Exploring the Interplay of Achievement Goals, Self-Efficacy, Prior Experience and Course Achievement

Hannu Pesonen
Aalto University
Espoo, Finland
hannu.pesonen@aalto.fi

Juho Leinonen
The University of Auckland
New Zealand
juho.leinonen@auckland.ac.nz

Lassi Haaranen
Aalto University
Espoo, Finland
lassi.haaranen@aalto.fi

Arto Hellas
Aalto University
Espoo, Finland
arto.hellas@aalto.fi

ABSTRACT

We explore achievement goal orientations, self-efficacy, gender, and prior experience, and look into their interplay in order to understand their contributions to course performance. Our results provide evidence for the appropriateness of the three-factor achievement goal orientation model (performance, mastery approach, mastery avoidance) over the more pervasive four-factor model. We observe that the aspects and the model factors correlate with course achievement. However, when looking into the interplay of the aspects and the model factors, the observations change and the role of, for example, self-efficacy as an aspect contributing to course achievement diminishes. Our study highlights the need to further explore the interplay of aspects contributing to course achievement.

CCS CONCEPTS

• **Social and professional topics → Computing education**.

KEYWORDS

achievement goal orientation, self-efficacy, course achievement

ACM Reference Format:
Hannu Pesonen, Juho Leinonen, Lassi Haaranen, and Arto Hellas. 2023. Exploring the Interplay of Achievement Goals, Self-Efficacy, Prior Experience and Course Achievement. In *The United Kingdom and Ireland Computing Education Research (UKICER) conference (UKICER 2023), September 07–08, 2023, Swansea, Wales Uk.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3610969.3611178

1 INTRODUCTION

Computing education research has explored a wide variety of aspects that contribute to students' behavior and achievement (see e.g. [22, 27, 32, 37, 38, 43]). These include students' achievement goal orientation, self-efficacy, gender, and so on. Much of this attention has been directed to introductory programming courses [32]. When seeking to understand aspects that contribute to students' behavior, however, it is also important to look into subsequent courses, as aspects such as achievement goal orientation and self-efficacy are all evolving processes [34, 42]. In this work, our focus is on four aspects that have received plenty of attention in prior computing education research: achievement goal theories, self-efficacy, prior experience, and gender. In brief, achievement goal theories have been classically described through *intrinsic* (i.e. *mastery* and learning) and *extrinsic* (i.e. *performance*) goals, discussing how students approach learning [34]. Self-efficacy refers to the beliefs of an individual about their own abilities to act in a way that they can succeed [3]. Prior experience in general describes existing knowledge and exposure to the topic. Finally, the role of gender has also been studied in computing education research, e.g. in seeking to understand the confidence gap between genders [8, 39]. All of these have been found to correlate with course achievement in prior computing education research studies.

In essence, our study explores topics of prior studies in new contexts and with new data [1, 25]. We surveyed students in a web software development course (primarily intended for second-year computer science students), asking for information related to the above four aspects. Examining the survey data and contrasting it with course achievement, in this article, we answer the following research question: *What is the relationship between course achievement, achievement goal orientations, self-efficacy, gender and prior experience?*

UKICER 2023, September 07–08, 2023, Swansea, Wales Uk

ACM ISBN 979-8-4007-0876-3/23/09.
https://doi.org/10.1145/3610969.3611178

2 BACKGROUND

2.1 Aspects Relating to Student Achievement

2.1.1 Achievement Goal Orientations. Achievement goal orientations (in short: *achievement goals*) are typically defined as schemata that organize and provide structure and meaning to achievement-related situations [34, 46]. Achievement goal orientation theories in essence are interested in why things are done rather than what is being done. They are typically combinations of two dimensions. The first dimension consists of a propensity towards either mastery of a skill for the sake of learning it (mastery goal orientation) or towards acquiring the skill because it entails some form of external reward (performance goal orientation). The second dimension is whether to approach a given task or to avoid it (see e.g. [2, 17]). The dimensions combine to form 4 distinct orientations (or strategies):

Performance-approach, performance-avoidance, mastery-approach, and mastery-avoidance.

There has, however, been considerable debate as to whether the four distinct strategies is the best model. For instance, a two-goal structure has been argued for, where the goals would entail a mastery-approach goal orientation and a performance goal orientation, where the latter would be a combination of both approach and avoidance [34].

Of these orientations, mastery orientation is related to perseverance, interest, adaptive help-seeking and achievement (e.g., [9, 24, 34, 41]). Pursuing mastery-approach goals has thus been found to be beneficial in fostering interest in academics and also related to deeper-processing cognitive and metacognitive strategies [34].

2.1.2 *Achievement goals and CS.*

In a multi-context study looking into achievement goals in introductory programming, Zingaro et al. [55] replicated earlier studies on the role of achievement goals in introductory programming. They highlighted that mastery goals – or other goals for that matter – predicted exam grades only in some of the course contexts. Also, contrary to earlier findings, normative goals – a subcategory of performance goals – correlated positively with exam outcomes in only one of the course contexts under study. These results highlight the importance of looking into contextual aspects such as the teaching approach, which may influence students' perceptions of their ability [54] and even the way students approach learning tasks [36].

