# Student Achievement Goals with Alternative and Traditional Exam Formats

**Connor Ott, University of Colorado Boulder**

Connor Ott is a recent graduate of the University of Colorado Boulder where he received a bachelor's and master's degree in Aerospace Engineering Sciences. His research primarily focuses on assessment in large engineering courses and ways to improve their effectiveness as feedback for students and instructors. He currently works as an aerospace engineer designing satellite missions at Advanced Space in Westminster, Colorado.

**Dr. Kathryn Anne Wingate, University of Colorado Boulder**

Dr. Kathryn Wingate is an instructor at University of Colorado Boulder, where she teaches design and mechanics courses. She holds her PhD in mechanical engineering, and worked at NGAS as a materials scientist.

**Dr. Aaron W. Johnson, University of Colorado Boulder**

Aaron W. Johnson is an Instructor in Smead Aerospace Engineering Sciences at the University of Colorado Boulder. Prior to this he was a postdoctoral research fellow at the University of Michigan and the Tufts University Center for Engineering Education and Outreach. He received his Ph.D. in Aeronautics and Astronautics from the Massachusetts Institute of Technology in 2014 and a bachelor's degree in aerospace engineering from the University of Michigan in 2008.

# Student Achievement Goals with Alternative and Traditional Exam Formats

**Abstract**
In an engineering course, students' motivation during an exam can be understood through the framework of their "achievement goals," which can be divided into performance and mastery goal orientations. Performance goals are associated with seeking high performance and avoiding low performance. Mastery goals are associated with developing skills and increasing understanding. We hypothesize that in large courses, the structure of exams—as high-value assessments—has a large influence on students' goal orientation. This study investigates the use of an alternative exam structure in a sophomore-level aerospace engineering course at the University of Colorado Boulder. The instructors gave students six quizzes and an optional final exam, rather than the traditional three midterm and comprehensive final exam format historically used in the course and similar courses in the curriculum. The optional final exam allowed students to correct mistakes and misconceptions from up to three of the midterm quizzes. The researchers queried students with a survey regarding both assessment formats to determine differences in perceived goal structures surrounding them. From 112 responses, survey results indicate a noticeable difference in student attitude towards the traditional and alternative assessment formats. Students generally perceived the alternative format to be more mastery-oriented than the traditional format and perceive the traditional format as being more performance-oriented than the alternative format. In our discussion, we point out how these findings could help instructors design more focused assessments that target different achievement goal structures, without sacrificing the efficiency and rigor of administering traditional exams to large engineering classes.

## Introduction

In an engineering course, students' motivation during an exam can be understood through the framework of "achievement goals," which can be divided into *performance* (demonstrating competence) and *mastery* (developing competence) goal orientations. In a broad sense, performance goals are associated with seeking high performance and avoiding low performance. Mastery goals are associated with developing skills and increasing understanding. Eppler et al. (2000) point out that the behavior patterns associated with performance and mastery goal orientations can be valuable when coordinated with one another, and too great an emphasis on either can be problematic [1]. Furthermore, Elliot & Dweck (1988) demonstrated that the perceived value of a task (in the context of achievement goals) factors into an individual's behavioral response to the task [2]. The interaction between instructors and students can help to influence students' goal orientation in an undergraduate engineering course, but this becomes difficult as class sizes increase and there is more detachment between student and instructor. Therefore, we hypothesize that in large courses, the structure of exams—as high-value assessments—can be modified to influence students' goal orientation.

This study investigates the use of an alternative exam structure in a sophomore-level aerospace engineering course at the University of Colorado Boulder during the Fall 2020 semester. The

second and third authors of this paper served as the co-instructors of this course, and the first author served as a graduate teaching assistant in the course. All three authors also served in these same roles the previous time the course was taught in Fall 2019. The instructors gave students six quizzes and an optional final exam, rather than the traditional three midterm and comprehensive final exam format historically used in the course and similar courses in the curriculum. This traditional exam format is not generally presented to students as specifically targeting performance achievement goals; however, its design tends to implicitly emphasize demonstrating over developing competence. The alternative six-quiz format entails more frequent, shorter assessments, each covering less material. The optional final exam consists of six questions, each corresponding to material on a quiz. Students could answer three questions on the final to replace up to three quiz scores. Students were told that the optional final could only improve their quiz score. That is, the higher score between the original quiz and final exam would count as their final grade for the quiz. The instructors did not allow students to make up quizzes during the semester; therefore, some students used the optional final to replace a quiz score of 0 due to extenuating circumstances. This study, however, focuses on a different, larger group of students—those who saw the optional final as an opportunity to revise their understanding and respond to their initial performance explicitly and directly.

