JYX JYVÄSKYLÄN YLIOPISTO UNIVERSITY OF JYVÄSKYLÄ

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Kaijanaho, Antti-Juhani; Tirronen, Ville

Title: Fixed versus Growth Mindset Does not Seem to Matter Much : A Prospective Observational Study in Two Late Bachelor level Computer Science Courses

Year: 2018

Version: Published version

Copyright: © the Authors, 2018.

Rights: In Copyright

Rights url: http://rightsstatements.org/page/InC/1.0/?language=en

Please cite the original version:

Kaijanaho, A.-J., & Tirronen, V. (2018). Fixed versus Growth Mindset Does not Seem to Matter Much : A Prospective Observational Study in Two Late Bachelor level Computer Science Courses. In L. Malmi, A. Korhonen, R. McCartney, & A. Petersen (Eds.), ICER '18 : Proceedings of the 2018 ACM Conference on International Computing Education Research (pp. 11-20). ACM. https://doi.org/10.1145/3230977.3230982

Fixed versus Growth Mindset Does not Seem to Matter Much

A Prospective Observational Study in Two Late Bachelor level Computer Science Courses

Antti-Juhani Kaijanaho
antti-juhani.kaijanaho@jyu.fi
University of Jyväskylä
Faculty of Information Technology
Jyväskylä, Finland

Ville Tirronen
ville.e.t.tirronen@jyu.fi
University of Jyväskylä
Faculty of Information Technology
Jyväskylä, Finland

ABSTRACT

Psychology predicts that a student's mindset—their implicit theory of intelligence—has an effect on their academic performance. We attempted to corroborate this in the computer science education context by asking the students on two bachelor-level courses, typically taken in the third year of studies, to fill out a standard mindset questionnaire, and analyzing their answers in relation to their grades on those courses. In a sample of 133 students, with only 24 (18 %) students with a clear fixed mindset, there is no detectable correlation between the students' mindsets and their course grades. An ordinal logistic regression estimates, at the 95 % confidence level, a statistically nonsignificant effect between a decrease by a factor of 0.46 and an increase by a factor of 2.03 in the odds of achieving a better course grade when moving from a strong fixed mindset to neutral mindset, or when moving from a moderate fixed mindset to a moderate growth mindset. This suggests that any effect the mindset has on the outcomes of these courses is small. We conclude that educational interventions targeting students' mindsets may not be worth the effort in late bachelor-level CS education, possibly because students who suffer from their fixed mindset have already dropped out by the third year.

CCS CONCEPTS

• **Social and professional topics → Computer science education**; **Adult education**;

KEYWORDS

mindsets, growth mindset, fixed mindset, implicit theories, entity theory, incremental theory, psychology

ACM Reference Format:
Antti-Juhani Kaijanaho and Ville Tirronen. 2018. Fixed versus Growth Mindset Does not Seem to Matter Much: A Prospective Observational Study in Two Late Bachelor level Computer Science Courses. In *ICER '18: 2018 International Computing Education Research Conference, August 13-15, 2018, Espoo, Finland*. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3230977.3230982

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
ICER'18, August 13–15, 2018, Espoo, Finland
© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.

https://doi.org/10.1145/3230977.3230982

1 INTRODUCTION

We, like many other teachers, are intuitively inclined to adopt the stance that *mindsets matter* and act accordingly in our teaching [2, 37, 40]. Here, *mindset* refers to Carol Dweck and colleagues' psychological theory [10, 11, 13] according to which one's *implicit theory* of one's own intelligence—either an *entity theory* (a *fixed mindset*) or an *incremental theory* (a *growth mindset*)—influences how one reacts to difficulties, leading to different levels of academic achievement. However, is it worth our time and effort to try to change our students' mindset? In this paper we investigate, in an observational setting, the size of any effect mindsets have on student achievement in late bachelor-level courses, and find, to our surprise, that this effect is too small for us to measure. We also find a relatively small number of students with a fixed mindset in our courses.

The context of our study is two core bachelor-level computer science courses at the University of Jyväskylä, Finland, targeting second and third year students but usually taken in the third year of studies: one on the theory of computing and one on functional programming. Both courses are essential for the field, but many students struggle with them. The subject matter in these courses is challenging and seems to require deep understanding only obtainable through long hours of practice causing students to face both *motivational* and *affective* difficulties (cf. Kinnunen and Malmi [20]). This highlighted to us the potential importance of the correct mindset as enabler of persistence.

The effect of mindset on student achievement has been experimentally demonstrated, but there are relatively few studies on the magnitude of such an effect, especially in the context of computer science education. We thus set out to investigate the size of the effect mindsets have on outcomes in our own courses. Knowing the effect size estimate is important in practice: the larger the effect, the more important it is for us to try to influence student mindset, while a small effect would make such interventions mostly irrelevant. Our research question then became:

RQ Assuming a causal connection between a student's implicit self-theory of intelligence and their academic achievement, how large is its contribution in the context of late bachelor-level computer science courses?

Notice that we make a causal assumption here; it is, we think, well established by prior research.

Table 1: Features associated with fixed and growth mindset (summarizing Dweck and Leggett [13]).

Category	Fixed mindset	Growth mindset
Implicit theory of intelligence	*Entity:* intelligence is something one was born with and cannot be changed	*Incremental:* intelligence can be developed by conscious action
Goals	*Performance:* trying to achieve the recognition of one's preexisting abilities and to hide one's preexisting weaknesses	*Learning:* trying to improve one's abilities
Behavior pattern	*Helpless:* difficulties are insurmountable barriers	*Mastery:* difficulties are challenges to be overcome
View of high required effort	A sign of lack of talent	A sign of learning
Task choice with low perceived ability	Seek easy tasks	Seek challenging tasks
Effect of difficulties on …		
…affect	Negative emotions	Neutral or positive emotions
…problem-solving performance	Decrease	Neutral or increase

2 MINDSETS

2.1 Basic theory

The psychological theory of *mindsets* [10, 11, 13] posits that people can be roughly sorted into two categories based on their (current) view on the nature of their own intelligence:

- People holding the *fixed mindset* believe that (their own) intelligence is something they were born with in a certain amount that they cannot affect.
- People holding the *growth mindset* believe that (their own) intelligence is something they can develop by conscious action.