2.1.3 *Self-efficacy.*

Self-efficacy relates to the beliefs of an individual on their ability to successfully execute tasks in a domain [3]. Over time, it evolves through performing tasks, observing others performing tasks, receiving encouragement and feedback; it is also influenced by physiology and mental states, including stress [4]. In CS education research, self-efficacy has been explored together with goal orientations, prior experience and gender (see, e.g., [30]), and it has been found to correlate with course performance [7, 44, 50, 51]. It has also been noted that self-efficacy can change during a course [51] and that the instructional approach likely contributes to self-efficacy [54]. When considering achievement goal orientations, Lishinski et al. [30] discovered that self-efficacy was the primary predictor of course performance in a CS1 context, with mastery goal (and not performance goal) orientation only exhibiting an indirect effect on course performance through self-efficacy.

2.1.4 *Prior Experience.*

As self-efficacy evolves over time through performing tasks [4], it is not surprising that prior experience has been studied when seeking to understand aspects that contribute to course achievement [6, 23, 51, 52]. In the introductory programming context, prior programming experience (and prior experience of ICT use) has been observed to positively correlate with course achievement [23, 51]. However, some studies reported no considerable correlation between prior programming experience and introductory programming course outcomes (see e.g. [6]).

2.1.5 *Gender & Confidence Gap.*

Prior research has identified an effect which has been dubbed the confidence gap [8, 31]. Male students tend to self-report higher rates of confidence in their ability compared to their female classmates. However, this difference in confidence does not translate into actual performance in terms of ability [8, 47]. In fact, some studies have reported women having slightly higher grade averages than men [11].

Duran et al. [16] investigated a CS1 course over multiple years offered both as traditional on campus and as an online course. They found that a gap in confidence between men and women existed in both local and online versions of the course. Pirttinen et al. [39] looked at multiple instances of a CS1 course and compared the self-efficacy of students from different majors. They found that overall, men were more confident in their ability to do well in the CS1 course regardless of their study major. At the same time, men in CS reported higher confidence compared to men in other majors while a similar effect for women did not exist.

Murphy et al. reported women having lower levels of prior programming experience before the introductory programming course [35]. However, they noted that after finishing the introductory course women reported nearly equal levels of mastery and did not find programming more challenging than men.

2.2 Contextual Aspects and Need for Replication

In line with the broader call for replication studies [25], we view that research into the relationship between programming experience, self-efficacy, achievement goals, and course achievement needs further evidence. In particular, prior studies on these topics in computing education research have primarily focused on introductory programming courses, and there is a need to understand whether the same or similar observations hold up in subsequent courses. We also view that further evidence is needed on whether mastery goals and self-efficacy are a *must-have* to improve course outcomes more broadly in CS studies. In line with this is the question of whether moving forward in CS studies might bring about a less performance- or ego-involved approach and a more task- and mastery-involved one. Hence, analysis of these aspects from a subsequent course can offer insight into this, when contrasted with results from CS1-related studies. As highlighted in [16, 29, 30, 39], there is also a need for further insight into the topic of confidence gap and whether prior programming experience possibly mediates the effect of gender.

3 METHODOLOGY

3.1 Context

The context of the study is Aalto University, which is a research-oriented university in Europe. The majority of the students at Aalto University are native Finnish speakers but fluent in English as well. The University also has a growing international student body. The study was conducted in a web software development course, which is typically taken during the second or third year of Bachelor studies in the Computer Science program. International students who enroll in graduate programs can also take the course depending on whether they have sufficient background in programming.

The course teaches the principles of designing and building web applications with a strong focus on server-side functionality including the design of application programming interfaces, securing web applications, and working with databases. The front end functionality of the course focuses on building user interfaces with template-based languages.

The course went through a major revision in 2020 (also in part to support students during the Covid-19 pandemic). During the major revision, an online ebook with automatically assessed programming exercises and quizzes and a larger programming project with a number of mandatory and optional features was created for the course. The course focuses on using JavaScript, HTML and CSS, and relies on JavaScript when building server-side functionality.

3.2 Study and Data Collection

The present study was conducted in the fall of 2020 during the first iteration of the revised course.

3.2.1 Course survey. At the beginning of the course, students were asked to complete a survey that asked about their self-efficacy, prior programming experience and types of goals. The questions related to types of goals were from the Achievement Goal Questionnaire – Revised (AGQ-R) [19] that explores students' mastery and performance goal orientations, with slightly varying items for gauging each of the four distinct types of achievement goals discussed in Section 2 to form four sum variables whose validity and reliability can then be tested in accordance with the basic principles of psychometrics. The AGQ-R was chosen over the original Achievement Goal Questionnaire [18] due to concerns about interpretations in some of the original questions (for a broader discussion, see [19]).

