Categorizing Research on Identity in Undergraduate Computing Education

Amanpreet Kapoor
University of Florida
kapooramanpreet@ufl.edu

Christina Gardner-Mccune
University of Florida
gmccune@ufl.edu

ABSTRACT

Researchers in education have explored identity constructs to solve a variety of problems such as improving retention, ameliorating diversity and inclusion, fostering learning, and gauging decision-making. However, literature in social sciences describes identity research as often fragmented, with researchers often building their work on siloed factions in identity literature. This paper aims to build a categorization model for classifying types of papers on identity in computing education research (CER). We categorized 55 papers that either investigated identity formation of students in computing undergraduate degree programs or suggested relationships of other constructs to identity using a systematic literature review. We first explored trends in the types of papers with respect to their demographics and then categorized the papers based on semantics and contributions using inductive content analysis. We observed a growing interest in identity over the last five years. The types of papers on identity in CER fell into two themes: *identity-centric studies* and *non-identity centric studies*. These themes included six categories of papers that *described* identity, *assessed* identity formation, *measured* identity construct, studied the *influence* of identity on a factor, *implied* another construct as identity, and *inferred* relationships of other constructs to identity. We shed light on our categorization scheme, provide a framework for positioning future research, and discuss opportunities for future work on identity in computing. Our model can support researchers to position their work or find appropriate literature when investigating work related to identity in computing at the undergraduate level.

CCS CONCEPTS

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

KEYWORDS

identity, computing education, systematic literature review, computing identity

ACM Reference Format:
Amanpreet Kapoor and Christina Gardner-Mccune. 2022. Categorizing Research on Identity in Undergraduate Computing Education. In *Koli Calling '22: 22nd Koli Calling International Conference on Computing Education Research (Koli 2022), November 17–20, 2022, Koli, Finland*. ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3564721.3565948

Koli 2022, November 17–20, 2022, Koli, Finland

ACM ISBN 978-1-4503-9616-5/22/11…$15.00
https://doi.org/10.1145/3564721.3565948

1 INTRODUCTION

Computing skills are required for a large number of jobs in the industry [18] and jobs in computing will continue to grow by 13% annually over the next decade [90]. While enrollments in computing majors [6] are rising to make up for this demand, we know little about how to equally engage students of different demographics and abilities in our field [18] or how students learn computing given the relative nascence of our field when compared to other STEM-related fields. Learning is a complex construct as evidenced across disciplines and even in computing education research (CER), researchers have used a variety of proxies, theories, and frameworks to investigate learning [46, 53]. One construct that has been suggested to have the potential to unify research across framings is that of identity [69]. Identity itself is a complex construct [69] and researchers have studied identity to solve a variety of problems [24, 50, 83] such as retention, diversity, learning, decision-making, etc. The Handbook of Identity Theory and Research defines *identity* as "*viewed through the lens of an individual person, identity consists of the confluence of the person's self-chosen or ascribed commitments, personal characteristics, and beliefs about herself; roles and positions in relation to significant others; and her membership in social groups and categories (including both her status within the group and the group's status within the larger context); as well as her identification with treasured material possessions and her sense of where she belongs in geographical space*" [69]. At a fundamental level, *identity* involves a person's implicit or explicit response to the question: "*Who are you*?". Identity formation can be studied with a focus on individual, relational, or collective processes and there are various debates and philosophical viewpoints in social science research on the structure, domains, and processes of identity [69]. For example, is identity fluid or stable? Is it discovered, personally constructed, or socially constructed? Should it be studied through the lens of psychological, sociocultural, or socio-cognitive theories? Is identity possessed or negotiated in interaction? While work on identity can be classified using these philosophical viewpoints or theoretical frameworks, our work in this paper takes a different view of classifying prior work on identity literature in CER. We aim to categorize papers on identity in CER based on focus and contributions rather than philosophical underpinnings. This categorization can support our community's understanding of pragmatic semantics of what kind of work has been done so far related to identity. Our categorization model can help in (1) identifying gaps in the literature, (2) supporting researchers, especially those who aren't acquainted with identity literature, to identify relevant prior work, and (3) assisting scientific agencies in identifying areas to allocate future funds for identity work in computing. This work is the first step in our larger project that aims to understand researcher motivations for working on identity, identity definitions and components, identity representations or proxies for identity, methodologies for identity explorations, and factors that impact identity formation through a structured synthesis of prior literature on undergraduate computing identity.

Identity literature is often described as fragmented and disintegrated with researchers often building their work on siloed pieces in the literature (pp.1, [69]). The problem this fragmentation leads to is redundant work, inconsistent nomenclatures, and bodies of knowledge that have similar ideas or constructs but different terminologies. Researchers have tried to reduce this fragmentation of identity through systematic literature reviews (SLR) in other disciplines such as reviews on professional identity in higher education [83], identity development in engineering education [50, 58, 66, 80], and identity in mathematics education [9]. These reviews often explain the state of a field or topic in a discipline, describe models or taxonomies to categorize work, and suggest gaps that need to be addressed in future work. For example, a review on professional identity literature in higher education by Trede et. al. [83] categorized identity-related papers in higher education based on professional identity definitions and theoretical frameworks and found that the research base is underdeveloped. They emphasized the need for clarity on the meaning and conceptualization of the professional identity construct in future work. Within CER, the only review that synthesizes identity-related work is the recent work that assessed identity studies in K-12 computing research [24]. While our paper is not a complete synthesis of work on identity research similar to other SLRs, it is a first step that will aid in the synthesis of literature and initiating a dialogue on what are the types of papers related to identity in CER. We aim to categorize research on identity for the unexplored undergraduate computing identity literature.

2 METHODS

The goal of this paper is to consolidate research on identity in undergraduate computing education so that researchers can identify prior work and situate future work. To conduct prior work, we followed guidelines from Kitchenham et al. [36] in order to synthesize the existing evidence on identity contents and processes [69] and develop a categorization model. Specifically, we aim to answer the following research question: *What are the types of papers on identity with respect to semantics and contributions in undergraduate computing identity literature?*

2.1 Database search process

To answer our research question, we identified literature from three sources: (a) scientific article databases, (b) Google scholar, and (c) through snowball sampling. For the database search, we used five digital libraries: ACM Digital Library, IEEE Xplore, ScienceDirect, SpringerLink, and Taylor & Francis. These libraries or databases were selected based on an exhaustive list used in recent systematic literature reviews [26, 29, 44] in Computing Education Research (CER) literature. We used two queries or equivalent search strings to identify relevant literature in the databases:

- Preliminary query used on ACM Digital Library: *("student") AND ("identity") AND ("comput*")*
- Subsequent query used on all other databases: ("undergraduate" OR "student" OR "education") AND ("identity formation" OR "identity development" OR "form* identity" OR "influenc* identity") AND ("CS" OR "comput*" OR "software engineering" OR "informati*")

These queries were iterated and refactored multiple times to capture known literature and accurate results. The second query specifically was constructed after skimming abstracts and keywords of papers retrieved from the first query to ensure relevant papers can be retrieved in fewer hits in other databases. The search terms were applied to all fields and the query was run on March 31, 2020. Software engineering, Informatics, and Information Technology (IT) were included in the search terms as undergraduate computing programs follow different taxonomies or nomenclature for classifying computing degree programs around the world [71]. The regular expression queries were excluded from databases that did not support regular expressions and an alternate equivalent query was used.

A total of 825 papers were hit based on the queries of which 38 were selected based on our inclusion and exclusion criteria described in the next section. In addition, 17 papers were further selected from the first ten pages of Google scholar and through snowballing references (backward snowballing) in the identified set of papers. Backward snowballing is the process of identifying additional relevant papers from the reference list or citations in a paper [91]. This technique helps in reducing the risk of exclusion of articles that might be undetected using the search queries due to keyword or terminology inconsistencies. We went through the citation list of each of the initial 38 papers that were identified based on the queries. The papers from snowballing were included in the corpus if they met the following criteria (checked in this order): (1) the paper was not a part of the existing corpus, (2) the relevance of a paper title to our review topic, (3) the papers' abstract aligned with the scope of our review, and (4) the inclusion and exclusion criteria as described in the next section were met.

2.2 Inclusion and exclusion criteria

The inclusion and exclusion criteria were decided based on our research question and were further refined after reading abstracts of potential paper hits using the preliminary query. The studies that met each of the following four inclusion criteria were added to the pool of papers that were selected for review:

1. Study demographics or source of data (one or more)
   a. all or a subset of study participants were enrolled in undergraduate computing, CS, or IT programs; or
   b. study participants were undertaking an undergraduate computing course/professional development activity such as REU programs or bootcamps; or
   c. practitioners were teaching an undergraduate computing course; or
   d. data consisted of documents related to computing undergraduate programs.
2. Study theme (one or more)

Table 1: Papers that met our inclusion and exclusion criteria

Database/Method	Hits	Selected	Comments
ACM Full-Text Collection	493	29	A preliminary query was used for ACM
IEEE Xplore	24	4	The database is skewed toward Engineering identity
Taylor and Francis	133	3	
Science Direct	35	1	
SpringerLink	140	1	The database is skewed toward STEM identity
Google scholar and snowballing	-	17	
Total.	825	55	

a. construct under study was identity or the construct under study had a relationship with identity which was derived systematically; or
b. intervention under study was designed for influencing identity or the intervention indirectly impacted identity.

3. Type of publication (one): journal article, conference paper, dissertation, workshop/work-in-progress paper.
4. Publication language was English or a translation in English was available.

