New Orleans Jazzed about Engineering Education

Paper ID #14546

More Comprehensive and Inclusive Approaches to Demographic Data Collection

Mr. Todd Fernandez, Purdue University, West Lafayette

Todd is a PhD Student in Engineering Education at Purdue University who's research is focused on entrepreneurship education and entrepreneurship education as a component of modern engineering education efforts.

Dr. Allison Godwin, Purdue University, West Lafayette

Allison Godwin, Ph.D. is an Assistant Professor of Engineering Education at Purdue University. Her research focuses what factors influence diverse students to choose engineering and stay in engineering through their careers and how different experiences within the practice and culture of engineering foster or hinder belongingness and identity development. Dr. Godwin graduated from Clemson University with a B.S. in Chemical Engineering and Ph.D. in Engineering and Science Education. She is the recipient of a 2014 American Society for Engineering Education (ASEE) Educational Research and Methods Division Apprentice Faculty Grant. She also was an NSF Graduate Research Fellow for her work on female empowerment in engineering which won the National Association for Research in Science Teaching 2015 Outstanding Doctoral Research Award.

Jacqueline Doyle, Florida International University
Dina Verdin, Purdue University, West Lafayette

Dina Verdín is an Engineering Education graduate student at Purdue University. She completed her undergraduate degree in Industrial and Systems Engineering at San José State University. Her research interest focuses on the first-generation college student population, which includes changing the perspective of this population from a deficit base approach to an asset base approach.

Hank Boone, University of Nevada - Reno

Hank Boone is a Graduate Research Assistant and Masters Student at the University of Nevada, Reno. His research focuses on First Generation engineering college students' engineering identity, belongingness, and how they perceive their college experience.He is also on a National Science Foundation project looking at non-normative engineering students and how they may have differing paths to success. His education includes a B.S. in Mechanical Engineering from University of Nevada, Reno.

Dr. Adam Kirn, University of Nevada - Reno

Adam Kirn is an Assistant Professor of Engineering Education at University of Nevada, Reno. His research focuses on the interactions between engineering cultures, student motivation, and their learning experiences. His projects involve the study of student perceptions, beliefs and attitudes towards becoming engineers, their problem solving processes, and cultural fit. His education includes a B.S. in Biomedical Engineering from Rose-Hulman Institute of Technology, a M.S. in Bioengineering and Ph.D. in Engineering and Science Education from Clemson University.

Dr. Lisa Benson, Clemson University

Lisa Benson is an Associate Professor of Engineering and Science Education at Clemson University, with a joint appointment in Bioengineering. Her research focuses on the interactions between student motivation and their learning experiences. Her projects involve the study of student perceptions, beliefs and attitudes towards becoming engineers and scientists, and their problem solving processes. Other projects in the Benson group include effects of student-centered active learning, self-regulated learning, and incorporating engineering into secondary science and mathematics classrooms. Her education includes a B.S. in Bioengineering from the University of Vermont, and M.S. and Ph.D. in Bioengineering from Clemson University.

Dr. Geoff Potvin, Florida International University

More Comprehensive and Inclusive Approaches to Demographic Data Collection

Introduction

The purpose of this paper is to synthesize existing research about collecting demographic data, and to propose better methods or questions about demographics that are useful to engineering education research. We have organized this paper into three parts. First, we review why demographic questions are important and how demographics influence our research. Second, we discuss the logistics and potential pitfalls of demographic data collection. Finally, we turn to common dimensions of demographic data collection. For each demographic dimension we analyze current data collection approaches, potential problems with these approaches, and suggestions to improve the collection of data.

Leaving ASEE's year of action on diversity[1], we think this topic especially timely. As the public discussion on issues of diversity progresses to consider more nuanced and complex descriptors (e.g., sexual orientation, gender identity, biological sex, and gender expression instead of simply "sex"), how we collect these pieces of information must adapt as well in order to be more comprehensive and inclusive. The result is research that is both more accurate and more useful, by avoiding simplistic characterizations of demographics in favor of descriptors that are more representative of, and aligned with, the dimensions of current and future engineering students.

Principles of and changes to demographic data collection

Demography is concerned with the following aspects of human populations: size, distribution, composition, components of population change, and determinants and consequences of population change[2(p1)]. Populations are dynamic and change over time. The *size* of a population quantifies people in a given location, *distribution* denotes the geographic location, and the *composition* of a population refers to ascribed characteristics (e.g., sex, age, or birthplace) and attained characteristics (e.g., occupation, education, or socioeconomic status). The *components of population change* are defined as births, deaths, migrations, and broadly speaking, any factors that may influence change[2(p2)]. Usually, researchers collect demographic information through surveys, census, or public records. Formal agencies have been established to formulate systematic methods of demographic data collection and analysis with a goal of unbiased, high-quality, data that rationally represents a population for decision-making purposes[3].

The collection of demographic data generally arises from a need to make data-driven decisions. As an example, the U.S. Census uses demographic data in order to track the diversity of the US population and inform policy decisions[4]. Collecting data about population demographics allows governments to more deeply understand populations, to monitor the fairness and efficacy of programs, to assess public services, and to identify areas of public need with better granularity[5]. The U.S. Department of Education uses demographic information to improve assessment of educational conditions, progress, and fairness; demographic data aids the United States in maintaining and improving the national education system[6(p1)].

However, the collection of demographic data does not rely on static questions or conceptions of social identities like race, ethnicity, gender, or socioeconomic status. Often, changes in demographic data collection mirror changes in society becoming either more diverse in accepted categories or more accepting of complex identities that may not fit a single category[7(p8)]. These changes reflect shifting social norms, and appropriate assumptions about the individuals being asked demographic questions[7,8]. For example, the first US census, conducted in 1790, counted both (White[a]) males and (White) females, which was a novel approach at the time. However, it took 180 years, until 1970, for the census to differentiate people of Hispanic or Latino origin from those who identified as White, a change introduced to help measure anti-discrimination compliance[9]. Beyond simply including new categories or dimensions of demographics, small changes in how questions are asked such as a shift from "select one" a response to "select all that apply" can subtly influence how demographics reflect or record diversity in populations[10].

In addition to revising typical measures of demography, our developing understanding of multiple and overlapping social identities (e.g., race, ethnicity, gender expression, sexual orientation, etc.) has led to the inclusion of not just more dimensions, but of more refined dimensions. Part of this shift is an attempt to develop measurements that are accurate and representative of the underlying experience or phenomenon researchers wish to capture[8], and which represent a more holistic conception of demography[11]. However, the focus on holistically representing students' multiple social identities can cause tension between quantity and quality of quantitative demography. In this paper, we present best practices for asking demographic questions, and offer suggestions to deal with the tension of concise as well as precise questions.

Part 1: Why to ask and how to frame

While the external message of engineering claims that all people can be engineers, the culture of engineering is such that students from backgrounds that are underrepresented in engineering programs often feel relegated to only peripheral participation in engineering[12]. Students who have differently-identified gender, race, ethnicity, sexual orientation, disability status, backgrounds, or attitudes may not feel that they can fully participate in engineering communities of practice when they see only normative (i.e., traditional, standard, or typical) populations represented within engineering education. The lack of representation in engineering programs severely limits the ability of individuals from underrepresented groups to form an authentic engineering identity, and reduces the likelihood that they will continue to study engineering.

