Assessing student behavior in computer science education with an fsQCA approach: The role of gains and barriers

Ilias O. Pappas

Department of Computer Science, Norwegian University of Science and Technology (NTNU), ilpappas@ntnu.no

Michail N. Giannakos

Department of Computer Science, Norwegian University of Science and Technology (NTNU), michailg@ntnu.no

Letizia Jaccheri

Department of Computer Science, Norwegian University of Science and Technology (NTNU), letizia.jaccheri@ntnu.no

Demetrios G. Sampson

School of Education, Curtin University, demetrios.sampson@curtin.edu.au

Abstract

This study uses complexity theory to understand the causal patterns of factors that stimulate students' intention to continue studies in computer science (CS). To this end, it identifies gains and barriers as essential factors in CS education, including motivation and learning performance, and proposes a conceptual model along with research propositions. To test its propositions, the study employs fuzzy-set qualitative comparative analysis (fsQCA) on a data sample from 344 students. Findings indicate eight configurations of cognitive and noncognitive gains, barriers, motivation for studies and learning performance that explain high intention to continue studies in CS. This research study contributes to the literature by: (1) offering new insights into the relationships among the predictors of CS students' intention to continue their studies and (2) advancing the theoretical foundation of how students' gains, barriers, motivation and learning performance combine to better explain high intentions to continue CS studies.

• Social and professional topics~Computer science education

Additional Key Words and Phrases: Higher education; Student behavior; Fuzzy-set qualitative comparative analysis; Configuration; Contrarian case

1. INTRODUCTION

Computer information science and technology—referred to here as computer science (CS)—has been receiving increased attention owing to a growing demand for CS professionals. Specifically, the US Bureau of Labor Statistics predicts that by 2020, half of STEM (science, technology, engineering, and mathematics) jobs will be in CS, and more than half will require significant CS skills and knowledge [2014]. In addition to the importance of the CS degree per se, nearly every field related to STEM has seen the growth of a computing counterpart (e.g., Bioinformatics, Computational Statistics, Computational Chemistry, Computational Biology). This rise in importance of computation with respect to the broader STEM fields has been recognized both by those within the STEM education communities and CS education organizations [ACM/IEEE curricula 2013]. The number of CS graduates has declined over the past 10 years, and only recently has this number begun to recover [Zweden 2014]. By 2020, estimates are that there will be a shortage of more than 800,000 CS professionals across Europe [1] [see also Ireland Department of Education and Skills, 2014]. Nonetheless, students seem reluctant to continue studies in CS and pursue this career [US Bureau of Labor Statistics 2014].

Students who experience high levels of academic and social integration will, in many cases, continue their studies in CS [Cohoon 2006; Rosson et al. 2011; Xenos et al. 2002]. Prior research reports a lack of awareness of the discipline coupled with negative attitudes towards CS among students [Carter, 2006; Grover et al., 2014]. Therefore, it appears that misconceptions and lack of the expected gains and career perspectives potentially hinder interest, which in turn results in students not opting to enroll in a CS course [Grover et al. 2014]. Students' disinterest in CS is related with a lack of familiarity with the subject [Carter, 2006], which may lead to poor educational decisions [Hewner, 2013]. Graduate students, compared with undergraduates, have shown more interest in the gains from the learning process and in their overall perceptions of the process, as well as in their evaluations of learning effectiveness [Lee and Tsai 2011]. It is thus essential to understand students' perceptions toward CS studies, including potential benefits and barriers, and to consider their motivations and performance in order to examine their behavior after graduation (i.e., intention to continue studies in the field) [Barker et al. 2014].

Prior studies contribute to understanding the importance of cognitive and noncognitive gains in students' future behavior [Pike et al. 2011]. *Cognitive gains* refer to students' general education, including writing and speaking skills, as well as their critical thinking [Toutkoushian and Smart 2001]. *Noncognitive gains* refer to benefits from working with others, developing ethical standards, and community engagement [Toutkoushian and Smart 2001]. Similarly, it has been found that performance and motivation are antecedents of graduates' future behavior [Araque et al. 2009]. *Learning performance* refers to students' academic achievements and is represented by their overall grade point averages (GPA). *Motivation to study* refers to students' intrinsic or extrinsic motivations to choose CS as their field of study. These motivations are career perspectives, reputation, and personal interest in CS.

Further, literature suggest a number of barriers to explain students' intentions to continue CS studies [Barker et al. 2009; Biggers et al. 2008]. *Barriers* refer to impediments to students continuing their CS studies, including *personal values, quality of teaching* and *satisfaction with learning effectiveness*. In detail, *personal values* refer to students' sense of belonging, fulfillment and social norms. *Quality of teaching* refers to the quality of teaching provided to students during their studies.

[1] http://ec.europa.eu/digital-agenda/en/skills-jobs

Satisfaction with learning effectiveness refers to students' satisfaction with their studies. However, it is less clear which of these factors drive students to continue studying CS and, more importantly, how their combinations better explain students' intentions to continue their CS studies. Existing studies on the antecedents of students' future behaviors either focus on the main effects of specific predictors [e.g., gains, barriers] on intention to continue CS studies, or they fail to examine the combined effects of gains, motivation, performance and barriers on that intention.

This research showed factors that influence interest in CS education, and builds on complexity theory with the aim of exploring the causal patterns of factors that stimulate students to continue their studies in CS. In particular, this study attempts to *elucidate how students' cognitive and noncognitive gains, learning performance, motivation and barriers combine to lead to increased intention to continue CS studies*. Instead of focusing on the main effects between students' intentions and their antecedents, the goal of this study is to detect specific configurations that explain students' intentions to continue their studies. Thus, the study addresses the following research question:

What configurations of cognitive and noncognitive gains, learning performance, motivation and barriers lead to high intention to continue studying CS?

Identifying these configurations should help universities and colleges to specify detailed patterns of factors that stimulate CS student behavior, and allow them to create more efficient CS programs and study conditions. To answer the research question, we employ fuzzy set qualitative comparative analysis (fsQCA) [Ragin, 2008]. Specifically, we bridge configurational analysis with complexity theory in the field of CS education. fsQCA has received increased attention recently because when it is applied together with complexity theory, researchers have the opportunity to gain deeper and richer perspectives on their data [Woodside 2014]. "Complexity theory is destined to be the dominant scientific trend of the 1990's…This revolutionary technique can explain any kind of complex system—multinational corporations, or mass extinctions, or ecosystems such as rainforests, or human consciousness. All are built on the same few rules" [Lewin 1992, back cover]. We expand on the contributions of other studies from the areas of sociology [Ragin 2008], business management [Pappas et al. 2017], information systems [Liu et al., 2015] and others.

The remainder of the paper is organized as follows. Section 2 presents related work and develops the propositions of this study. Section 3 presents the applied measures for data collection, and section 4 describes the research methodology. Section 5 presents the empirical results derived, and the final section of the paper discusses the findings and conclusions, highlighting theoretical and practical implications.

2. RELATED WORK, CONCEPTUAL MODEL AND RESEARCH PROPOSITIONS

CS education enables creative problem-solving capabilities and helps students identify opportunities for innovations that may lead to a broad range of career paths (e.g., high-tech companies, healthcare industries). Towards this end, CS offers a wide range of knowledge that can aid students in gaining those skills and competences that are required by the industry (e.g., project management, progress monitoring and communication, problem solving, understanding human behavior). Such competences will give CS students the opportunity to design and build meaningful artifacts. Further, skills such as computational thinking, which includes problem solving and evaluation, as well as understanding human behavior, are crucial in CS education.

Many previous studies identify factors related to CS enrollment, retention and career choice. In this study we focused on peer reviewed empirical articles from the CS/Informatics discipline, which identify factors associated with students' retention in higher education. Thus, a review search was conducted by following the Critical

Appraisal Skills Programme (CASP)[2] quality assessment, in order to evaluate the relative strength of research rigor, as well as the empirical evidences reported. Next, after identifying the main articles, we briefly review factors that have been found to influence students' interest in CS education and career, and finally we select the appropriate measures based on the literature. In particular, the measures used in this study are carefully singled out from prior related studies [Biggers et al. 2008; Joo et al. 2013; Pike et al. 2011; Pirker et al. 2014; Rosson et al. 2011; Xenos et al. 2002].

2.1 Factors that influence students' behavior

The factors that may influence students to continue to pursue a certain degree or choose a different one are varied, but may be divided into three categories: (1) academic environment and resources, (2) perceptions of the discipline and career and (3) experience [Hein et al. 2012]. Academic environments and resources include lectures and laboratories; faculty and teaching assistants; university services and others. Perceptions refer to ideas about oneself, including confidence, self-efficacy and determination to succeed; perceptions of the major and career include the opportunities and advancements provided by the field, as well as the society's perceptions of the field. Experience refers to personal experiences, such as discrimination based on stereotypes and the university environment.

