Introduction

Despite the fact that women constitute about half (46.6%) of the United States workforce, the proportion of women in computer science and information technology (CSIT) careers has been declining. From a high of 40% in 1986, it dropped to 29% in 1996, and further declined to about 20% in 2004 according to the Information Technology Association of America (Ramsey and McCorduck, 2005). Although outsourcing and economic downturn have been initial reasons for the general decline in CSIT jobs within the U.S., a report by the Bureau of Labor Statistics show that the overall employment of computer network, systems, and database administrators is projected to increase by 30 percent from 2008 to 2018, much faster than the average for all occupations. In addition, this occupation will add 286,600 new jobs over that period (BLS, 2011). While women are realizing parity in other career paths, the continuing decline of women in CSIT professions is troubling and a cause for concern. Studies have shown that diverse teams, including gender diversity, that are well managed tend to produce better solutions than teams that are homogeneous in nature (Johnson, 1998). Hence, efforts to understand the factors that affect this disparity in the computing field is important so appropriate intervention mechanisms can be put in place to improve the representation of women in CSIT.

Education in a relevant discipline is generally a strong indicator of the respective career path a person may choose. The enrollment of women in computing education has also been steadily decreasing in the recent decades. Consider that in 1983, 36.8% of bachelor's degrees in computer science went to women, roughly matching the percentage of bachelor's degrees awarded to women across all science and engineering (S&E) fields (38.8%). However, since that time the percentage of S&E degrees awarded to women has steadily grown, reaching 50.2% in 2007, while the percentage of computing degrees awarded to women has steadily decreased, to 18.6% in 2007 (National Science Foundation, 2002; National Science Foundation, 2010). In fact, every other science and engineering field for which NSF reports degrees awarded since 1983 shows an increase in percentage of degrees awarded to women, while the fraction awarded in computing is less than half of what it was in 1983. The National Science Foundation statistics just described are specific to computer science, but similar trends are seen in other computing fields, such as Information Systems (IS) and Information Technology (IT) (Davis et al. 2008). Given the opposite direction of trends in computing as compared to other S&E fields, there are clearly gender issues that are specific to computing which deserve study and understanding (Ahuja 2002). Because the enrollment of women in CSIT education is decreasing, it is no surprise that women are underrepresented in CSIT.
Given that CSIT related job prospects are growing significantly according to Bureau of Labor Statistics, it is astounding to discover that fewer women are participating in this opportunity. The number of women who majored in computing related areas declined by 80% in the last decade, which ultimately represents a 93% decrease since its peak in 1982. These staggering drop in numbers suggest that girls are not identifying with computing fields and professions, and the proposed study aims to gain insight into both broad identity issues related to computing fields, as well as more fine-grained issues that consider differences in computing fields (e.g., computer science versus information systems) and gender differences in how these fields are perceived and identified with. Thus, broadening the role of women in CSIT areas requires a systematic and close examination of factors that shape a student’s CSIT identity as occupational identity formation is shown to inform equity in education, learning, and career choice (Carlone and Johnson, 2007; Hazari et al., 2010). Identity-based studies in science and engineering show how identity-related measures affect a student’s education and career persistence. While there are studies that systematically investigate different aspects of the gender gap in IT, there is a lack of IS research that uses the identity lens to explain the gender disparity in CSIT.

This study uses an integrative approach to examine gender differences within the central notion of CSIT identity formation by drawing upon social cognitive career theory (SCCT), theory of planned behavior (TPB), and the theory of identity. Using both quantitative and qualitative data from undergraduate students, we will address the following research questions:

(1) What personal and environmental factors predict student CSIT identity formation?

(2) How does student CSIT identity formation mediate personal and environmental factors to influence the pursuit of CSIT education and careers?

(3) Does the interaction between student’s CSIT identity formation and personal and environmental factors differ across gender? And how?

We hope that the theoretical model we develop for CSIT identity will help shed light on the factors that positively impact retention in computer science and information technology. In addition, we will examine if the factors differ between students who are retained and those who are not? External factors such as societal norms, career aspects' impact, if any, will also be explored.

To test the model, we have developed an instrument which has been pilot tested with data from students in CSIT and not CSIT areas. We present the details of instrument validation, potential implications and plans for further research.

**Foundation Work and Development of Research Model**

The integrative research model (see Figure 1) for the study is based on the theory of planned behavior (TPB) (Ajzen, 1991), social cognitive career theory (SCCT) (Bandura, 1986; Lent et al., 2001) and the theory of identity (Carlone and Johnson, 2007; Hazari et al., 2010). The constructs used in the study, their definitions and key sources from CSIT and non-CSIT research are summarized in Table 1.