The papers that met any of the following exclusion criteria were discarded from the corpus:

1. Posters, doctoral consortium articles, and any publication less than or equal to two pages,
2. Non-peer reviewed articles,
3. Studies focusing on K-12 education,
4. Opinion papers with no empirical evidence or papers that proposed an opinion based on synthesis of prior work. Note that we define empirical evidence as the evidence derived or inferred from direct observation or experimentation. This evidence was obtained or was in the process of getting obtained from qualitative or quantitative studies/experiments conducted by a researcher(s). We excluded papers that answered research questions using previous literature and did not have empirical data directly collected from participants;

2.3 Filtering process, data extraction, and data analysis

We found 825 papers based on our search queries and we read the title and skimmed the abstract of these potential hits. After eliminating duplicates, 62 unique papers were found that met our inclusion and exclusion criteria. Of the 62 potential matches, 24 papers were further discarded after carefully reading sections in the paper as they failed to meet our inclusion and exclusion criteria. Such papers:

- did not meet participant requirements (n=9, four papers did not have students in CS programs, e.g. [34, 88], and five papers were targeted towards K-12 CS programs, e.g. [1, 30, 59]); or
- had no empirical data collected from participants but rather opinions backed up by synthesis of prior work (n=10, e.g., [31, 48, 65, 72, 84]); or
- did not follow study theme in our inclusion criteria (n=5, e.g., [3, 15]);

Thus, the data extraction process resulted in a corpus with 38 papers. In addition, 17 papers were further selected from the first ten pages of Google scholar and through snowballing references in the identified set of papers that met the inclusion and exclusion criteria. The snowballed papers were chosen based on reference titles and their alignment to the topic of our review. The frequency of papers obtained from different databases during our filtering process is shown in Table 1. Through this process, we obtained a corpus of 55 papers.

For the first 27 papers, the author read the paper twice, first highlighting relevant text and then extracting snippets from the highlighted text that answered our research question. This extracted data was organized using documents and spreadsheets and was analyzed using inductive content analysis [5], following a constant comparison technique [8, 73] to annotate data snippets into codes. The codes were based on commonalities across studies and emerged from the surface meaning of data or using a semantic method [5]. As we were reading new papers we compared if they fitted in one of the existing codes or else a new code was generated. The codes were further merged into six categories and these categories were abstracted into two themes. Our iterative process led to a persistent merging of codes into an existing category or reorganization of existing categories. We were confident that we had reached theoretical saturation [8] after coding 27 papers as no new category or theme emerged when we coded papers 22-27. We developed a preliminary categorization scheme consisting of categories and themes and used this model to categorize the remaining papers (n=28). These remaining papers fell into one of our categories and themes which strengthens the validity of our categorization model. An example of our content analysis process is shown in Table 2. We also use frequency analysis to discuss the broader trends that emerged from our analysis for each of the categories and themes.

2.4 Limitations

There are many limitations to our systematic review. First, we focused on exact search terms related to *identity* in our search strings. Identity is a vague concept and previous systematic literature reviews on identity in K-12 computing [24], engineering education [50], and higher education [83], have incorporated different thesaurus terms such as "*self-concept*", "*self-perception*", "*self-image*", "*sense of belonging*", "*reflective practice*", "*professional values*", etc. for finding papers related to identity. However, such review papers included identity-centric papers, had fewer papers, and excluded papers where another construct was the primary focus although

Table 2: Inductive content analysis example from a sample extracted text

RQ: What are the types of papers on identity with respect to semantics and contributions in undergraduate computing identity literature?			
Paper Title	Developing a Computing Identity Framework: Understanding Computer Science and Information Technology Career Choice [45]	Increasing Technical Excellence, Leadership and Commitment of Computing Students through Identity-Based Mentoring [4]	Developing Communities of Practice to Serve Hispanic Students: Supporting Identity, Community, and Professional Networks [28]
Raw Data	"This paper expands on knowledge of computing identity by building on what is known about prior identity models in science and mathematics education. The model theorizes three primary sub-constructs that contribute to the development of a computing identity [. . .]. Drawing on data from a nationally representative survey [. . .], *the study tested the alignment of the theorized model to the measures on the survey*."	We *present Computing Identity Mentoring, an intervention designed to increase commitment to computing while enhancing students' technical and leadership skills*. [. . .] This paper presents early findings on the effectiveness of the approach and illustrates Computing Identity Mentoring in the context of three of the seven institutions where it has been implemented.	Results from S-STEM program indicate scholars experience greater retention and higher achievement than their peers, yet *little is known about how S-STEM scholarship programs shape students' professional identities in their fields*. [. . .] The research questions that drive this project are: What evidence suggests the *Cybersecurity S-STEM program supports minority students' development of science identities* through access to performance, competence, and recognition? [. . .]
Primary Code	measure computing identity	assess the efficacy of computing identity mentoring program	assess S-STEM programs' role in identity formation
Category	Quantitative operationalization of identity to develop a measurement instrument	Assessing interventions for identity development/formation	
Theme	Identity-centric studies		

it had an association with identity. We included the papers that were not about identity in order to identify factors that influence identity indirectly. We recommend that future researchers explore the use of some of the terms from other fields to extract work related to identity as well as explore identity formation of computing professionals who are not necessarily computing undergraduates (e.g. conversational programmers or technical executives). In addition, our review focused on CS, IT, or computing students' identity formation. Thus, there is a possibility that work on identity formation in sub-disciplines of CS such as Data Science or Cybersecurity might not be included based on our search queries. Given the recent uptrend of student enrollments in majors specific to these sub-disciplines and the nascency of prior work on identity in these sub-areas, we decided to focus on the broader computing discipline. Another limitation is that the author ignored exploring work from identity research in other disciplines and/or general identity before extracting and analyzing data from the selected papers. This was intentional as the author wanted to be agnostic to prior work to ensure that the inductive data coding process for generating frameworks for positioning future research was not influenced by literature and described their viewpoint as an outsider. In addition, we excluded articles that were not written in English and conducted searches in five databases. In the future, other researchers can include articles in other languages and from other databases. Lastly, qualitative analysis can be influenced by a researcher's theoretical stance and interpretation. We attempt to address the validity of our inductive analysis through the transparency of our process and provide detailed exhaustive examples with references of the situated work in our categorization model.

3 FINDINGS

We first examine the trends in the types of papers with respect to their demographics and then describe or categorization model for classifying work on identity based on semantics and contributions. The description of demographics of papers is described in terms of types of publications, venues, and timeline.

3.1 Demographics of publications in our corpus

A majority of scientific articles that described research related to identity in computing were conference papers (n=43, 78%). This is in line with research in engineering education [50], however different than identity research in K-12 CS literature where journal articles are more prominent [24]. This was followed by scientific articles published as journal articles (n=9, 16%), and a small presence of peer-reviewed articles in dissertations and workshops as shown in Table 3.

We also observed a growing interest in identity in computing since 2014 (see Figure 1). 80% of the 55 papers (n=44) were published in or after 2014. The earliest paper in our corpus was published in 1997 and most papers were published in 2018 (n=13). Finally, in terms of publication demographics, we found that the venues where researchers published their work on identity in computing followed a long tail distribution with 55 papers published in 22 venues. Most papers (n=13, 24%) were published in unique venues [7, 25, 35, 37, 38, 43, 49, 54, 78, 79, 81, 82, 85]. Next, 16% of the 55 papers on identity

Table 3: Types of publications related to identity in computing at the undergraduate level

Types of publications	Count (N=55)
Conference paper	43
Journal article	9
Dissertation	2
Workshop papers	1

in computing were published at the ACM International Computing Education Research (ICER) conference (n=9), followed by 13% at the ACM Technical Symposium on Computer Science Education (SIGCSE) conference (n=7). All other venues with two or more publications are shown in Figure 2. Only one venue for publishing journal articles, ACM Transactions on Computing Education, had more than one paper (n=2).

3.2 Types of papers on identity in computing with respect to a paper's focus, semantics, and contributions

Two themes emerged from our review corpus related to the types of papers or studies on identity in computing with respect to the focus, semantics, and contributions of the papers: Theme 1 (*identity-centric studies*) and Theme 2 (*non-identity centric studies: nexus between the central construct under study and identity*). The themes are mutually exclusive and exhaustive, denoting that each paper fell into one of the two themes. Four categories of papers were found under Theme 1 and two in Theme 2. The categories are exhaustive, demonstrating that each paper in our corpus fell into one or more of these categories within one theme. These categories are conceptualized in Figure 3.