Literature in the area of CS offers evidence on various factors that may influence students' behavior in STEM-related studies, such as learning performance, motivation and different learning styles, demographics, experience [Marra et al. 2012; McGill et al., 2016, Pappas et al., 2016b]. These factors may be either external or internal to the students; for example, their gains [Li et al. 2009] and their self-perceptions regarding their basic science and math skills, their ability to work effectively in teams and their capacity to apply theories to real-world problems [Bjorklund et al. 2004]. Self-reported gains may help in understanding student retention and academic success [Bjorklund et al. 2004]. In addition, studies show that gains influence students' decisions regarding pursuing STEM studies, in that some students might be disappointed with their curricula or their knowledge gained [Seymour and Hewitt 1997].

Researchers find students' *gains* (both cognitive and noncognitive) to be among the most important reasons for choosing a particular study program. Previous studies identify gains as essential for academic achievement and choice behavior in the demanding STEM subjects [Chow et al. 2012; Watt et al. 2012]. *Motivation to study*, which describes students' reasons for pursuing CS, also influences graduates' behavior [Pirker et al. 2014]. Furthermore, learning outcomes may change significantly if students put more effort into improving their academic performance [Yu and Jo 2014].

2.2 Barriers to studying CS

There are many potential barriers to students' continuing their studies in CS [Rosson et al. 2011; Xenos 2002], including negative or inaccurate beliefs about the nature of CS professions, and stereotypes and biases derived from cultural differences [Rosson et al., 2011]. The main reasons for not continuing CS studies include professional, academic, family, health-related and personal issues [Xenos, 2002]. In addition, Morton [2005] suggests that students who choose computer programming as a career realize the actual nature of the profession only after they begin working, indicating that they do not always have a clear view of CS during their studies. Similarly, Carter [2006] posits that students with an evident aptitude toward CS careers do not have sufficient knowledge about what becoming a computer professional might entail.

[2] http://media.wix.com/ugd/dded87_a02ff2e3445f4952992d5a96ca562576.pdf

Perceived values related to CS programs influence students' feelings and sense of belongingness, and may have both negative and positive effects on their behavior regarding the choice of the same program or a different one [Marra et al. 2012]. Further, students' behavior may be positively affected by a welcoming atmosphere [Walden and Foor 2008] or when they feel part of the intellectual and social CS community [Barker et al. 2009]. In contrast, a negative effect may occur when students feel that the group they belong to does not reflect them [Barker et al. 2009]. Adding to this, high levels of social support, such as studying with friends, also contribute to students' overall sense of belonging in their programs. Teaching activities may act as a disincentive to successful CS studies [Blickenstaff 2005]. For example, the format and quality of lectures, which predominate in many CS courses, especially during the first two years, may act as barriers to students identifying with the course subjects, creating a distance between the students and the program. Furthermore, common pedagogical activities in CS education may not fit with students' personal approaches and styles to learning course material.

Literature has widely studied the role of difficult and excessive learning material, as well as that of demanding exams and assignments, in CS and STEM programs [Jacobs 2005]. The key introductory undergraduate courses in CS study programs may be perceived as barriers, as although the degree requires them, they have high withdrawal and failure rates [e.g., calculus, physics], and performance in these courses is related to CS persistence [Suresh, 2006]. To this end, Araque et al. [2009] suggest that the difficulty of a course influences CS persistence. Mau [2003] identifies academic preparation and self-efficacy in difficult subjects as the only reliable predictors of persistence when examining intentions to continue studies in STEM.

Previous studies in the area of CS offer evidence on various factors that influence students' behavior, including their perceptions, beliefs and previous experiences with their studies [Barker et al. 2009; Barker et al. 2014; Biggers et al. 2008; Giannakos et al., 2016]. Nonetheless, the existence of different results throughout the literature, as mentioned above, suggests that more research is necessary on students' intentions to continue studying CS. Moreover, the various factors identified in the literature, as well as the changes in industry demands combined with the increased need for CS graduates, highlight the need for further research on the area, along with new methods that will offer fresh insight into the existing CS literature. Thus, the present study takes a different methodological approach by implementing configurational analysis, which studies find explains the contradicting results from commonly used regression-based symmetric tests [Fiss 2011; Pappas et al. 2017; Woodside 2013].

2.3 Conceptual model

Prior research on learning and CS implements symmetric tests to examine the hypotheses and calculate net effects on the desired outcomes. The main focus of tests such as multiple regression analysis is to estimate the significance of the effects between two variables or to compare the effects among the variables between two or more models. However, as Woodside [2013; 2014] posits, focusing on net effects may be misleading, usually because the observed net effects do not apply to all of the cases in a dataset. Thus, we suggest quite a different approach from the commonly used structural equation modeling in order to show the various combinations that may occur among the variables, thereby increasing the contribution of the research. In other words, two variables in a dataset, as with most relationships in real life, are likely to present positive effects for part of the sample, but also negative effects for a different part of the same sample.

Relationships between two variables (e.g., A, B) are complex, and the presence of one (i.e., A) may lead to the presence of the other (i.e., B), suggesting sufficiency.

However, at the same time, variable B may be present even when variable A is absent, suggesting that the presence of A is a sufficient but unnecessary condition for variable B to occur. Sufficiency and necessity describe subset relationships among variables [Glaesser and Cooper 2012]. Similarly, especially when additional variables exist, variable A may be necessary but insufficient for B to occur. For example, a student who is highly interested in CS may have high intentions to continue studying CS regardless of any perceived gains, performance or barriers, suggesting that motivation to study is a sufficient condition for high intentions. Furthermore, we expect that students with high interest will also be motivated by the perceived gains of studying CS, indicating that their combination will be a sufficient condition for high intentions. Similarly, multiple relationships exist among variables that, depending on how they combine, may or may not explain students' high intention to continue studies in CS. Thus, in order to conceptualize these relationships, we propose a Venn diagram (Fig. 1) that accurately reflects them.

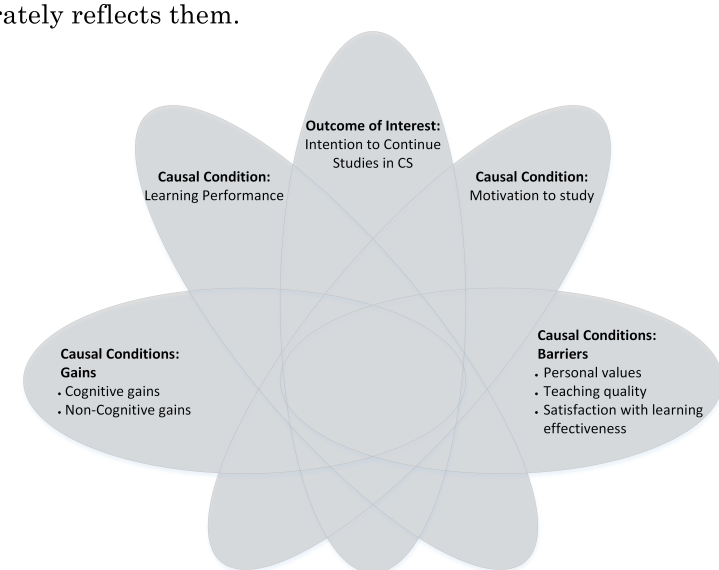

Fig. 1. Venn diagram illustrating the conceptual model

The Venn diagram illustrates five sets of constructs and their intersections. The five sets of constructs reflect the outcome of interest (dependent variable) of this study and four sets of causal conditions to predict the outcome (independent variables). Specifically, the outcome of interest is students' intention to continue studies in CS, and the four sets of causal conditions are gains (i.e., cognitive and noncognitive], motivation to study (i.e., interest in CS), learning performance [i.e., actual grades] and barriers (i.e., feelings, personal values, quality of teaching, satisfaction with learning effectiveness). The intersections illustrate factor configurations, which are higher-level interactions. Specifically, the overlapped areas represent the possible combinations among the factors, that is areas on which a distinct factor may co-exist with the rest. For example, all combinations that explain high intention to continue studies in CS are included within the outcome of interest area (Fig.1).

2.4 Research propositions

Extant research in the area of CS education identifies multiple factors that affect students' behavior and their intention to continue studying CS [e.g., Barker et al. 2009; Pappas et al., 2016b; Rosson et al. 2011]. Previous studies indicate different effects of cognitive and noncognitive gains on students' behavior. For example, one study detects meaningful differences in self-reported gains and attitudes among students in inquiry-based learning [Laursen et al. 2014]. Also, cognitive gains have been found to increase the usefulness of a CS degree, while non-cognitive gains reduce it [Giannakos et al., 2016]. An examination of the results for first-year students' cognitive gains reveals statistically significant variation in perceived gains across institutions. Furthermore, statistically significant variation is noted in first-year students' noncognitive gains across institutions [Pikes et al. 2011].

Furthermore, studies cite many barriers as important antecedents of students' behavior, although their results have been mixed [Barker et al., 2014; Marra et al. 2012;]. Similarly, research finds that high learning performance on its own is not always able to explain students' behavior because students may maintain high GPAs, but their interest in CS may be low. In addition, the motivation to study CS varies, and when students have little or incorrect knowledge about the subject, they are likely to be uninterested in continuing CS studies [Carter 2006]. For example, students with low interest in CS may still be motivated to continue their studies by the career opportunities that CS offers. Thus, in order to better understand students' intention to continue their CS studies, a configurational analysis of related factors is more appropriate than examining individual causal factors. As presented in Fig. 1, this perspective leads to more complex causal patterns and higher-level interactions between the constructs.