Most students in this class were simultaneously enrolled in another aerospace course that kept the traditional three-exam and comprehensive final format. In the time between each class's final midterm assessment and before their final exams, volunteers enrolled in both courses were asked to complete a survey with questions adapted from the Patterns of Adaptive Learning Scales (Midgley et al., 2000), a tool used to examine student motivations in the context of an achievement goal framework [3]. To assess the influence of exam structure, survey items were tailored to the alternative (six-quiz) or traditional (three-exam) assessment formats, and quantitative analyses were performed to examine differences in student goal orientation. The work presented here may provide insight as to whether alternative assessment methods can be used to orient students towards a specified goal orientation. This would allow instructors of large courses more control over their students' goal orientations and provide a foothold for further research investigating student motivation in large classes.

**Methods**
Students' perceptions of classroom goal structures were evaluated using the Patterns of Adaptive Learning Scales (PALS). Each question is answered on a 5-point Likert scale rating how much respondents agreed with a statement. Questions regarding classroom goal structures were adapted to target the alternative and traditional assessment formats. As an example, the PALS question, "In our class, how much you improve is really important." was modified to ask: "On the ASEN 2001 quizzes, how much you improve is really important" to target achievement goals surrounding the alternative assessment format. Questions were modified to target each class individually, so each question about one course's midterm assessments has an identical question targeting the other course.[1] Questions on the survey specifically target the midterm assessments

---

[1] The alternate version of this question is, "On the ASEN 2002 exams, how much you improve is really important."

(quizzes and exams) rather than the final exams. However, for the alternative assessment format, the role of optional final exam is explicitly related to the quizzes as a chance to revise scores. This was made clear to students throughout the semester, and it is expected that this influenced their response to questions about the quizzes. The PALS questions are subdivided into three subcategories: performance-avoid, where questions are framed around avoiding negative judgement, performance-approach, where questions are framed around seeking positive judgement, and mastery, where questions are framed around seeking growth in capability. Each subcategory contains independent sets of questions. Higher scores on performance questions indicate a performance-oriented goal structure, and the same is true for mastery questions. For this study's analysis, performance-avoid and performance-approach responses were grouped together. To provide additional context to this research, the survey included questions regarding students' personal motivations in the context of achievement goal framework. An example personal mastery goal question is: "It's important to me that I thoroughly understand my class work."

The survey was approved by the university's Institutional Review Board (IRB) and students were informed of the purpose and format of the survey, as well as that their responses would remain anonymous. The survey was presented during the final lab period in the alternative course, where student attendance was expected. This took place before the students' final exams in both courses, but after the last midterm assessment (i.e., after the 6th quiz and 3rd exam). Students were provided with a link to complete the survey online. Incomplete surveys and surveys in which students indicated they either did not intend to take the optional final exam, or that they intended exclusively to take the final for a reason other than to improve a poor quiz score, were omitted from analysis. Of 323 students enrolled, 134 responses were collected. Of these, 112 responses were included in the data analyses according to the above criteria.

Response data were coded as integer values from 1 to 5 based on lowest (1) to highest (5) levels of agreement with each statement. Each response was grouped as either performance or mastery questions. Analogous questions in these groups were tested against each other using a two-tailed repeated measures t-test. In total, the survey consists of 9 personal (5 mastery and 4 performance) and 16 course structure (4 mastery and 4 performance for each class) questions for 25 total questions.