The questions of the survey are outlined in Table 1. In addition, the survey included an introduction to the study, options for providing (or not providing) research consent, contact details for mapping answers to course outcomes, guidelines for answering the questions, gender and year of birth.

Table 1: Course Survey outlining Self-efficacy, Prior experience, and Types of Goals. Answers to each of the questions were given on a 5-point Likert scale ranging from Strongly disagree to Strongly agree. The labels MAP, MAV, PAP, and PAV correspond to Mastery-approach, Mastery-avoidance, Performance-approach, and Performance-avoidance and have been added here for clarity.

Self-efficacy
I believe that I will do well in this course on Web Application Development
Prior Experience
I have completed a course on programming.
I have completed a course on databases.
I have completed a course on web development.
I have used HTML and CSS.
I have programmed using JavaScript.
I know how web applications work.
I know how web servers work.
I can already simple write web applications.
I can already write complex web applications.
I have worked as a software developer.
Types of Goals
My aim is to completely master the material presented in this class. (MAP1)
I am striving to do well compared to other students. (PAP1)
My goal is to learn as much as possible. (MAP2)
My aim is to perform well relative to other students. (PAP2)
My aim is to avoid learning less than I possibly could. (MAV1)
My goal is to avoid performing poorly compared to others. (PAV1)
I am striving to understand the content of this course as thoroughly as possible. (MAP3)
My goal is to perform better than the other students. (PAP3)
My goal is to avoid learning less than it is possible to learn. (MAV2)
I am striving to avoid performing worse than others. (PAV2)
I am striving to avoid an incomplete understanding of the course material. (MAV3)
My aim is to avoid doing worse than other students. (PAV3)

3.2.2 Grading. The course used automatic assessment for the programming assignments and relied on peer and staff assessment for the course project. The grade of the course was based on a combination of completed assignments and the course project. To pass the course, the student had to complete the mandatory parts of the project (approx 50% of all features) and to complete at least 70% of the programming assignments. For achieving the highest possible grade, the student had to complete at least 90% of the outlined features of the course project and at least 90% of the programming assignments in the course. The course featured weekly deadlines for the course assignments and the course project had an end-of-course deadline. No deadline extensions were given.

3.2.3 Ethics and incentives. Participating in the study was voluntary and a small amount of course points was awarded for completing the course survey, regardless of whether students provided research consent. The two options that students could choose from when giving research consent were as follows: (1) "I allow the use of my responses for research purposes. (Yes, my responses can be used for research purposes. I will answer the questions in this survey to the best of my ability.)" and (2) "No, my responses cannot be used for research purposes. I might not even try to answer all the questions in this survey to the best of my ability.". All research data was stored on computing facilities offered by Aalto University.

3.3 Data

In total, 351 students started the course and answered the course survey which was given as the first task in the course. 43 students did not provide research consent, which led to a sample of 308 students (87.7%). We then discarded all students who completed less than 5% of the course assignments (n = 52) or did not complete the survey, leaving us with data from 256 students (73% of the initial course population). Of the 256 students, 48 self-identified as women, and 208 as men[1]. All of the participants were 18 or over (i.e. not minors in terms of the ethics review). To assess the students' *course achievement* for the purposes of this study, we used their course points.

[1]The survey also included options for "other" and "do not wish to disclose", which were not used by the 256 included students. In the correlational analysis, men were encoded with number 1, while women were encoded with number 2.

3.3.1 Self-efficacy. For self-efficacy, we use the single question on students' beliefs in doing well in the course. A similar approach was taken in [39] and also in [29], albeit in a different setting.

3.3.2 Prior usage and know-how. For prior experience, we formed two (ad-hoc) sum variables based on the questions in the survey: (1) prior programming usage and (2) prior programming know-how. The first consisted of statements related to using web-related technologies (e.g. "I have used HTML and CSS"), while the second consisted of statements outlining more in-depth knowledge (e.g. "I know how web servers work", "I can already write complex web applications.").

3.3.3 Achievement goal orientations. To assess whether our data matched the four-factor model of AGQ-R and to further validate the AGQ-R survey, we conducted an exploratory factor analysis (EFA [20]) on the data[2]. Kaiser-Meyer-Olkin (KMO) test [26] showed that the collected data was adequately suitable for factor analysis (KMO test value 0.857), and Bartlett's test of Sphericity [5] showed that the data was not an identity matrix ($p < 0.001$). When inspecting the underlying factors, we used the Scree test [13] to assess the appropriate number of factors. As a rotational method, we used varimax [15], which is an orthogonal rotation method, with Kaiser normalization.

The exploratory factor analysis showed that our data fits a three-factor model instead of the four factor model. The resulting loadings are outlined in Table 2. We observed three factors: (1) Performance (goal) orientation (including both performance approach and avoidance goals), (2) Mastery-approach orientation, and (3) Mastery-avoidance orientation. The only item that loaded with some meaningful extent to two factors was MAV3 – we've included the loading of this into the Table 2 for both Mastery Approach and Mastery Avoidance. In the subsequent analyses, we use the identified three-factor model when studying achievement goal orientations.