Complexity theory incorporates the principle of *equifinality*, based on which the outcome of interest can be explained equally by alternative sets of causal conditions that combine in sufficient configurations for the outcome [Fiss 2011; Woodside 2014]. Gains, barriers, motivation to study and learning performance are essential causal conditions for understanding students' intentions to continue studies in CS, and they may combine in different configurations. For example, Pike et al. [2011] suggest that different gains [i.e., cognitive or noncognitive] and student characteristics may have different relationships with learning outcomes. Thus, configurations may include various combinations, leading to the first proposition:

Proposition 1. *No single configuration of students' gains, motivation to study, learning performance and barriers leads to high intention to continue CS studies; rather, there exist multiple, equally effective configurations of causal factors.*

Complexity theory further proposes the occurrence of causal asymmetry. Causal asymmetry means that for an outcome to occur, the presence and absence of a causal condition depend on how this condition combines with one or more others [Leischnig and Kasper-Brauer 2015; Woodside 2014]. For example, although barriers are likely to have a negative influence on students' behavior [Mara et al. 2011], students who face high barriers may still have high intention to continue studying CS depending on how the barriers combine with gains, motivation or performance. Further, the absence of barriers might not always lead to high intention to continue studying CS, given that students are also affected by cognitive and noncognitive gains, and may have different motivations or show different performance. Thus, the second and third propositions will be formed as follows:

Proposition 2. *Single causal conditions may be present or absent within configurations for students' high intention to continue studies in CS, depending on how they combine with other causal conditions.*

Proposition 3. *Configurations, with the presence of at least one barrier, that lead to high intention to continue studies in CS will also require the presence of at least gains or motivation as causal conditions.*

3. RESEARCH METHODOLOGY

3.1 Sample

The research methodology used a survey composed of questions on background and on the identified constructs. We used a number of different methods to attract respondents; we distributed questionnaires in university areas and sent e-mails to CS student mailing lists. The survey was open during the last three weeks of February 2015. We aimed to engage approximately 1100 Norwegian CS students, 438 of whom responded, giving a response rate of approximately 40%. Female participation was very high, at 57% [82 out of the approximately 150 female students], compared with 38% of males [353 out of the approximately 950 male students]. Surveys with over 5% missing responses were removed from the sample. Thus, 344 (32.58%) valid surveys were used for the analysis. The vast majority of the respondents (81.1%) were males. Further, most of the respondents were younger than 26 (94.5%), with the majority being between 20 to 23 years old (64.3%). These ratios indicate good representativeness in that they accord with most CS study programs [Zweben 2014].

3.2 Measures

The questionnaire consisted of three parts. The first part comprised questions on the demographics of the sample (e.g., age, gender). The second part consisted of measures of constructs that the previous literature has identified as important: (1) gains, (2) motivation and (3) performance; in addition, we included a factor indicating students' intention to continue their studies in CS. The third part comprised measures of the various reasons identified in the literature that lead to students' decisions to leave CS studies (barriers).

In the second part, student gains were either cognitive or noncognitive. The cognitive gains referred to three items; specifically, the questions asked students to indicate the extent to which their college experiences had contributed to their progress in general education, writing and speaking effectively and critical thinking [Toutkoushian and Smart 2001]. The noncognitive gains were measured based on the students' responses to two questions concerning self-understanding, working with others, developing ethical standards and civic/community engagement [Toutkoushian and Smart 2001]. Motivation and learning performance were measured with single items. Specifically, regarding motivation, the survey asked students about their reasons for pursuing CS studies; e.g., interest in the subject, career perspectives, reputation. The study assessed learning performance using the students' actual GPAs. Table I lists the operational definitions of the constructs in this theoretical model, as well as the studies from which we adopted the measures. In all cases except for the two single-item questions, items were rated on seven-point Likert scales [1 = Not at all to 7 = Very much]. Appendix A lists the questionnaire items used to measure each construct, along with descriptive statistics and loadings. All questions were formed in way that made it clear that we refer to CS studies, that is "Please rate the following qualities based on your personal experience while studying CS".

Table I. Construct definitions

<table>
<tr><th>Construct</th><th>Definition</th><th>Source</th></tr>
<tr><td colspan="3">Gains</td></tr>
<tr><td>Cognitive</td><td>General education, writing and speaking effectively and critical thinking</td><td rowspan="2">Toutkoushian and Smart [2001]</td></tr>
<tr><td>Noncognitive</td><td>Working with others, developing ethical standards, and civic/community engagement</td></tr>
<tr><td colspan="3">Barriers</td></tr>
</table>

Personal Values	Students sense of belonging, fulfillment and social norms	Biggers et al. [2008]
Teaching Quality	The quality of teaching provided to students during their studies	
Satisfaction with Learning Effectiveness	Students' satisfaction with their studies	
Intention to Continue Studies in CS	Students' intention to continue studying CS	Barker et al. [2014]
Motivation to Study	Reason for taking studies in CS	Xenos et al. [2002]
Learning Performance	The students' overall grade point average (GPA) when they study CS as a major.	

3.2.1 Reliability and validity of the measures

This study evaluated its constructs in terms of reliability and validity. Reliability testing based on Cronbach's alpha showed acceptable indices of internal consistency in that all constructs exceeded the cut-off threshold of 0.70. Establishing validity requires that average variance extracted (AVE) be greater than 0.50 and that the correlations between the different variables in the confirmatory models not exceed 0.8 points, the latter because exceeding 0.8 suggests low discrimination; in addition, the square root of each factor's AVE must be larger than its correlations with other factors [Fornell and Larcker 1981]. The AVEs for all constructs ranged between 0.55 and 0.80, all correlations were lower than 0.80, and the square root AVEs for all constructs were larger than their correlations. Table II displays the findings.

Table II. Descriptive statistics and correlations of latent variables

				Construct							
Construct	Mean (SD)	CR	AVE	1	2	3	4	5	6	7	8
Cognitive Gains	4.96 (1.21)	.79	.56	**.748**							
Noncognitive Gains	3.63 (1.43)	.71	.55	.343	**.742**						
Learning Performance	2.73 (.78)	-	-	-.141	-.006	**-**					
Motivation	2.50 (.92)	-	-	-.019	-.128	-.066	-				
Personal Values	2.44 (1.38)	.84	.58	.008	.259	.084	-.122	**.762**			
Teaching Quality	3.48 (1.35)	.82	.53	-.057	.036	.019	-.040	.511	**.728**		
Satisfaction with Learning Effectiveness	3.22 (1.51	.85	.66	-.005	-.166	.273	-.099	.640	.569	**.812**	
Intention to Continue Studies in CS	5.38 (1.83)	.94	.80	.101	-.010	-.022	-.011	-.096	-.144	.078	**.894**

Note: Diagonal elements [in bold] are the square roots of the AVE. Off-diagonal elements are the correlations among constructs (correlations of 0.1 or higher are significant, p< 0.01). For discriminant validity, diagonal elements should be larger than off-diagonal elements. Learning performance and motivation are single-item variables.

Further, the study tested for multicollinearity [O'Brien 2007] along with the potential common method bias by utilizing Harman's single-factor test [Podsakoff et al. 2003]. The variance inflation factor for each variable was below 3, indicating that multicollinearity was not an issue. The results also suggest an absence of common method bias in that the first factor did not account for the majority of the variance and no single factor occurred from the factor analysis. Next, we assessed the research model's goodness of fit using multiple indices. The chi-square statistic, the comparative fit index (CFI), and the root mean square error of approximation (RMSEA) served to assess the overall measurement model fit. All values were within the recommended range; specifically, x^2/df: 2.65, CFI: 0.93 and RMSEA: 0.06.

3.2.2 Internal and external validity

Internal validity refers to "the validity of inferences about whether observed covariation between A [the presumed treatment] and B [the presumed outcome] reflects a causal relationship from A to B as those variables were manipulated or measured" [Shadish, 2010, p.4]. Thus, concerns regarding the internal validity of an experiment may arise when changes in B have effects other than the manipulation of A. In order to establish internal validity, the participants in the experiment should experience similar stimuli with only a minimum of distortion [Campbell and Stanley 1966]. The threat regarding factors outside of the experiment was the same for all participants in this study. The participants were all active CS students in higher education from different CS study programs; thus, all of them had the same conditions regarding their overall experiences during their studies. One could argue that it might have been better to include students from only one study program; however, this study used all of the available programs to ensure that the sample consisted of students with different experiences, and also that the sample would be large and representative [344 respondents]. In addition, all of the study programs conformed to the international standards [ACM/IEEE curricula 2013] and were accredited by international authorities.