3.2.1 Theme 1: Identity-centric studies. The first theme denoted *identity-centric studies* where each paper had (i) *identity* as the central construct under study; and/or (ii) explicitly determined the relationship of a construct to identity through one or more research questions. In this theme, researchers often situated their work in

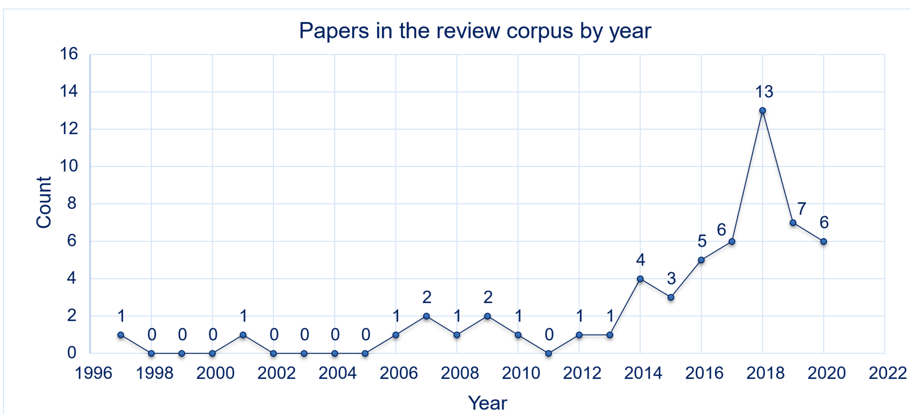


Figure 1: Papers published related to identity in computing at an undergraduate level by year before March 2020

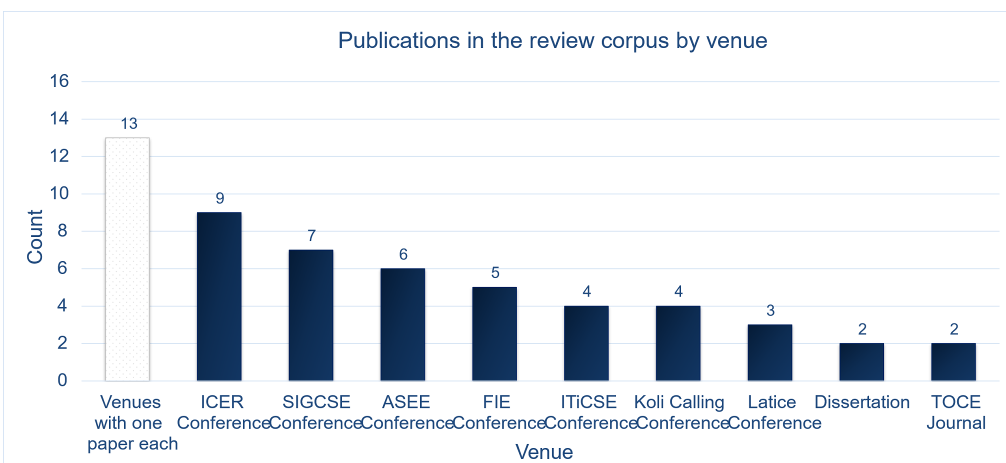
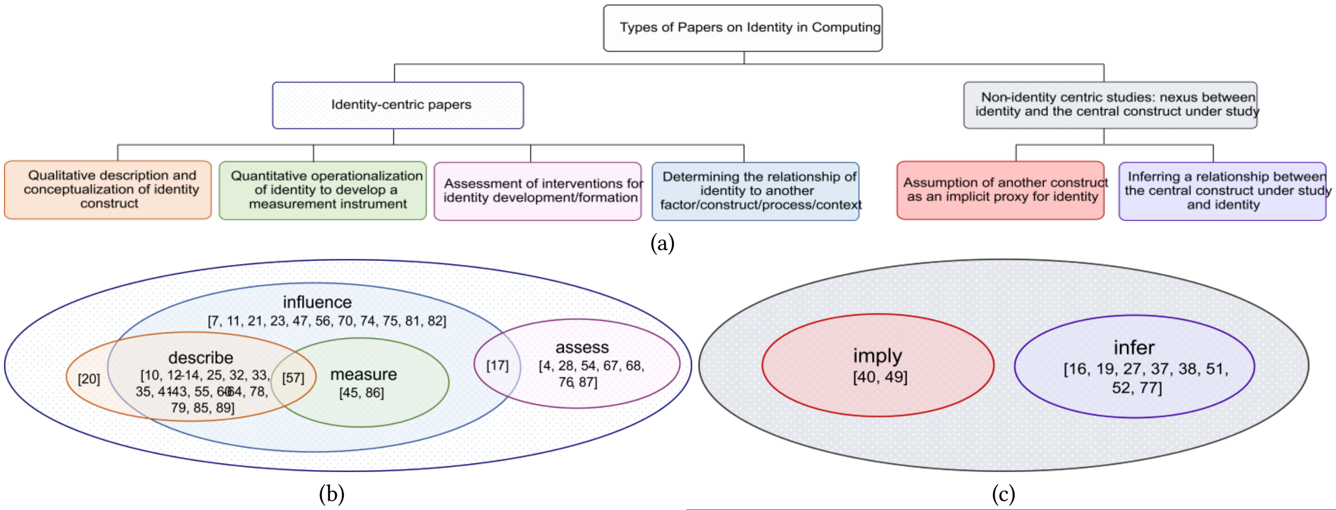


Figure 2: Papers published related to identity in computing at undergraduate level by venue



Category abbv.	Category Name	Count
describe	Description, conceptualization, and qualitative operationalization of identity in computing	23
measure	Quantitative operationalization of theoretical identity construct in computing	3
assess	Assessment of interventions for identity development/formation	9
influence	Determining the relationship of identity to another factor	36

(d)

Category abbv.	Category Name	Count
imply	Assumption of another construct as an implicit proxy for identity	2
infer	Inferring a relationship between the central construct under study and identity	8

(e)

Figure 3: (a) Categorization of papers on identity in computing with respect to focus, semantics, and contributions (N=55); (b) Venn diagram demonstrating focus of 45 identity-centric papers; (c) Venn diagram demonstrating focus of 10 non-identity centric papers; (d) Frequency table for categorization of identity-centric papers; (e) Frequency table for categorization of non-identity centric papers

prior work on identity and grounded their findings in identity literature. There were four categories of papers in this theme which either systematically studied identity construct or an intervention meant for fostering the development of identities in computing. 82% of the 55 papers (n=45) fell into this theme.

Qualitative description and conceptualization of identity construct in computing (describe): Within this category, researchers conducted fundamental research on the identity construct or the process of identity formation/development in computing. The papers focused on (i) describing a type or component of identity such as professional identity in computing [25, 32, 57], narrative identity or life stories [12–14], nerd identity [10], etc.; and/or (ii) explaining processes that can aid in the representation or conceptualization of the identity formation such as participation in the discipline [60–64], identity statuses [33], engagement, imagination, and alignment [42, 43], etc. The descriptions of the identity construct or processes were either inferred from different disciplines or were derived inductively and researchers examined these conceptualizations in the context of computing undergraduate programs. The abstract conceptualizations were further concretized using operational markers or proxies in contextualized qualitative data. 42% of the 55 papers (n=23) fell into this category [10, 12, 42, 43, 55, 57, 60–13, 14, 20, 25, 32, 33, 35, 41, 64, 78, 79, 85, 89].

Papers in this category that described a type or component of identity included Harrington et al.'s work [25] that answered the RQ "*What is the nature of computing students' professional identity*?", Kapoor and Gardner-McCune's [32] description of "*What are CS students' professional identities*?", and Parker's [57] findings on "*How well does a future role in the computing profession describe the current professional identity of undergraduate CS majors*?". The three former studies described the professional identities of students which is a type of a person's identity. However, there were differences in how they conceptualized professional identity. While Harrington et al.'s paper [25] presumed and conceptualized professional identity in computing as a student's association with a "*nerd*" or "*geek*" stereotype, Kapoor and Gardner-McCune [32] inductively determined that students described their computing professional identities in line with computing jobs or subdisciplines in computing. Parker [57] employed a conceptualization of identity similar to Kapoor and Gardner-McCune [32] and used future profession perceptions as a measure to operationalize CS students' professional identities. Parker, however, sought students' self and social (peers and supervisors) views on professional alignment as a second measure (in addition to what professional roles students see themselves in) to understand the extent of students' CS professional identity formation. Another example of papers in this category is Lundberg et al.'s work [42, 43] which conceptualized the pre-professional identity or employability of computing students. They asked students to imagine the future professional roles they see themselves in (similar to Parker [57] and Kapoor and Gardner-McCune [32]) and determined that the three processes of engagement, imagination, and alignment can be used to gain insight into students' pre-professional identity.

We also observed that researchers used two prominent terms to describe identity in computing: computing/CS identity [e.g., 4, 45, etc.] and computing/CS professional identity [e.g., 32, 33, 57, etc.]. Based on our interpretation of how the papers defined the terminology, the terms "*computing identity*" and "*CS identity*" were broader terms that implied the same meaning, "a person who sees themselves doing or using computing or computer science". The terms "*computing professional identity*" and "*CS professional identity*" however implied "doing computing professionally" and was a more restrictive term than "*computing*" or "*CS*" identity. Thus, the researchers differed in the terminology based on a person's association with computing. For instance, Bettin et al. [2] presumed "*computing identity*" as a part of students' "*engineering identity*" and they intended that students "*use or do*" computing as a professional engineer rather than "*do*" computing as a computing professional. The term "*computing identity*" was also sometimes associated with doing computing as a computing professional and hence there were overlaps. For example, Boyer et al. [4] defined "*computing identity*" as "*a sense of pride and belonging to the community of researchers and practitioners in CS and IT*" implying the person sees themselves as a computing professional belonging to research or practice. However, the opposite of this was not true - "*computing professional identity*" or "*computer science professional identity*" never implied using computing but rather the act of viewing one's self as a computing professional. For example, Kapoor and Gardner-McCune [32, 33] defined "*computing professional identity*" as "*the transformation of a person's interest in computing into a person's self-identification as someone who engages in computing professionally through one or more computing subdisciplines and career paths*", emphasizing a person's association with a computing profession. Thus, the terminologies related to identity in computing differ from engineering education literature where engineering identity is viewed as equivalent to professional identity or working as an engineer professionally [50] rather than using engineering outside of disciplines.