Table 2: EFA loadings for the three factors from the data. The table includes scale items (referred to using labels presented in Table 1 for space purposes), factor loadings, and (interpreted) factor labels.

Item	Performance	Mastery Appr.	Mastery Avoid.
PAP3	0.889	–	–
PAV2	0.883	–	–
PAP2	0.878	–	–
PAP1	0.866	–	–
PAV3	0.862	–	–
PAV1	0.833	–	–
MAP2	–	0.833	–
MAP3	–	0.808	–
MAP1	–	0.790	–
MAV2	–	–	0.842
MAV1	–	–	0.836
MAV3	–	0.386	0.663

[2]We used SPSS version `28.0.1.0 (142)` for EFA analyses.

3.4 Approach

Our analysis for answering the research question is two-fold. First, we correlate (1) self-efficacy, (2) prior programming usage, (3) prior programming know-how, (4) mastery avoidance, (5) mastery approach, (6) performance orientation, (7) course achievement, and (8) gender. In this, we conduct $n = 28$ pairwise correlations. For the correlations, we use Spearman's r, which is a standard, oft-used rank-based method. Then, to highlight the need to explore the interplay of the aspects and to unveil potential overinterpretation of effects from pair-wise correlations, we explore partial correlations for each variable pair, controlling for the effect of the other variables. Partial correlation is a linear regression of the residuals of two variables, left over from their respective regressions with what forms the controlled variable (i.e. the covariate.)

When reporting results, we use p-values as one component that jointly with the correlations contribute to our understanding of the data [49], and following [21], we interpret $0.1 \leq r < 0.2$ as a small effect size, $0.2 \leq r < 0.3$ as a medium effect size, and $r \geq 0.3$ as a large effect size[3]. Finally, we do not make threshold-based claims of statistical significance [12] and do not perform corrections for multiple testing to avoid arriving at overly stringent interpretation of study outcomes [10, 40].

4 ANALYSES AND RESULTS

We divide the analyses into two parts when answering our RQ, *What is the relationship between course achievement, achievement goal orientations, self-efficacy, gender and prior experience?*. We first conduct correlational analyses of the data, which is followed by partial correlation analyses. When reporting results, we report p values using three significant numbers and correlations with two significant numbers. Non-rounded test results, including 95% confidence interval values for correlations, are available in an online appendix[4].

4.1 Correlational Analysis

Table 3 outlines the results of the correlational analysis. To summarize, out of the 28 pairs, 17 pairs had a p-value smaller than 0.001, and 23 pairs had a correlation of 0.1 or higher, indicating at least a small effect size. Correlations between achievement and the other variables, with the exception of gender, had at least a medium effect size and statistical significance of $p < 0.001$. The same holds true also for self-efficacy and mastery approach goal orientation.

For self-efficacy, strong effect sizes are observed between the two prior programming variables, usage and know-how ($r = .44$ for both), as well as for mastery approach ($r = .42$). When further considering prior programming experience, which was divided into two sum variables (usage and know-how), we observe a strong effect size between usage and know-how ($r = .75$), which is to be expected as they are likely strongly related. We also observe a strong effect between mastery approach and usage and know-how ($r = .30$ and $r = .31$, respectively), and a strong effect size between usage and achievement ($r = .33$).

[3]We acknowledge that these thresholds are arbitrary and other interpretations exist [14]; the interpretations are intended to facilitate the discussion and we encourage the reader to focus on the coefficients and their differences between the correlation and partial correlation analyses.

[4]https://osf.io/kjgfx/?view_only=e9594d71938d4910a154005d3d81b991

Table 3: Spearman's r correlations between self-efficacy, prior programming usage, prior programming know-how, the three achievement goal orientation factors (mastery avoidance, mastery approach, and performance orientation), course achievement, and gender.

	Usage	Know-how	M. Avoidance	M. Approach	Perf. Or.	Achievement	Gender
Self-efficacy	r=.44, p<.001	r=.44, p<.001	r=.22, p<.001	r=.42, p<.001	r=.24, p<.001	r=.24, p<.001	r=-.14, p=.030
Usage	-	r=.75, p<.001	r=.15, p=.018	r=.30, p<.001	r=.11, p=.089	r=.33, p<.001	r=-.05, p=.432
Know-how		-	r=.15, p=.017	r=.31, p<.001	r=.06, p=.313	r=.25, p<.001	r=-.10, p=.099
M. Avoidance			-	r=.51, p<.001	r=.40, p<.001	r=.21, p<.001	r=-.01, p=.890
M. Approach				-	r=.29, p<.001	r=.34, p<.001	r=.03, p=.613
Perf. Or.					-	r=.25, p<.001	r=-.02, p=.743
Achievement						-	r=-.10, p=.125

Table 4: Partial correlations (Spearman's r) between self-efficacy, prior programming usage, prior programming know-how, the 3 achievement goal factors (mastery avoidance, mastery approach, and performance orientation), achievement, and gender.