External validity refers to "the validity of inferences about whether the cause–effect relationship holds over variation in persons, settings, treatment variables, and measurement variables" [Shadish, 2010, p.4]. In other words, it concerns the extent to which an effect in research may be generalized. The results of this study should be used in higher education to examine CS students' behavior and may be applied to a variety of teaching subjects. Nonetheless, the full sample consisted of CS students from Norway, and the findings are based on a single survey; hence, generalization should be performed with caution.

3.3 Data analysis

3.3.1 Contrarian case analysis

When examining main relations between two variables, stating that a variable positively or negatively affects an outcome suggests that most cases in a sample verify this relationship. Nonetheless, the opposite relationship will occur for some of the cases in the sample; thus, researchers should test their data for such contrarian cases [Pappas et al. 2016a; Woodside 2014]. In other words, two variables may have positive, negative and no effect in the same dataset, regardless of the main effect of one on the other. In order to identify such opposite relationships, studies employ contrarian case analysis; contrarian cases occur regardless of the significance of the main effects [Woodside 2014].

We followed the method described by Pappas et al. [2016a,c]. First, the sample needed to be divided in order to investigate the relationships among the examined variables. Continuous independent variables should not be split using methods such as median split, because it may lead to a reduction of statistical power as well as to false results when the variables are correlated [Fitzsimons 2008]. Thus, we created quintiles [i.e., dividing the sample into five equal groups] by ranking the cases using the SPSS Rank Cases corresponding function with the Ntiles option. Next, we performed cross-tabulations across the quintiles, using the SPSS Crosstabs function, between every independent variable and the dependent variable. The result for any two variables is a 5x5 table that presents all combinations for all of the cases in the sample between the two variables. The overall results for the contrarian case analysis are presented in Appendix B. The findings indicate the existence of various relationships between the variables, separate from the main effect. To this end, the results support the importance of configurational analysis for explaining these relationships [Woodside 2014].

3.3.2 Fuzzy Set Qualitative Comparative Analysis

This study used fsQCA using fs/QCA 2.5 [Ragin and Davey 2014]. fsQCA was developed by integrating fuzzy set and fuzzy logic [Zadeh 1965] with QCA [Ragin, 2000]. Fuzzy sets and fuzzy logic principles apply in engineering and control theory, as well as in social sciences [Liu et al. 2015]. fsQCA identifies patterns between independent and dependent variables, which leads to outcomes and goes a step further from analyses of variance, correlations and multiple regression models. A variable that affects the outcome in only a small subset of cases cannot be identified by regression analysis [Liu et al. 2015; Vis 2012]. Further, fsQCA offers two types of configurations: necessary and sufficient. Such configurations may be marked by their presence, their absence, or a "do not care" condition. The necessary and the sufficient conditions create a distinction among core and peripheral elements. Core elements are those with strong causal relationships with the outcome, and peripheral elements are those with weaker ties [Fiss 2011].

The first step in fsQCA is to define the outcome and the independent measures. The next is to calibrate all measures into fuzzy sets with values ranging from 0 to 1 [Ragin 2008]. Data calibration may be either direct or indirect. In the direct method, the researcher chooses three qualitative breakpoints, whereas in the indirect method, the measurements require rescaling based on qualitative assessments. The researcher may choose either method depending on the data and the underlying theory [Liu et al. 2015; Ragin 2008]. Studies recommend the direct method of setting three values that correspond to full-set membership, full-set non-membership and intermediate-set membership [Ragin 2008].

This study follows the direct method of data calibration. In detail, the value of 1 stands for full-set membership and that of 0 stands for non-set membership. Thus, all variables are continuous from 0 to 1, which defines the level of their membership. Variables were transformed into calibrated sets with the fsQCA software [using the Calibrate function] by setting three meaningful thresholds: full membership, full non-membership and the cross-over point, which describes how much the case belongs to a set [Ragin 2008]. Calibration then followed the procedure employed by Ordanini et al. [2014]. With this method, the three qualitative anchors for the calibration were based on the survey scale (seven-point Likert scale). The full membership threshold was fixed at the rating of 6; the full non-membership threshold was fixed at 2; and the crossover point was fixed at 4. The values of every variable were calibrated based on a linear function to fit into the three aforementioned thresholds.

Following the calibration, the fsQCA algorithm produced a truth table of 2^k rows, with k representing the number of outcome predictors and each row representing each possible combination. For example, a truth table between two variables (i.e., conditions) would provide four possible logical combinations between them. For every combination, the minimum membership value was calculated; that is, the degree to which every case supports the specific combination. fsQCA uses the threshold of 0.5 to identify the combinations that are acceptably supported by the cases [Liu et al., 2015]. Thus, all combinations that were not supported by at least one case with membership over the threshold of 0.5 were removed from further analysis.

The final step is to sort the truth table based on frequency and consistency [Ragin 2008]. Frequency describes the number of observations for each possible combination. Consistency refers to "*the degree to which cases correspond to the set-theoretic relationships expressed in a solution*" [Fiss 2011, p. 402]. A frequency cut-off point needs to be set in order to ensure that a minimum number of empirical observations is obtained for the assessment of subset relationships. For small and medium-sized samples, a cut-off point of 1 is appropriate, but for large-scale samples [e.g., 150 or more cases], the cut-off point should be set higher [Ragin 2008]. Thus, the minimum acceptable observation frequency was set at 3 [Fiss 2011], and the lowest acceptable consistency for observations was set at >.80, higher than the recommended threshold of 0.75 [Ragin 2006].

4. FINDINGS

4.1 Results from the Fuzzy Set Qualitative Comparative Analysis

Table III shows the outcomes of the fuzzy set analysis for high intention to continue studying CS. Specifically, black circles (●) denote the presence of a condition, whereas crossed-out circles (⊗) indicate its absence [Fiss, 2011]. Blank spaces suggest a do not care situation, in which the causal condition may be either present or absent. Large circles indicate core elements of a configuration, and small ones indicate the peripheral elements. The solution table includes set-theoretic consistency values for each configuration as well as for the overall solution, with all values being above the threshold (>0.75). Consistency measures the degree to which a subset relationship has been approximated, whereas coverage assesses the empirical relevance of a consistent subset [Mendel and Korjani 2012; Ragin 2006]. The overall solution coverage indicates the extent to which high intentions can be determined based on the configurations, and is comparable to the R-square value reported in correlational methods [Woodside 2013]. The results show an overall solution coverage of .569, which suggests that the eight solutions accounted for a substantial proportion of the outcomes.

Table III. Configurations for high intention to continue studies in CS

	Solution							
Configuration	1	2	3	4	5	6	7	8
Gains								
Cognitive		●	●	●	●		●	●
Noncognitive	⊗			⊗	⬤	⊗	●	⬤
Motivation to Study								
Interest in CS			●		⬤	⬤	⨂	⬤
Learning Performance								
Excellent/Very Good (Grades)	⨂	⨂	⨂	⊗	●	●	⨂	
Barriers								
Personal Values	⊗	⊗	⊗	⊗	⊗	⊗		●
Teaching Quality	⨂	⨂	⨂			⬤	●	●
Satisfaction with Learning Effectiveness	⊗	⊗		⊗	⊗	⨂	●	●
Consistency	0.880	.889	.906	0.881	.817	.849	.930	.953
Raw Coverage	0.259	.307	.183	.254	.087	.074	.068	.091
Unique Coverage	0.031	.069	.018	.022	.046	.035	.029	.036
Overall Solution Consistency	0.856							
Overall Solution Coverage	0.569							

Note: Black circles (●) indicate the presence of a condition, and circles with "x" (⊗) indicate its absence. Large circles indicate core conditions, and small ones represent peripheral conditions. Blank spaces indicate "don't care."

For high intention to continue studies in CS, solutions 1–5 reflect combinations in which barriers are absent. Teaching quality and grades were core constructs, highlighting the importance of these factors (solutions 1–3). In detail, the absence of all three barriers, along with the absence of noncognitive gains and low grades, led to high intention to continue studying CS regardless of cognitive gains or motivation (solution 1). Similarly, the combination of increased cognitive gains with low grades and barriers led to the same outcome regardless of students' motivation for studying CS (solution 2). Further, the combination of cognitive gains for students with low grades who studied CS because of interest led to high intention to continue their studies in the absence of the barriers of teaching quality and satisfaction with learning effectiveness (solution 3). Next, in the absence of the barriers of personal values and satisfaction with learning effectiveness, students showed high intention to continue studying CS with either [1] high cognitive gains or low noncognitive gains with low grades, regardless of the motive for studying CS (solution 4) or (2) the combination of cognitive and noncognitive gains for students who were interested in CS and had high grades (solution 5). Solution 5 also highlights the importance of both cognitive and noncognitive gains as core factors.