This trend further reduces the diversity of students who stay in engineering thereby propagating diversity issues into the engineering profession. A less diverse population of professional engineers, in turn, limits the ability of underrepresented groups to identify with engineering and makes them less likely to choose engineering in college. Students with non-normative identities who do choose engineering despite these barriers still face the issue of acting within an engineering culture defined by normative identities. This may cause many to leave, further

[a] Throughout our paper, we have consciously chosen to capitalize both Black and White when they serve as racial, ethnic, or cultural descriptors. We note that this is a somewhat fraught issue, but agree with Merrill Perlman, formerly of the New York Times, who notes "[l]anguage can reflect and foster bias and even invite violence, so respectfulness should always trump style or linguistic ambiguities. There may be contexts where bias is appropriately intentional, but absent that, equality should rule"[66].

exaggerating this negative feedback loop, and may also cause those who stay to adopt elements of the normative identities either by necessity, choice, or lack of alternative models[13].

A significant problem in quantitative measures of diversity is the "binning" of students' identities into categories defined by others rather than by their own personal identification[14]. Typical demographic variables include gender, race, ethnicity, age, residence (including zip code which can be used as a proxy for socio-economic information), parental education level, parental income level, and high school/high school district (which indicates the academic resources available to the student). These data points allow researchers to look for correlations or interactions of demographic variables.

These variables also enable researchers to explore not only the quantified diversity of student populations, but also how that diversity effects a student's experience in engineering education. Qualitative studies that seek to understand students' experiences in engineering benefit from the multiple lenses that come from different genders, ethnicities, and races, as well as intersections between different demographic subcomponents. Overall, demographic data is a powerful tool for understanding the relationship between how students experience engineering education, and how engineering education treats students.

Framing an effective demographic data mindset

Traditional visual markers of students bring value-laden assumptions about students' past experiences and current intentions, strengths and weaknesses, and what variables influence them. While diversity that can be seen may influence the way that others interact with a student, they do not necessarily represent how the student experiences, interprets, or responds to those interactions. The diversity of engineering students is not entirely driven by what we can directly observe. Demographic variables capture not only elements of a student's *social* identity (i.e., the identity with which the external world interacts) but also their *personal* identity (i.e., the identity with which the student views the world). A complete accounting of student diversity in engineering must capture dimensions that influence how students process their experiences, not just dimensions that may affect their experiences with others.

Additionally, demographic variables do not act in isolation. An individual has multiple social identities that comprise how he/she sees him/herself in the world[15]. The identities interact in complex ways that are not separable. For example, the experiences of being Black (both male and female) and being a woman (aggregate of all female experiences or responses) do not add together to form the experiences of being a Black woman. In engineering, if a researcher disaggregates data by underrepresented minority students, men will be overrepresented in the sample, and the experiences of minority women will be lost in the average responses. If the data are bifurcated on gender (based on a binary representation of male and female), then the responses of White women will overwhelm the responses of women of color based on the representation of each group in the sample. For these reasons, we caution quantitative researchers about the ways they disaggregate and represent findings in quantitative work when using demographic information; intersectionality examines the connections between social constructs of race and gender together rather than separately[16]. A holistic approach better captures how specific groups within engineering negotiate belonging to, and contributing their perspective and talent to, engineering.

In sum, it is important to ask demographic questions to capture the diversity of experiences and identities of students in engineering education. This is especially true of efforts to track the experience of students who do not meet the normative engineering identities. Asking demographic questions allows researchers to prevent the normative experience from overwhelming underrepresented students' experiences. Further, it allows for researchers and educators to understand, appreciate, and react to differences between normative and non-normative groups in order to help improve and/or understand diversity. However, we must be careful to frame the collection and use of demographic data appropriately. The use of multiple univariate approaches to identity (e.g., male vs. female *and then* Black vs. White) does not appreciate the very real and important intersectionalities that exist. When researchers ask and discuss demographics it is important to unpack the richness of all the groups and demographic clusters, not simply the dyads and binning which come from single measures treated separately.

Part 2: How to collect and when to ask

Appreciating non-normative and holistic identities in engineering begins by understanding the methods that researchers use to gather demographic information. Methods include ensuring participant anonymity, avoiding bias simply through collecting data, and balancing efficiency with accuracy. However, as will become apparent in the remainder of the paper, decisions about demographic data collection are highly situated and rarely have a single answer.

At the core of efforts to collect demographic data is a single, but vital, point: The planning and identification of demographic data collection must occur prior to data collection. Rather than as an afterthought, researchers must plan and prepare for the collection of demographic data with the same attention to detail that they pay to the "main" part of a study. Identifying what demographic variables are relevant and important serves as a prelude to a discussion of how to ask questions about demographics. For example, researchers might decide that their interest is not whether a student speaks fluent English, but instead what dialect of English they learned or who they learned English from. To support that discussion, part 2 describes the technical details that are useful for researchers to consider when we explore specific dimensions of demographic data collection in part 3.The appreciation of the non-normative and holistic identities in engineering begins with an accounting of the methods that researchers use to gather demographic information. This includes efforts to ensure participant anonymity, avoid creating bias simply through the collection of demographic information, and balancing efficiency with accuracy. However, as will become apparent in the remainder of the paper, decisions about demographic data collection are highly situated and rarely have a single answer.

Avoiding the creation of identifiable records

It is important that researchers acknowledge that demographic questions contain personal information that students may view as sensitive even if a single question seems innocuous. While we specifically advocate more *comprehensive* approaches to demographics, it is important to appreciate the potential impact of the data created. Researchers must carefully consider whether they are unintentionally creating personally identifiable records through the combination of demographic questions they ask[17].

Asking multiple, individually innocuous, demographic questions may inadvertently identify the one or two students at the intersection of those questions. While a single Black student, a single female student, or a single gay student may not be identifiable, the combination of multiple pieces of demographic information can easily identify single students with surprisingly high accuracy—a problem known as reidentification[18].

It is the researcher's responsibility to minimize the amount of sensitive or re-identifiable information they collect. While the opportunity to collect all potentially useful data in a single occasion is tempting given the realities of research, if one is not planning to use demographic questions for future analysis it is inappropriate and potentially problematic to collect that data. An effective research design will address such considerations before data collection in order to minimize the amount of "extra" information gathered; this lowers the risk of data tying an individual to a specific response. More importantly, it lowers the chance that an individual will modify their responses (because of social desirability or the complexities of identity) to mitigate the perceived risk of providing such information on a questionnaire.

When to ask about demographics in the course of a study

In addition to minimizing extra data collection, researchers should almost universally place demographic questions at the end of studies. Literature on the phenomenon of "stereotype threat"[19–21] clearly shows the risk of bias when students are asked to provide demographic information **before** completing a survey or test. For example, asking students to identify their gender at the beginning of a mathematics test risks activating a stereotype about women's mathematics performance, which is a phenomenon known as "priming." Priming causes significant decreases in performance[20]. Such effects are not specific to any particular stereotype, assessment, or population; rather, such problems arise in many generic situations that activate a contextually relevant stereotype.

Thus, to avoid stereotype threat or related priming, it is most appropriate to place demographic questions at the *end* of an instrument. The cost of such a choice is that it raises the risk of non-response rates—if an instrument is sufficiently long or complex that respondents are fatigued, they may not complete the survey. However, given the abundance of robust and well-developed data imputation methodologies[22], elevated non-response rates are a more tolerable problem with a lower fundamental impact on research results.