We use here terminology apparently first used in a popular exposition published in 2006 [11]. The seminal article [13], along with much of the older literature, used the general term *implicit theories* together with the specific terms *entity theory* (fixed mindset) and *incremental theory* (growth mindset). Other terms seen in the literature include *self-theories*, *lay theories* and *naive theories*.

The theory of mindsets predicts a number of differences between individuals based on their mindset, as summarized in Table 1. The educational import of the theory is the prediction that a fixed mindset leads one to suboptimal study strategies and to give up studying a topic early, as challenges start to appear, resulting in worse educational outcomes, while a growth mindset encourages working through difficulties, resulting in better educational outcomes.

Lüftenegger and Chen [25] recommend using the *implicit theories* terminology in academic publications. However, as mindsets are more familiar for the nonpsychologist audience, we continue to use them. In our usage, mindsets are (mutable) categories of people, and the implicit theories of intelligence are psychological constructs measured by specific instruments, which— according to the theory— determine mindsets.

Intelligence is not the only thing that people have implicit theories of. Dweck [10] mentions studies on implicit theories of other people's personality and morality as well as of the malleability of the world. Researchers have also studied implicit theories of programming aptitude [35, 36, 37]. However, our interest is in the original theory of mindsets, involving the claim that the implicit theory of intelligence predicts particular behaviors and thus influences outcomes; for our purposes, then, these alternative constructs are beside the point.

2.2 Empirical evidence

The theory of mindsets grew out of experimental phenomena. Carol Dweck and colleagues showed in psychological experiments in the 1970s that children's ability to persist in the face of failures depends on whether they take responsibility for their own success and failure, and that training children to take such responsibility improves persistence [9, 14]. Subsequent experiments uncovered much of the associated features summarized in Table 1, and eventually it was noticed that the theory of intelligence that a person holds (often without realizing it) is a reliable predictor of these two categories. Dweck and Leggett [13] summarize these developments and the associated experiments up to the late 1980s, by which time the theory was largely complete and experimentally corroborated in many contexts. Dweck further published at the turn of the century an academic monograph [10] on the theory, and later another book for the popular audience [11].

Since the seminal work by Dweck, the theory has been examined in field conditions and in specific contexts such as mathematics or in transitions between educational levels. This research has discovered new associations, such as the association between self-esteem and growth mindset [33] in young adults.

Although field research has provided some evidence supporting the theory in real world context [4], there are a number of studies that cast doubt on using mindsets as predictors of academic success. For example, Macnamara and Rupani [27] found no evidence that growth mindset would predict higher academic achievement in their study of first year psychology students. Instead, the authors observed the opposite, though non-significant, association. Likewise, Clevenger [7] finds in her thesis no apparent association between mindset and academic achievement in K12 students and their parents. However, growth mindset did predict performance goals in the way described by Dweck [10].

Similar results were also obtained in context of mathematics by Priess-Groben and Hyde [30], who studied mindsets during transition from high school to college. They found that, although growth mindset was a significantly associated with achievement, controlling factors such as prior success effectively eliminated this association. They suspect that mindset is just one factor of many affecting success in this field. Relatedly, Zonnefeld [45] studied mindsets connected to learning university level statistics. Here also, the result hints that student mindset does not affect the measured learning outcome.

Regardless of several negative studies, a meta-analysis of experimental and observational studies by Burnette et al. [1] shows a small positive association of a growth mindset with goal achievement.

In the context of computer science education, mindsets have been studied fairly little. Experiments involving mindset interventions on first year university students [8, 40] show mixed results in changes of mindset and no measurable effect in course outcomes. Observational studies, mainly in the CS1 context, have found either weak or unmeasurable effect of mindsets on course outcomes [15, 24, 38, 41].

All of the above mentioned studies have been conducted in a WEIRD (Western, Educated, Industrialized, Rich, and Democratic) context [17], mostly in the United States and mostly with college and university level students. Although the original theory does not posit a cultural dependency on the effect of mindset, it is possible that one exists. For example Chen and Wong [5] has studied the mindset theory within Chinese culture, finding that their results were consistent some of the time with some differences the authors identify as cultural.

2.3 Measurement

The standard instruments for measuring implicit theories (mindsets) are given in an appendix to Dweck's monograph [10]. We focus on her self-theories of intelligence form for adults (on p. 178), as it is the most relevant for this study. It consists of eight items scored from 1 (strongly agree) to 6 (strongly disagree) with no neutral option. Four items are assertions consistent with an entity theory of intelligence (fixed mindset), e. g.,

> *"1. You have a certain amount of intelligence, and you can't really do much to change it."*

and the rest are assertions consistent with an incremental theory (growth mindset), e. g.,

> *"5. You can always substantially change how intelligent you are."*

Dweck [10] advises (p. 176) that the entity theory items can be used alone; in such a case, a rejection of the entity theory would be taken as the acceptance of the incremental theory.

The entity theory and the incremental theory item responses can be straightforwardly summed or averaged separately. This creates two scores, the entity theory score and the incremental theory score. Most empirical studies tend to either use the entity theory score alone or a combined score where the incremental theory items have been reverse scored [1, 25], considering in both cases a low score as indicating an entity theory. This seems logical, as the entity theory and the incremental theory appear to be logically inconsistent, and it seems difficult to see how one can endorse both of the example items quoted above. However, Chen and Tutwiler [4], Lüftenegger and Chen [25], and Tempelaar et al. [42] argue that there is not sufficient correlation between the two scores to justify combining them; in effect, they say, the entity theory and the incremental theory appear to be separate constructs that are highly related instead of two faces of the same coin.

Older studies seem to simplify their analyses by dichotomizing the scores. For example, Dweck et al. [12, p. 269] created two groups by categorizing all participants with a score of at most 3 as entity theorists and all participants with a score of at least 4 as incremental theorists. The remaining participants—with a score between 3 and 4—were then excluded from analysis. They reported that only 15 % of their participants were thus excluded. However, dichotomization of scores is generally disfavored by methodologists (see, e. g., MacCallum et al. [26] and Rucker et al. [34]).