**Study Context**
The courses studied in this research are two required sophomore-level aerospace courses which students generally take concurrently during the fall semester of their sophomore year. These are: ASEN 2001: Introduction to Statics, Structures, and Mechanics (Statics/Structures), and ASEN 2002: Introduction to Thermodynamics and Aerodynamics (Thermo/Aero). Statics/Structures combines a traditional statics course with the first half of a traditional mechanics course, and covers force equilibrium, truss analysis, stress, strain, and beam analysis. The Thermo/Aero course combines the first half of a traditional thermodynamics course with the first half of an aerodynamics course, and covers the conservation of energy, one-dimensional compressible and incompressible flows, two-dimensional flow, lift, and drag. To understand student baseline

impressions of these courses, the publicly available university-administered end-of-semester evaluations from Fall 2019, the year preceding this study, were examined for each course. The faculty instructors of record were the same during the Fall 2019 and Fall 2020 semesters for both courses. Students' evaluations of the courses from Fall 2019 indicated that students had a higher personal interest prior to enrollment in the Thermo/Aero class (4.8/6) than the Statics/Structures class (3.9/6). Additionally, students rated the intellectual challenge of the Thermo/Aero course to be a 5.1/6, which was higher than the intellectual challenge rating of the Statics/Structures course (4.7/6). In Fall 2019, the instructor overall ratings for the Statics/Structures and Thermo/Aero courses were very close. For the Fall 2020 semester, the University changed the end-of-semester evaluation questions and they did not include questions on instructor effectiveness or course difficulty. We include this discussion of student evaluations to note a limitation of the study, which is that we compare two fundamentally different courses. Though the survey questions do not ask students about personal interest, intellectual challenge, or the instructors, differences in these areas exist and contribute to students' overall perception of a class and may influence how they perceive the goal structures we have asked them to reflect on.

In addition to the goal structures a classroom presents to students, students themselves may identify more closely with some motivations than others. The box-and-whisker plots in Figure 1 indicates how students self-identify in the context of achievement goals. The box-and-whisker plots mark the minimum, 1st quartile, median, 3rd quartile, and maximum of each distribution. The whiskers are subject to being shortened for distributions where the minimum or maximum are greater than 1.5 times the interquartile range from the median.
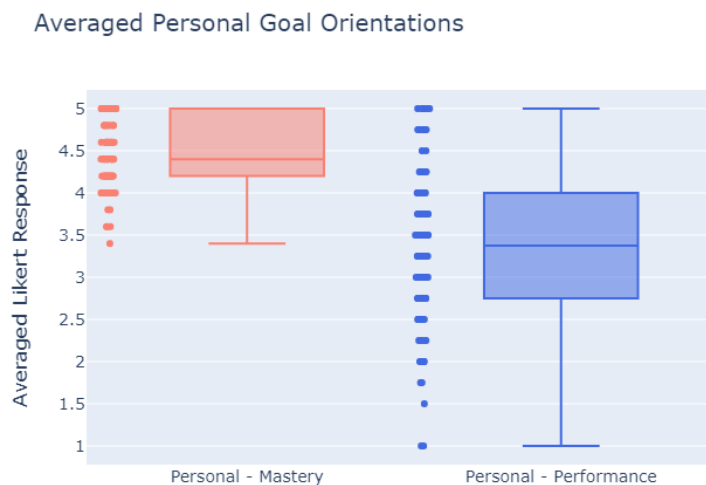


*Figure 1: Averaged responses to personal goal orientation questions.*

Each point in these distributions represents a student's average response to the 4 personal-mastery questions (red) and 5 personal-performance questions (blue). This gives some context to how students self-identify within the achievement goals framework. We gather from this plot that students generally identify with a mastery goal orientation but are less consistent in whether they identify with a performance goal orientation. This aligns with previous work studying student goal orientation [1].

