypothesis, and the possibility that there is no benefit to student aids use at all. The present study fell short of being able to definitively support the student engagement hypothesis, but it was the one--of the three hypotheses based on an assumed beneficial effect of aids use--most consistent with the observed pattern of results. And while the student engagement hypothesis might not have been fully supported, the placebo effect hypothesis and the dependency hypothesis were unable to account for the present pattern of results and were subsequently dismissed. A study that both comprehensively and economically addresses all the competing hypotheses pertaining to the potential benefits of aids use, to the extent that this one does, is unprecedented in the available research.

Limitations

The present experiment makes use of a quasi-experimental design in that students pre-assigned to existent class sections of an applied statistics course served as participants in the study. One limitation associated with the use of quasi-experimental research is that the pre-established groups may not be as equivalent in statistics ability prior to the implementation of the treatment conditions as we would like to assume. These potential pre-existing differences in student ability could exist along with any treatment effects, and could distort the measures of the treatment effects. However, the fact that each class section experienced each of the four testing conditions, in a randomly determined order, should help to control for this potential extraneous variation. Yet, being a quasi-experimental design, these concerns cannot entirely be ruled out.

Another concern is that the different instructors may have had differential impact on student performance, independent of treatment effects. This effect could be due to different instructional styles and different instructor personalities and level of rapport with students. Also relevant here is the fact that this was not a blind study. The instructors were aware of the testing conditions students in their classes were to experience, and which exams featured which testing conditions for each class section.

CONCLUSION

The first question this investigation sought to address is whether or not there are beneficial effects of student aids use on examinations in a college level applied statistics course. While the pattern of mean differences that emerged matched the a priori expectations, unfortunately one of the key differences did not reach the level of statistical significance. The study was not able to clearly demonstrate that student-generated aids in particular resulted in superior exam performance relative to standard testing conditions where aids use is not allowed. This finding is inconsistent with a number of other studies that have found such beneficial effects of self generated aids (Cherim, 1981; Hindman, 1980; Skidmore & Aagaard, 2004; & Wachsman, 2002), leaving some question as to why this pattern of results was not observed in the current investigation, and exactly how the present study fits into the larger pattern of research findings in this area. While some studies have failed to demonstrate clear benefits of self-generated aids use relative to No aids control conditions (e.g. Dorsel & Cundiff, 1979), these studies were also plagued with methodological and interpretive problems that the present study sought to avoid. In the end, while there are a number of possibilities as to why this key difference between the self-generated aids condition and the condition where aids were not allowed did not emerge as a significant one in the present study, the answers to these questions and possibilities await further investigation. Yet one conclusion about the effectiveness of aids use is clear from the present study: If college level statistics instructors plan to allow their students to use some form of aids during course examinations, the self-generated aids is the mnemonic aids of choice. It was clearly superior to the instructor provided aids and open textbook conditions.

The present study also examined, for the first time simultaneously, a number of theoretical accounts of why aids use may benefit student learning and performance. This included a theoretical possibility heretofore unproposed and unexplored: the placebo effect. From the pattern of results obtained in the present research, it appears that if student use of aids during exams does have a beneficial effect on exam performance, then the student engagement hypotheses is the most likely candidate to account for such an effect. The competing explanations—that aids only create the illusion of beneficial effects as students become overly dependent upon them, or that their effects are due to a simple placebo effect—were not supported in the present research.

REFERENCES

- Aronson, E., Stephan, C., Blaney, N., Sikes, J., & Snapp, M. (1978). *The jigsaw classroom*. Beverly Hills, CA: Sage.
- Burger, J. M. (1987) The effects of desire for control on attributions and task performance. *Basic and Applied Social Psychology*, 8, 309-320.
- Burger, J. M. (1992) Desire for control and academic performance. *Canadian Journal of Behavioural Science*, 24, 147-155.
- Burger, J. M. (2005) Personality and Control. In Derlega, Winstead, & Jones (Eds.) *Personality: Contemporary theory and research. Third Edition*. (pp. 309-331). Belmont, CA: Thomson Wadsworth.
- Burger, J.M., & Cooper, H. M. (1979) The desirability of control. *Motivation and Emotion*, 3, 381-393.
- Cherim, S. (1981) A philosophy of teaching preparatory chemistry. Paper presented at the Annual Two-Year College Chemistry Conference, Atlanta, Ga.
- Dickson, K. L., & Bauer, J. J. (2008) Do students learn course material during crib sheet construction? *Teaching of Psychology*, 35(II), 117-120.

- Dorsal, T., & Cundiff, G. (1979) The cheat-sheet: Efficient coding device or indispensable crutch? *Journal of Experimental Education*, 48, 39-42.
- Endler, N. S., Speer, R. L., Johnson, J. M., Flett, G. L. (1999) General self-efficacy and control in relation to anxiety and cognitive performance. *Current Psychology*, 20, 36-52.
- Fernandez, G. C., & Liu, L. (1999) A technology-based teaching model that stimulates statistics learning. *Computers in the Schools*, 16, 173-191.
- Findley, M. J. & Cooper, H. M. (1983) Locus of control and academic achievement: A literature review. *Journal of Personality and Social Psychology*, 44, 419-427.
- Garfield, J., Hogg, B., Schau, C., & Whittinghill, D. (2002) First courses in statistical science: The status of educational reform efforts. *Journal of Statistics Education*, 10, 19.
- Hindman, C. D. (1980) Crib notes in the classroom: Cheaters never win. *Teaching of Psychology*, 7, 166-168.
- Ioannidou, M. K. (1997) Testing and life-long learning: Open-book and closed-book examination in a university course. *Studies in Educational Evaluation*, 23(II), 131-139.
- Lefcourt, H. M. (1992) Durability and impact of the locus of control construct. *Psychological Bulletin*, 112, 411-414.
- Loftman, G. R. (1975) Study habits and their effectiveness in legal education. *Journal of Legal Education*, 27, 418-472.
- Onwuegbuzie, A. J. (2004) Academic procrastination and statistics anxiety. *Assessment & Education in Higher Education*, 29, 3-19.
- Rotter, J. B. (1966) Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 80 (1, Whole No. 609).
- Skidmore, R. L., & Aagaard, L. (2004) The relationship between testing condition and student test scores. *Journal of Instructional Psychology*, 31, 304-313.
- Thompson, S. C. (1991) Intervening to enhance perceptions of control. In Snyder, C. R., & Forsyth, D. R. *Handbook of Social and Clinical Psychology*. (pp.589-609). NY: Pergamon Press.
- Wachsman, Y. (2002) Should cheat sheets be used as study aids in economics tests? *Economics Bulletin*, 1, 1-11.
- Ware, M. E., & Chastain, J. D. (1991) Developing Selection Skills in Introductory Statistics. *Teaching of Psychology*, 18, 219-222.
- Webb, N. M. (1997) Assessing students in collaborative groups. *Theory Into Practice*, 36, 205-213.
- Yesilcay, Y. (2000) Research Project in Statistics: Implications of a Case Study for the Undergraduate Statistics Curriculum. *Journal of Statistics Education*, 8, 14.
- Zimmerman, B. J., & Pons, M. (1986) Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, 23, 614-628.