Dweck and Leggett [13, fn. 5 on p. 263] reported that she and her colleagues have obtained bimodal distributions for scores measuring implicit theories of intelligence. In other works [12, 22], Dweck and colleagues have reported various distributions of dichotomized implicit theory variables, with both incremental and entity theories having substantial (but not always equal) support, and a fairly low exclusion rate; this indirectly suggests a fairly bimodal distribution of the underlying scores. More recently, Tempelaar et al. [42] found a roughly normal distribution of implicit theory scores, with 64 % of their sample lying within a one deviation around the mean; however, both theories were approximately equally endorsed. Most reports of studies that we are aware of do not give sufficient detail on their score distributions for similar analyses. There thus is some doubt on whether there actually is a bimodal underlying distribution that would justify classifying people into two groups in this manner; however, both theories seem to be generally endorsed in the population.

3 METHOD

3.1 Participants

We recruited participants in two bachelor-level computer science courses typically taken in the third year—one in functional programming (TIEA341 Functional Programming 1, henceforth FP1) and the other in the theory of computing (TIEA241 Automatons and Formal Languages, henceforth TCS)—taught in Fall 2017 at the Faculty of Information Technology of the University of Jyväskylä, Finland. In one course (FP1), completion of the informed consent form form (either giving or refusing consent) was presented as required though not enforced, and in the other course (TCS), completion of the form counted as one of 59 exercises that together could contribute to the final grade up to two grade points out of five.

As to demographic data, we obtained the age, the number of credits attained, and the number of years enrolled in our department for each student enrolled in our courses as part of course completion data. They thus represent a snapshot after the courses had been completed. We had no ready access to students' sex or gender: legal sex would have only been available by derivation from each students' national personal identity code, but we had no legal or moral basis to access this information, which is legally considered sensitive data; and neither the course completion data nor our questionnaire data included self-reported gender.

Our sampling stopping rule [39] was simple: we recruited only from the aforementioned two courses, and we included all consenting students who answered all questions in our questionnaire. Thus, there was no need to set any arbitrary sample size, and no need for pre-data power calculations.

3.2 Measurements

The implicit theories of the students were measured using the self-theories of intelligence scale as described by Dweck [10]. In addition to the original English, a Finnish version translated by one of us and checked by the other was used. We scored each item on the implicit theory form on an ordinal scale of 1–6 with *strongly agree* as 1 or, when reverse scoring, with *strongly disagree* as 1.

We created three scores, which we treat as interval data, for each student: the *entity theory score* averages the entity theory items; the *incremental theory score* averages the reverse-scored incremental theory items; and the *combined theories score* averages the entity theory items and the reverse-scored incremental theory items. All scores were then rescaled to range between -1 and 1, so that a one-unit change corresponds to a move from a strongly held theory to ambivalence (or vice versa), or from one weakly held theory to another. A high score indicates an incremental theory and a low score indicates an entity theory.

The FP1 course awarded passing students a whole number of 1–5 study credits with a vacuous grade ("pass"), and TCS awarded passing students exactly 5 credits and a grade on the ordinal scale 1–5. The FP1 course was assessed based on homework exercises successfully completed, with a mandatory brief oral interview at the end to discourage fraudulent submissions. There was no final exam in FP1. The general deadline for submitting work for this course was January 15, 2018.

The TCS course could be completed by taking a written final exam, or by completing three course topics separately. Each of the three topics could be completed separately by taking a written partial exam soon after the relevant material was covered in lectures; alternatively, two of the topics could be completed separately by a programming project and the third topic by submitting satisfactory answers to a specific set of homework exercises. In all cases, if a student attempted more than one way of completion, the best results prevailed. For all passing students in TCS, the successful completion of homework exercises counted toward the grade, up to two grade points. The general deadline for the programming project and homework exercises was January 8, 2018; and the first opportunity to take the final exam was January 19, 2018.

3.3 Procedure

The informed consent form and the implicit theories form were administered together as an online questionnaire that was open for the students throughout the course period (from October 23 to December 15, 2017). The questionnaire was offered in Finnish or English based on the student's language preference.

We took a snapshot of course completion data for both courses on January 26, 2018. Thus, our completion data is based on the FP1 and TCS credits and grades as they stood in the official registry on that date. Any grades and credits later awarded are ignored by our analysis.

Throughout the courses and the grading period (until January 26, 2018), we deliberately kept ourselves ignorant of the students' answers to both the informed consent form and the implicit theories form; we only obtained a list of students who had answered the questionnaire, so that we could mark the corresponding exercises as completed. We acquired the full data on the answers after taking the snapshot of the completion data.

3.4 Analyses

We excluded from analysis students who did not answer the informed consent form or expressedly declined consent. We further excluded from analysis those students who did not answer all items on the implicit theory form.

We assessed the internal reliability of these scores by computing Coefficient Alpha for each. Nunnally and Bernstein [29, p. 265] consider values near $\alpha = .70$ acceptable only for early stages of measure development, while $\alpha \approx 0.80$ is sufficient for group-based basic research, and $\alpha > 0.90$ is necessary only for individual testing affecting decisions of importance.

We examined the resulting scores in a scatter plot of entity theory score versus incremental theory score. We expected to find a visually apparent correlation, and perhaps two visually obvious distinct clusters; we further computed the Pearson product–moment correlation for these two scores. Given the strong correlation we observed (as discussed later) and the difficulties collinearity gives to regression analysis, we elected to pursue our analyses with the combined theories score.

We grouped students into three groups: those who registered only for FP1, those who registered only for TCS, and those who registered for both courses. For the FP1 group, we used as the outcome variable the number of credits awarded, treating noncompletion (for whatever reason) as a zero credit outcome; since credits nominally measure hours worked, we treat this as an interval variable. For the TCS group, we used as the outcome variable the grade awarded, treating noncompletion (for whatever reason) as a zero grade outcome. Grades are most naturally modeled as ordinal variables, though treating them as interval variables is not uncommon.

The outcome variable for the both-courses group necessarily is a combination of the outcomes of both courses. Both FP1 credits and TCS grades range from 0 to 5 and both can be regarded as measuring educational achievement. Thus, we can combine them into a single interval variable by averaging them. For analyses better done using an ordinal variable, we rounded the result upward so that passing one course is counted as 1.