Other works describing components of identity include Kinnunen et al.'s work [35] which described students' CS identities through the lens of students' perception of the discipline, role in computing, and expectations from degree programs. Davis et al. [10] described "*nerd identity*" of CS students through the lens of "*discourse elements related to: competence, performance, and recognition*" in students' actions in a CS1 course. They broadened the description of "*nerd identities*", highlighted the potential benefits of such identities, and presented a case that a myriad of diverse nerd identities should be included in broader CS identities as some students who participate in CS imagine themselves as nerds. Dziallas and Fincher [12–14] used learning stories as a proxy to understand the narrative identities of students in computing with a focus on determining "*aspects of graduateness*" [12], instances of "*accountable disciplinary knowledge*" [14], "*learning life*" stories [13], and "*acquisition and use of disciplinary knowledge within and beyond their undergraduate education*" [13]. They stated, "*we are not concerned with the whole identity . . ., but with participants' "learning life" and with how the stories they tell about their learning experiences change over time*" [13]. Other descriptive work in this category included Friedensen et al.'s [20] work on how a computer engineer's identity is understood and used by an electrical and computer engineering department in departmental documents, Zander et al.'s [89] work that described a classification of students' identities in relation to the computing discipline, and Walker's paper [85] that described engineering students' feminine and masculine identities through experiences in CS courses.

Research on the description of processes that impact identity formation includes Peters' (et al.'s) work [60–64] that determined ways in which CS and IT students experience participation in the discipline. In this work, Peters' assumed that participation is a process through which identity is negotiated and this process was based on Wenger's interpretation of identity formation (social theory of learning [39]). She found seven hierarchical ways in which students experience participation in the CS/IT discipline: "*using, learning about technology, creating, problem solving, problem solving for others, creating new knowledge, and contributing to societal endeavors*". In another work on identity formation processes, Kapoor and Gardner-McCune [33] identified that the extent of computing professional identity formation can be represented through identity statuses which correspond to variation in students' exploration in the discipline or commitment to a discipline. Tomer and Mishra [79] investigated the process in which software engineering students "*construct their professional identities*" and found students resolving conflicts of their identities through the process of "*identity morphing*". Thiry and Hug's work [78] explained the process of "*identity production*" in Latina undergraduate computing students and Lewis et al. [41] described the process of students' assessment of their "*fit*" between identity and selection of a CS major.

Quantitative operationalization of identity to develop a measurement instrument (measure): Papers in this category quantitatively operationalized a conceptualized theoretical identity construct to develop instruments for measuring identity-related to computing at a time point. The operationalized identities were identified through surveys. We found three papers (5% of the 55 papers, [45, 57, 86]) in this category of which one instrument was exploratory and not validated [57]. These papers included Mahadeo et al.'s study [45] that determined the efficacy of a STEM model for measuring "*computing identity*" using three sub-constructs: belief in one's performance, interest, and recognition in computing. Other work in this category included Parker's [57] exploratory study to develop a quantitative survey instrument for "*operationalizing CS professional identity*" and Washington et al.'s study [86] that developed and validated the "*Computer Science Cultural Attitude and Identity Survey (CSAIS)*" for measuring ethnic identity and four other constructs (confidence, interest, gender, professional) which influence the attitudes of undergraduate CS students of color.

Assessing interventions for identity development/formation (assess): In this category, papers consisted of studies that tested the efficacy of an intervention for promoting identity development using operationalized identity indicators. We consider an activity as an intervention (and not a context/process/factor) if the researchers were involved in designing or assessing a activity, had agency to manipulate the activity, or prospectively collected data during an on-going activity. For instance, "course" can be both an intervention (as in this category) as well as a context (as in the next category). The difference between the two usages lies in whether a researcher is determining the efficacy of a specific course on identity development (intervention) or a researcher is seeking to understand the efficacy of generic coursework on identity development (context). 16% of the 55 papers (n=9) fell into this category [2, 4, 17, 28, 54, 67, 68, 76, 87].

The role of the intervention was to foster identity development in a student and the interventions included mentoring programs in undergraduate research [4, 67, 68], a professional conference [87], a scholarship program [28], or activities in formal coursework [2, 17, 54]. Most of these studies were quantitative and followed a quasi-experimental [2, 4] or pre-post survey design [17, 54, 67, 87] setup. Three studies were qualitative and descriptive [28, 68, 76]. Work in this category included Bettin et al.'s study [2] that investigated the inclusion of a pedagogical intervention exhibiting practical computing scenarios that could foster the development of computing identities in engineering students. Similar to Bettin et al.'s study for an intervention in a course, Novakocvich et al. [54] incorporated the use of social media tools in a professional writing course taken by majors from eclectic disciplines including CS to understand if social media technologies can be used for facilitating professional identity development. Similarly, Erdil and Ronan [17] determined if a rotation-based CS survey course can be used for fostering career identity formation. Outside of coursework, Boyer et al. [4] assessed their Computing Identity Mentoring program for promoting the development of students' computing identities in the context of undergraduate research. They assessed the efficacy of their intervention through a quasi-experiment using self-reported computing identity metrics and compared aggregate scores of study participants and non-participants. Santolucito and Piskac [68] extended Boyer et al.'s work [4] and assessed identity formation in an undergraduate research program in which students were mentored for formal methods and program synthesis research in computing. The goal of the program was to foster the formation of students' computing identities. Santolucito and Piskac [68], however, used a qualitative approach different from Boyer et al. [4] and described the efficacy of the program using rich case studies. Along the lines of undergraduate research experiences, Rorrer et al. [67] also determined the role of the National Science Foundation's Research Experiences for Undergraduates (NSF – REU) program in fostering scientific identities. In addition to the evaluation of identity formation in research experiences, Wei et al. [87] evaluated the role of a professional conference for women in engineering for the development of undergraduate students' engineering identities. Hug [28] explored how NSF's "*Cybersecurity S-STEM program supports minority students' development of science identities*" and Taylor-Smith et al. [76] determined if apprentices' employment during a Graduate Apprenticeship program (full-time work and study) can support the development of IT professional identity.

Determining the relationship of identity to another factor/construct/process/context (influence): This category had a majority of the papers, with 65% of the 55 papers (n=36). Within this category, papers described studies to determine the relationship of an identity construct to another factor, construct, process, or context [7, 10, 32, 33, 35, 41–11, 43, 45, 47, 55–57, 60–12–14, 17, 21, 23, 25, 64, 70, 74, 75, 78, 79, 81, 82, 85, 86, 89]. Note that these contexts were not planned interventions as in the previous category (*Assessing interventions for identity development/formation*) and researchers did not have opportunities to manipulate the context under study. Rather, the studies were observational and retrospective. However, both papers in this category and the ones in the previous category discussed mechanisms that impact identity. The relationships were examined quantitatively (e.g. [21, 57]) or qualitatively (e.g. [12, 33, 61]) using operational markers. The papers were situated in identity literature and the researchers either explicitly explored relationships between a factor/construct/process/context and identity (e.g. [33, 57]), or implicitly explained a factor/construct/process/context's role in identity formation (e.g. [12, 61]). There was a high degree of cohesion between papers in this category and Category *describe* or Category *measure* as researchers typically described, conceptualized, or operationalized identity and then found its variations across the study population or determined relationships of other factors to their described identity construct (see Figure 3). The explicit exploration of relationships either assumed no directions (*is* $factor_x$ *related to identity?*) or one direction (*does* $factor_x$ *influence identity?* or *does identity influence* $factor_x$?). The research questions typically were framed as: (a) How does a $factor_x$ influence $identity_y$? (b) How does $identity_y$ influence $factor_x$? (c) What is the relationship between $factor_x$ and $identity_y$? (d) How does $identity_y$ vary across a $factor_x$? (e) What factors influence $identity_y$?. Here, $factor_x$ denotes any factor/construct/process/context, e.g., coursework, self-efficacy, gender, etc., and $identity_y$ denotes a type or component of identity, e.g., professional identity, etc.

Next, we present examples of "*what*" types of associations were examined. Papers in this category included Harrington et al.'s work [25], which explored the relationship among "*computing students' professional identity*", "*students' reaction to stereotypes*", and "*experiences with stressors*". Davis et al. [10] assessed the relationship between "*nerd identity*" and "*engagement with CS course content*". Kapoor and Gardner-McCune studied the variation of professional identities across gender and academic year [32] as well as assessed the factors and processes that influence "*professional identity formation in CS undergraduate students*" [33]. McCartney and Sanders [47] also determined "*significant school- and job-related events that affect computing students' professional identity*". Similar to Kapoor and Gardner-McCune [33] and McCartney and Sanders [47], Tomer and Mishra [79] determined the factors that influenced the formation of students' professional identities. Parker [55, 56] examined the role of emotion in influencing "*professional identity formation*" and identified the variation in CS students' professional identity with respect to gender, degree program, academic year, and participation in internships or capstone projects [57]. Lundberg and Ness [43] sought to determine the people who shape students' "*understanding of jobs they qualify for*" in order to understand the source of students' imagination ability. Mahadeo et al. [45] and Demsey et al. [11] also studied the variation of computing or CS identity across gender similar to Parker [57] and Kapoor and Gardner-McCune [32], but they also determined if computing/CS identity can be used to predict students' "*choice*" or "intention" of a computing career or major. Sigurdson and Petersen [70] determined the role of mathematical curriculum in shaping students' identity while Kinnunen et al. [35] determined how students' expectations from degree programs, reasoning about the choice of major, and career aspirations influence identity.