SCCT is based on social cognitive theory (SCT) and seeks to explain how environmental and cognitive factors affect a student’s intended and actual academic persistence (Lent et al., 2001). SCT (Bandura, 1986) is a widely accepted and empirically validated model to predict and explain individual behavior. It describes a “triadic reciprocality” relationship among environmental influences, cognitive factors, and individual behavior. In particular, individual behavior is affected by cognitive factors such as self-efficacy and outcome expectations and environmental influences such as social pressures or unique situational characteristics, and individual behavior in turn affects these environmental and cognitive factors. Akbulut and Looney (2007) and Lent et al. (2008) found SCCT to be useful in predicting student intentions to choosing a computing major.

Self-efficacy is an influential cognitive factor to individual behavior. Bandura (1986) defines self-efficacy as “People’s judgments of their capability to organize and execute courses of action required to attain designated types of performances.” It relates to an individual’s beliefs or expectations about one’s ability
to successfully perform particular behaviors or courses of actions. This set of beliefs influence individual choices of behaviors to undertake as well as the corresponding effort to exert. Compeau and Higgins (1995) developed a measure of computer self-efficacy (i.e., an individual perception of his or her capability to use a computer to accomplish a job task).

Another important cognitive force guiding individual behavior is outcome expectations. Outcome expectation reflects an individual’s beliefs or expectations about future outcomes. Individuals are more likely to undertake behavior that they believe leads to favorable outcomes. Self-efficacy has been shown to positively impact outcome expectation (Compeau and Higgins, 1995; Lent et al., 2001). Thus, we argue that the following hypothesis can be made:

H1: A student’s CSIT self-efficacy positively impacts their CSIT outcome expectation.

Table 1: Constructs and Definitions as adapted to current study and Key Sources.

<table>
<thead>
<tr>
<th>Constructs used in our study</th>
<th>Definition adopted for the study</th>
<th>Theoretical Source</th>
<th>Key Sources adopted from</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attitudes toward pursuing IT</td>
<td>A student’s predisposition towards CSIT</td>
<td>Theory of planned behavior (TPB)</td>
<td>Ajzen (1991); Heinze and Hu (2009); Engelman, et al. (2011)</td>
</tr>
<tr>
<td>Social norm</td>
<td>A student’s perceptions of social pressure to perform or not perform a behavior</td>
<td>Theory of planned behavior (TPB)</td>
<td>Heinze and Hu (2009); Joshi and Kuhn (2001)</td>
</tr>
<tr>
<td>CSIT Perceived behavioral control</td>
<td>A student’s perception of the difficulty in pursuing IT career</td>
<td>Theory of planned behavior (TPB)</td>
<td>Heinze and Hu (2009)</td>
</tr>
<tr>
<td>CSIT Self-efficacy</td>
<td>A student’s perception of his or her capability to use a computer to accomplish a job task</td>
<td>Social Cognitive Career Theory (SCCT)</td>
<td>Compeau and Higgins (1995); Heinze and Hu (2009)</td>
</tr>
<tr>
<td>CSIT Outcome expectation</td>
<td>How satisfied a student thinks he or she would be in an IT career</td>
<td>Social Cognitive Career Theory (SCCT)</td>
<td>Heinze and Hu (2009); Lent et al. (1993)</td>
</tr>
<tr>
<td>Social support</td>
<td>Social support and encouragement, financial resources, instrumental assistance, and access to role models or mentors</td>
<td>Social Cognitive Career Theory (SCCT)</td>
<td>Lent et al. (2001)</td>
</tr>
<tr>
<td>Social barriers</td>
<td>Negative social or family influences, financial constraints, instructional barriers, and gender and race discrimination</td>
<td>Social Cognitive Career Theory (SCCT)</td>
<td>Lent et al. (2001)</td>
</tr>
<tr>
<td>IT job availability</td>
<td>A student’s expectation concern job availability</td>
<td></td>
<td>Heinze and Hu (2009)</td>
</tr>
<tr>
<td>CSIT identity</td>
<td>A student’s perception of himself or herself as a CSIT person.</td>
<td>Identity theory</td>
<td>Carlone and Johnson (2007); Hazari et al. (2010) [non-CSIT studies]</td>
</tr>
<tr>
<td>Intention to pursue CSIT education</td>
<td>A student’s intention to pursue CSIT education</td>
<td>Theory of planned behavior (TPB)</td>
<td>Ajzen (1991), Pedroni et al. (2010)</td>
</tr>
<tr>
<td>Intention to pursue CSIT career</td>
<td>A student’s intention to pursue CSIT career</td>
<td>Theory of planned behavior (TPB)</td>
<td>Ajzen (1991)</td>
</tr>
</tbody>
</table>