Since the outcome can be treated as an interval variable, we tried to use linear regression with the course group as well as the combined theory score as the explanatory variables and the outcome variable as the response variable. We expected the course group to be needed because each course was graded using course-specific criteria, potentially introducing statistical dependence among group members; however, we also tried an analysis without the group variable. Model misspecification was tested (see, e. g., Chatterjee and Simonoff [3], p. 15) by using

Table 2: Enrollment in the courses and exclusions from the study. The FP1 and TCS rows in this table include the overlap; thus, the total equals the sum of FP1 and TCS minus the overlap.

	Students	No answer	Refused consent	Missing data	Included participants
FP1	214	94	10	11	99
TCS	139	44	21	6	68
overlap	59	14	8	3	34
Total	294	124	23	14	133

- a scatter plot of residuals versus predictors or fitted values to detect nonlinearity, heteroscedasticity, and outliers; and
- a normal probability plot for residuals to detect nonnormal errors.

An alternative, should the linear regression not be satisfactory, was ordinal logistic regression. Here, the explanatory variables would be the same as in the linear regression, and the response variable would be the ordinal version of the outcome variable. This has the proportional odds assumption; this can be checked by running separate binary outcome logits for each threshold.

We should briefly note that we considered the use of structural equation modeling (SEM) to analyze our data; however, our sample size is too small for a worth-while SEM model of our data, especially considering the ordinal nature of our data.

We generally report confidence intervals (CIs) at the 95 % level instead of statistical significance tests and p values, because we feel that CIs are more informative. However, CIs correspond to significance tests in a very simple manner: a 95 % CI consists of exactly those values for which the hypothesis that the true value equals that value is not rejected at $\alpha = 0.05$; in particular, a null hypothesis is rejected if and only if the null value is outside the CI. Thus, we are licensed to rule out any value that is outside a CI to the same extent that rejection of a hypothesis licenses us to rule out the value specified by the hypothesis. We do not conduct post-data power analyses, as they add nothing useful to confidence intervals [18, 19].

All statistical analyses were conducted using the statistical programming language R [31], with the aid of the packages moments [21], ordinal [6], plyr [43], and psych [32]. All plots were produced using pgfplots in LaTeX.

4 RESULTS

Course enrollment, exclusions due to missing informed consent or incomplete questionnaire data, and the number of students included in this study are summarized in Table 2. In addition, one student was not registered in either course (and thus is not included in this study) yet had granted informed consent; they must have been registered to at least one of the courses earlier.

Basic demographic information available to us is shown in Table 3. Credits are nominally a measure of time spent in studies, with 60 credits nominally equaling one year of study. Thus, 138 credits are consistent with a student on their third full-time study year.

Table 3: Demographic profile of the participants, reported as means and standard deviations. The FP1 and TCS rows in this table include the overlap.

	Age	Credits	Year
Consenting students	25 (4.3)	138 (95.8)	3 (2.6)
Included participants			
— all	25 (4.4)	138 (97.6)	3 (2.6)
— FP1	24 (3.5)	137 (95.1)	3 (2.6)
— TCS	25 (5.1)	136 (96.3)	4 (2.5)

Table 4: Odds ratios (and their 95 % CIs) given by successive logistic regressions with various outcome thresholds, where the combined theory score is the explanatory variable and the outcome is dichotomized at each successive level.

≤ 0	≤ 1	≤ 2	≤ 3	≤ 4
0.79	0.96	1.02	1.39	1.72
[0.34, 1.82]	[0.41, 2.23]	[0.42, 2.52]	[0.49, 4.13]	[0.47, 6.86]

However, credits are attached to courses and do not directly correspond to an individual student's time use.

The Cronbach coefficient alpha is, over the whole sample, for the entity theory score $\alpha = 0.92$, 95 % CI [0.88, 0.94], for the incremental theory score $\alpha = 0.94$, 95 % CI [0.91, 0.96], and for the combined theory score $\alpha = 0.95$, 95 % CI [0.92, 0.97]. The entity and incremental theory scores are plotted against each other in Figure 1. The distribution of the combined scores is shown in Figure 2; it has a mean of 0.16, a standard deviation of 0.41, a skewness of −0.36, and a kurtosis of 3.24. Attempting a Dweck et al. [12, p. 269] dichotomization, using the thresholds −0.2 and 0.2, leads to a highly problematic result: in our sample, there are 24 clear entity theorists and 60 clear incremental theorists, while 49 participants would be excluded as ambivalent.

The distributions of the course outcomes are plotted in Figure 3.

Course outcomes plotted against the combined theory score, see Figure 4, reveal no apparent functional relationship. Thus, it seems doubtful that an association is present in this data. We corroborate this by attempting to demonstrate one using regression analysis.

A linear regression with the group and the combined theory score as explanatory variables and the averaged outcome as the response variables results in diagnostic plots indicating both non-normality of residuals and heteroscedasticity, thus, a misspecification. Respecifying the model without the group as an explanatory variable gives better residuals, but even here, the distribution of residuals is far from normal. The pattern suggests that a logistic regression might be more appropriate.

Successive logistic regressions with the combined theory score as the explanatory variable and the binary response variable of under/over a threshold result with the odds ratios for the theory score given in Table 4; the confidence intervals have a large overlap, and the point estimates all fit inside the intersection of the confidence intervals. Thus, it seems that the assumption of proportional odds is approximately met. Thus, we adopt a proportional-odds model

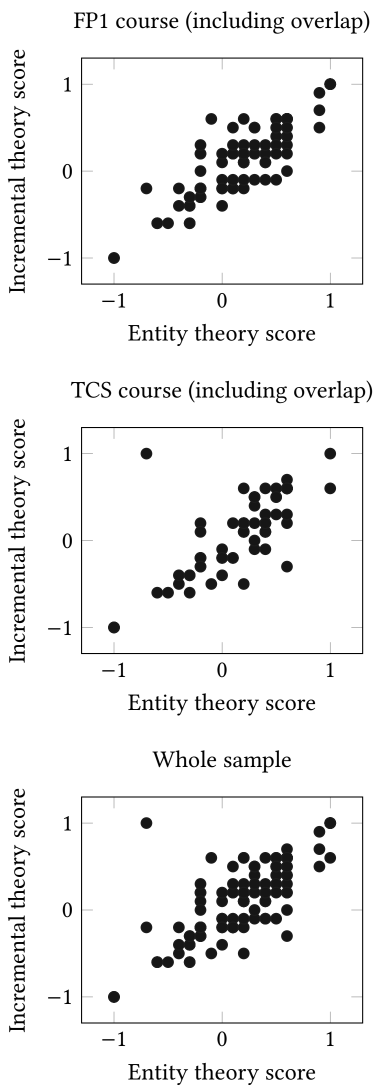

Figure 1: Scatter plot of entity and incremental theory scores. Note that a low entity score indicates an entity theory, and a high incremental theory score indicates an incremental theory; thus, the positive correlation is expected. The Pearson product-moment correlation coefficient is $r = 0.79$ for the whole sample, $r = 0.87$ for the FP1 attendees (including overlap), and $r = 0.76$ for the TCS attendees (including overlap).