	Usage	Know-how	M. Avoidance	M. Approach	Perf. Or.	Achievement	Gender
Self-efficacy	r=.15, p=.018	r=.15, p=.021	r=-.03, p=.626	r=.27, p<.001	r=.14, p=.024	r=-.01, p=.927	r=-.14, p=.032
Usage	-	r=.69, p<.001	r=-.01, p=.918	r=-.02, p=.763	r=.01, p=.881	r=.21, p<.001	r=.08, p=.211
Know-how		-	r=.01, p=.904	r=.11, p=.093	r=-.08, p=.205	r=-.05, p=.455	r=-.09, p=.140
M. Avoidance			-	r=.42, p<.001	r=.30, p<.001	r=.00, p=.990	r=-.02, p=.700
M. Approach				-	r=.03, p=.636	r=.20, p=.002	r=.13, p=.048
Perf. Or.					-	r=.16, p=.012	r=.01, p=.913
Achievement						-	r=-.10, p=.107

For the achievement goal orientations, we observe that mastery avoidance and mastery approach as well as mastery avoidance and performance orientation have strong effect sizes ($r = .51$ and $r = .40$ respectively). Further, mastery approach has a strong effect size with achievement ($r = .34$), whereas the effect size for achievement and mastery avoidance and performance orientation is medium ($r = .21$ and $r = .25$ respectively).

4.2 Partial Correlations

Table 4 outlines the results of the partial correlation analysis whereby we measured the association between each variable pair, controlling for the effect of the remaining variables. To summarize, in the partial correlation analysis, out of the 28 pairs, 5 pairs had a p-value smaller than 0.001, and 14 pairs had a correlation of 0.1 or higher, indicating at least a small effect size.

Strong effect sizes ($r \geq 0.3$) are observed between prior programming usage and know-how ($r = .69$), mastery avoidance and mastery approach ($r = .42$), and mastery avoidance and performance orientation ($r = .30$). Medium effect sizes are observed between self-efficacy and mastery approach ($r = .27$), prior programming usage and achievement ($r = .21$), and mastery approach and achievement ($r = .20$).

5 DISCUSSION

5.1 Achievement Goal Orientations

When starting the data analysis and conducting EFA to assess the validity of AGQ-R, we observed the best fit with a three-factor model (Mastery approach, Mastery Avoidance and Performance Goal Orientation) instead of the expected four-factor model discussed in conjunction with AGQ-R [19]. Defining the number of achievement goal orientations is not a particularly straightforward question in research, and in fact having approach and avoidance orientations at the same time has been rather scantly examined [34]. While mastery avoidance rose as a factor in our data, the factor has received less attention than the other factors in the two-by-two achievement goal orientation framework. One could even argue that it has been regarded as the least accepted one ([34] cit., e.g., [28]). Overall, the two-by-two framework has also been viewed to be ambiguous and complicating the picture [33]; our results seem to corroborate that the two-by-two framework may not represent the reality of students. More broadly, considering our observations in the light of other computing education researchers who have observed that fitting scales to existing models might not be as straightforward as one would think (see e.g. [53]), our results highlight the need to assess the use of scales in context before applying them, and perhaps even – as aptly discussed in [45] – question the direct adaptation of theories from other fields.

5.2 Correlations and Partial Correlations

Our correlation and partial correlation analyses represent two narratives, both of which have been present in computing education research. The naive correlation analysis outlined in Table 3 corroborates the approach present in the field whereby individual aspects have been studied to identify correlating variables. Such studies have highlighted similar observations to ours, including correlations between self-efficacy and course outcomes [7, 44, 50, 51], prior programming experience and course outcomes [23, 51], and achievement goal orientations and course outcomes [55]. We also observed a small negative effect size between gender and self-efficacy, which could be interpreted as evidence of the confidence gap [8, 31] existing also in courses after CS1.

When looking into partial correlations, outlined in Table 4, the narrative changes. As partial correlation analysis conducted between two variables takes the effect of the other variables into account, it effectively removes the underlying effect of the other variables that may bias the results of the simple correlational analyses. We saw, e.g. that when accounting for other variables, the classic view of self-efficacy contributing to course achievement no longer holds. Instead, we observe the strongest effects on achievement from prior programming usage (r = .21), mastery approach (r = .20), and performance orientation (r = .16), effectively highlighting that some prior knowledge of the topic helps, an approach where one seeks to deeply understand the topic helps, and an approach where one seeks to perform well helps. As the teaching approach can be used to influence students' learning approaches [36], and as courses can be adjusted to match the skill level of students, these results can be seen as somewhat soothing – at least when compared to the view in which success stems from self-efficacy that could even be (mis)interpreted as "students will do well when they really start believing that they can do well". In the broader context, our results from the partial correlation analysis thus highlight that any construct (including self-efficacy) should not be looked at in isolation, rather one should seek to explain and understand the findings in context.