How to ask demographic questions

In addition to offering clear guidance on how much and where to ask demographic information in studies, we feel it is also important to address the specifics of data collection. The format of demographic items can influence data whether researchers collect demographic data verbally, on a paper form, or electronically. A good strategy is to collect demographic information that is consistent with a theoretical framing of social identities. Optimally, survey items are comprehensive of all potential choices; however, the individualized nature of demographic variables necessitates the collection of self-described identities as well.

Speaking pragmatically, a good approach is to structure such questions not as either/or questions but as "select all that apply" questions. Another approach is to provide open-ended "write-in"

responses that allow for students who do not fit within the predefined choices to document their identities accurately. Both approaches respect that a set of mutually exclusive categories cannot always faithfully represent a respondent's race, ethnicity, gender, etc. However, current approaches often treat demographic information as such (e.g., a student may identify as White and Black, but data are often collected such that an individual cannot identify as both). The advantage of "select all that apply" items is that researchers are not forcing respondents to put themselves into a single, researcher-chosen box or skip the question entirely.

As always, researchers should pilot and validate the response options. Items should be reasonably interpretable and representative of the spectrum of possible responses in the population under study. Open-ended responses have the advantage of providing even more flexibility for individuals to self-identify according to whatever responses come most naturally to them. However, this benefit exists at the (not insignificant) cost of the collected data being challenging to analyze in subsequent quantitative analysis -- there may be a temptation on the part of the researcher to subsequently "bin" responses, which reproduces the very problem it was intended to avoid. Electronic survey tools, which allow researchers to insert follow-up questions automatically, are an advantage for these cases.

In summary, researchers should ask only the demographic questions necessary to their study, and do so with an understanding of the risks of creating reidentifiable records. To counter the potential bias from "priming" demographic questions, researchers should ask demographic questions last while employing tools to maximize response rates. Lastly, the traditional framing of demographic questions often leads to a series of closed choices for participants that forces them to choose the closest response to their identification, or causes them to skip the item if they are unsure how to respond. Providing participants with the ability to self-identify responses to a demographic category allows them to provide any response they may wish. For researchers, this infinite level of responses may create challenges in the interpretation of results. However, a good middle ground is to use approaches that blend quantitative items with options that provide space for more qualitative clarification.

Part 3: How to ask demographics questions

Having discussed the reasons and best practices for collecting demographic information at a high level, we now turn to specific guidance for collecting individual dimensions of demographic identities. In this section, we present examples of how we have begun to address the multifaceted and complex nature of student demographic questions by developing questions that are more inclusive of students' identities (e.g., race, ethnicity, sexual orientation, gender identity, etc.). Most of the examples presented come from the Intersectionalities of Non-normative Identities in the Cultures of Engineering (InIce). We developed the InIce in spring 2015 and piloted it with 371 second semester, first-year engineering students at three U.S. institutions. After gathering evidence confirming the validity of the instrument to capture students' affective profiles (e.g., measures including belongingness, motivation, identity, personality, grit, and career intentions), we deployed the instrument at the beginning of the Fall 2015 semester in introductory engineering courses at four U.S. institutions with 2,966 student responses. The sections below describe demographic questions that we developed and asked in the data collection process. We developed these items with a careful balance between length of the survey, ease of administration, and flexibility of participants to self-identify. These questions are examples of

options that we have employed to address the issues raised in this paper. These options provide a starting point other researchers can use to move beyond binary choices for measuring and analyzing students' diverse backgrounds and attitudes. In this work, we do not argue that we have perfectly measured students' identities; instead, we offer these questions as one way to improve demographic data collection by changing the thinking behind the demographic questions that researchers use. We developed these items with a careful balance between length of the survey, ease of administration, and flexibility of participants to self-identify.

Asking about components of gender identity

Typical approach

Many surveys that include a question about the respondent's gender do so as a binary choice between "Male" and "Female." Often, researchers use the terms "gender" and "sex" interchangeably when asking about gender. For most students, the ambiguity in this question has no impact on their answer. However, for those who do not fit a binary, single-dimension gender identity (e.g., transgender people, genderqueer, gender fluid, and others) the lack of possible responses both erases their identities and makes responding to the question difficult. When the survey is aimed specifically at an LGBT+ population, or is about LGBT+ topics, the question tends to expand dramatically both in the number of options presented and the types of questions asked[15,16]. Using several questions to establish the respondent's biological sex, gender identity, gender expression, and other sub-dimensions creates more accurate data. However, these clarifying questions are also a potential cause of overly long surveys. Especially when combined with attempts to expand other demographic questions, long surveys can lead to faster survey fatigue, or confusion as to why all of the questions are relevant (especially for students who identify all components of their gender through a single, traditional, binary choice). Furthermore, long surveys can increase dropout or lower the quality of responses[23].

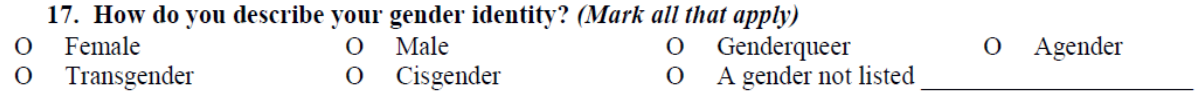
Problems with typical approach

The typical approach suffers by only allowing a single choice. For example, does a person who identifies as a trans man mark "Male" or "Transgender" for a question asking, "How do you describe your gender" that only allows for a single choice[24]? Further, including both "Male" and "Transgender Male" in an attempt to correct this problem immediately "othering" (i.e., intentionally or unintentionally classifying a group as non-normative) the trans people answering the survey. Adding a modifier to the normative male communicates that people who identify as trans are not part of the "normal" categories; this is especially true when nonbinary genders appear after the binary options seemingly as an afterthought. Another example is when binary categories appear alongside trans and "cisgender"[b] options. While an improvement over questions which list "Transgender" as a mutually exclusive category for gender (as opposed to "Male" or "Female"), it is better to avoid othering entirely if possible. In addition, to problems of an implied normal, the ambiguity that single choice questions introduce is a potential source of error if members of the same subpopulation respond differently.

[b] Cisgender is a term used to indicate people who identify with the physical sex assigned at birth. The "cis" modifier serves as a complimentary term to "trans" and seeks to reduce feelings of othering by making the modification of male and female non-exclusive to the trans community.

Improved approach

Our approach, seen in Figure 1, attempts to overcome all of these limitations while retaining a slim profile in the survey by asking only one question instead of spreading the query out over several questions. For students unaware of gender identity topics, the familiar choices of "Male" and "Female" are present as are other non-binary options. Additionally, both "Transgender" and "Cisgender" appear as gender modifiers that extend the students' ability to self-identify. Realistically, most respondents will likely choose one of the two traditional binary options. However, phrasing the question in this way can help increase respondent confidence in the survey and in the survey designers. Rather than presenting transness as the "other" choice, this phrasing presents both trans and cis categories on equal footing. Further, presenting the catchall fill-in-the-blank option (usually called "Other") as "A gender not listed above," validates any gender identity the respondent brings to the survey instead of relegating it to the status of Other to the "normal" gender identities which have their own bubbles to select.


17. How do you describe your gender identity? *(Mark all that apply)*

O Female O Male O Genderqueer O Agender
O Transgender O Cisgender O A gender not listed _____

Figure 1 InIce gender identity question.