(ordinal logistic regression) with the combined theory score as the explanatory variable and the ordinal outcome as the response variable; it results in a non-significant odds ratio of 0.96, 95 % CI [0.46, 2.03], for the theory score. The model coefficients (which are logarithms of either odds or odds ratios, depending on the coefficient) are shown in Table 5.

Figure 5 is an simple effect display inspired by Fox and Andersen [16], visualizing the behavior of the proportional-odds model we chose. It shows, for each ordinal outcome, the 95 % confidence interval (as a gray band) for the predicted probability of a student to

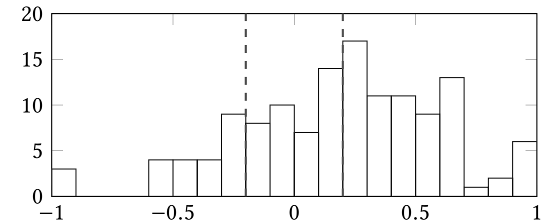

Figure 2: Histogram of combined theory scores across the sample. The dashed vertical lines indicate the cutoff points for Dweck et al. [12, p. 269] dichotomization (see Subsection 2.3).

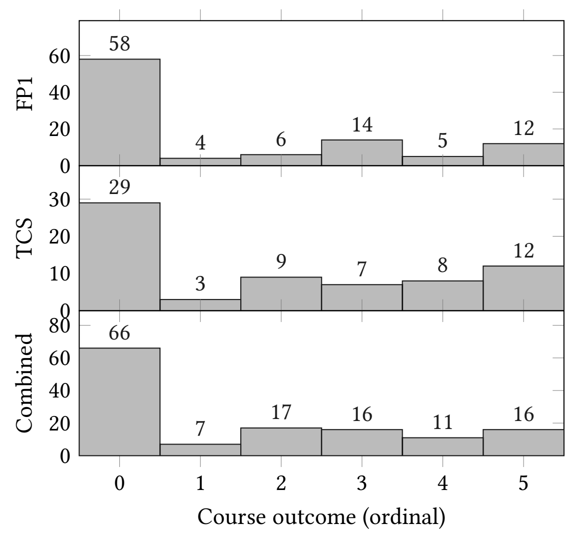

Figure 3: Distribution of course outcomes on FP1 only, TCS only, and in the whole sample. In the combined case, the outcomes of both courses have been averaged and rounded up.

Table 5: Coefficients in ordinal logistic regression with the combined theory score as the explanatory variable.

	Coefficient	95 % CI
Combined theory score	−0.04	[−0.78, 0.71]
Threshold 0 | 1	−0.02	[−0.38, 0.34]
Threshold 1 | 2	0.19	[−0.17, 0.55]
Threshold 2 | 3	0.73	[0.36, 1.11]
Threshold 3 | 4	1.36	[0.93, 1.80]
Threshold 4 | 5	1.98	[1.45, 2.52]

achieve that particular outcome, conditional on and as a function of that student's combined theory score. It also superimposes the corresponding empirical conditional probability (as a bar graph) derived from our data; note that the distribution is somewhat influenced by the choice of bins, and the choice here was influenced by a balance between avoiding small bins where individual data points dominate and avoiding too few bins.

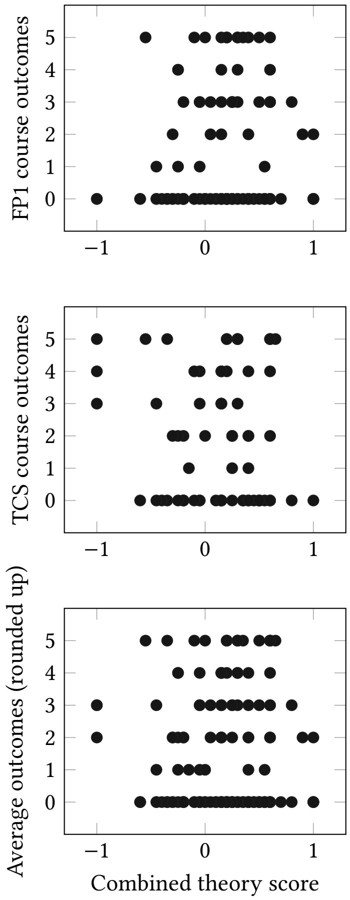

Figure 4: Scatter plot of combined theory score with ordinal course outcomes. The Spearman rank correlation coefficient is $\rho = 0.03$ ($p = 0.76$) for the whole sample, $\rho = 0.11$ ($p = 0.27$) for the FP1 attendees (including overlap), and $\rho = -0.02$ ($p = 0.89$) for the TCS attendees (including overlap). The p values are bootstrapped.