Solutions 6–8 present different configurations in which barriers were present and they combined with the other examined factors. In solution 6, students' motivation for studying CS and the barriers of teaching quality and satisfaction with learning effectiveness were important [core] factors. In detail, students who chose to study CS because of high interest in the subject and who had high grades showed high intention to continue their studies even with the barrier of poor teaching quality as long as the personal values and satisfaction with learning effectiveness barriers were absent and noncognitive gains were low. Cognitive gains played a minor role in this solution. Next, solution 7 suggests that the motivation to study CS and learning performance are core factors in explaining high intention to continue studies in the area. Specifically, the combination of cognitive and noncognitive gains with the barriers of teaching quality and satisfaction with learning effectiveness increased the intent to continue among students with low grades who chose CS because of its reputation or the career opportunities, rather than interest in the subject. Finally, in solution 8, noncognitive gains and interest in CS were core constructs that combined with cognitive gains and all three barriers to explain students' continuance intention regardless of their GPAs.

The findings provide support for all three propositions. First, more than one configuration led to high intention to continue studying CS, which indicates equifinality (proposition 1). Second, the results reveal configurations of high intention to continue studying CS in which one condition could be either present or absent depending on its combination with the other conditions, indicating causal asymmetry (proposition 2). Third, when at least one barrier was present, gains or interest in CS needed to be present in order to explain high intention to continue CS studies (proposition 3).

4.2 Testing for predictive validity

This study also tested predictive validity in order to examine how well the model predicted the dependent variable in additional samples [Gigerenzer and Brighton 2009; Pappas et al. 2016a,c; Woodside 2014]. Predictive validity is important because achieving only good model fit does not necessarily mean that the model offers good predictions. Following the guidelines of previous studies [Gigerenzer and Brighton

2009; Pappas et al. 2016a,c; Woodside 2014], we divided this study's sample into a subsample and a holdout sample, and again ran the analyses for each sample. A holdout sample can be determined by removing respondents from the original sample used for the model estimation in order to increase the robustness of the results [Venkataraman, 1989], thus the same sample may be used once it is randomly divided. In predictive validity testing the overall solution consistency and coverage for the subsample (Table IV) should be similar with the ones for the whole sample (Table III), however the configurations for the subsample are not expected to be the same. Table IV shows that the patterns of complex antecedent conditions were consistent indicators (overall solution consistency was 0.877, and overall solution coverage was 0.565) of high scores for intention to continue studies in CS, using the subsample.

The results presented in Table IV then had to be tested against the second sample; that is, the holdout sample. Each configuration in Table IV represents a model that was plotted against the outcome variable (i.e., intention to continue studies in CS). Thus, for each of the models from subsample 1 (Table IV), the value of coverage and consistency should be similar with their value when testing the same models using data for the holdout sample (Fig. II). Performing that function in the fsQCA software requires modeling each configuration as a variable. To this end, we first we used the function "*fuzzynot(x)*" for each variable that was absent (~) in the configurations. This function computes the negation (1-x) of a variable (fuzzy set). Next, in order to create model 1, we used the function "*fuzzyand(x,..,)*", which takes as input all of the variables that are present in each configuration and the new variables that occurred as the outcome of the "*fuzzynot(x)*" function. The "*fuzzyand(x,…,)*" function returns a minimum of two variables [fuzzy sets]. Finally, the new variable result (model 1) was plotted against the study outcome of interest.

Table IV. Complex configurations indicating high intention to continue studies in CS for the subsample

Models from Subsample 1	Raw Coverage	Unique Coverage	Consistency
Model 1. CG*~NCG*~PV*~TQ*~SLE	0.383	0.114	0.907
Model 2. CG*~LP*~PV*~TQ*~SLE	0.306	0.071	0.877
Model 3. CG*~LP*~MS*~PV*~SLE	0.081	0.021	0.923
Model 4. ~NCG*~LP *MS*~PV*~TQ*~SLE	0.192	0.019	0.888
Model 5. CG*NCG*MS*PV*TQ*SLE	0.096	0.052	0.960
Model 6. ~CG*~NCG*LP*MS*~PV*~TQ*~SLE	0.050	0.019	0.910
Overall Solution Consistency	0.877		
Overall Solution Coverage	0.565		

CG; Cognitive gains, NCG; Noncognitive gains, MS; Motivation to study, LP; Learning performance, PV; Personal values, TQ; Teaching quality, SLE; Satisfaction with learning effectiveness

As presented in Fig. II, the findings for testing model 1's predictions with data from the holdout sample indicate high consistency (0.831) and coverage (0.350), similar to the consistency and coverage of model 1 for the subsample (Table IV). Predictive tests for all models suggest that the highly consistent models for the subsample had high predictive abilities for the holdout sample and vice versa. All results are available upon request.

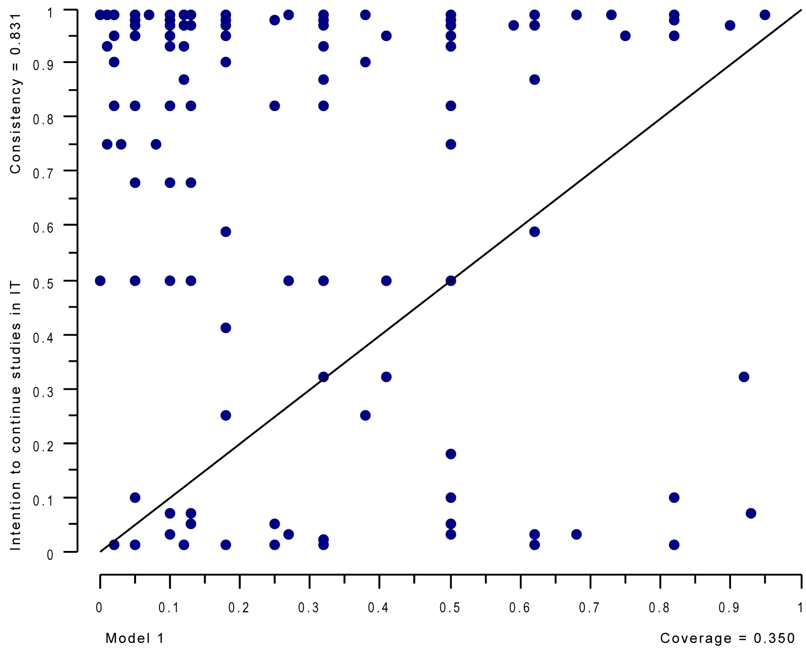

Fig. II. Testing model 1 on the subsample using data from the holdout sample

5. DISCUSSION, IMPLICATIONS AND CONCLUSION

This study proposes that in CS education, students' gains, barriers, motivation and learning performance combine to form configurations for predicting their intention to continue studies in CS. To this end, we constructed a conceptual model that served as the basis for identifying the aforementioned configurations. Of particular interest in the findings was the role of barriers in students' behavior. In fact, the absence of barriers led to high intention to continue studies in CS with either the presence or absence of gains, motivation, or learning performance (solutions 1–5). This means that when students feel that they belong in CS, are fulfilled by a career in CS; social norms and the sense of belongingness (e.g., have friends studying CS) increase students' intention to continue their studies in the field. This is one of the most important barriers in CS students, especially in females [Sankar et al., 2015]. Similarly, if students are satisfied with the quality or the effectiveness of their courses, they have higher intentions to study CS. Consequently, it is critical to reduce the aforementioned barriers for CS students in order to significantly increase the chances that they stay in the field.

Interestingly, students with poor learning performance and low noncognitive gains showed high intention to continue studying CS as long as they also experienced low barriers (solution 1). This solution highlights the importance of minimizing the identified barriers, which seems to be the main reason for keeping (or not) students in CS. On solutions 2-5, there are other factors that can explain students' intentions, without reducing the important role of barriers, instead, they are able to explain different types of students. In particular, students that expect to develop critical thinking and increase their problem solving skills have higher intentions to continue their studies in CS (solutions 2,4). In addition, students' high intentions may be also influenced by their interest in studying CS, or their expectations about developing ethical standards and increasing their engagement with the community (solutions 3,5).

In contrast, students that experienced more barriers in CS needed to have at least cognitive or noncognitive gains from CS studies or to be highly motivated to study CS, rather than being attracted by career perspectives or reputation (solutions 6–8). Thus, when students feel that the quality of the courses is low or the classes are too big and boring; they need to be very motivated, high performers and interested in the subject for high intentions to exist (solution 6). It is interesting to note that when the barrier of teaching quality was present as a core (i.e., important) factor, interest in CS also had to be present as a core factor for high intention to occur (solution 6). This suggests that students who are already interested in the subject might be able to overcome problems with poor teaching quality and that this is why the students in this study had high learning performance.

Next, when the low teaching quality is combined with students' low performance, unhappiness with their grades, low interest in CS, and high workload, students' intention can be raised with focusing in critical thinking, increasing their problem solving skills, developing ethical standards and ultimately raising their engagement with the community (solution 7). Hence, community building and support in CS education, can allow students to overcome typical problems, like high workload and low grades. The benefits of a vibrant learning community in CS students have been documented in previous studies [Angelaina & Jimoyannis, 2012], however their capacity to help students to overcome other barriers is new.