Asking about race, ethnicity, and culture

Typical approach

The 2010 U.S. Census addressed race and ethnicity as two questions: Asking if the person was of Hispanic origin; and asking the person's race with a "select all that apply" strategy[5]. Although individuals could select multiple races, the available categories of race and ethnicity have always reflected "current politics, science and public attitudes"[25]. Reflecting caution or acceptance of the norms of data collection, most studies collecting race, ethnicity, and culture data adopt existing approaches like those employed by the U.S. Census or Department of Education.

Due to the growing inclusion of multiracial individuals in the U.S., the 2000 census was the first census where people could select multiple races for self-classification. It was also the first census to include a separate a question for individuals of Hispanic origin. Because individuals were confused by the existing race categories and limited racial sub-categories[26] in the 2010 Census, changes to the demographic measures on the U.S. Census will be implemented in 2020. Proposed questions would not even include the word race, but would describe racial/ethnic options as categories with which the person identifies instead. Another proposed idea for the 2020 census would allow individuals to check the racial group they identify with and then write the countries of origin for each of their races[27].

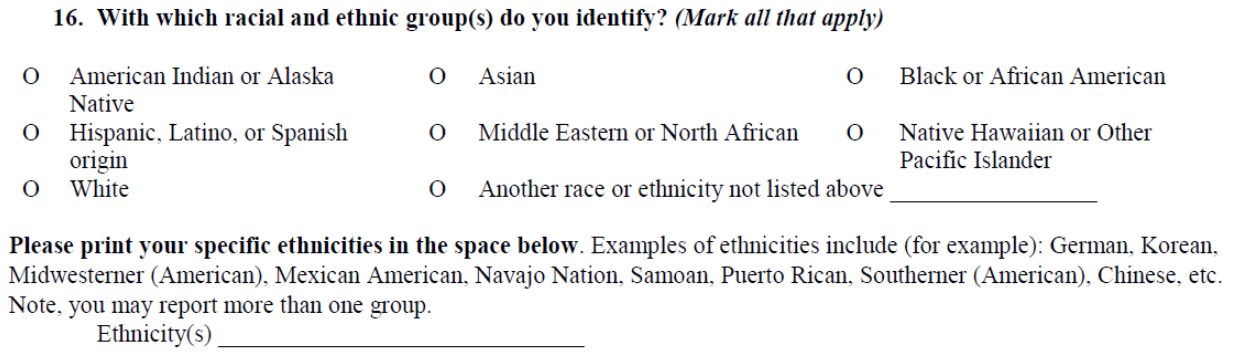
Problems with typical approach

The proposed changes to future U.S. Censuses may improve the self-classification of individuals' race, but these changes do not solve the problem of subtle differences between race and ethnicity, which may be further influenced by an individuals' country of origin. In the U.S., international students from at least six of the top ten home countries could select "Asian" as a

racial and ethnic group[28]. This categorization would tie students from countries as culturally diverse as India and Japan in the same demographic category.

Improved approach

The InIce survey handled the issue of self-classification by allowing participants to select multiple racial and ethnic groups in one question, and also by allowing participants to fill in their specific ethnicity as a written response, as presented below in Figure 2.



16. With which racial and ethnic group(s) do you identify? ***(Mark all that apply)***

O American Indian or Alaska Native	O Asian	O Black or African American
O Hispanic, Latino, or Spanish origin	O Middle Eastern or North African	O Native Hawaiian or Other Pacific Islander
O White	O Another race or ethnicity not listed above _____	

Please print your specific ethnicities in the space below. Examples of ethnicities include (for example): German, Korean, Midwesterner (American), Mexican American, Navajo Nation, Samoan, Puerto Rican, Southerner (American), Chinese, etc. Note, you may report more than one group.

Ethnicity(s) _____

Figure 2 InIce racial and ethnic identity item.

The combined wording of race and ethnicity into one question simplifies the confusion between race and ethnicity. Furthermore, the written response allows participants the option of self-classification which many past surveys did not permit. However, the question we used still relies primarily on measures typified through externally visible phenomenon. Further work is necessary to allow a separation of race, ethnicity, and cultural background because many individuals (specifically multi-racial individuals) may be of a similar genetic-race, but have a totally different cultural-race[29]. We theorize that this may include not just racial or ethnic classifiers, but also a country of birth and country/language of high school instruction. This approach allows participants to explain their perceptions of race, ethnicity, and the differences between the two rather than defaulting to federal standards to choose the broadness of the race and ethnicity questions.

Asking about parents and acknowledging the post-nuclear family-age[30]

Typical approach

From our experience, the typical approach to asking about family measures still assumes the traditional (i.e., one mother and one father) family arrangement. This assumption appears in the use of ‘mother’ and ‘father’ as identifiers throughout questions about parents and guardians.

Problems with typical approach

The shift of traditional family structures has been well documented[31], as has the changing shape of the modern family[30]. However, documentation of the shift in familial demographics has not always been adapted to quantitative instruments seeking to understand more about college students’ parents/guardians. Items seeking information about parent(s)/guardian(s) are used to understand students’ family life, to understand their access to education in the form of social capital[32], and as a proxy for SES. Such items often assume a traditional familial structure: One father or male guardian, and one mother or female guardian. Often these items provide little to

no instruction to students on how to fill them out, or they instruct students to only fill out the information that may be applicable to them (e.g., students from single parent/guardian households). While this information has been used to demonstrate the role of parents/guardians in engineering education[33–35], it is becoming less and less useful because the construction of the term 'family' is evolving more rapidly than at any point in human history[30].

In addition to the increase in single-parent and adoptive parent households[30], there has been an increase in the development or identification of households with same-sex couples. The latest census numbers indicate that there are at least 549,000 same-sex households in the U.S. with 115,000 of those households reporting having children. With the legalization of gay marriage the number of same-sex households is likely to rise, and as gay rights activists continue to push for equal adoptive rights, in addition to increased options for surrogacy and *in vitro* fertilization, the number of same-sex households with children is also likely to rise[36]. As a result of these changes, improved demographic measures are needed[36].

The rise of the new definition of family necessitates the need for an improved demographic measure to understand students' familial experiences. Despite these noted changes in the American household, change to familial measures has been slow and often met with resistance. An attempt by the State Department to replace the words "mother" and "father" on formal documents with "parent 1" and "parent 2"[37] was met with resistance and the changes were never adopted[38]. As demographic measures in education tend to mirror federal standards, for reasons of both policy and practice, little to no change has occurred in current demographic measures related to parents/guardians.

Improved approach

Improving demographics around familial structure requires limiting the pre-definitions imposed by the researcher. Asking the student to identify their parent(s)'/guardian(s)' gender identities aids in deconstructing the binary of male and female parent(s)/guardian(s). In addition, it avoids constructing a new binary question related to same-sex or mixed-sex household makeup (which could also otherize students from non-traditional families). Using student driven parental gender identification allows the demographic item to evolve with society rather than in response to it. When working with parent/guardian questions, assigning a number or identifier to the parent/guardian allows for each parent/guardian that students identify to be tracked through subsequent questions related to education level, type of employment, etc. However, this number assignment could create an additional cognitive load for students by asking them to remember how they identified their parent's gender identity.

Figure 3 below demonstrates one way to help identify parent/guardian status in the student household while allowing the student to acknowledge their parent's/guardian's identity. When constructing questions related to parent/guardian status, it is often important to ask a series of questions in order to obtain all the necessary information. Figure 4 represents one way to ask about the educational status of parents/guardians that builds off the trend set forth in Figure 3.