While most of the previously described work determined relationships of identity to another construct when the participants were enrolled in the university, Dziallas and Fincher [12–14] followed a novel approach where they asked participants to reflect on learning stories longitudinally - during and after graduating from CS programs - and assessed the changes in these stories. They determined events, contexts, boundary objects, and turning points that influence students' learning trajectories [13], graduateness [12], and acquisition of accountable disciplinary knowledge [14]. These learning stories, elements of graduateness, or cases of accountable disciplinary knowledge acquisition influence a student's identity and eventually become a part of students' narrative identities [12–14]. Other relationships that were determined included (1) Grande et al.'s work [23] on the association between role models and identity; (2) Thiry and Hug's work [78] on processes influencing identity formation of Latina undergraduate computing students; (3) Peter (et al.'s) work [60–64] on understanding contexts in which CS and IT students experience different meanings of participation (a process of identity negotiation); (4) Garcia et al.'s paper [21] which demonstrated the variation in high-achieving underserved students computing identities across gender, academic year, and degree program; (5) Erdil and Ronan's work [17] on factors that influence career identities of students; (6) Cheryan et al.'s work [7] which explored sources of "*identity expression threat*" and studied the relationship between "*identity expression threat*" and women's expression of interest in CS; (7) Taheri et al.'s paper [74, 75] which determined the relationship of computing identity and persistence and also studied this relationship controlling gender and academic standing; (8) Trauth et al.'s work [81, 82] on the relationship among gender identity, ethnic identity, and stereotypes about the skills and knowledge in the IT profession; (9) Zander et al.'s paper [89] that described the relationship between CS students' identities and threshold concepts or transformative events; (10) Walker's work [85] on processes or factors that influence feminine and masculine identity construction; and (11) Lewis et al.'s work [41] which described factors that influence students' perception of "*fitting*" into CS, i.e., a process of alignment of a person's identity with values and cultural expectations in CS.

3.2.2 Theme 2: Non-identity centric studies: nexus between identity and the central construct under study. The second theme was *non-identity centric studies* and papers in this theme studied a construct other than identity that was the center of attention but the researchers either assumed or inferred a relationship of this other construct to identity. The researchers studied, conceptualized, or operationalized this other construct and assumed that this construct is an implicit proxy [40, 49] to identity, suggesting an outright or partial relationship, or they inferred a partial relationship between this other construct and identity [16, 19, 27, 37, 38, 51, 52, 77]. These papers rarely cited identity literature and defined research questions in terms of the central construct under study rather than identity. Also, the researchers grounded their results in the literature related to the construct under study and not in *identity* literature. These studies were included in our review corpus as such studies either inferred a relationship of the construct under study to identity or used a similar measure for the central construct as used in identity-centric papers for measuring identity (e.g. sense of belonging in [40] has a measure "*I see myself as a computing person*" which is commonly used as a proxy for identity such as in [22, 45, 57]). In essence, these papers were not studying *identity* construct per se, and the authors of these papers rarely discuss the relationship of their work to identity or the significance of their work for identity literature in computing, thus having a birds-eye view on identity. 18% of the 55 papers (n=10) fell into this theme. There were two categories of papers that belonged to this theme:

Assumption of another construct as an implicit proxy for identity (imply): The two papers in this category, assumed another construct as a proxy for identity and focused on identifying the relationship of this other construct to a factor. We determined this implicit relationship through cues in the paper because an explicit connection between identity and the construct under study was not made [40, 49]. In addition, the authors never discussed the significance of their findings through the lens of identity and often used measures that overlapped with identity constructs from *identity-centric* papers. Consider Lewis et al.'s work [40] which focused on understanding the relationship between sense of belonging, students' communal goals, and perception of computing, but did not define an explicit connection between identity and sense of belonging. When measuring the sense of belonging, Lewis et al. used a composite measurement item consisting of an indicator of identity "*I see myself as a computing person*" which is often used as a unidimensional indicator for measuring engineering identity in the engineering education discipline [22]. This usage of overlapping constructs to signify identity has been previously reported in computing education by McGill et al. [48] who found that researchers used identity and sense of belonging interchangeably in computing studies and recommended researchers clarify terminology in future studies. Thus, we established a relationship between a sense of belonging and identity construct based on such cues.

Inferring a relationship between the central construct under study and identity (infer): 15% of the 55 papers (n=8, [16, 19, 27, 37, 38, 51, 52, 77]) were primarily focused on studying another construct, phenomenon, or intervention and while understanding this construct/phenomenon/intervention, they found an influence on a person's identity. For instance, Thayer and Ko [77] conducted a study to understand the barriers faced by coding bootcamp students and found that students who enrolled in bootcamps and joined the software industry, had "*difficulties in claiming an identity as a software developer*" and felt imposter syndrome. Similarly, Eckerdal et al. [16] conducted a study to understand academics' perspectives on massive open online courses (MOOCs) and found that teachers had mixed responses on the MOOCs' ability to afford students' professional identity development. Other constructs or phenomena that were under study included students' perceptions of the university to work transition [38], women's experiences in computing with a focus on culture [19], students' conceptions of programming [51], students' association with computing stereotypes [27], students' interpretation on how CS is linked to other disciplines (CS in context) [37], and transformational incidents or threshold concepts in undergraduate programs [52]. The papers in this category were different from the studies in Theme 1: Category *influence* (Determining the relationship of identity to another factor) as these studies derived a relationship to identity while studying another construct while those in Theme 1 determined the relationship of a construct to identity explicitly.

4 DISCUSSION AND CONCLUSION

We explored trends in identity in the context of undergraduate computing education and developed a categorization model that will allow future researchers to situate their work in the CER community or identify relevant literature. Researchers have explored eclectic identity contents and processes in computing using different methodologies and theoretical frameworks. In addition, several researchers have determined relationships of another construct/context/process to identity. However, there is a lack of work that unifies what we know about identity in CER at an undergraduate level. Our future work will describe these identity definitions and terminologies, methodologies for studying identity and synthesize the factors that influence identity formation. We will be using our categorization model to organize this synthesis and determine how the definitions, proxies, units of measurements, and analysis vary across our categories of papers. We also observed that most papers that reported relationships of a construct/context/process/factor's influence on identity did not state the strength of association of these relationships. For instance, we know that contexts such as undergraduate research experiences, conferences, coursework, internships, etc. influence identity in computing. But how much do they matter when compared to each other for the development of computing identities? In which context, should researchers or practitioners spend resources to maximize students' identity formation? Which contexts or factors have a greater impact on a certain demographic of students? Several of these questions need to be answered and as a field, we need a theory to identify the relationship between these factors so that we can have a better understanding of how identity formation takes place in computing. Future work can also explore how to quantify or measure the strength of these associations.

Also, papers in Theme 2 - Category *imply* (Assumption of another construct as an implicit proxy for identity), specifically never made an explicit connection between identity and the other construct. This can be ambiguous for the readers or researchers and an explicit connection can be suggested if implied. We also observed that most studies in our corpus were observational and exploratory. Future work can investigate causal evidence on a factor's role in identity change through controlled experiments or deploy behavioral mechanisms (such as leveraging data captured from socio-technical systems) to gain insights into one's identity which could complement what we know about one's identity from self-reported measures. Our findings also provide evidence that researchers used similar or overlapping constructs leading to inconsistent terminology (e.g. identity [22] and sense of belonging [40]). A more consistent vocabulary could be developed in the future or at least clarification can be added to understand how researchers view or define constructs. As a community, we have not yet reached a consensus on how we define or measure these constructs. One way in which we can mitigate this issue is through a rigorous peer-review process that ensures future articles ground their work in literature. Another way to mitigate this issue is to reuse the codes and categories obtained from inductive content analysis in prior work if adequate. Researchers also contributed two survey instruments for measuring computing professional identity [45] and ethnic identity [86]. We recommend other researchers to use these instruments as well as develop more instruments rooted in sociocultural theories as the instruments developed so far are heavily influenced by socio-cognitive theories on identity formation.

ACKNOWLEDGMENTS

Thanks to Dr. Kristy Boyer, Dr. Sharon Chu, Dr. Mehmet Celepkolu, Cheryl Resch and the anonymous reviewers of the Koli Calling conference for their feedback on preliminary drafts.

REFERENCES

[1] Kera Bell-Watkins, Tiffany Barnes, and Nathan Thomas. 2009. Developing Computing Identity as a Model for Prioritizing Dynamic K-12 Computing Curricular Standards. J. Comput. Sci. Coll. 24, 3 (January 2009), 125–131.

[2] Briana Bettin, Michelle Jarvie-Eggart, Kelly S. Steelman, and Charles Wallace. 2020. Infusing Computing Identity into Introductory Engineering Instruction. In Proceedings of 2020 IEEE Frontiers in Education Conference (FIE), (October 2020). DOI: https://doi.org/10.1109/FIE44824.2020.9274290

[3] Jennifer M Blaney. 2018. Looking Beyond Representation: How Women Develop Their Leadership Perceptions During the Introductory Computing Course. In Proceedings of the 49th ACM Technical Symposium on Computer Science Education (SIGCSE '18), Association for Computing Machinery, New York, NY, USA, 453–458. DOI: https://doi.org/10.1145/3159450.3159458

[4] Kristy Elizabeth Boyer, E Nathan Thomas, Audrey S Rorrer, Deonte Cooper, and Mladen A Vouk. 2010. Increasing Technical Excellence, Leadership and Commitment of Computing Students through Identity-Based Mentoring. In Proceedings of the 41st ACM Technical Symposium on Computer Science Education (SIGCSE '10), Association for Computing Machinery, New York, NY, USA, 167–171. DOI: https://doi.org/10.1145/1734263.1734320

[5] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. Qual. Res. Psychol. 3, 2 (2006), 77–101.