**Results**

Each survey question regarding course goals is the same for the Statics/Structures course and the Thermo/Aero course, with slight modifications to target one course or the other. Figure 2 shows the average mastery and average performance score distribution for each class. Each point represents the average of the 4 questions in each category for a student.
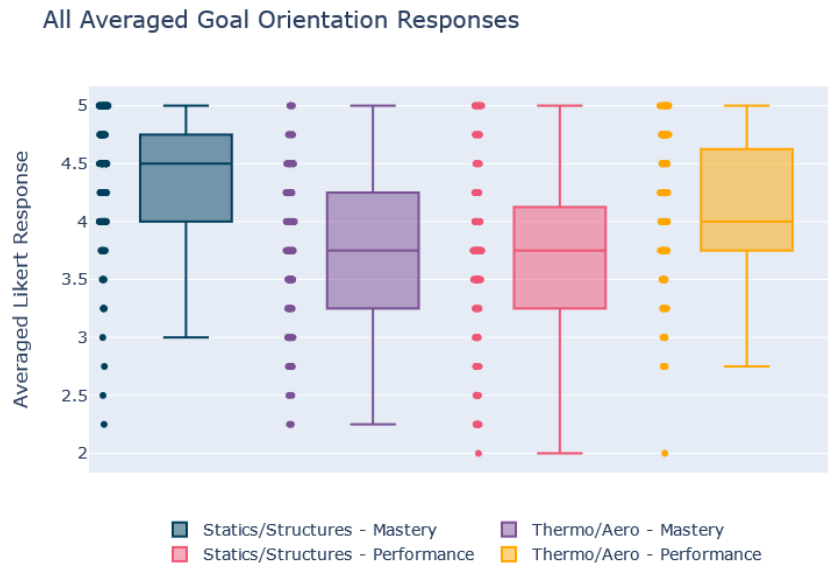


*Figure 2: Mastery and performance question responses averaged for each student.*

To determine whether students perceive a difference in goal structure between the Statics/Structures course and the Thermo/Aero course, we compare student responses in a within-subjects format. Figure 3 below shows the differences in each student's response to analogous questions, subtracting the scores for the Thermo/Aero class from those for the Statics/Structures class. Questions 1-4, highlighted in red, correspond to mastery questions, while Questions 5-8, highlighted in blue, correspond to performance questions. Positive values indicate a higher score for the statics and structures course, while negative values indicate a higher score for the aero and thermo course.
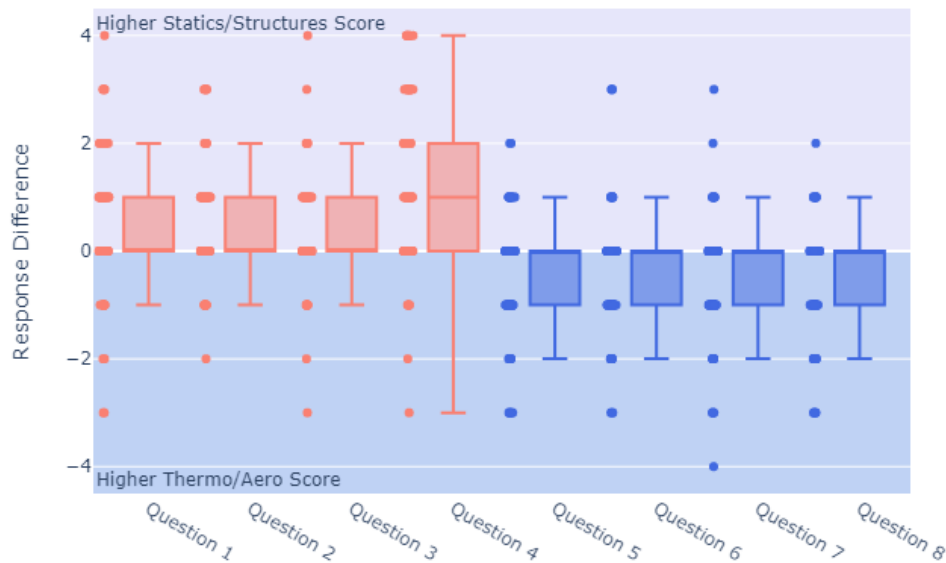
Mastery and Performance Question Delta

*Figure 3: Differences in response for each question. Mastery questions highlighted in red, performance questions highlighted in blue.*

The following table shows the results of testing these groups of response differences against the null hypothesis that they are zero mean. This is equivalent to a two-tailed repeated measures t-test.