19. How would your parent(s) / guardian(s) describe their gender identities? *(Mark all that apply)*

Parent / Guardian # 1

O Female	O Male	O Genderqueer	O Agender
O Transgender	O Cisgender	O A gender not listed _____	

Parent / Guardian # 2

O Female	O Male	O Genderqueer	O Agender
O Transgender	O Cisgender	O A gender not listed _____	

Figure 3 InIce parent/guardian gender identity demographic item.

20. What was the highest level of education for your parent(s) / guardian(s)?

	Less than high school diploma	High school diploma / GED	Some college or associate / trade degree	Bachelor's degree	Master's degree or higher	Don't know
Parent / Guardian # 1	O	O	O	O	O	O
Parent / Guardian # 2	O	O	O	O	O	O

Figure 4 Parent/guardian education level item.

Additionally, allowing the student to identify the correct number of guardians allows the correct characterization of students from a single-parent household, *or* from familial situations where parenting occurs across multiple households and more than two parents. Furthermore, when researchers present surveys electronically, the use of 'logic flow' options helps appropriately document multi-generational households, families with step-parent(s), single parent households, and other child-rearing arrangements with more or less than two parental options.

Asking about socioeconomic status

Typical approach

A significant body of research indicates that socioeconomic status (SES) impacts students' progress and experience within higher education[39]. However effectively characterizing SES is notably difficult. NCES and the associated National Forum on Educational Statistics (NFES) have both made significant efforts to accurately collect and represent SES data[7,8].

Approaches to SES data collection include asking students to self-identify their socioeconomic status, or their perception of socioeconomic status[7]. These self-assessments are prevalent, in varying forms, in research as a contributing measure to an SES composite or as a pure SES measure. Researchers often use proxies because students are rarely able to report their household income accurately. One approach involves asking students to self-report familial income and then using that self-report to calculate an approximation of SES[40]. Other metrics for measuring SES include home zip code, high school, parent or guardian education, neighborhood SES, and students' eligibility for programs such as free or reduced-price school meals[7]. Parental education, especially the mother's education, has become a common research tool approximating SES[41]. However, the changes in common family structure that we note earlier should be accounted for in question formats. Researchers typically use these data to approximate SES either through

singular measures or through the creation of a composite metric that attempts to correct for problems in approximating[8].

Donaldson and colleagues[41] created a formula that weights direct self-reports and proxy measures equally. Their goal was to retain components of researcher approximation, but also to include respondent perceptions of SES, an important component of identity, as a component. While they note that it is not without shortcomings, they suggest that the formula is an improvement because it seeks to merge methods based on research, and methods based on self-perception about SES.

Problems with typical approach

What Donaldson and colleagues[41] highlight by using a composite approach merging both research and experience is one of several problems with measuring SES. The first group of problems with measures of SES is technical and logistical. Studies find that these questions have higher skip rates than other demographic variables[42,43]. Further, a yearly study by Fidelity Investments[44] indicates that only a third of parents have spoken with their college-ready children about money. Despite the importance of family income to SES, students often lack precise knowledge of the collected variable used to make up SES making a direct request problematic. Moreover, students (and many adults) often lack perspective on SES. A recent survey found that 51% of families who reported an income of greater than $100k/year indicated that they were middle class while 8% self-reported themselves as lower-middle class[45]. These subjective characterizations exist in contrast to U.S. census data which shows the median household income at $53,657/year with breakdowns that put households with income greater than $100k/year above the 75th percentile[46].

Improved approach

Our suggestions for improving the measurement of SES focus on improving the potential quality of data. First, select questions that have a factual basis that is reasonably likely to be known to the students, and that capture a comprehensive multi-faceted composite of SES[7,8]. Second, select SES components that primarily align with the study and its goals, as opposed to components that primarily align with measuring SES: In other words, what questions are the inclusion of SES trying to answer within a research study? Are you trying to measure students' perception of their wealth and status in relation to their peers, tangible experiences that inform their identity, or constraints from their past that may influence their future decision-making? Lastly, we note that future research is necessary to integrate the social status components of SES more effectively.

These suggestions focus on collecting data that has a higher level of accuracy and relevance, and which allows exploration of SES as a composite measure correlated and interrelated with many other demographic variables. We would also suggest a focus in higher education on measures that are particularly relevant to undergraduate students' SES experiences and conceptualizations. These include asking whether students have had to take out loans or work during the semester in order to fund their education in addition to other tangible measures of SES that are directly within the purview of students. This approach may not just include objective components SES, but also attempts to identify or appreciate students' subjective feelings of peer referenced SES.

Each operationalization of SES likely requires a different set of questions; however, all improve if they are based on information students are likely to know.

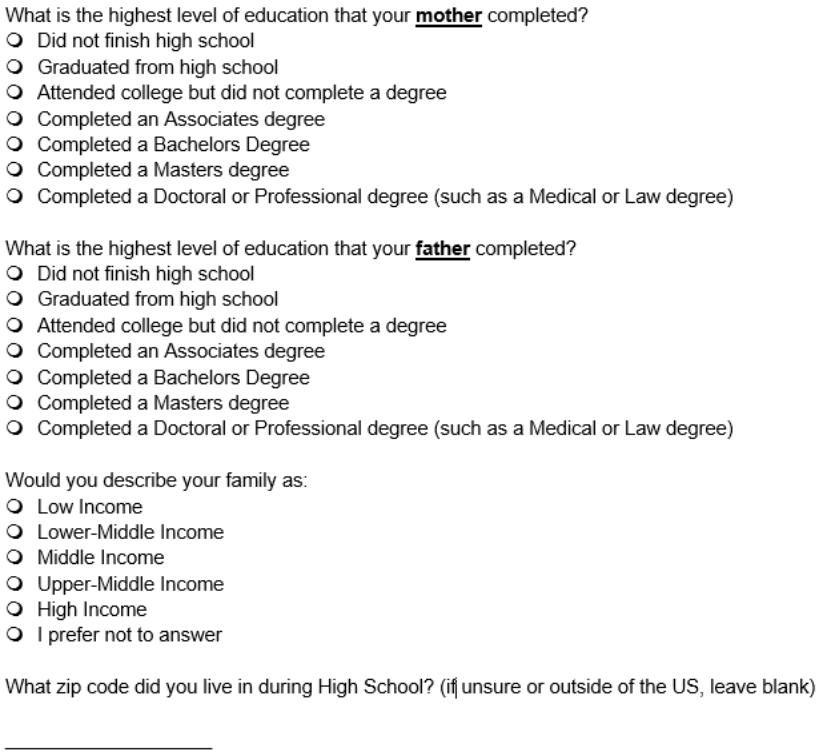


Figure 5 Example SES composite approach based on Donaldson et al[41].

Asking about sexual orientation

Typical approach

As with gender identity, encompassing questions on sexual orientation often appear only when the LGBT+ context is relevant to the research at hand, limiting the information we have on queer students outside of such contexts. Best practices recommend sets of questions to probe the respondents' romantic and sexual attraction, as well as their self-identified sexual orientation[24]. Respondents are often willing to answer questions about their sexuality[47], and are not more likely to stop the survey just because they reached a question about sexuality[48]. However, as with questions about gender, and as discussed in Part 2 of this paper, these questions should serve a clear research purpose by improving accuracy or contributing to research goals.