[6] Tracy Camp, Stuart Zweben, Jane Stout, and Duncan Buell. 2016. Booming enrollments. ACM Inroads 7, 2 (May 2016), 12–14.

[7] Sapna Cheryan, Ella J Lombard, Lauren Hudson, Kengthsagn Louis, Victoria C Plaut, and Mary C Murphy. 2019. Double isolation: Identity expression threat predicts greater gender disparities in computer science. Self Identity 0, 0 (2019), 1–23. DOI: https://doi.org/10.1080/15298868.2019.1609576

[8] John W. Creswell and J. David Creswell. Research design: qualitative, quantitative, and mixed methods approaches.

[9] Lisa Darragh. 2016. Identity research in mathematics education. Educ. Stud. Math. 93, 1 (2016), 19–33.DOI: https://doi.org/10.1007/s10649-016-9696-5

[10] Don Davis, Timothy Yuen, and Matthew Berland. 2014. Multiple Case Study of Nerd Identity in a CS1 Class. In Proceedings of the 45th ACM Technical Symposium on Computer Science Education (SIGCSE '14), Association for Computing Machinery, New York, NY, USA, 325–330. DOI: https://doi.org/10.1145/2538862.2538960

[11] Jennifer Dempsey, Richard T. Snodgrass, Isabel Kishi, and Allison Titcomb. 2015. The Emerging Role of Self-Perception in Student Intentions. In Proceedings of the 46th ACM Technical Symposium on Computer Science Education - SIGCSE '15, ACM Press, New York, New York, USA, 108–113. DOI: https://doi.org/10.1145/2676723.2677305

[12] Sebastian Dziallas and Sally Fincher. 2016. Aspects of Graduateness in Computing Students' Narratives. In Proceedings of the 2016 ACM Conference on International Computing Education Research - ICER '16, ACM Press, New York, New York, USA, 181–190. DOI: https://doi.org/10.1145/2960310.2960317

[13] Sebastian Dziallas and Sally Fincher. 2018. "I Told You This Last Time, Right?": Re-Visiting Narratives of STEM Education. In Proceedings of the 2018 ACM Conference on International Computing Education Research (ICER '18), Association for Computing Machinery, New York, NY, USA, 223–231. DOI: https://doi.org/10.1145/3230977.3230989

[14] Sebastian Dziallas and Sally Fincher. 2019. Accountable Disciplinary Knowledge in Computing Education: A Case-Comparative Approach. In Proceedings of the 2019 ACM Conference on International Computing Education Research (ICER '19), Association for Computing Machinery, New York, NY, USA, 1–9. DOI: https://doi.org/10.1145/3291279.3339403

[15] Sebastian Dziallas, Sally Fincher, Colin G Johnson, and Ian Utting. 2017. A First Look at the Year in Computing. In Proceedings of the 2017 ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE '17), Association for Computing Machinery, New York, NY, USA, 275–280. DOI: https://doi.org/10.1145/3059009.3059049

[16] Anna Eckerdal, Päivi Kinnunen, Neena Thota, Aletta Nylén, Judy Sheard, and Lauri Malmi. 2014. Teaching and Learning with MOOCs: Computing Academics' Perspectives and Engagement. In Proceedings of the 2014 Conference on Innovation & Technology in Computer Science Education (ITiCSE