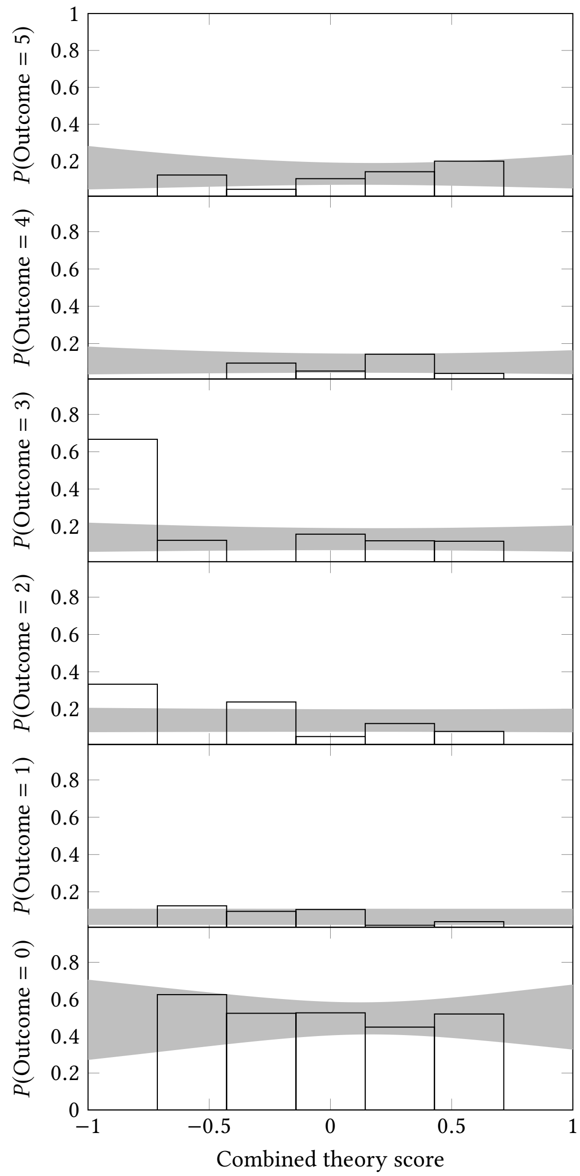

Figure 5: Confidence intervals of fitted probabilities (gray areas) and empirical probabilities (bars) of each outcome level conditional on each combined theory score. The model is an ordinal logistic regression with the combined theory score as the explanatory variable and the ordinal outcomes as the response.

As final corroboration, we categorized each student as entity theorist, incremental theorist, or ambivalent, based on the dichotomization proposed by Dweck et al. [12, p. 269], as discussed in Subsection 2.3. A Kruskal-Wallis rank sum test fails to see any difference in the three groups' averaged outcomes ($\chi^2(2) = 0.133$, $p = 0.936$).

5 DISCUSSION

We can detect no association between a student's mindset and their course outcomes in this study with 133 participants. Assuming, as we do, a causal relationship between them, this means that the effect of mindset on outcomes is too small to be measured using this sample size and study design. Our results are as consistent with a decrease in odds down to roughly one third as with an increase of odds up to roughly doubling them, associated with a one unit of difference in mindset. In the mindset scale we used, one unit difference means moving from a strongly held mindset to a neutral one or vice versa, or moving from one weakly held mindset to another.

The main source of this uncertainty appears to be the pronounced normal distribution of theory of intelligence scores in our data set, contrary to what Dweck et al. suggest: using the Dweck et al. [12, p. 269] dichotomization—which in their study resulted in only 15 % of their participants being categorized as neutral—would discard 49 (37 %) of our participants because they do not endorse either theory clearly. In our data set, we have only 24 (18 %) entity theorists (students with a fixed mindset) by this criterion.

Similarly, there was a large number of noncompletions in our data set, which means that only a small number of students receive any particular passing grade; thus, the amount of data for fixed mindset students who pass our courses is fairly limited—only 13 fixed mindset students passed at least one course. Accordingly, there are very few events per variable in the ordinal regression. However, the events per variable count is adequate in the binary logistic regression with fail/pass as the response variable. As our regressions aim to verify that there is no detectable association in our data, we regard these problems as noncritical.

The key observation is that even our modest data set is able to reject most effect sizes at the conventional $\alpha = 0.05$ level. We are justified in rejecting odds ratios below 0.46 and above 2.03 for a unit change in the combined theory score. Thus, we can say with confidence that the odds of increasing one grade point are not lower than 0.46 times as low, and not higher than 2.03 times as high, with a moderate incremental theory (growth mindset) as with a moderate entity theory (fixed mindset). The remaining effect sizes, which we are unable to reject here, are not trivial, but they are small. This result is consistent with the results of previous observational studies in both computing and other fields.

We do not believe our study to suffer from construct validity issues. We used a well known instrument to measure mindsets, one that has been in use for decades and that has a record of no serious validity issues. While our Finnish translation is new, we have no reason to suspect it to have caused problems. Further, our own data exhibit excellent Coefficient Alpha reliability. Similarly, our measure of course outcomes is the gold standard (grades and credits awarded): while it can be argued, with good reason, that grades and credits do not accurately reflect true skill and knowledge, they do reflect academic achievement at this level by definition. Our course designs do not provide more precise measures of academic success.

Our study also mostly complies with the requirements for authors that Simmons et al. [39] proposed for mitigating the problems of false positive reports. We report our stopping rule; we report all variables that we measured; we report all experimental conditions performed (there were none); we did not eliminate observations based on an analysis of our data; and we report covariate-less analyses. The only problem this study has is in the cell size: our data includes only 11 clear entity theorists who did not complete either course, while Simmons et al. recommend a cell size of at least 20.

The greatest limitation of this study is its observational nature. We cannot rule out the possibility that another factor, not controlled for in this study, acts to counterbalance the effect of the mindset and thus masks it from our view. A properly conducted experiment would be able to eliminate any such confounding. We think such confounding is unlikely, however; far more likely is that the effect of mindsets are simply too low to stand out from the noise with this sample size.

A related limitation of this study is that we did not generate more psychological measurements beyond the implicit theory data. For example, the mindset literature suggests that a high confidence in one's intelligence, which we did not measure, makes one's mindset irrelevant in determining goal orientation. Thus, it is possible that our entity theorists had high self-confidence, removing the disadvantage of their mindset. In future studies, we suggest that researchers always measure self-confidence in intelligence as well as implicit theories; we plan to follow our own suggestion.

It is possible that there is a causal effect of changing mindsets that goes beyond the simple difference in preexisting mindsets; thus, while we could not demonstrate a large effect, we cannot rule out the possibility that an intervention experiment would see effect sizes in the range that we rule out in this study.

One of our coworkers, having heard us discuss these results at the coffee table, suggested an explanation for our results: both of our courses discussed in this paper occur typically in the third year of study, and it is possible that those students who tend to choose to avoid challenges (due to their mindset or for other reasons) have already dropped out of our degree program before they encounter either of those courses. However, Macnamara and Rupani [27] found no influence of mindsets to education levels achieved. Still, even if that explanation was true, it would only mean that our results may not be applicable to first-year student populations.