Problem with typical approach

In engineering education, the lack of collected data on student sexuality makes it difficult to make many claims about the experiences of non-heterosexual students (including gay or lesbian students as well as bisexual, asexual, etc.) especially concerning classroom outcomes. While the influence of race and gender are becoming an increasingly frequent topic of study, the influence of sexual orientation is not. Especially when studying stereotypes, a topic that has seen growing research[49], sexual orientation can easily become compounded with issues of gender expression and perceptions of gender identity. However, the idea that sexual orientation is a dimension of

identity that ends at the bedroom door can result in researchers ignoring sexuality as a relevant demographic variable.

Improved approach

We focus on self-identification of sexuality rather than attempting to construct it from responses about romantic and sexual attraction or activity. Though our answer choices seem more mutually exclusive than other options, we continue our pattern of allowing students to select multiple responses, and the "A sexuality not listed" option allows participants to construct a more nuanced option than those presented. Other research has offered solely write-in blanks to allow for complete self-identification, but this leads to either completely qualitative results or to researchers imposing their own interpretation onto students' self-identification as they categorize responses. Some have suggested a scaled rather than categorical approach, but understanding the nuance of the distance individuals mark between poles on paper-and-pencil surveys adds to the complexity of measuring these demographic items.

18. **How do you describe your sexual identity?** ***(Mark all that apply)***

O Heterosexual / straight O Homosexual / gay / lesbian O Bisexual O Asexual
O A sexuality not listed _____

Figure 6 InIce sexual orientation question.

Asking about ability and disability status

Typical approach

Often when collecting demographic information on survey instruments disability status is not included; furthermore, the number of publications that focus on this aspect of diversity in engineering education is minimal. A search through the last 10 years of the *Journal of Engineering Education* reveals only five mentions of disability[50–54], two of which[51,53] only reference disability as a characteristic of the "client" of a design project. Because of the lack of research in this area, we suggest that a typical approach to asking about ability and disability in engineering education is relatively non-existent.

Problem with typical approach

According to the National Science Foundation, students with disabilities, female students, and students of color from three racial/ethnic groups—Blacks, Hispanics, and American Indian/Alaska Native—are underrepresented in engineering and science. These students constitute a smaller percentage of engineering and science degree recipients than their respective percentages within the U.S. population[55]. A significant amount of research has focused on women, people of color, and individuals at the intersection of these social identities, women of color. However, little research in engineering education has focused on students with disabilities. The few studies that have focused on this group are skewed towards students with physical disabilities. Thirteen percent of all U.S. school-age students receive disability services, with 41% of those services related to learning disabilities[56]. However, the number of students with disabilities may be even higher. Estimates of the proportion of the student population with disabilities vary depending on the definition of the term "disability." Persons with disabilities may or may not require accommodation, and their disabilities do not necessarily limit their

ability to participate in educational experiences or be productive in an occupation. Despite the possible underreporting of disability status, the fact that a significant portion (greater than 10%) of the population reports having some level of disability emphasizes the need for strong demographic measures of disability status, and an increased focus on understanding how these students learn and develop as engineers.

Disability status is one of the least consistently asked questions about social identity across different organizations. This is probably, in part, due to the complex and sensitive nature of these types of questions. Disability takes many different forms, and measuring it satisfactorily is difficult. Additionally, different administrative and political priorities influence the focus and form of disability questions, as well as how the information is used. The National Center for Education Statistics (NCES) defines disability as, "a physical or mental condition that causes functional limitations that substantially limit one or more major life activities, including mobility, communication (seeing, hearing, speaking), and learning"[57].

The federal government has tracked disability status through census data and through surveys administered by other federal agencies. However, the American Community Survey (ACS) and the Survey of Income and Program Participation (SIPP) have replaced questions in the decennial census, and there is no information on disability status in the U.S. 2010 Census. Instead, these surveys present participants with lists of specific behavioral and medical questions to determine disability status rather than relying on self-reports of disability status[5,58].

Improved approach

The National Center for Education Statistics has grouped disabilities into twelve categories: difficulty hearing; difficulty seeing; difficulty speaking or language impairment; mobility limitation/orthopedic impairment; traumatic brain injury; specific learning disabilities; ADD or ADHD; Autism Spectrum Disorders; cognitive difficulties or intellectual disability; health impairment/condition including chronic conditions; mental illness/ psychological or psychiatric condition; and other[57]. Reporting disability status is measured by directly asking about these specific categories of disability[59].

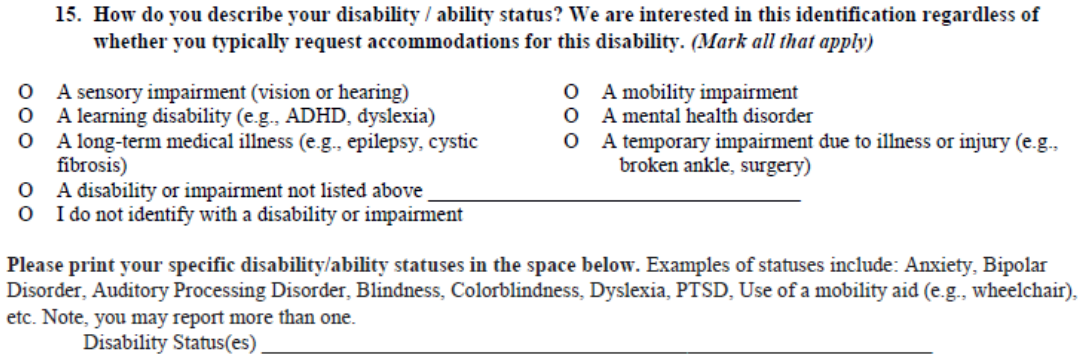
Since 2000, the annual National Survey of Student Engagement (NSSE) has obtained information about various aspects of undergraduate experience from random samples of first-year and senior students at four-year colleges and universities nationwide. The 2013 survey revised questions about disability status and, for the first time, reported the results of institutions. Questions asked, "Have you been diagnosed with any disability or impairment?" and, "If 'Yes' Which of the following have been diagnosed? (Mark all that apply)"[60]. These choices included: "A sensory impairment (vision or hearing)," "A mobility impairment," "A learning disability (e.g., ADHD, dyslexia)," "A mental health disorder," "A disability or impairment not listed above," and "Have not been diagnosed with a disability or impairment."

One of the problems with current approaches to asking students about disability status is the complexity of categorizing a number of different types of disabilities by fully-able researchers. Often, in our culture, and especially in STEM, disability status is invisible or rarely discussed. Madriaga describes the root causes of this unawareness[61], "Lacking disability awareness can be attributed to occupational cultures which reflect general societal attitudes towards disability"(p.410). In an article in the *Chronicle of Higher Education* titled, "Why is Disability

Missing from the Discourse on Diversity?" Lennard Davis notes key reasons why disability is not discussed in current diversity initiatives[62]. First, disability is not commonly considered an identity, and only identities are included in diversity discourses. Second, we do not celebrate disability as we do ethnic or cultural diversity – identities used to empower students, faculty, and staff. Finally, the current model of our society falsely dichotomizes able versus not able, and does not acknowledge the spectrum of individual's disability status[63].

Additionally, detailed questionnaires that have historically been used in the U.S. census are long, and therefore add a significant cognitive load and time burden on students taking voluntary surveys. Shorter versions of disability status measurements (like the one from the NSSE) focus on medical diagnoses and may not encompass the full range of students' self-identified disability statuses. The challenge of asking questions about disability status is to ask detailed enough questions to give students the ability to describe themselves while not parsing the data of this underrepresented group into identifiable data.