’14), Association for Computing Machinery, New York, NY, USA, 9–14. DOI: https://doi.org/10.1145/2591708.2591740
[17] D. Cenk Erdil and Darcy Ronan. 2020. Impact on Computing Attitudes and Career Intentions in a Common First Year Program Survey Course. In ASEE Virtual Annual Conference Content Access, American Society for Engineering Education. DOI: https://doi.org/10.18260/1-2--34760
[18] Sally A Fincher and Anthony V Robins. 2019. An Important and Timely Field. In The Cambridge Handbook of Computing Education Research, Sally A Fincher and Anthony V.Editors Robins (eds.). Cambridge University Press, 1–8. DOI: https://doi.org/10.1017/9781108654555.001
[19] Allan Fisher, Jane Margolis, and Faye Miller. 1997. Undergraduate Women in Computer Science: Experience, Motivation and Culture. In Proceedings of the Twenty-Eighth SIGCSE Technical Symposium on Computer Science Education (SIGCSE ’97), Association for Computing Machinery, New York, NY, USA, 106–110. DOI: https://doi.org/10.1145/268084.268127
[20] Rachel E. Friedensen, Erin E. Doran, and Sarah Rodriguez. 2018. Documenting Engineering Identity: Electrical and Computer Engineering Departmental Documents and Student Identity. ASEE Annu. Conf. Expo. Conf. Proc. 2018-June, (June 2018). DOI: https://doi.org/10.18260/1-2--30343
[21] Atalie Garcia, Monique S Ross, Zahra Hazari, Mark A Weiss, Tiana Solis, and Mohsen Taheri. 2018. Examining the Computing Identity of High-Achieving Underserved Computing Students on the Basis of Gender, Field, and Year in School. In CoNECD - The Collaborative Network for Engineering and Computing Diversity Conference, ASEE Conferences, Crystal City, Virginia. Retrieved November 17, 2021 from https://monolith.asee.org/public/conferences/113/papers/21493/view
[22] Allison Godwin, Geoff Potvin, Zahra Hazari, and Robynne Lock. 2016. Identity, Critical Agency, and Engineering: An Affective Model for Predicting Engineering as a Career Choice. J. Eng. Educ. 105, 2 (2016), 312–340. DOI: https://doi.org/10.1002/jee.20118
[23] Virginia Grande, Anne Kathrin Peters, Mats Daniels, and Matti Tedre. 2018. “Participating under the Influence”: How Role Models Affect the Computing Discipline, Profession, and Student Population. In Proceedings of 2018 IEEE Frontiers in Education Conference (FIE), 1-9. DOI: https://doi.org/10.1109/FIE.2018.8658944
[24] Gregor Große-Bölting, Dietrich Gerstenberger, Lara Gildehaus, Andreas Mühling, and Carsten Schulte. 2021. Identity in K-12 Computer Education Research: A Systematic Literature Review. In Proceedings of the 17th ACM Conference on International Computing Education Research. Association for Computing Machinery, New York, NY, USA, 169–183. Retrieved from https://doi.org/10.1145/3446871.3469757
[25] K. Vernard Harrington, Tracy L. Lewis, Wanda J. Smith, and France Bélanger. 2007. A Qualitative Analysis of Computing Students’ Professional Identity and its Relationship to Strategies for Coping with Stressors in the Computing Disciplines. In Proceedings of 13th Americas Conference on Information Systems (AMCIS), 1386–1395. Retrieved from http://aisel.aisnet.org/amcis2007/492
[26] Arto Hellas, Petri Ihantola, Andrew Petersen, Vangel V Ajanovski, Mirela Gutica, Timo Hynninen, Antti Knutas, Juho Leinonen, Chris Messom, and Soohyun Nam Liao. 2018. Predicting Academic Performance: A Systematic Literature Review. In Proceedings Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE 2018 Companion), Association for Computing Machinery, New York, NY, USA, 175–199. DOI: https://doi.org/10.1145/3293881.3295783
[27] Michael Hewner and Maria Knobelsdorf. 2008. Understanding Computing Stereotypes with Self-Categorization Theory. In Proceedings of the 8th International Conference on Computing Education Research (Koli ’08), Association for Computing Machinery, New York, NY, USA, 72–75. DOI: https://doi.org/10.1145/1595356.1595368
[28] Sarah Hug. 2018. Developing Communities of Practice to Serve Hispanic Students: Supporting Identity, Community, and Professional Networks. In ASEE Annual Conference & Exposition, ASEE Conferences, Salt Lake City, Utah. DOI: https://doi.org/10.18260/1-2--30305
[29] Petri Ihantola, Arto Vihavainen, Alireza Ahadi, Matthew Butler, Jürgen Börstler, Stephen H Edwards, Essi Isohanni, Ari Korhonen, Andrew Petersen, Kelly Rivers, Miguel Ángel Rubio, Judy Sheard, Bronius Skupas, Jaime Spacco, Claudia Szabo, and Daniel Toll. 2015. Educational Data Mining and Learning Analytics in Programming: Literature Review and Case Studies. In Proceedings of the 2015 ITiCSE on Working Group Reports (ITICSE-WGR ’15), Association for Computing Machinery, New York, NY, USA, 41–63. DOI: https://doi.org/10.1145/2858796.2858798
[30] Betsy James DiSalvo, Sarita Yardi, Mark Guzdial, Tom McKlin, Charles Meadows, Kenneth Perry, and Amy Bruckman. 2011. African American Men Constructing Computing Identity. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ’11), Association for Computing Machinery, New York, NY, USA, 2967–2970. DOI: https://doi.org/10.1145/1978942.1979381
[31] Iris A Junglas, Norman A Johnson, Douglas J Steel, D Chon Abraham, and Paul Mac Loughlin. 2007. Identity Formation, Learning Styles and Trust in Virtual Worlds. SIGMIS Database 38, 4 (October 2007), 90–96. DOI: https://doi.org/10.1145/1314234.1314251
[32] Amanpreet Kapoor and Christina Gardner-McCune. 2018. Understanding Professional Identities and Goals of Computer Science Undergraduate Students. In Proceedings of the 49th ACM Technical Symposium on Computer Science Education - SIGCSE ’18, ACM Press, New York, New York, USA, 191–196. DOI: https://doi.org/10.1145/3159450.3159474
[33] Amanpreet Kapoor and Christina Gardner-McCune. 2019. Understanding CS undergraduate students’ professional identity through the lens of their professional development. In Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE ’19, Association for Computing Machinery, 9–15. DOI: https://doi.org/10.1145/3304221.3319764
[34] Junghyun Kim. 2009. “I Want to Be Different from Others in Cyberspace” The Role of Visual Similarity in Virtual Group Identity. Comput. Hum. Behav. 25, 1 (January 2009), 88–95. DOI: https://doi.org/10.1016/j.chb.2008.06.008
[35] Päivi Kinnunen, Matthew Butler, Michael Morgan, Aletta Nylen, Anne-Kathrin Peters, Jane Sinclair, Sara Kalvala, and Erkki Pesonen. 2018. Understanding initial undergraduate expectations and identity in computing studies. Eur. J. Eng. Educ. 43, 2 (2018), 201–218. DOI: https://doi.org/10.1080/03043797.2016.1146233
[36] Barbara Ann Kitchenham, David Budgen, and Pearl Brereton. 2015. Evidence-based software engineering and systematic reviews. CRC press.
[37] Maria Knobelsdorf and Carsten Schulte. 2007. Computer Science in Context: Pathways to Computer Science. In Proceedings of the Seventh Baltic Sea Conference on Computing Education Research - Volume 88 (Koli Calling ’07), Australian Computer Society, Inc., AUS, 65–76.
[38] Marjatta Lairio and Leena Penttinen. 2006. Students’ career concerns: Challenges facing guidance providers in higher education. Int. J. Educ. Vocat. Guid. 6, 3 (2006), 143–157. DOI: https://doi.org/10.1007/s10775-006-9107-z
[39] Jean Lave and Etienne Wenger. 1991. Situated learning: legitimate peripheral participation. Cambridge University Press.
[40] Colleen Lewis, Paul Bruno, Jonathan Raygoza, and Julia Wang. 2019. Alignment of Goals and Perceptions of Computing Predicts Students’ Sense of Belonging in Computing. In Proceedings of the 2019 ACM Conference on International Computing Education Research (ICER ’19), Association for Computing Machinery, New York, NY, USA, 11–19. DOI: https://doi.org/10.1145/3291279.3339426
[41] Colleen Lewis, Ruth E. Anderson, and Ken Yasuhara. 2016. “I Don’t Code All Day”: Fitting in Computer Science When the Stereotypes Don’t Fit. In Proceedings of the 2016 ACM Conference on International Computing Education Research - ICER ’16, ACM Press, New York, New York, USA, 23–32. DOI: https://doi.org/10.1145/2960310.2960332
[42] Gunhild M Lundberg and Birgit R Krogstie. 2020. Employability Through Imagination, Alignment, and Engagement - Students’ Prospects and Change During Their First Year in Computing Education. In Proceedings of the 20th Koli Calling International Conference on Computing Education Research (Koli Calling ’20). Association for Computing Machinery, New York, NY, USA, Article 34, 1–5. https://doi.org/10.1145/3428029.3428049
[43] Gunhild M Lundberg and Ingunn Johanne Ness. 2020. First Year Students’ Imagination of Future Employment Identity as an Important Employability Aspect. In Proceedings of the 9th Computer Science Education Research Conference, ACM, New York, NY, USA. DOI: https://doi.org/10.1145/3442481
[44] Andrew Luxton-Reilly, Simon, Ibrahim Albluwi, Brett A Becker, Michail Giannakos, Amruth N Kumar, Linda Ott, James Paterson, Michael James Scott, Judy Sheard, and Claudia Szabo. 2018. Introductory Programming: A Systematic Literature Review. In Proceedings Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE 2018 Companion), Association for Computing Machinery, New York, NY, USA, 55–106. DOI: https://doi.org/10.1145/3293881.3295779
[45] Jonathan Mahadeo, Zahra Hazari, and Geoff Potvin. 2020. Developing a Computing Identity Framework: Understanding Computer Science and Information Technology Career Choice. ACM Trans. Comput. Educ. 20, 1 (January 2020). DOI: https://doi.org/10.1145/3365571
[46] Lauri Malmi, Judy Sheard, Päivi Kinnunen, Simon, and Jane Sinclair. 2019. Computing Education Theories: What Are They and How Are They Used? In Proceedings of the 2019 ACM Conference on International Computing Education Research (ICER ’19), Association for Computing Machinery, New York, NY, USA, 187–197. DOI: https://doi.org/10.1145/3291279.3339409
[47] Robert McCartney and Kate Sanders. 2015. School/Work: Development of Computing Students’ Professional Identity at University. In Proceedings of the eleventh annual International Conference on International Computing Education Research - ICER ’15, ACM Press, New York, New York, USA, 151–159. DOI: https://doi.org/10.1145/2787622.2787732
[48] Monica M McGill, Adrienne Decker, Tom McKlin, and Kathy Haynie. 2019. A Gap Analysis of Noncognitive Constructs in Evaluation Instruments Designed for Computing Education. In Proceedings of the 50th ACM Technical Symposium on Computer Science Education (SIGCSE ’19), Association for Computing Machinery, New York, NY, USA, 706–712. DOI: https://doi.org/10.1145/3287324.3287362
[49] Catherine Mooney, Brett A. Becker, Lana Salmon, and Eleni Mangina. 2018. Computer science identity and sense of belonging: A case study in Ireland. In Proceedings of the 1st International Workshop on Gender Equality in Software Engineering, Association for Computing Machinery, New York, NY, USA, 1–4. DOI: https://doi.org/10.1145/3195570.3195575

[50] John R Morelock. 2017. A systematic literature review of engineering identity: definitions, factors, and interventions affecting development, and means of measurement. Eur. J. Eng. Educ. 42, 6 (2017), 1240–1262. DOI: https://doi.org/10.1080/03043797.2017.1287664

[51] Adon Christian Michael Moskal, Joy Gasson, and Dale Parsons. 2017. The "Art" of Programming: Exploring Student Conceptions of Programming through the Use of Drawing Methodology. In Proceedings of the 2017 ACM Conference on International Computing Education Research (ICER '17), Association for Computing Machinery, New York, NY, USA, 39–46. DOI: https://doi.org/10.1145/3105726.3106170

[52] Jan Erik Moström, Jonas Boustedt, Anna Eckerdal, Robert McCartney, Kate Sanders, Lynda Thomas, and Carol Zander. 2009. Computer Science Student Transformations: Changes and Causes. In Proceedings of the 14th Annual ACM SIGCSE Conference on Innovation and Technology in Computer Science Education (ITiCSE '09), ACM, New York, NY, USA, 181–185. DOI: https://doi.org/10.1145/1562877.1562935

[53] Greg L Nelson and Amy J Ko. 2018. On Use of Theory in Computing Education Research. In Proceedings of the 2018 ACM Conference on International Computing Education Research (ICER '18), ACM, New York, NY, USA, 31–39. DOI: https://doi.org/10.1145/3230977.3230992

[54] Jeanette Novakovich, Sophia Miah, and Steven Shaw. 2017. Designing curriculum to shape professional social media skills and identity in virtual communities of practice. Comput. Educ. 104, (January 2017), 65–90. DOI: https://doi.org/10.1016/j.compedu.2016.11.002

[55] Richard J Parker. 2019. Developing Software Engineers: Investigating the Influence of a Computer Science Capstone on Professional Identity Formation. University of Colorado at Boulder. Retrieved from https://www.proquest.com/docview/2234287045

[56] Rick Parker. 2017. How Do You Feel: Affective Expressions from Computer Science Senior Capstone Projects. In 2017 International Conference on Learning and Teaching in Computing and Engineering (LaTICE), IEEE, 35–42. DOI: https://doi.org/10.1109/LaTiCE.2017.14

[57] Rick Parker. 2018. Who I Am Becoming, Now: Toward a Computer Science Professional Identity Instrument. In 2018 IEEE Frontiers in Education Conference (FIE), 1–9. DOI: https://doi.org/10.1109/FIE.2018.8658921

[58] Anita D Patrick and Maura Borrego. 2016. A Review of the Literature Relevant to Engineering Identity. In ASEE Annual Conference & Exposition, ASEE Conferences, New Orleans, Louisiana. DOI: https://doi.org/10.18260/p.26428

[59] Kylie A Peppler and Yasmin B Kafai. 2008. Developing a Design Culture in a Computer Clubhouse: The Role of Local Practices and Mediators. In Proceedings of the 8th International Conference on International Conference for the Learning Sciences - Volume 2 (ICLS'08), International Society of the Learning Sciences, 196–203.

[60] Anne-Kathrin Peters. 2017. Learning Computing at University: Participation and Identity: A Longitudinal Study. Uppsala University.