5.3 Limitations of Work

Our study comes with multiple limitations. First, we acknowledge that the data is collected from a single course at a single university, meaning that we do not know to what extent the results would generalize. In discussing the results in the broader context of the computing education literature, we have drawn from literature where the context is an introductory programming class, simply because this is where most of the evidence stems from. We also acknowledge that we do not know the effect of the web software course on the observations; it is conceivable that similar observations might have been made with data from another course at Aalto University.

Second, self-efficacy and our sum variables for prior programming experience were constructed ad-hoc, which could influence the credibility of the findings. In subsequent studies, self-efficacy, similarly to other psychological concepts, might be better studied with a scale rather than a single item. This could entail either drawing from generic self-efficacy scales or adjusting CS-specific self-efficacy scales such as [44], or following [22], in a more elaborate setting to gain a more comprehensive picture of self-efficacy.

Third, we acknowledge that we intentionally did not conduct corrections for multiple testing, as discussed in our methodology. If we had used the strict Bonferroni correction and a starting threshold of $p < 0.05$ for statistical significance, based on having a total of $n = 56$ tests, the threshold would have been marginally smaller than $p < 0.001$. Even in this case, the majority of our highlighted observations in fact hold.

Finally, we acknowledge that the course under study had just undergone a major revision, which might have influenced the observed results. As an example, prior research on approaches to teaching programming [48] has highlighted that change in general tends to yield better learning outcomes, at least in reported research. At the same time, it is a good question whether there were flaws in the previous version of the course, which could have influenced some students negatively.

6 CONCLUSION

The research question for this work was as follows: *What is the relationship between course achievement, achievement goal orientations, self-efficacy, gender and prior experience?*

We set out to answer the question by collecting data in a web software development course at Aalto University. Our analysis started with a validation of the achievement goal orientation scale (AGQ-R [19]) whereby we observed that a three-dimensional model fit our data the best instead of the four-dimensional one outlined in the article discussing the scale. This observation highlighted that it is as yet unclear which of the models in the achievement goal theories best describes the reality (as discussed e.g. in [34]).

Using the three-dimensional model and data gathered from the web software development course, we conducted both normal Spearman correlations and partial correlations to first paint a classic view of how one might interpret the results, and then to examine our findings in more detail. While the Spearman correlation analysis was in line with much of the prior research that has looked into various aspects and course outcomes, including studies that have underscored the link between self-efficacy and course outcomes, the partial correlation analysis painted a different picture. Our results highlighted that the strongest contributors to course achievement were prior programming experience, leaning towards a mastery approach and a performance orientation in studying. These results emphasize the need to study any underlying aspects when seeking to understand what contributes to students' behavior and performance.

Future work should re-explore our findings in other contexts, including in introductory programming. There is arguably a general need in CER for comprehensive replication studies regarding the topics of our study. We also call for studies that use repeated measurements (similar to e.g. [29, 30, 38]) both in introductory programming courses and in subsequent ones. Such studies could provide more insight into fluctuations in measurements, which – when combined with information from the respective course contexts – in turn could help us better understand the effects of the study contexts.

REFERENCES

[1] Alireza Ahadi, Arto Hellas, Petri Ihantola, Ari Korhonen, and Andrew Petersen. 2016. Replication in computing education research: researcher attitudes and experiences. In *Proc. of the 16th Koli calling int. conf. on computing education research*. 2–11.
[2] Marta Andreatta and Paul Pauli. 2015. Appetitive vs. aversive conditioning in humans. *Frontiers in behavioral neuroscience* 9 (2015), 128.
[3] Albert Bandura. 1977. Self-efficacy: toward a unifying theory of behavioral change. *Psychological review* 84, 2 (1977), 191.
[4] Albert Bandura. 1982. Self-efficacy mechanism in human agency. *American psychologist* 37, 2 (1982), 122.

[5] Maurice Stevenson Bartlett. 1937. Properties of sufficiency and statistical tests. *Proc. of the Royal Society of London. Series A-Mathematical and Physical Sciences* 160, 901 (1937), 268–282.

[6] Susan Bergin and Ronan Reilly. 2005. Programming: factors that influence success. In *Proc. of the 36th SIGCSE technical symp. on Computer science education.* 411–415.

[7] Susan Bergin, Ronan Reilly, and Desmond Traynor. 2005. Examining the role of self-regulated learning on introductory programming performance. In *Proc. of the first int. workshop on Computing education research.* 81–86.

[8] Sylvia Beyer, Kristina Rynes, Julie Perrault, Kelly Hay, and Susan Haller. 2003. Gender differences in computer science students. In *Proc. of the 34th SIGCSE technical symposium on Computer science education.* 49–53.

[9] Ruth Butler. 1993. Effects of task-and ego-achievement goals on information seeking during task engagement. *J. of pers. and social psych.* 65, 1 (1993), 18.

[10] Marc Buyse, SA Hurvitz, F Andre, Z Jiang, HA Burris, M Toi, W Eiermann, M-A Lindsay, and D Slamon. 2016. Statistical controversies in clinical research: statistical significance – too much of a good thing... *Annals of Oncology* 27, 5 (2016), 760–762.