TPB has been used to predict and explain behavioral intent regarding technology adoption (Brown and Venkatesh, 2005) and career choice (Arnold et al. 2006; Giles and Rea, 1999; Shevlin and Millar, 2006). TPB proposes that behavior is influenced by specific personal traits and environmental factors (Ajzen, 1991). It considers two sets of personal traits—attitudes and perceived behavior control (PBC). The attitude toward the behavior refers to an individual predisposition towards the behavior and PBC is an individual’s perception of their ability to perform a given behavior (Ajzen, 1991). Social norm, as the environmental factor in TPB, refers to the social pressures an individual perceives to engage or not to engage in a behavior (Ajzen, 1991). Prior study combines self-efficacy and outcome expectation with TPB.
to better predict behavior. Pavlou and Fygenson (2006) found that self-efficacy and PBC are significantly related. Research has also found that outcome expectations are especially predictive of the attitude towards the choice of major and anticipated academic achievement (Diegelman and Subich, 2001; Hackett et al. 1992; Kuechler et al. 2009; Lent et al. 1993). Hence, we project two additional hypotheses:

H2: A student’s CSIT self-efficacy positively impacts their CSIT perceived behavioral control.
H3: CSIT outcome expectations positively affects their attitude towards CSIT.

Identity-based studies have established the connection between identity-related measures and career persistence in the domain of Science. Carlone and Johnson (2007) identified three interrelated key contributors to an individual’s identification with science: competence, performance and recognition. A person with the science identity should demonstrate meaningful knowledge and understanding of science content; should have requisite skills to perform scientific practices; and recognize herself or himself and gets recognized by other as a science person. Hazari et al. (2010) added another key contributor: interest, which is defined as an individual’s personal desire to learn/understand more subject matter and voluntary activity in the focus area.

The motivation for combining these the identity theory with SCCT and TPB comes from the importance of the identity-related measures in predicting persistence/engagement (Carlone and Johnson, 2007; Chinn, 2002; Hazari et al., 2010; Hughes, 2001; Morgan et al., 2001). In addition, the identity indicator construct partially overlaps with the core TPB constructs and SCCT constructs. For example, self-efficacy, outcome expectation, and perceived behavior control reflect the performance expectations of the identity
construct, social norm reflects the recognition expectations (recognition by others) of the identity construct, and attitude corresponds to the interest expectations of the identity construct. Thus, we have the following hypotheses:

H4: A student’s CSIT perceived behavioral control is related to CSIT Identity.
H5: Attitude towards CSIT is positively related to CSIT identity.
H6: The subjective norm (recognition by others important to the student) regarding CSIT is related to CSIT identity.

In addition to cognitive factors, SCCT also includes two environmental influences, social support and social barriers. Social barriers include negative social or family influences, financial constraints, instructional barriers, and gender and race discrimination (Lent et al., 2001). Social support includes social support and encouragement, financial resources, instrumental assistance, and access to role models or mentors (Lent et al., 2001). Social support and barriers are two countervailing forces affecting an individual’s behavioral intent (Lent et al., 2001; 2008). An individual’s identity formation is likely to be positively influenced by social support and negatively influenced by social barriers. Hence, we will test the following hypotheses:

H7: Social barriers negatively impact CSIT Identity.
H8: Level of social support is related to CSIT identity.

In this study, we specifically look at how CSIT identity will shape students’ intention to pursue CSIT education or career path. Hazari et al. (2010) found the physics identity indicator is significantly related to the choice to pursue a physics career. In the computing fields, students with CSIT identity are more likely to pursue a computing career. Career goals that a student has developed can influence their choice of major. So, we also examine the effect of intention to choose a CSIT career on intention to major in CSIT discipline. An external factor relevant in this context is perceptions of job availability. Students are more likely to choose the career with good job prospects. We also examine if perceived job availability has an impact on either intention to choose a CSIT education or career. Hence, we have the following hypotheses:

H9: CSIT identity is related to intention to choose CSIT career.
H10: Perceived CSIT job availability positively affects intention to choose CSIT career option.
H11: Perceived CSIT job availability positively affects intention to choose CSIT education.
H12: CSIT identity is related to intention to choose CSIT education.
H13: Intention to choose CSIT career will positively affects intention to choose IT education.

While the hypotheses above test the base research model, our ultimate intent is to examine gender differences in factors that influence CSIT identity formation and hence on intention to choose an IT career and education. Research shows that gender difference plays a significant role in one’s career’s path. Carlone and Johnson (2007) found that female scientists and those with “disrupted scientist identities” encountered more difficulty in their career paths. Since we will be collecting other demographics variables we can potentially examine how identity formation varies across those demographic variables. The study’s final hypothesis:

H14: There is a significant difference in the interaction between the CSIT identity formation and personal and environmental factors between male and female students.