*Table 1: Quantitative results of statistical tests.*

| Number | Question | $t$ | $p$ | Effect Size |
|---|---|---|---|---|
| 1 | On the [alternative/traditional assessment], how much you improve is really important. | 4.3 | $\ll 0$ | 0.4 |
| 2 | For the [alternative/traditional assessment], really understanding the material is the main goal. | 4.7 | $\ll 0$ | 0.5 |
| 3 | For the [alternative/traditional assessment], it's important to understand the work, not just memorize it. | 3.2 | 0.002 | 0.3 |
| 4 | On the [alternative/traditional assessment], it's OK to make mistakes as long as you are learning. | 7.9 | $\ll 0$ | 0.8 |
| 5 | For the [alternative/traditional assessment], getting good grades is the main goal. | -4.4 | $\ll 0$ | 0.4 |
| 6 | On the [alternative/traditional assessment], getting right answers is very important. | -3.4 | 0.001 | 0.3 |
| 7 | On the [alternative/traditional assessment], it's important to get high scores. | -4.5 | $\ll 0$ | 0.4 |
| 8 | On the [alternative/traditional assessment], it's important not to do worse than other students. | -4.9 | $\ll 0$ | 0.5 |

**Discussion**
This research is concerned with whether the structure of high-value assessments can orient students towards either mastery or performance goal orientations. To address this, we focus on the difference in student responses to survey questions about the assessment formats in two classes with drastically different midterm assessment formats. Figure 3 gives the first insight to whether students perceive differences in the overall goal of the assessments in each class. At a glance, this figure indicates that students generally interpret the alternative assessment format as being more oriented towards mastery than the traditional format, and vice versa for the traditional assessment format. We interrogate this further using a two-tailed repeated measures t-test, resulting in Table 1. Each t-test tests the null hypothesis that the distribution of response differences from the alternative course (Statics/Structures) to the traditional course (Thermo/Aero) is zero mean. For $p < 0.01$, we reject the null hypothesis and accept the alternative hypothesis that the differences in responses are not zero mean, meaning that there was a statistically significant difference in how students responded to that question when considering Statics/Structures vs. Thermo/Aero. Each question yielded significantly different responses between the alternative assessment format and the traditional assessment format $t(112)$, $p < 0.01$. Mastery questions (numbers 1-4) all resulted in $t > 0$, indicating a higher score for the alternative assessment format. Performance questions (numbers 5-8) all resulted in $t < 0$, indicating higher scores for the traditional assessment format. From these results, we gather that students generally perceive the alternative assessment format as more mastery-oriented, and the traditional format as more performance-oriented.

Question number 4, "On the [alternative/traditional assessment], it's OK to make mistakes as long as you are learning," shows the greatest difference between the alternative and traditional formats. This is attributed to the fact that the alternative format is uniquely tailored towards correcting mistakes through the optional final exam where students can replace up to three quiz grades. Though the traditional format provides some room for improvement via a final comprehensive exam, the connection to previous mistakes is often less explicit, and the avenue to correct them, less direct.

While all results are significant, questions generally vary between low and medium effect sizes, except for Question 4, for reasons explained above. We take this to lend credence towards the idea that students see a genuine difference between the goal orientations laid out by the alternative and traditional assessment formats. Effect size should be considered with the advantages and drawbacks of implementing this assessment format discussed later in the discussion.

Though the difference between each assessment format is clear, we take care to note that these differences should be contextualized by the overall scores seen by the mastery and performance questions for both classes. Figure 3 shows the distribution of responses swapping between higher mastery for the alternative format and higher performance for the traditional format. However, no individual distribution is particularly low. Students generally agree that both classes exhibit characteristics of mastery and performance goal orientations. This is important to us because it