Taking all that into account, we conclude that mindsets do not seem to matter much. Certainly, our results cast doubt on the wisdom of subjecting all students at the late-stage bachelor level to mindset interventions. The expected payoff is simply not large enough to justify it. In contrast, addressing other student self-beliefs may have a larger effect. For example, self-efficacy interventions have a stronger basis in the literature [28] and their effect is established in context of computer science [23, 44].

Our participants are Finnish students and thus WEIRD (Western, Educated, Industrialized, Rich, and Democratic) people [17]. One should be cautious about transferring our results to a different population. Likewise, it should be noted that there are important differences in educational systems between Finland and, for example, the United States. Nevertheless, our results are, for the most part, compatible with those reported in the prior research, which was primarily based on American participants.

Our conclusion only applies to mass interventions at a single-course context. As another of our colleagues pointed out to us in private conversation, our results say nothing about the usefulness of targeted interventions based on a teacher's expert assessment that a particular student might be hampered specifically by their fixed mindset. Similarly, it may be that mindsets only affect the level at which one stops studying altogether, not the level of achievement at a particular course.

Indeed, we feel that further study of this topic can still be fruitful. We think it may be worthwhile to study the effect of student mindsets on peak achievement (such as dropping out versus graduating from a degree program) in computing. It may also be useful for studies to aim for identifying student subpopulations for which mindset interventions might be worthwhile, and design and evaluate effective interventions for the population of computing student population at various education levels.

ACKNOWLEDGMENTS

We thank Jukka-Pekka Santanen and Ville Isomöttönen for discussions, and Jonne Itkonen for maintaining the questionnaire while we were keeping ourselves ignorant of its results during the courses. We also thank the anonymous reviewers for their useful feedback.

REFERENCES

[1] Jeni L. Burnette, Ernest H. O'Boyle, Eric M. VanEpps, Jeffrey M. Pollack, and Eli J. Finkel. 2013. Mind-sets matter. A meta-analytic review of implicit theories and self-regulation. *Psychological Bulletin*, 139, 3, 655–701. DOI: 10.1037/a0029531.

[2] John Cato. 2011. Mindset matters. *The Physics Teacher*, 49, 1, 60–60. DOI: 10.1119/1.3527763.

[3] Samprit Chatterjee and Jeffrey S. Simonoff. 2013. *Handbook of Regression Analysis*. Wiley, Hoboken, NJ.

[4] Jason A. Chen and M. Shane Tutwiler. 2017. Implicit theories of ability and self-efficacy. *Zeitschrift für Psychologie*, 225, 2, 127–136. DOI: 10.1027/2151-2604/a000289.

[5] Wei-Wen Chen and Yi-Lee Wong. 2015. Chinese mindset. Theories of intelligence, goal orientation and academic achievement in hong kong students. *Educational Psychology*, 35, 6, 714–725. DOI: 10.1080/01443410.2014.893559.

[6] R. H. B. Christensen. 2015. Ordinal: regression models for ordinal data. R package version 2015.6-28. (2015). https://cran.r-project.org/web/packages/ordinal/.

[7] Erin Clevenger. 2013. *The relation of theory of intelligence to academic motivation and academic outcomes*. PhD thesis. The Florida State University. http://purl.flvc.org/fsu/fd/FSU_migr_etd-7331.

[8] Quintin Cutts, Emily Cutts, Stephen Draper, Patrick O'Donnell, and Peter Saffrey. 2010. Manipulating mindset to positively influence introductory programming performance. In *SIGCSE '10 Proceedings of the 41st ACM Technical Symposium on Computer Science Education*, 431–435. DOI: 10.1145/1734263.1734409.

[9] Carol S. Dweck. 1975. The role of expectations and attributions in the alleviation of learned helplessness. *Journal of Personality and Social Psychology*, 31, 4, 674–685. DOI: 10.1037/h0077149.

[10] Carol S. Dweck. 2000. *Self-theories. Their Role in Motivation, Personality, and Development*. Psychology Press, New York.

[11] Carol S. Dweck. 2016. *Mindset. The New Psychology of Success*. (Updated edition). Random House, New York.

[12] Carol S. Dweck, Chi-yue Chiu, and Ying-yi Hong. 1995. Implicit theories and their role in judgments and reactions. A world from two perspectives. *Psychological Inquiry*, 6, 4, 267–285. DOI: 10.1207/s15327965pli0604_1. http://www.jstor.org/stable/1448940.

[13] Carol S. Dweck and Ellen L. Leggett. 1988. A social–cognitive approach to motivation and personality. *Psychological Review*, 95, 2, 256–273. DOI: 10.1037/0033-295X.95.2.256.

[14] Carol S. Dweck and N. Nickon Reppucci. 1973. Learned helplessness and reinforcement responsibility in children. *Journal of Personality and Social Psychology*, 25, 1, 109–116. DOI: 10.1037/h0034248.

[15] Abraham E. Flanigan, Markeya S. Peteranetz, Duane F. Shell, and Leen-Kiat Soh. 2015. Exploring changes in computer science students' implicit theories of intelligence across the semester. In *ICER '15 Proceedings of the Eleventh Annual International Conference on International Computing Education Research*, 161–168. DOI: 10.1145/2787622.2787722.

[16] John Fox and Robert Andersen. 2006. Effect displays for multinomial and proportional-odds logit models. *Sociological Methodology*, 36, 1, 225–255. DOI: 10.1111/j.1467-9531.2006.00180.x.

[17] Joseph Henrich, Steven J. Heine, and Ara Norenzayan. 2010. The weirdest people in the world? *Behavioral and Brain Sciences*, 33, 2–3, 61–135. DOI: 10.1017/S0140525X0999152X.

[18] John M. Hoenig and Dennis M. Heisey. 2001. The abuse of power. The pervasive fallacy of power calculations for data analysis. *American Statistician*, 55, 1, 19–24. DOI: 10.1198/000313001300339897.

[19] Michael R. Jiroutek and J. Rick Turner. 2018. Why it is nonsensical to use retrospective power analyses to conduct a postmortem on your study. editorial. *Journal of Clinical Hypertension*, 20, 2, 408–410. DOI: 10.1111/jch.13173.

[20] Päivi Kinnunen and Lauri Malmi. 2006. Why students drop out cs1 course? In *ICER '06 Proceedings of the second international workshop on Computing education research*, 97–108. DOI: 10.1145/1151588.1151604.