For a small part of the sample, the absence of belongingness, social support and personal values (i.e., all three barriers are present), can be overcomed and result positive intentions to CS studies if students are interested in the subject. From the three barriers examined in this study, personal values and satisfaction with learning effectiveness appeared (present or absent) in seven of eight solutions, followed by teaching quality, which appeared in six solutions. Personal values were the most important barrier because it was absent in six of eight solutions, followed by satisfaction with learning effectiveness. This means that students' sense of belongingness in CS, feeling fulfilled by their CS studies, and being among friends in the class is very important for their future decisions towards studying CS. Maloney [2007] has indicated the importance of personal interaction and connectedness among the students; as well as the construction of relationships around shared interests in order to improve learning. Our study goes one step further by verifying how these elements can help students not only to improve their learning but also continue their studies in the area of CS.

Finally, the findings indicate that none of the examined factors is a sole indicator of students' intention neither to dropout nor to continue their studies in CS, instead different types of students exist with different characteristics that should be addressed differently in order to reduce dropout rates. For example, when adopting a typical performance measure (i.e., GPA), the findings show that even students with low grades will continue studying CS, based on how the examined factors combine with each other, while one would expect that low performers might not be interested in the subject or find it too difficult, thus leading to higher chances to dropout. However, the results show that in most of the cases even if students do not have a very good GPA, they are not unhappy with the teaching quality or learning effectiveness. This suggests that students acknowledge the difficulty in CS, and at the same time they may be intrigued by it since their intentions to continue in the field remain high.

5.1 Implications

These findings contribute to the literature in a number of ways. The study adds to CS education literature by presenting conditions for explaining students' intention to continue their studies in CS. Previous studies explain students' CS-related behavior by analyzing the effects of various antecedents [e.g., Giannakos et al., 2016; Rosson et al. 2011;]. However, these studies mostly use multiple regression analysis and focus on the main effects of these antecedents on one or more dependent variables while neglecting the interdependencies and interconnected causal structures between the variables [Woodside 2014]. This research study builds on complexity theory, takes a configurational approach toward CS education and contributes to the literature by explaining how four sets of causal conditions, that is, gains, barriers, motivation, and learning performance, combine to form configurations that predict students' behavior. In addition, this method offers a better understanding regarding the specific patterns of the gains, barriers, motivation and learning performance that lead to increased intention to continue studies in CS. It also adds to the literature by providing specific conditions under which barriers, gains and motivation coexist. The results suggest that on certain occasions, barriers can be diminished if students are motivated or have cognitive and noncognitive gains.

Regarding its methodology, this paper is one of the first to perform configurational analysis based on individual-level data from CS students. The implementation of complexity theory in individual phenomena may be appropriate for theory building. Hence, this paper makes propositions based on complexity theory in order to explain students' intention to continue studies in CS. Further, the paper tests these propositions using fsQCA, a commonly used approach that is receiving increased attention in recent studies [Leischnig and Kasper-Brauer 2015; Ordanini et al. 2014; Woodside 2014]. The study confirms the importance of examining complex causal patterns of predictors, contrarian cases and asymmetric relationships between antecedents and outcomes. As mentioned above, the aim of fsQCA is to identify combinations of factors that are able to explain a specific outcome. Hence, multiple combinations of independent factors can potentially explain the same outcome. In addition, because the methodology examined combinatorial effects, the influence of every independent factor on the outcome was not quantified [Liu et al. 2015; Woodside 2013].

This study offers useful insights for instructors, administrators and CS policy makers because it helps them to identify important factors for CS students and explains how these factors may combine to better predict students' intention to continue studying CS. They should be aware of their students' different types, in order to address them based on their expectations, motivations and feelings towards studying CS. The results of this study may pave the ground for practitioners to redesign CS courses, by paying particular attention in building learning communities in their courses, focusing in different students, especially the ones who are not self-motivated and self-interested with the subject, this will allow CS students to overcome certain barriers and persist with the major in CS and even continue with a master degree. In addition, developing critical thinking and increasing problem solving skills as early as possible is another strategy in order to reduce dropout rates, especialy during the first year.

Another contribution of this study is the identification of how different expectations and motivations of students combine with each other and explain behavioral intentions; hence there are multiple student types in CS education with different needs. The patterns that create these student types give to practitioners a roadmap on how to address more effectively their students in CS classes, since not the same factors are equally important for all students. With the analysis of our empirical data, it is clear that we should strive CS education into more personalized manner (e.g., intensifying project based learning and adopting adaptive learning systems). Furthermore, the findings suggest that high-quality teaching has a crucial influence on whether students persist in CS studies. Teaching quality heavily relies in students' perceptions and their expectations ; however, CS students' expectations are in many times much higher compared to other STEM disciplines [Schmitt et al., 2013], with significant in many times differences between their teachers' expectations [Utting et al., 2013]. Hence, instructors need to provide students with the required awareness and self-reflection mechanisms during their coursework [Trætteberg et al., 2016].

Finally, the study offers empirical evidence to support the assumption that even with high barriers such as dissatisfaction with learning effectiveness, students are likely to continue studying CS as long as cognitive and noncognitive gains are visible/high. In such cases, students may not feel that they belong in CS or feel that the course is too difficult, but may still continue studying CS if, for example, they expect to develop critical thinking and to increase their problem solving. Nonetheless, the results suggest that there is still further work to be done in order to understand and better explain students' behavior in CS.

5.2 Limitations and future research

As mentioned above, there are some limitations in this study. First, the respondents were Norwegian CS students who had experienced the Norwegian higher education system, and thus generalization of the results may be limited. Nonetheless, the Norwegian CS study programs conform to the international standards defined by the ACM's Computer Science 2013 curriculum guidelines for undergraduate programs in computer science [ACM/IEEE-CS Joint Task Force 2013]. Also, one of the MSc programs is international, thus a very small part of the sample (i.e., 10 students), may be international students, which however is not enough to influence significantly the findings of this study. It would be interesting for future studies to examine the difference between such samples. In addition, Norway has adhered to the objectives of the Bologna process in the European Higher Education Area. National generic learning outcomes descriptions' levels for the bachelor's, master's and doctoral degrees were defined by the instructions on the Norwegian Qualifications Framework for Higher Education in accordance with the European Qualifications Framework. Last but not least, yearly survey studies[3] from the Norwegian Agency for Quality Assurance in Education indicate that Norwegian students experience teaching quality and performance similarly with other countries [Wiebe, 2003].

Another limitation is that these analyses are based on one set of data collected from a single survey study; this places significant limitations on how strongly we can interpret and generalize the reported patterns. Future research is needed not only to replicate and verify the patterns we have reported, but also to determine whether these results characterize students who have left their studies [dropouts] or others who have graduated and now work in the industry (e.g., via an alumni survey). This study is distinct from the majority of previous work on CS education, which has focused on multiple regression analysis [e.g., Giannakos et al., 2016], and confirms the importance of complexity theory and configurational analysis. However, more studies are needed in a variety of contexts to enhance the usefulness of the current findings.

Finally, the present study is missing some qualitative insights. Although a qualitative analysis has various benefits, we have chosen in this study a quantitative approach in order to be able to perform fsQCA, due to its advantages over the

[3] Yearly survey studies about CS education under the Information and computer technology category: http://www.studiebarometeret.no/en/

traditional statistical analyses. Also based on the limitations of the qualitative approach, we chose the quantitative approach because we wanted to gather data from a relatively large number of students, to identify the effects between the variables, and to be able to generalize the findings to a broader population. Future studies should take also a qualitative approach in order to complement and extend our findings.

5.3 Conclusion

This research study examined combinations of students' gains, motivation, learning performance, and barriers in order to explain and predict their intention to continue studying CS. The students ranked the latter as being significant predictors of their decision to continue studying the subject. Towards this aim, we employed complexity theory and highlighted the importance of analyzing complex patterns of predictors, contrarian cases and asymmetric relationships. Cognitive and noncognitive gains, motivation to study the subject, learning performance, and the barriers of personal values, teaching quality, and satisfaction with learning effectiveness do not all have to combine to stimulate intention to continue studies in CS. Complex but parsimonious patterns occurred in which the various antecedents could be present or absent, suggesting that different factors may combine to explain CS students' behavior.

ACKNOWLEDGEMENTS

The authors would like to thank all the students at the Department of Computer and Information Science of NTNU that took part and responded in this study. This work was carried out during the tenure of an ERCIM "Alain Bensoussan" Fellowship Programme. This work was funded by the Norwegian Research Council under the projects FUTURE LEARNING (number: 255129/H20).