indicates that, though the alternative format is decidedly different from the traditional format, there is still an appropriate emphasis on getting good grades and correct answers, in addition to understanding mistakes and revising misconceptions. Dweck & Leggett's (1988) original model reminds us that the overemphasis on performance goals can lead to learned helplessness, a response to failure that teaches individuals to doubt their ability to grow and learn in the face of failure. On the other hand, too great an emphasis on mastery goals can lead to maladaptive overpersistence, with no room to acknowledge or accept failure in the first place [4]. Heyman & Dweck (1992) acknowledge that an individual who solely pursues mastery goals, but never meets performance requirements such as grades or workplace productivity, may be hurting themselves more than their mastery orientation helps them. They elaborate further on the interplay of mastery and performance goals, pointing out that an individual may hold many goals aligning with either mastery or performance orientations. They posit that mastery and performance goals should be coordinated with one another to promote adaptive learning patterns in a way that encourages growth and development and at the same time recognizes achievement [5]. We believe the alternative assessment format follows this line of thinking, by providing explicit routes to demonstrate growth and improvement while not straying away from performance-oriented goals.

Lastly, there are practical benefits of implementing the alternative assessment format. From a faculty perspective, we found numerous benefits to the alternative assessment format. In a large class (>300 students), tracking and validating the documentation of excused exam absences due to participation in university-sanctioned events, medical and family emergencies, or unexpected life events can become a tremendous burden on the teaching staff. Further, scheduling and writing multiple make up exams is simply not feasible with large numbers of students. This alternative assessment format allowed us to implement a strict "no quiz make-up" policy as students could easily "make-up" a missed quiz by completing the corresponding final exam question. This policy was also helpful for non-traditional students who may have childcare emergencies, work emergencies, or other constraints. We were able to transform from "policing" exam make-up requests to providing students with a flexible assessment format that they could then implement to fit their life. In terms of faculty load, we found that implementing a six quiz with optional final exam format was about the same amount of work as the traditional assessment format in terms of exam development and grading.

We also acknowledge the potential drawbacks to the alternative assessment format. Some faculty prefer to write a final exam where each question combines several topics covered in the course. This would be difficult to implement in the six-quiz and optional final format, as each final exam question must map to a specific quiz topic. If this policy is implemented across several courses in one semester, students taking four courses could have up to 24 quizzes in a single semester, which may be overwhelming. Finally, we want to acknowledge that this work did not examine students' long-term retention of information, and therefore it is unclear if one of these assessment formats is better for long-term retention.

**Limitations**
We note that results presented here are contingent on the assumption that the major difference between the alternative and traditional assessments is their format. However, the two courses host very different content, and student feedback from prior years indicates some difference between intellectual challenge and personal interest in each class. Though the results here are not necessarily related to intellectual challenge or interest, mastery and performance goal orientations are a way of describing student motivation, which can certainly be influenced by intellectual challenge and personal interest. We deliberately chose and modified questions that focus on assessment format, but even this may not completely control for these other factors.

Additionally, questions used in this survey are a subset of the full PALS survey from which they were adapted. There is a question as to whether the subset of questions used provide consistent enough results to indicate a mastery or performance orientation (or lack thereof). We're confident that the high number of responses, along with the significance of test results, helps to outweigh this limitation.

Finally, we acknowledge that this study was performed during the Fall 2020 semester while the United States faced the brunt of the COVID-19 pandemic. Both courses engaged in asynchronous lectures with virtual lab periods and take-home assessments. These complications may confound this research and should be noted for faculty and researchers attempting to use this study in the future.

**Future Work**
We hope this work serves as a valuable resource for instructors looking to encourage more mastery-oriented practices in their classes. However, we believe more work is warranted to prove the efficacy of the alternative assessment format. A first step would be to expand the study over more semesters and collect additional survey data about students' perceptions of traditional and the alternate exam format. This expanded data set would provide data from semesters not impacted by the COVID-19 pandemic, thereby increasing the generalizability of the results.

We also appreciate that, while students may find this format more flexible and forgiving, it's not clear whether this is an accurate metric for learning. We have no reason to suspect it not to be, but we maintain a healthy skepticism of all assessment formats until they're shown to be good indicators of learning and retention. In addition to assessing the alternative format's use as a measure of learning, it would be interesting to further investigate the relationship between personal goal orientation, assessment goal orientation, and long-term retention. Also, an analysis of the usefulness of this assessment format for lower- and higher-performing groups of students may be interesting. We suspect that this format is especially helpful for students near the cutoff for a passing grade whose initial midterm performance may be discouraging to the point of withdrawal.