[21] Lukasz Komsta and Frederick Novomestky. 2015. *moments: Moments, cumulants, skewness, kurtosis and related tests*. R package version 0.14. https://cran.r-project.org/web/packages/moments/.

[22] Sheri R. Levy, Steven J. Stroessner, and Carol S. Dweck. 1998. Stereotype formation and endorsement. The role of implicit theories. *Journal of Personality and Social Psychology*, 74, 6, 1421–1436. DOI: 10.1037/0022-3514.74.6.1421.

[23] 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 *Proceedings of the 2016 ACM Conference on International Computing Education Research* (ICER '16). ACM, New York, NY, USA, 211–220. DOI: 10.1145/2960310.2960329.

[24] Dastyni Loksa, Andrew J. Ko, Will Jernigan, Alannah Oleson, Christopher J. Mendez, and Margaret M. Burnett. 2016. Programming, problem solving, and self-awareness. Effects of explicit guidance. In *CHI '16 Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 1449–1461. DOI: 10.1145/2858036.2858252.

[25] Marko Lüftenegger and Jason A. Chen. 2017. Conceptual issues and assessment of implicit theories. *Zeitschrift für Psychologie*, 225, 2, 99–106. DOI: 10.1027/2151-2604/a000286.

[26] Robert C. MacCallum, Shaobo Zhang, Kristopher J. Preacher, and Derek D. Rucker. 2002. On the practice of dichotomization of quantitative variables. *Psychological Methods*, 7, 1, 19–40. DOI: 10.1037/1082-989X.7.1.19.

[27] Brooke N. Macnamara and Natasha S. Rupani. 2017. The relationship between intelligence and mindset. *Intelligence*, 64, 52–59. DOI: 10.1016/j.intell.2017.07.003.

[28] Karen D. Multon, Steven D. Brown, and Robert W. Lent. 1991. Relation of self-efficacy beliefs to academic outcomes. A meta-analytic investigation. *Journal of Counseling Psychology*, 38, 1, 30–38. DOI: 10.1037/0022-0167.38.1.30.

[29] Jum C. Nunnally and Ira H. Bernstein. 1994. *Psychometriic Theory*. (3rd edition). McGraw-Hill, New York.

[30] Heather A. Priess-Groben and Janet Shibley Hyde. 2017. Implicit theories, expectancies, and values predict mathematics motivation and behavior across high school and college. *Journal of youth and adolescence*, 46, 6, 1318–1332. DOI: 10.1007/s10964-016-0579-y.

[31] R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/.

[32] William Revelle. 2017. *psych: Procedures for Psychological, Psychometric, and Personality Research*. R package version 1.7.8. Northwestern University. Evanston, Illinois. https://cran.r-project.org/web/packages/psych.

[33] Richard W. Robins and Jennifer L. Pals. 2002. Implicit self-theories in the academic domain. Implications for goal orientation, attributions, affect, and self-esteem change. *Self and identity*, 1, 4, 313–336. DOI: 10.1080/1529886029010680.

[34] Derek D. Rucker, Blakeley B. McShane, and Kristopher J. Preacher. 2015. A researcher's guide to regression, discretization, and median splits of continuous variables. *Journal of Consumer Psychology*, 25, 4, 666–678. DOI: 10.1016/j.jcps.2015.04.004.

[35] Michael J. Scott and Gheorgita Ghinea. 2014. On the domain-specificity of mindsets. The relationship between aptitude beliefs and programming practice. *IEEE Transactions on Education*, 57, 3, 169–174. DOI: 10.1109/TE.2013.2288700.

[36] Michael James Scott and Gheorghita Ghinea. 2013. Implicit theories of programming aptitude as a barrier to learning to code. Are they distinct from intelligence. In *ITiCSE '13 Proceedings of the 18th ACM conference on Innovation and technology in computer science education*. ACM, New York, 347. DOI: 10.1145/2462476.2462515.

[37] Michael James Scott and Gheorghita Ghinea. 2014. Measuring enrichment. The assembly and validation of an instrument to assess student self-beliefs in cs1. In *ICER '14 Proceedings of the tenth annual conference on International computing education research*, 123–130. DOI: 10.1145/2632320.2632350.

[38] Duane F. Shell, Leen-Kiat Soh, Abraham E. Flanigan, and Markeya S. Peteranetz. 2016. Students' initial course motivation and their achievement and retention in college cs1 courses. In *SIGCSE '16 Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, 639–644. DOI: 10.1145/2839509.2844606.

[39] Joseph P. Simmons, Leif D. Nelson, and Uri Simonsohn. 2011. False-positive psychology. Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22, 11, 1359–1366. DOI: 10.1177/0956797611417632.

[40] Beth Simon, Brian Hanks, Laurie Murphy, Sue Fitzgerald, Renée McCauley, Lynda Thomas, and Carol Zander. 2008. Saying isn't necessarily believing. Influencing self-theories in computing. In *ICER '08 Proceedings of the Fourth International Workshop on Computing Education Research*, 173–184. DOI: 10.1145/1404520.1404537.

[41] F. Boray Tek, Kristin S. Benli, and Ezgi Deveci. 2018. Implicit theories and self-efficacy in an introductory programming course. *IEEE Transactions on Education*. DOI: 10.1109/TE.2017.2789183.

[42] Dirk T. Tempelaar, Bart Rienties, Bas Giesbers, and Wim H. Gijselaers. 2015. The pivotal role of effort beliefs in mediating implicit theories of intelligence and achievement goals and academic motivations. *Social Psychology of Education*, 18, 1, 101–120. DOI: 10.1007/s11218-014-9281-7.

[43] Hadley Wickham. 2011. The split-apply-combine strategy for data analysis. *Journal of Statistical Software*, 40, 1, 1–29. http://www.jstatsoft.org/v40/i01/.

[44] Daniel Zingaro and Leo Porter. 2016. Impact of student achievement goals on cs1 outcomes. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education* (SIGCSE '16). ACM, New York, NY, USA, 279–296. DOI: 10.1145/2839509.2844553.

[45] Valorie L Zonnefeld. 2015. *Mindsets, attitudes, and achievement in undergraduate statistics courses*. PhD thesis. University of South Dakota. https://digitalcollections.dordt.edu/faculty_work/199.