APPENDIX A

Construct and Scale Items	Mean	SD	Loading
Please rate the following qualities based on your personal experience while studying CS:			
Gains in Cognitive Learning and Development			
1. Acquiring a broad general education*	5.38	1.07	0.525
2. Acquiring job or work-related knowledge and skills*	5.08	1.23	0.522
3. Thinking critically and analytically	5.23	1.22	0.758
4. Analyzing quantitative problems	5.03	1.30	0.651
5. Solving complex real-world problems	4.83	1.29	0.670
Gains in Noncognitive Learning and Development			
1. Working effectively with others*	4.76	1.32	0.521
2. Developing a personal code of values and ethics	4.32	1.45	0.814
3. Developing a deepened sense of spirituality	2.96	1.73	0.693
Personal Values			
1. I do not feel as if I belonged in CS	2.41	1.80	0.847
2. A non-computer science career would be more fulfilling to me	2.48	2.48	0.859
3. Classes were unfriendly	2.39	2.39	0.639
4. Few of my friends are studying CS	2.62	1.80	0.657
Teaching Quality			
1. Poor teaching by CS faculty or teaching assistants	3.65	1.78	0.836
2. Classes were boring	3.80	1.67	0.792
3. The classes are too big	2.89	1.64	0.606
Satisfaction with Learning Effectiveness			
1. I am unhappy with my grades	3.27	1.75	0.676
2. Excessive workload	3.45	1.72	0.819
3. Overall curriculum was too difficult or too lengthy	3.06	1.72	0.918
Intention to continue your studies in CS			
1. I plan to study in CS in the future	5.23	2.11	0.906
2. I intend to continue my studies in CS in the future	5.45	2.02	0.971
3. If I have to select where to study in the future, I will choose CS.	5.41	1.73	0.700
4. I expect to continue my studies in CS in the future	5.38	2.07	0.972
*Item deleted due to low loading			

APPENDIX B

		Intention to Continue Studies in CS				
		1	2	3	4	5
CG [phi^2 =.07, NS]	1	*18 [5.4%]*	*19 [5.7%]*	11 [3.3%]	**5 [1.5%]**	**24 [7.1%]**
	2	*18 [5.4%]*	*8 [2.4%]*	10 [3%]	**7 [2.1%]**	**15 [4.5%]**
	3	11 [3.3%]	24 [7.1%]	20 [6%]	5 [1.5%]	21 [6.3%]
	4	**5 [1.5%]**	**6 [1.8%]**	7 [2.1%]	*3 [0.9%]*	*16 [4.8%]*
	5	**14 [4.2%]**	**13 [3.9%]**	17 [5.1%]	*8 [2.4%]*	*31 [9.2%]*
PV [phi^2 =.14, p < .01]	1	**19 [5.7%]**	**6 [1.8%]**	6 [1.8%]	*4 [1.2%]*	*30 [9%]*
	2	**6 [1.8%]**	**10 [3%]**	9 [2.7%]	*7 [2.1%]*	*28 [8.4%]*
	3	14 [4.2%]	11 [3.3%]	19 [5.7%]	4 [1.2%]	21 [6.3%]
	4	*12 [3.6%]*	*16 [4.8%]*	16 [4.8%]	**8 [2.4%]**	**18 [5.4%]**
	5	*14 [4.2%]*	*27 [8.1%]*	15 [4.5%]	**5 [1.5%]**	**10 [3%]**
SLE [phi^2 =.11, p < .01]	1	**29 [8.7%]**	**9 [2.7%]**	8 [2.4%]	*7 [2.1%]*	*27 [8.1%]*
	2	**6 [1.8%]**	**11 [3.3%]**	13 [3.9%]	*5 [1.5%]*	*27 [8.1%]*
	3	9 [2.7%]	14 [4.2%]	19 [5.7%]	8 [2.4%]	18 [5.4%]
	4	*10 [3%]*	*17 [5.1%]*	9 [2.7%]	**3 [0.9%]**	**18 [5.4%]**
	5	*11 [3.3%]*	*19 [5.7%]*	16 [4.8%]	**5 [1.5%]**	**17 [5.1%]**

		Intention to Continue Studies in CS				
		1	2	3	4	5
NCG [phi^2 =.08, N.S.]	1	*10 [3%]*	*8 [2.4%]*	11 [3.3%]	**7 [2.1%]**	**29 [8.6%]**
	2	*23 [6.8%]*	*13 [3.9%]*	16 [4.8%]	**9 [2.7%]**	**20 [6%]**
	3	19 [5.7%]	19 [5.7%]	13 [3.9%]	4 [1.2%]	27 [8%]
	4	**4 [1.2%]**	**8 [2.4%]**	4 [1.2%]	*2 [0.6%]*	*10 [3%]*
	5	**10 [3%]**	**22 [6.5%]**	21 [6.3%]	*6 [1.8%]*	*21 [6.3%]*
TQ [phi^2 =.1, p < .01]	1	**15 [4.5%]**	**10 [3%]**	8 [2.4%]	*10 [3%]*	*30 [9%]*
	2	**11 [3.3%]**	**10 [3%]**	12 [3.6%]	*5 [1.5%]*	*33 [9.9%]*
	3	7 [2.1%]	14 [4.2%]	9 [2.7%]	4 [1.2%]	17 [5.1%]
	4	*17 [5.1%]*	*22 [6.6%]*	21 [6.3%]	**5 [1.5%]**	**18 [5.4%]**
	5	*15 [4.5%]*	*14 [4.2%]*	15 [4.5%]	**4 [1.2%]**	**9 [2.7%]**

Items in **bold** represent contrarian cases. Items in *italics* represent main effect.

The sets of contrarian cases are counter to the main effect size [phi^2 range from .01 to .14].

CG: Cognitive gains, NCG: Noncognitive gains, PV: Personal Values, TQ: Teaching Quality, SLE: Satisfaction with Learning Effectiveness

REFERENCES

ACM/IEEE-CS Joint Task Force 2013. Computer science curricula 2013: Curriculum guidelines for undergraduate degree programs in computer science. Technical report. Association for Computing Machinery/IEEE Computer Society. DOI: 10.1145/2534860

Angelaina, S., & Jimoyiannis, A. 2012. Analysing students' engagement and learning presence in an educational blog community. *Educational Media International*, 49, 3, 183-200.

Araque, F., Roldan, C., & Salguero, A. 2009. Factors influencing university dropout rates. *Computers and Education*, 53, 3, 563–574.

Barker, L., Hovey, C. L., & Thompson, L. D. 2014. Results of a large-scale, multi-institutional study of undergraduate retention in computing. In Frontiers in Education Conference, FIE, 2014 IEEE pp. 1–8. IEEE.

Barker, L. J., McDowell, C., & Kalahar, K. 2009. Exploring factors that influence computer science introductory course students to persist in the major, SIGCSE Bull., 41, 1, 153–157.

Biggers, M., Brauer, A., & Yilmaz, T. 2008. Student perceptions of computer science: A retention study comparing graduating seniors with CS leavers. ACM SIGCSE Bulletin, 40, 1, pp. 402–406.

Bjorklund, S. A., Parente, J. M., & Sathianathan, D. 2004. Effects of faculty interaction and feedback on gains in student skills. *Journal of Engineering Education*, 93, 2, 153–160.

Blickenstaff, J. C. 2005. Women and science careers: Leaky pipeline or gender filter? *Gender and Education*, 17, 4, 369–386

Carter, L. 2006. Why students with an apparent aptitude for computer science don't choose to major in computer science. *ACM SIGCSE Bulletin*, 38, 1, 27–31.

Campbell, D. T., & Stanley, J. C. 1966. *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

Chow, A., Eccles, J. S., & Salmela-Aro, K. 2012. Task value profiles across subjects and aspirations to physical and IT-related sciences in the United States and Finland. *Developmental Psychology*, 48, 6, 1612.

Cohoon, J. 2006. Just get over it or get on with it: Retaining women in undergraduate computing. *Women and Information Technology*, 205–238.

Fiss, P. C. 2011. Building better causal theories: A fuzzy set approach to typologies in organization research. *Academy of Management Journal*, 54, 2, 393–420.

Fitzsimons, G. J. 2008. Death to dichotomizing. *Journal of Consumer Research*, 35, 1, 5–8.

Fornell, C., & Larcker, D.F. 1981. Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18, 1, 39–50.

Giannakos, M. N., Pappas, I. O., Jaccheri, L., & Sampson, D. G. 2016. Understanding student retention in computer science education: The role of environment, gains, barriers and usefulness. *Education and Information Technologies*, 1-18. doi:10.1007/s10639-016-9538-1

Gigerenzer, G., & Brighton, H. 2009. Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science,* 1, 1, 107–143.

Glaesser, J., & Cooper, B. 2012. Gender, parental education, and ability: Their interacting roles in predicting GCSE success. *Cambridge Journal of Education*, 42, 4, 463–480.

Grover, S., Pea, R., & Cooper, S. 2014. Remedying misperceptions of computer science among middle school students. In Proceedings of the 45th ACM technical symposium on Computer science education pp. 343-348. ACM.

Hein, G. L., Bunker, K. J., Onder, N., Rebb, R. R., Brown, L. E., & Bohmann, L. J. 2012. University studies of student persistence in engineering and computer science. Washington, DC: American Society for Engineering Education.

Hewner, M. 2013. Undergraduate conceptions of the field of computer science. In Proceedings of the ninth annual international ACM conference on International computing education research pp. 107-114. ACM

Ireland Department of Education and Skills 2014. ICT skills action plan 2014–2018: Government, education and industry working together to make Ireland a global leader in ICT talent. Resource document. http://cork.etb.ie/wp-content/uploads/sites/20/2014/08/14042014-ICT_Skills_Action_Plan-Publication.pdf. Accessed 19 February.