**References**

[1]     M. A. Eppler, C. Carsen-Plentl, and B. L. Harju, "Achievement Goals, Failure Attributions, and Academic Performance in Nontraditional and Traditional College Students," *J. Soc. Behav. Pers.*, vol. 15, no. 3, pp. 353–372, 2000.

[2]     E. S. Elliot and C. S. Dweck, "Goals: An Approach to Motivation and Achievement," *J. Pers. Soc. Psychol.*, vol. 54, no. 1, pp. 5–12, 1988, doi: 10.1080/02109395.1989.10821105.

[3]     C. Midgley *et al.*, "Manual for the Patterns of Adaptive Learning Sciences (PALS)," *Pals*, pp. 734–763, 2000.

[4]     C. S. Dweck and E. L. Leggett, "A Social-Cognitive Approach to Motivation and Personality," *Psychol. Rev.*, vol. 95, no. 2, pp. 256–273, 1988, doi: 10.1037/0033-295X.95.2.256.

[5]     G. D. Heyman and C. S. Dweck, "Achievement goals and intrinsic motivation: Their relation and their role in adaptive motivation," *Motiv. Emot.*, vol. 16, no. 3, pp. 231–247, 1992, doi: 10.1007/BF00991653.

## Appendix – Survey Questions

These are the questions given to students. The first set of questions include the personal-mastery and personal-performance questions. The second set includes the course-specific questions.

*The following questions are about yourself as a student. Please rate how much you agree with each statement.*

| | Strongly disagree (1) | Disagree (2) | Neither agree nor disagree (3) | Agree (4) | Strongly agree (5) |
|---|---|---|---|---|---|
| It's important to me that I learn a lot of new concepts this year. (1) | ○ | ○ | ○ | ○ | ○ |
| One of my goals in my classes is to learn as much as I can. (2) | ○ | ○ | ○ | ○ | ○ |
| One of my goals is to master a lot of new skills this year. (3) | ○ | ○ | ○ | ○ | ○ |
| It's important to me that I thoroughly understand my class work. (4) | ○ | ○ | ○ | ○ | ○ |
| It's important to me that I improve my skills this year. (5) | ○ | ○ | ○ | ○ | ○ |
| One of my goals is to show others that I'm good at my class work. (6) | ○ | ○ | ○ | ○ | ○ |
| It's important to me that other students in my classes think I am good at my class work. (7) | ○ | ○ | ○ | ○ | ○ |
| It's important to me that I don't look stupid in my classes. (8) | ○ | ○ | ○ | ○ | ○ |
| One of my goals in class is to avoid looking like I have trouble doing the work. (9) | ○ | ○ | ○ | ○ | ○ |

*The following questions are about the course assessment format. Please rate how much you agree with each statement.*

Note, [assessment format] is a placeholder for the interchangeable course numbers and assessment format for each class. For example, "On the ASEN 2002 exams, how much you improve is really important" or "On the ASEN 2001 quizzes, how much you improve is really important"

| | Strongly disagree (1) | Disagree (2) | Neither agree nor disagree (3) | Agree (4) | Strongly agree (5) |
|---|---|---|---|---|---|
| On the [assessment format] how much you improve is really important. (1) | ○ | ○ | ○ | ○ | ○ |
| For the [assessment format], really understanding the material is the main goal. (2) | ○ | ○ | ○ | ○ | ○ |
| For the [assessment format], it's important to understand the work, not just memorize it. (3) | ○ | ○ | ○ | ○ | ○ |
| On the [assessment format], it's OK to make mistakes as long as you are learning. (4) | ○ | ○ | ○ | ○ | ○ |
| For the [assessment format], getting good grades is the main goal. (5) | ○ | ○ | ○ | ○ | ○ |
| On the [assessment format], getting right answers is very important. (6) | ○ | ○ | ○ | ○ | ○ |
| On the [assessment format], it's important to get high scores. (7) | ○ | ○ | ○ | ○ | ○ |
| On the [assessment format], it's important not to do worse than other students. (8) | ○ | ○ | ○ | ○ | ○ |