[61] Anne-Kathrin Peters. 2018. Students' Experience of Participation in a Discipline—A Longitudinal Study of Computer Science and IT Engineering Students. ACM Trans. Comput. Educ. 19, 1 (September 2018), 1–28. DOI: https://doi.org/10.1145/3230011

[62] Anne-Kathrin Peters, Anders Berglund, Anna Eckerdal, and Arnold Pears. 2014. First Year Computer Science and IT Students' Experience of Participation in the Discipline. In 2014 International Conference on Teaching and Learning in Computing and Engineering, 1–8. DOI: https://doi.org/10.1109/LaTiCE.2014.9

[63] Anne-Kathrin Peters, Anders Berglund, Anna Eckerdal, and Arnold Pears. 2015. Second Year Computer Science and IT Students' Experience of Participation in the Discipline. In Proceedings of the 15th Koli Calling Conference on Computing Education Research (Koli Calling '15), Association for Computing Machinery, New York, NY, USA, 68–76. DOI: https://doi.org/10.1145/2828959.2828962

[64] Anne-Kathrin Peters and Arnold Pears. 2013. Engagement in Computer Science and IT – What! A Matter of Identity? In 2013 Learning and Teaching in Computing and Engineering, 114–121. DOI: https://doi.org/10.1109/LaTiCE.2013.42

[65] Sarah L. Rodriguez and Kathleen Lehman. 2017. Developing the next generation of diverse computer scientists: the need for enhanced, intersectional computing identity theory. Comput. Sci. Educ. 27, 3–4 (October 2017), 229–247. DOI: https://doi.org/10.1080/08993408.2018.1457899

[66] Sarah L Rodriguez, Charles Lu, and Morgan Bartlett. 2018. Engineering Identity Development: A Review of the Higher Education Literature. Int. J. Educ. Math. Sci. Technol. (2018).

[67] Audrey Smith Rorrer, Joseph Allen, and Huifang Zuo. 2018. A National Study of Undergraduate Research Experiences in Computing: Implications for Culturally Relevant Pedagogy. In Proceedings of the 49th ACM Technical Symposium on Computer Science Education (SIGCSE '18), Association for Computing Machinery, New York, NY, USA, 604–609. DOI: https://doi.org/10.1145/3159450.3159510

[68] Mark Santolucito and Ruzica Piskac. 2020. Formal Methods and Computing Identity-Based Mentorship for Early Stage Researchers. In Proceedings of the 51st ACM Technical Symposium on Computer Science Education (SIGCSE '20), Association for Computing Machinery, New York, NY, USA, 135–141. DOI: https://doi.org/10.1145/3328778.3366957

[69] Seth J. Schwartz, Koen Luyckx, and Vivian L. Vignoles. 2011. Handbook of Identity Theory and Research. Springer Science & Business Media.

[70] Nikki Sigurdson and Andrew Petersen. 2017. Student Perspectives on Mathematics in Computer Science. In Proceedings of the 17th Koli Calling International Conference on Computing Education Research (Koli Calling '17), Association for Computing Machinery, New York, NY, USA, 108–117. DOI: https://doi.org/10.1145/3141880.3141888

[71] Simon, Alison Clear, Janet Carter, Gerry Cross, Atanas Radenski, Liviana Tudor, and Eno Tõnisson. 2015. What's in a Name? International Interpretations of Computing Education Terminology. In Proceedings of the 2015 ITiCSE on Working Group Reports (ITICSE-WGR '15), Association for Computing Machinery, New York, NY, USA, 173–186. DOI: https://doi.org/10.1145/2858796.2858803

[72] Jane Sinclair and Sara Kalvala. 2015. Exploring Societal Factors Affecting the Experience and Engagement of First Year Female Computer Science Undergraduates. In Proceedings of the 15th Koli Calling Conference on Computing Education Research (Koli Calling '15), Association for Computing Machinery, New York, NY, USA, 107–116. DOI: https://doi.org/10.1145/2828959.2828979

[73] Anselm Strauss and Juliet Corbin. 2008. Basics of Qualitative Research. SAGE Publications, Inc.

[74] Mohsen Taheri, Monique Ross, Zahra Hazari, Mark Weiss, Michael Georgiopoulos, Ken Christensen, Tiana Solis, Atalie Garcia, and Deepa Chari. 2018. A structural equation model analysis of computing identity sub-constructs and student academic persistence. In 2018 IEEE Frontiers in Education Conference (FIE), 1-7. DOI: https://doi.org/10.1109/FIE.2018.8658378

[75] Mohsen Taheri, Monique S. Ross, Zahra Hazari, Mark Weiss, Michael Georgiopoulos, Ken Christensen, Tiana Solis, Deepa Chari, and Zahra Taheri. 2019. Exploring Computing Identity and Persistence Across Multiple Groups Using Structural Equation Modeling. ASEE Annu. Conf. Expo. Conf. Proc. (June 2019). DOI: https://doi.org/10.18260/1-2--32803

[76] Ella Taylor-Smith, Sally Smith, and Colin Smith. 2019. Identity and Belonging for Graduate Apprenticeships in Computing: The Experience of First Cohort Degree Apprentices in Scotland. In Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE '19), Association for Computing Machinery, New York, NY, USA, 2–8. DOI: https://doi.org/10.1145/3304221.3319753

[77] Kyle Thayer and Amy J Ko. 2017. Barriers Faced by Coding Bootcamp Students. In Proceedings of the 2017 ACM Conference on International Computing Education Research (ICER '17), ACM, New York, NY, USA, 245–253. DOI: https://doi.org/10.1145/3105726.3106176

[78] Heather Thiry and Sarah Hug. 2014. "we should all help each other": Latina undergraduates' practices and identities in the figured world of computing. Proc. Int. Conf. Learn. Sci. ICLS 1, January (2014), 214–221.

[79] Gunjan Tomer and Sushanta Kumar Mishra. 2016. Professional identity construction among software engineering students: A study in India. Inf. Technol. People 29, 1 (2016), 146–172. DOI: https://doi.org/10.1108/ITP-10-2013-0181

[80] Karen L Tonso. 2014. Engineering Identity. In Cambridge Handbook of Engineering Education Research, Aditya Johri and Barbara M.Editors Olds (eds.). Cambridge University Press, 267–282. DOI: https://doi.org/10.1017/CBO9781139013451.019

[81] Eileen M Trauth, Curtis C Cain, K D Joshi, Lynette Kvasny, and Kayla Booth. 2012. Embracing Intersectionality in Gender and IT Career Choice Research. In Proceedings of the 50th Annual Conference on Computers and People Research (SIGMIS-CPR '12), Association for Computing Machinery, New York, NY, USA, 199–212. DOI: https://doi.org/10.1145/2214091.2214141

[82] Eileen M Trauth, Curtis C Cain, K D Joshi, Lynette Kvasny, and Kayla M Booth. 2016. The Influence of Gender-Ethnic Intersectionality on Gender Stereotypes about IT Skills and Knowledge. SIGMIS Database 47, 3 (August 2016), 9–39. DOI: https://doi.org/10.1145/2980783.2980785

[83] Franziska Trede, Rob Macklin, and Donna Bridges. 2012. Professional identity development: a review of the higher education literature. Stud. High. Educ. 37, 3 (2012), 365–384. DOI: https://doi.org/10.1080/03075079.2010.521237

[84] Lars Ulriksen, Lene Møller Moller Madsen, and Henriette T. Holmegaard. 2010. What do we know about explanations for drop out/opt out among young people from STM higher education programmes? Stud. Sci. Educ. 46, 2 (2010), 209–244. DOI: https://doi.org/10.1080/03057267.2010.504549

[85] Melanie Walker. 2001. Engineering Identities. Br. J. Sociol. Educ. 22, 1 (2001), 75–89. DOI: https://doi.org/10.1080/01425690020030792

[86] Alicia Nicki Washington, Shaefny Grays, and Sudipta Dasmohapatra. 2016. The Computer Science Attitude and Identity Survey (CSAIS): A Novel Tool for Measuring the Impact of Ethnic Identity in Underrepresented Computer Science Students. In ASEE Annual Conference & Exposition, ASEE Conferences, New Orleans, Louisiana. DOI: https://doi.org/10.18260/p.26110

[87] Belle Wei, Amy Strage, and Jinny Rhee. 2018. Silicon Valley Women in Engineering Conference — Creating Community and Nurturing Engineering Identity. In 2018 IEEE Frontiers in Education Conference (FIE), 1–4. DOI: https://doi.org/10.1109/FIE.2018.8659187

[88] Angela M. White, Jessica T. DeCuir-Gunby, and Seongtae Kim. 2019. A mixed methods exploration of the relationships between the racial identity, science

identity, science self-efficacy, and science achievement of African American students at HBCUs. Contemp. Educ. Psychol. 57, (April 2019), 54–71. DOI: https://doi.org/10.1016/j.cedpsych.2018.11.006
[89] Carol Zander, Jonas Boustedt, Robert McCartney, Jan Erik Moström, Kate Sanders, and Lynda Thomas. 2009. Student transformations: Are they computer scientists yet? ICER'09 - Proc. 2009 ACM Work. Int. Comput. Educ. Res. (2009), 129–140. DOI: https://doi.org/10.1145/1584322.1584337
[90] Computer and Information Technology Occupations: Occupational Outlook Handbook:: U.S. Bureau of Labor Statistics. Retrieved January 21, 2018 from https://www.bls.gov/ooh/computer-and-information-technology/home.htm
[91] Claes Wohlin. 2014. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering (EASE '14). Association for Computing Machinery, New York, NY, USA, Article 38, 1–10.DOI: https://doi.org/10.1145/2601248.2601268