As shown in Figure 7, our approach to developing a question that collects data about students' disability status was threefold. First, we wanted to develop the most comprehensive question possible that was detailed enough to give students a variety of authentic responses. Second, we wanted to move away from a medical diagnosis approach and toward an identity approach to understanding the spectrum of disability status. Third, we consulted with individuals who self-identified as students with disabilities, experts from several university Disability Resource Centers, and researchers who work on culture and social constructs related to disability (e.g., Crip Theory research). We modified the wording of the question on the NSSE to meet our goals and refined the question through our consultations. We believe that our approach addresses the most critical concerns about current approaches to asking students questions about disability status.



15. How do you describe your disability / ability status? We are interested in this identification regardless of whether you typically request accommodations for this disability. ***(Mark all that apply)***

- O A sensory impairment (vision or hearing)
- O A learning disability (e.g., ADHD, dyslexia)
- O A long-term medical illness (e.g., epilepsy, cystic fibrosis)
- O A disability or impairment not listed above _____
- O I do not identify with a disability or impairment
- O A mobility impairment
- O A mental health disorder
- O A temporary impairment due to illness or injury (e.g., broken ankle, surgery)

Please print your specific disability/ability statuses in the space below. Examples of statuses include: Anxiety, Bipolar Disorder, Auditory Processing Disorder, Blindness, Colorblindness, Dyslexia, PTSD, Use of a mobility aid (e.g., wheelchair), etc. Note, you may report more than one.

Disability Status(es) _____

Figure 7 InIce disability/ability status question.

Summarizing our suggestions

To conclude our presentation of more comprehensive and inclusive demographic questions, we present a summary of our suggestions in Table 1. The unifying theme of the suggestions is to have an appreciation of the meaning of underlying identity constructs, and to make an effort to exhaustively include all potential options. While we have covered only a subset of the potential demographic questions of interest to researchers, the same principles apply generally.

Table 1 Comparison of approaches for demographic data collection

Dimension	Typical	Improved
Ability and disability	*Rarely asked, when asked typically focuses on physical disability*	*When relevant, creates disability situationally appropriate sub-dimensions.* Components of disability might include: Mobility Sensory Learning Mental health A disability not listed above
Gender identity	*Assumes a binary gender or provides a single non-binary option* Gender: Male Female Other*	*Separates components of gender identity. Provides more comprehensive options. Does not otherize* Gender identity: Woman Man Cisgender Transgender Genderqueer Agender A gender not listed above Gender Expression (*potentially scaled*) Feminine Masculine
Sexual orientation	*When asked, generally assumes a singular dimension with binary categories* Sexual orientation Heterosexual (or straight) Homosexual (or gay/lesbian) Bisexual*	*Integrates orientation based on attraction, gender identity, and cultural experience* Heterosexual / Straight Homosexual / Gay or Lesbian Bisexual Asexual Pansexual Asexuality not listed above *Alternate:* I am attracted to (select all that apply):
Family measures	*Assumes classical mother and father pair for reporting parental variables*	*Adapts questions to flexible familial units that do not assume gender, relationship, or number of parents and guardians.*
Race, ethnicity, and culture	*Treats race and ethnicity as nearly interchangeable. Ignores culture.*	*Includes opportunities for students to note specific cultural or ethnic identities. Separates ethnicity from the country of birth or residency.*
Socioeconomic status	*Uses direct measures:* What is your socioeconomic status? What is your parents' income? What is your parents' education? What is your high school zip code?*	*Builds latent, composite, measures. Uses measures that students have knowledge of* Are you or your parents taking out loans to fund your education? What is parent #1's maximum education?

* Options which sometimes appear as attempts to improve over typical demographic data measures

Discussion and Conclusions

The suggestions and examples above detail opportunities for researchers to adjust and improve the collection of demographic data in engineering education research. We have included several specific examples of ways to improve demographic data collection, but we strongly note that the suggestions must be treated within the unique context of each research project. Though often treated as an afterthought, demographic data collection requires researchers to make a serious effort to understand the nuances of a student's individual identity. Researchers must make sure

that they correctly frame items that capture the diversity of students' paths to, and through, engineering. Researchers must also respect differences between normative and non-normative groups, and not treat multiple dimensions of identity and demography as unlinked or binary. Demographic information should be collected at the end of studies to issues related to stereotype threat. When at all possible (e.g., using electronic survey tools) items and prepopulated categories should be flexible to allow for modern conceptions of demographics, such as more complex familial histories and personal identities.

In addition to offering suggestions for the overarching use and implementation of demographic questions, we also suggest improvements for including multiple dimensions of demographic identity. We have listed these suggestions in Table 1. In general, the suggestions focus on creating flexible questions that do not implicitly bias or otherize demographic identities outside of our disciplinary norms. We suggest that whenever possible researchers not only attempt to be comprehensive in their potential categories, but also allow students to adapt responses to fit their personal identity as closely as possible. In many cases the shifts will affect only small portions of research populations; however, these changes provide an opportunity to demonstrate greater respect and empathy for research participants while simultaneously improving the accuracy and validity of studies. As engineering education continues to improve its inclusion and development of diverse populations, we must make use of tools to account accurately for the rich and valuable complexity that diversity brings to engineering.

Implications

While we initially framed this paper through a discussion of formal processes of demography, there are specific implications of more comprehensive and inclusive data collection for researchers in engineering education. At a general level, we are arguing for an awareness that these issues, and student identities, are more complex than they typically appear in research. This complexity, as reflected in historical demographic trends, is only increasing over time. Perhaps more importantly than 'keeping with the times,' the educational implications of these demographics are broader than simply counting people accurately. In this final section, we will discuss both concrete and more abstract implications of this work for educators and researchers. We also touch on the implications for the analysis of these demographics, but a full discussion is beyond the scope of this paper and will be addressed in our future work.

The immediate reaction of many researchers to these new demographic data collection methods may be that more categories make research more complicated. We believe the opposite is true. More categories for students to self-identify their backgrounds increase accuracy, reduce error, and (by extension) reduce complexity. Demographic questions that force students into certain groups are especially problematic. These forced responses can cause participants to skip a question, creating data that is missing not at random which must be treated as inherently biased[64]. Alternatively, students can select from within the available options to which they do not fully align, which fundamentally increases error.

A practical consideration of the outcomes of this paper is how these more comprehensive and inclusive demographic data will be analyzed. Many of the approaches traditionally used in educational research to compare different research outcomes by demographics can be used with these categorical responses. For example, in a regression analysis of an outcome, all of the

demographic questions could be used as dummy variables to control for all possible responses on demographic questions. This approach can have some limitations if there are not enough individuals in the sample within each dummy variable.

Another approach to analyzing these types of complex demographic questions is taking a person-centered approach to analyzing student responses on questions through data mining methods like cluster analysis, latent class analysis, or topological data analysis. These approaches allow grouping of students based on their responses to questions rather than by demographic data. After the analysis is complete, the demographic data can aid in illustrating the types of students that fall into each response type. We have taken this approach in our current research and have found that this method of analyzing demographic data allows for a more inclusive understanding of the types of students in engineering.