We next discuss the research method and data analysis techniques.

**Method and Data Analysis**

**Procedure and Instrument Development**

The study will be conducted in three stages. In the first stage, we develop an instrument for the study based on the integrative model and pilot test it. The instrument will be further refined through formative research comprised of semi-structured student interviews and focus groups and a test-retest protocol to ensure high levels of face and construct validity and reliability. The second stage will involve administering the instrument to undergraduate students in CSIT and non-CSIT areas at a mid-sized university in the southeast as well as participating STARS Alliance (www.starsalliance.org) institutions.
The study’s final stage will be to conduct individual interviews and focus groups across representative groups to enrich the empirical findings and interpretation of results with qualitative data.

**Computer Science and Information Technology Identify Formation (CSITIF) Scale**

The initial CSITIF scale has been designed and developed through an extensive review of the computing and IT literature (Engelman et al., 2011; Heinze and Hu, 2009; Joshi et al., 2010; Pedroni et al., 2009; Trevisan et al., 2011) and self-efficacy and career identity formation (Ajzen, 1991; Lent et al., 2001; 2003; Hazari, et al., 2010) literature. The current instrument contains 263 identified items examining eight participant variables – demographics (9 items), prior technology experience (21 items), computing and IT skills (33 items), self-efficacy (29 items), theory of planned behavior (TPB) (37 items), theory of identity (42 items), computing and IT self-efficacy (72 items), and career aspects (20 items). The University’s Institutional Review Board (IRB) approval was obtained to conduct a pilot test of the instrument.

A pilot test of the CSITIF scale involved 123 college students comprised of CS undergraduates (36.6%, n=45), undergraduates who were non-majors (32.5%, n=40), IS undergraduates (22.8%, n=28), IS graduate students (4.1%, n=5), CS graduate students (2.4%, n=3), and undergraduate information systems pre-majors (1.6%, n=2). The sample was predominately male (65.6%, n=80; female, 34.4%, n=42), white (68.3%, n=84; black, 22.0%, n=27; Asian/Pacific Islander, 9.8%, n=12; Multi-racial, 4.9%, n=6); Native American/Alaskan Native, 3.3%, n=4), and 25 years old or younger (83.6%, n=104). Participants were asked to respond to 263 identified items examining eight different variables – demographics, prior technology experience, computing and IT skills, self-efficacy, theory of planned behavior, theory of identity, computing and IT self-efficacy, and career aspects. Overall completion rate for the entire scale was 91.3%.

Reliability tests (See Table 2) using Cronbach’s Alpha were highly positive for all construct items: CSIT Self Efficacy (Computer and IT self-efficacy=.933, Competence=.901, Recognition=.960), CSIT Outcome Expectation (Performance=.935), CSIT Perceived Behavioral Control (.966), Attitude Towards CSIT (Interests=.978), Social Barriers (.891), Social Norms (.936), Intention to Pursue CSIT Career (.986), and Intention to Pursue CSIT Education (.986) suggesting that there is high internal consistency reliability of the items identified for each construct.

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Cronbach’s Alpha</th>
<th>Constructs</th>
<th>Cronbach’s Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attitudes toward pursuing IT</td>
<td>.978</td>
<td>Social barriers</td>
<td>.891</td>
</tr>
<tr>
<td>Social norm</td>
<td>.936</td>
<td>IT job availability</td>
<td>.926</td>
</tr>
<tr>
<td>CSIT Perceived behavioral control</td>
<td>.966</td>
<td>CSIT identity</td>
<td>.933</td>
</tr>
<tr>
<td>CSIT Self-efficacy</td>
<td>.933</td>
<td>Intention to pursue CSIT education</td>
<td>.986</td>
</tr>
<tr>
<td>CSIT Outcome expectation</td>
<td>.935</td>
<td>Intention to pursue CSIT career</td>
<td>.986</td>
</tr>
<tr>
<td>Social support</td>
<td>.921</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Preliminary comparison of means through analysis of variance found a significant difference for gender F (1, 115) = 12.31, p<.001 when asked to react to the statement, “I consider myself to be a computing or information technology person.” Males had a stronger CSIT identity (M=4.8, SD=1. 74) rating this statement significantly higher on a seven point likert scale (1= strongly disagree, 7 = strongly agree) than females (M=3.53, SD=1.88). In general, IS graduate (M=5.80, SD=1.79), IS undergraduate (M=4.46, SD=1.9), and CS undergraduate (M=4.15, SD=1.85) students had the strongest CSIT identity, while non CSIT undergraduates (M=3.97, SD=1.80) and CS graduates (M=3.3, SD=2.08) had the weakest identity.