[11] Pat Byrne and Gerry Lyons. 2001. The effect of student attributes on success in programming. *Acm sigcse bulletin* 33, 3 (2001), 49–52.

[12] Ronald P Carver. 1993. The case against statistical significance testing, revisited. *The J. of Experimental Education* 61, 4 (1993), 287–292.

[13] Raymond B Cattell. 1966. The scree test for the number of factors. *Multivariate behavioral research* 1, 2 (1966), 245–276.

[14] Jacob Cohen. 1992. Statistical power analysis. *Current directions in psychological science* 1, 3 (1992), 98–101.

[15] Statistics Corner. 2009. Choosing the right type of rotation in PCA and EFA. *JALT testing & evaluation SIG newsletter* 13, 3 (2009), 20–25.

[16] Rodrigo Duran, Lassi Haaranen, and Arto Hellas. 2020. Gender differences in introductory programming: Comparing MOOCs and local courses. In *Proc. of the 51st ACM Technical Symposium on Computer Science Education.* 692–698.

[17] Andrew J Elliot. 1999. Approach and avoidance motivation and achievement goals. *Educational psychologist* 34, 3 (1999), 169–189.

[18] Andrew J Elliot and Holly A McGregor. 2001. A 2× 2 achievement goal framework. *J. of personality and social psychology* 80, 3 (2001), 501.

[19] Andrew J Elliot and Kou Murayama. 2008. On the measurement of achievement goals: Critique, illustration, and application. *J. of educational psychology* 100, 3 (2008), 613.

[20] Leandre R Fabrigar and Duane T Wegener. 2011. *Exploratory factor analysis.* Oxford University Press.

[21] Gilles E Gignac and Eva T Szodorai. 2016. Effect size guidelines for individual differences researchers. *Personality and individual differences* 102 (2016), 74–78.

[22] Jamie Gorson and Eleanor O'Rourke. 2020. Why do cs1 students think they're bad at programming? investigating self-efficacy and self-assessments at three universities. In *Proc. of the 2020 ACM Conf. on Int. Computing Education Research.* 170–181.

[23] Dianne Hagan and Selby Markham. 2000. Does It Help to Have Some Programming Experience before Beginning a Computing Degree Program? *SIGCSE Bull.* 32, 3 (Jul 2000), 25–28.

[24] Judith M Harackiewicz, Kenneth E Barron, Suzanne M Carter, Alan T Lehto, and Andrew J Elliot. 1997. Predictors and consequences of achievement goals in the college classroom: Maintaining interest and making the grade. *J. of Personality and Social psychology* 73, 6 (1997), 1284.

[25] Petri Ihantola, Arto Vihavainen, Alireza Ahadi, Matthew Butler, Jürgen Börstler, Stephen H Edwards, Essi Isohanni, Ari Korhonen, Andrew Petersen, Kelly Rivers, et al. 2015. Educational data mining and learning analytics in programming: Literature review and case studies. *Proc. of the 2015 ITiCSE on Working Group Reports* (2015), 41–63.

[26] Henry F Kaiser and John Rice. 1974. Little jiffy, mark IV. *Educational and psychological measurement* 34, 1 (1974), 111–117.

[27] Päivi Kinnunen and Beth Simon. 2012. My program is ok–am I? Computing freshmen's experiences of doing programming assignments. *Computer Science Education* 22, 1 (2012), 1–28.

[28] Lisa Linnenbrink-Garcia, Diana F Tyson, and Erika A Patall. 2008. When are achievement goal orientations beneficial for academic achievement? A closer look at main effects and moderating factors. *Revue internationale de psychologie sociale* 21, 1 (2008), 19–70.

[29] Alex Lishinski and Joshua Rosenberg. 2021. All the Pieces Matter: The Relationship of Momentary Self-efficacy and Affective Experiences with CS1 Achievement and Interest in Computing. In *Proc. of the 17th ACM Conf. on Int. Computing Education Research.* 252–265.

[30] Alex Lishinski, Aman Yadav, Jon Good, and Richard Enbody. 2016. Learning to program: Gender differences and interactive effects of students' motivation, goals, and self-efficacy on performance. In *Proc. of the 2016 ACM Conf. on Int. Computing Education Research.*

[31] Mary A Lundeberg, Paul W Fox, Amy C Brown, and Salman Elbedour. 2000. Cultural influences on confidence: Country and gender. *J. of Educational Psychology* 92, 1 (2000), 152.

[32] Andrew Luxton-Reilly, Ibrahim Albluwi, Brett A Becker, Michail Giannakos, Amruth N Kumar, Linda Ott, James Paterson, Michael James Scott, Judy Sheard, and Claudia Szabo. 2018. Introductory programming: a systematic literature review. In *Proc. Companion of the 23rd Annual ACM Conf. on Innovation and Technology in Computer Science Education.* 55–106.

[33] Martin L Maehr. 2001. Goal theory is not dead—Not yet, anyway: A reflection on the special issue. *Educational Psychology Review* 13, 2 (2001), 177–185.