Jacobs, J. E. 2005. Twenty-five years of research on gender and ethnic differences in math and science career choices: What have we learned? *New Directions for Child and Adolescent Development*, 110, 85–94.

Joo, Y. J., Lim, K. Y., & Kim, J. (2013). Locus of control, self-efficacy, and task value as predictors of learning outcome in an online university context. Computers & Education, 62, 149-158.

Lee, S. W. Y., & Tsai, C. C. 2011. Students' perceptions of collaboration, self-regulated learning, and information seeking in the context of Internet-based learning and traditional learning. *Computers in Human Behavior*, 27, 2, 905–914.

Leischnig, A., & Kasper-Brauer, K. 2015. Employee adaptive behavior in service enactments. *Journal of Business Research*, 68, 2, 273–280.

Li, Q., Swaminathan, H., & Tang, J. 2009 Development of a classification system for engineering student characteristics affecting college enrollment and retention. *Journal of Engineering Education*, 98, 4, 361–376.

Liu, Y., Mezei, J., Kostakos, V., & Li, H. 2015. Applying configurational analysis to IS behavioural research:

A methodological alternative for modelling combinatorial complexities. *Information Systems Journal*. doi:10.1111/isj.12094.

Maloney, E. J. 2007. What Web 2.0 Can Teach Us about Learning. *Chronicle of Higher Education*, 53, 18, B26

Marra, R. M., Rodgers, K. A., Shen, D., & Bogue, B. 2012. Leaving engineering: A multi-year single institution study. *Journal of Engineering Edu*cation, 101, 1, 6–27.

Mau, W. C. 2003. Factors that influence persistence in science and engineering career aspirations. *Career Development Quarterly*, 51, 3, 234–243.

McGill, M. M., Decker, A., & Settle, A. (2016). Undergraduate Students' Perceptions of the Impact of Pre-College Computing Activities on Choices of Major. *ACM Transactions on Computing Education (TOCE)*, *16*(4), 15.

Mendel, J. M., & Korjani, M. M. 2012. Charles Ragin's fuzzy set qualitative comparative analysis (fsQCA) used for linguistic summarizations. *Information Sciences*, 202, 1–23.

Morton, E. 2005. Beyond the barriers: What women want in IT. Resource document. TechRepublic. http://www.techrepublic.com/article/beyond-the-barriers-what-women-want-in-it/.

O'Brien, R. M. 2007. A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41, 5, 673–690.

Ordanini, A., Parasuraman, A., & Rubera, G. 2014. When the recipe is more important than the ingredients: A qualitative comparative analysis (QCA) of service innovation configurations. *Journal of Service Research*, 17, 2, 134-149.

Pappas, I. O., Kourouthanassis, P. E., Giannakos, M. N., & Chrissikopoulos, V. 2016a. Explaining online shopping behavior with fsQCA: The role of cognitive and affective perceptions. *Journal of Business Research*, 69, 2, 794-803.

Pappas, I. O., Kourouthanassis, P. E., Giannakos, M. N., & Lekakos, G. 2017. The interplay of online shopping motivations and experiential factors on personalized e-commerce: A complexity theory approach. *Telematics and Informatics*. doi: 10.1016/j.tele.2016.08.021

Pappas, I., Giannakos, M., & Jaccheri, L. 2016b. Investigating Factors Influencing Students' Intention to Dropout Computer Science Studies. In *Proceedings of the 2016 ACM annual conference on Innovation and technology in computer science education (ITiCSE '16), ACM Press*.

Pappas, I. O., Giannakos, M. N., & Sampson, D. 2016c. Making Sense of Learning Analytics with a Configurational Approach. In *Proceedings of the workshop on Smart Environments and Analytics in Video-Based Learning (SE@ VBL), LAK2016*.

Pike, G. R., Kuh, G. D., McCormick, A. C., Ethington, C. A., & Smart, J. C. 2011. If and when money matters: The relationships among educational expenditures, student engagement and students' learning outcomes. *Research in Higher Education*, 52, 1, 81–106.

Pirker, J., Riffnaller-Schiefer, M., & Gütl, C. 2014. Motivational active learning: Engaging university students in computer science education. In Proceedings of the 2014 Conference on Innovation & Technology in Computer Science Education, 297–302.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. 2003. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88, 5, 879.

Ragin, C. C. 2000. Fuzzy-set social science. University of Chicago Press.

Ragin, C. C. 2006. Set relations in social research: Evaluating their consistency and coverage. *Political Analysis*, 14, 3, 291–310.

Ragin, C. C. 2008. Redesigning social inquiry: Fuzzy sets and beyond. Chicago, IL: University of Chicago Press.

Ragin, C.C, and Davey, S. 2014. fs/QCA Computer Programme, Version 2.5. Irvine, CA: University of California.

Rosson, M. B., Carroll, J. M., & Sinha, H. 2011. Orientation of undergraduates toward careers in the computer and information sciences: Gender, self-efficacy and social support. *ACM Transactions on Computing Education (TOCE)*, 11, 3, 14.

Sankar, P., Gilmartin, J., & Sobel, M. 2015. An examination of belongingness and confidence among female computer science students. *ACM SIGCAS Computers and Society*, 45, 2, 7-10.

Schmitt, K. R., Badawy, A. H. A., Kramer, S., Hrapczynski, K., Larsen, E., Andrew, A., ... & Miller, M. 2013. Student expectations from CS and other stem courses: they aren't like CS-majors! or (CS!= Stem-CS). *Journal of Computing Sciences in Colleges*, 28, 6, 100-108.

Seymour, E. and Hewitt, N. 1997. Talking about leaving: Why undergraduates leave the sciences. Boulder, CO: Westview Press.

Shadish, W. R. 2010. Campbell and Rubin: A primer and comparison of their approaches to causal inference in field settings. *Psychological Methods*,15, 1, 3.

Suresh, R. 2006. The relationship between barrier courses and persistence in engineering. *Journal of College Student Retention*, 8, 2, p. 215–239.

Toutkoushian, R. K., & Smart, J. C. 2001. Do institutional characteristics affect student gains from college? *Review of Higher Education*, 25, 39–61.

Trætteberg, H., Mavroudi, A., Giannakos, M., & Krogstie, J. 2016. Adaptable Learning and Learning

Analytics: A Case Study in a Programming Course. In *European Conference on Technology Enhanced Learning* (pp. 665-668). Springer International Publishing.

US Bureau of Labor Statistics 2014. Employment Projections 2010–2020. Available at http://www.bls.gov/emp/

Venkatraman, N. (1989). The concept of fit in strategy research: Toward verbal and statistical correspondence. *Academy of management review*, 14, 3, 423-444.

Utting, I., Tew, A. E., McCracken, M., Thomas, L., Bouvier, D., Frye, R., Paterson, J., Caspersen, M., Kolikant, Y.B.D., Sorva, J., & Wilusz, T. 2013. A fresh look at novice programmers' performance and their teachers' expectations. In *Proceedings of the ITiCSE working group reports conference on Innovation and technology in computer science education-working group reports* (pp. 15-32). ACM.

Vis, B. 2012. The comparative advantages of fsQCA and regression analysis for moderately large-N analyses. *Sociological Methods & Research*, 41, 1, 168–198.

Walden, S. E., & Foor, C. 2008. What's to keep you from dropping out? Student immigration into and within engineering. *Journal of Engineering Education*, 97, 2, 191–205.

Watt, H. M., Shapka, J. D., Morris, Z. A., Durik, A. M., Keating, D. P., & Eccles, J. S. 2012. Gendered motivational processes affecting high school mathematics participation, educational aspirations, and career plans: A comparison of samples from Australia, Canada, and the United States. Developmental Psychology, 48, 6, 1594.

Wiebe, E., Williams, L., Yang, K., & Miller, C. 2003. Computer science attitude survey. Technical Report. North Carolina State University at Raleigh, Raleigh, NC, USA

Woodside, A. G. 2013. Moving beyond multiple regression analysis to algorithms: Calling for adoption of a paradigm shift from symmetric to asymmetric thinking in data analysis and crafting theory. *Journal of Business Research*, 66, 4, 463–472.

Woodside, A. G. 2014. Embrace• perform• model: Complexity theory, contrarian case analysis, and multiple realities. *Journal of Business Research*, 67, 12, 2495–2503.

Xenos, M., Pierrakeas, C., & Pintelas, P. 2002. A survey on student dropout rates and dropout causes concerning the students in the Course of Informatics of the Hellenic Open University. *Computers & Education*, 39, 4, 361–377.

Yu, T., & Jo, I. H. 2014. Educational technology approach toward learning analytics: Relationship between student online behavior and learning performance in higher education. In Proceedings of the Fourth International Conference on Learning Analytics and Knowledge pp. 269–270. ACM.

Zadeh, L. A. 1965. Fuzzy sets. Information and control, 8, 3, 338-353.

Zweben, S. 2014. Computing degrees and enrollment trends: From the 2012–2014 CRA Taulbee Survey. Washington, DC: Computing Research Association.