Preliminary regression analyses have created a path analysis found statistically significant relationships at the p<.001 level between 12 of the 14 hypothetical relationships suggesting the theoretical constructs as proposed have potential validity. The relationship between social barriers and CSIT identity was also significant but only at the p<.05 level. The final hypothesis H14 was supported. Statistically significant differences were found for gender across three primary scale items – “I would be able to complete a major in computing or IT” (Males, M=5.29; Females, M=4.21, p=0.001), “Majoring in computing or IT would be
easy for me” (Males, M=4.49; Females, M=3.77, p = 0.027), and “I have the resources and the knowledge
and the ability to major in computing or IT” (Males, M=5.04; Females, M=3.97, p=0.003).

While we are encouraged by the results of our initial pilot test, refinement and completion of the
instrument will involve several additional phases. Formative research will be conducted with
representative CSIT undergraduate students around the eight variables and associated items to ensure
they are valid (Ajzen, 2010). Exploratory factor analysis (EFA) will be used to identify the strongest latent
variables and confirmatory factor analysis (CFA) will be used to determine which combination of items
represent the strongest representation of the variable (Heinze and Hu, 2009; Joshi et al., 2010). The
refined instrument will be retested and only items found to be high in reliability will be retained;
Cronbach’s alpha will be used to measure the internal consistency reliability of the items retained for each
construct (Hazari et al., 2010).

Data analysis will involve general descriptive statistics, multiple regression analyses models (Baron and
Kenny, 1986; Engelman et al., 2011; Hazari et al., 2010), and factor and path analyses utilizing
comparative fit index (CFI), standardized root-mean-square residual (SRMR), and root-mean-square
error of approximation (RMSEA) (Lent et al., 2003; Joshi et al. 2010). Structural Equation Modeling
(SEM) will also be used to test for best fit models and illustrate significant relationships between primary
and latent variables and whether each of the study’s hypotheses was supported by the results (Heinze and
Hu, 2009; Joshi et al., 2010). In order to compare differences in identity formation factors, three groups
of undergraduate students will be studied: 1) pre-majors in computer science and information technology
(freshman and sophomore level), 2) majors in computer science and information technology (junior and
senior level), and 3) non-majors in computer science and information technology (freshman-senior level).

Study Implications and Plans for Further Research

The findings from this research will inform our understanding of factors that help shape CSIT identity
formation and how these factors differ across gender and other demographic information. Better
understanding of identity issues could suggest interventions that educational institutions could
implement to bring gender equity in the CSIT field, with significant benefits to individuals and the nation.
The fact that girls and women seem to identify less and less with computing fields is a very large factor in
projected workforce shortages in computing careers (Alesha and Theda, 2006; Burger et al., 2004; Kuhn
and Joshi, 2009; Kvasny et al., 2009; Scott et al., 2009; Trauth et al., 2009). In particular, the U.S.
Bureau of Labor Statistics projects 144,500 average annual job openings in computing, while the National
Center for Education Statistics shows 88,161 computing degrees earned annually, a shortfall of around
56,000 (NCWIT, 2010). Simply holding male CSIT degrees constant and increasing female degrees
earned to parity would add over 54,600 new CSIT degrees annually – in other words, the significant
shortfall of people for job openings could be overcome by raising the representation of women in CSIT.

A more inclusive CSIT workforce has many important and valuable consequences. According to the
National Association of Colleges and Employers, in 2010 computer science was the fourth highest-paying
major and IS was the tenth highest-paying major, and these two majors were the only non-engineering
majors to make the top-10 (National Association of Colleges and Employers, 2010). Increasing the
number of women in these majors can potentially result in opening more high-paying careers for women,
empowering women in their careers and lives. In addition to benefits to the women involved, a more
balanced CSIT workforce has consequences for innovation and national competitiveness: a 2007 study of
patents showed that patents from mixed-gender teams had citation rates 26-42 percent higher
(depending on technology area) than the norm (Ashcraft and Breitzman, 2007). Women are also seen to
be underrepresented in information technology jobs in all member states of the European Union
countries and other non-member states, such as Germany, Italy, Belgium and UK (Panteli et al. 1999;
Panteli et al. 2003; Wilson, 2003). Further research can examine focus on the role of culture and their
effects on CSIT identify formation.
References


National Center for Women and Information Technology (NCWIT). 2010. *Projected computing jobs and CIS degrees earned.*