[34] Martin L Maehr and Akane Zusho. 2009. Achievement Goal Theory: The Past, Present, and Future. In *Handbook of motivation at school.* Routledge, 91–118.

[35] Laurie Murphy, Brad Richards, Renée McCauley, Briana B Morrison, Suzanne Westbrook, and Timothy Fossum. 2006. Women catch up: gender differences in learning programming concepts. In *ACM SIGCSE Bulletin*, Vol. 38. ACM, 17–21.

[36] Claudia Ott, Brendan McCane, and Nick Meek. 2021. Mastery Learning in CS1-An Invitation to Procrastinate?: Reflecting on Six Years of Mastery Learning. In *Proc. of the 26th ACM Conf. on Innovation and Technology in Computer Science Education V. 1.* 18–24.

[37] Hannu Pesonen and Arto Hellas. 2022. On Things that Matter in Learning Programming: Towards a Scale for New Programming Students. In *2022 IEEE Frontiers in Education Conf. (FIE).* IEEE, 1–9.

[38] Hannu Pesonen and Arto Hellas. 2022. Weekly Fluctuations in Motivation in Introductory Programming. In *Proc. of the 2022 Conf. on United Kingdom & Ireland Computing Education Research.* 1–7.

[39] Nea Pirttinen, Arto Hellas, Lassi Haaranen, and Rodrigo Duran. 2020. Study Major, Gender, and Confidence Gap: Effects on Experience, Performance, and Self-Efficacy in Introductory Programming. In *2020 IEEE Frontiers in Education Conf. (FIE).* IEEE, 1–7.

[40] Priya Ranganathan, CS Pramesh, and Marc Buyse. 2016. Common pitfalls in statistical analysis: the perils of multiple testing. *Perspectives in clinical research* 7, 2 (2016), 106.

[41] Allison M Ryan, Margaret H Gheen, and Carol Midgley. 1998. Why do some students avoid asking for help? An examination of the interplay among students' academic efficacy, teachers' social–emotional role, and the classroom goal structure. *J. of educational psychology* 90, 3 (1998), 528.

[42] Richard M Ryan and Edward L Deci. 2009. Promoting Self-Determined School Engagement. In *Handbook of Motivation at School.* Routledge, Chapter 9.

[43] Sami Sarsa, Arto Hellas, and Juho Leinonen. 2022. Who Continues in a Series of Lifelong Learning Courses?. In *Proc. of the 27th ACM Conf. on on Innovation and Technology in Computer Science Education Vol. 1.* 47–53.

[44] Phil Steinhorst, Andrew Petersen, and Jan Vahrenhold. 2020. Revisiting self-efficacy in introductory programming. In *Proc. of the 2020 ACM Conf. on Int. Computing Education Research.* 158–169.

[45] Matti Tedre and John Pajunen. 2022. Grand theories or design guidelines? Perspectives on the role of theory in computing education research. *ACM Transactions on Computing Education (TOCE)* (2022).

[46] Timothy Urdan. 1997. Achievement goal theory: Past results, future directions. *Advances in motivation and achievement* 10 (1997), 99–141.

[47] Philip R Ventura Jr. 2005. Identifying predictors of success for an objects-first CS1. *Computer Science Education* 15, 3 (2005), 223–243.

[48] Arto Vihavainen, Jonne Airaksinen, and Christopher Watson. 2014. A systematic review of approaches for teaching introductory programming and their influence on success. In *Proc. of the tenth annual conf. on Int. computing education research.*

[49] Ronald L Wasserstein and Nicole A Lazar. 2016. The ASA statement on p-values: context, process, and purpose.

[50] Christopher Watson, Frederick WB Li, and Jamie L Godwin. 2014. No tests required: comparing traditional and dynamic predictors of programming success. In *Proc. of the 45th ACM tech. symp. on Computer science education.* 469–474.

[51] Susan Wiedenbeck, Deborah Labelle, and Vennila NR Kain. 2004. Factors affecting course outcomes in introductory programming.. In *PPIG.* 11.

[52] Brenda Cantwell Wilson and Sharon Shrock. 2001. Contributing to success in an introductory computer science course: a study of twelve factors. *Acm sigcse bulletin* 33, 1 (2001), 184–188.

[53] Albina Zavgorodniaia, Rodrigo Duran, Arto Hellas, Otto Seppala, and Juha Sorva. 2020. Measuring the cognitive load of learning to program: A replication study. In *United Kingdom & Ireland Computing Education Research conf.* 3–9.

[54] Daniel Zingaro. 2014. Peer instruction contributes to self-efficacy in CS1. In *Proc. of the 45th ACM tech. symp. on Computer Science Education.* 373–378.

[55] Daniel Zingaro, Michelle Craig, Leo Porter, Brett A Becker, Yingjun Cao, Phill Conrad, Diana Cukierman, Arto Hellas, Dastyni Loksa, and Neena Thota. 2018. Achievement goals in CS1: Replication and extension. In *Proc. of the 49th acm technical symposium on computer science education.* 687–692.