Similar implications for the conditioning and appreciation of demographics apply to qualitative studies as well. Going beyond the suggestions that we provide and engaging complete, student driven, descriptions of demographic identity provides an important opportunity to understand the richness that is at the heart of qualitative research. For example, a researcher seeking to understand students creativity may find that exploring SES more deeply (e.g., work during school, scholarship GPA expectations, parental career role and progression, or social hierarchical characteristics of a student's previous schools[65]) helps identify a better outcome space to find what influences students choices related to creativity. In the case of SES, understanding the different subdimensions and the complex interrelation between the social and economic status factors can have a large and lasting impact on analysis and findings as researchers identify commonalities and dissimilarities between students who may appear similar on a surface level.

Identifying diverse populations may be useful, beyond statistical analysis, for identifying opportunities for additional research. If a research study intended to examine the influence of gender on retention of engineering students, understanding that trans students may experience engineering differently than a cis student who identifies with their gender due to many factors (e.g., a faculty members lack of respect for their identity) is important to an effective and accurate analysis.

In collecting more inclusive demographic data, researchers much balance the need for more comprehensive and complex data with quantitative analysis. Our goal in this paper was to present some ways in which we have worked to measure demographics that achieve this balance and offer practical ways for researchers to use these measures in their own work.

References

1. ASEE Board of Directors. ASEE Statement on Diversity. 2014. http://www.asee.org/about-us/the-organization/our-board-of-directors/asee-board-of-directors-statements/diversity. Accessed July 12, 2015.
2. Yusuf F, Martins JM, Swanson DA. *Methods of Demographic Analysis*. Amsterdam, Netherlands: Springer Netherlands; 2014. doi:10.1007/978-94-007-6784-3.
3. Lew JJ. Statistical programs of the United States government. 2012. https://www.whitehouse.gov/sites/default/files/omb/assets/information_and_regulatory_affairs/12statprog.p

df. Accessed January 12, 2016.

4. United States Census Bureau. Demographic questions on the 2010 census. 2010. https://www.census.gov/history/www/programs/demographic/. Accessed January 8, 2016.
5. United States Census Bureau. The questions on the form. 2010. https://www.census.gov/2010census/text/text-form.php. Accessed January 8, 2016.
6. Snyder TD. *120 Years of American Education: A Statistical Portrait*.; 1993. http://nces.ed.gov/pubs93/93442.pdf.
7. Cowan CD, Hauser RM, Kominski RA, et al. *Improving the Measurement of Socioeconomic Status for the National Assessment of Educational Progress: A Theoretical Foundation*.; 2012. http://nces.ed.gov/nationsreportcard/pdf/researchcenter/Socioeconomic_Factors.pdf.
8. National Forum on Education Statistics. *Forum Guide to Alternative Measures of Socioeconomic Status in Education Data Systems*. Washington, D.C.; 2015. http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2015158. Accessed April 3, 2016.
9. Census History Staff. 1790 Overview - History. https://www.census.gov/history/www/through_the_decades/overview/1790.html. Accessed March 15, 2016.
10. US Census Bureau. About the census. https://www.census.gov/topics/population/race/about.html. Accessed March 15, 2016.
11. Riley D. *Engineering and Social Justice*. Vol 3. 2008. doi:10.2200/S00117ED1V01Y200805ETS007.
12. Tonso KL. Engineering gender - gendering engineering: A cultural model for belonging. *J Women Minor Sci Eng*. 1999;5(4):365-405.
13. Downey GL, Lucena JC. Knowledge and professional identity in engineering: code-switching and the metrics of progress. *Hist Technol*. 2004;20(4):393-420. doi:10.1080/0734151042000304358.
14. Slaton AE, Alice LP. The power and politics of STEM research design: Saving the "small N." In: *ASEE Annual Conference and Exposition*. Vol Seattle, WA; 2015.
15. Burke PJ, Stets JE. *Identity Theory*. Oxford, UK: Oxford University Press; 2009.
16. Crenshaw K. Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stanford Law Rev*. 1991:1241-1299. http://www.jstor.org/stable/1229039.
17. Schwartz PM, Solove DJ. Pii Problem: Privacy and a new concept of personally identifiable information, The. *New York Univ Law Rev*. 2011;86:1814-1829.
18. Narayanan A, Shmatikov V. Myths and fallacies of "personally identifiable information." *Commun ACM*. 2010;53(6):24. doi:10.1145/1743546.1743558.
19. Steele CM. A threat in the air: how stereotypes shape intellectual identity and performance. *Am Psychol*. 1997;52(6):613.
20. Spencer SJ, Steele CM, Quinn DM. Stereotype threat and women's math performance. *J Exp Soc Psychol*. 1999;35(1):4-28.
21. Steele CM, Aronson J. Stereotype threat and the intellectual test performance of African Americans. *J Pers Soc Psychol*. 1995;69(5):797.
22. Rubin DB. Multiple imputation after 18+ years. *J Am Stat Assoc*. 1996;91(434):473-489.
23. Meade AW, Craig SB. Identifying careless responses in survey data. *Psychol Methods*. 2012;17(3):437-455. doi:10.1037/a0028085.
24. Sexual Minority Assessment Research Team (SMART). *Best Practices for Asking Questions about Sexual Orientation on Surveys*.; 2009. http://williamsinstitute.law.ucla.edu/wp-content/uploads/SMART-FINAL-Nov-2009.pdf.
25. Pew Research Center. What census calls us: A historical timeline. *Pew Res Cent*. 2015.

http://www.pewsocialtrends.org/interactives/multiracial-timeline/.

26. Cohn D. Census considers new approach to asking about race- by not using the term at all. *Pew Research Center,*. http://www.pewresearch.org/fact-tank/2015/06/18/census-considers-new-approach-to-asking-about-race-by-not-using-the-term-at-all/. Published June 18, 2015.

27. Thompson JH. Preparing for the 2020 census: Measuring race and ethnicity in America. *Dir Blog*. 2014. http://directorsblog.blogs.census.gov/2014/10/06/preparing-for-the-2020-census-measuring-race-and-ethnicity-in-america/.

28. International students in the US. *Open Doors Rep*. 2016. http://www.iie.org/Services/Project-Atlas/United-States/International-Students-In-US.

29. Pew Research Center. *Multiracial in America: Proud, Diverse, and Growing in Numbers*.; 2015. http://www.pewsocialtrends.org/2015/06/11/multiracial-in-america/. Accessed January 28, 2016.

30. Keller G. The new demographics of higher education. *Rev High Educ*. 2001;24(3):219-235. doi:10.1353/rhe.2001.0004.

31. Popenoe D. American family decline 1960-1990: A review and appraisal. *J Marraige Fam*. 1993;55(3):527-542. doi:10.2307/353333.

32. Simmons DR, Martin JP. Developing effective engineering fictive kin to support undergraduate first-generation college students. *J Women Minor Sci Eng*. 2014;20(3):279-292.

33. Martin JP, Simmons DR, Yu SL. Family roles in engineering undergraduates' academic and career choices: Does parental educational attainment matter. *Int J Eng Educ*. 2014;30(1):136-149.

34. Hsu M-C, Cardella ME, Purzer S. Parents' Perceptions of and Familiarity with Engineering Introduction. In: *ASEE Annual Conference and Exposition*. Vol Vancouver, BC; 2011.

35. Godwin A, Potvin G, Hazari Z. Do Engineers Beget Engineers? Exploring Connections Between the Engineering-related Career Choices of Students and Their Families. In: *ASEE Annual Conference and Exposition*. Vol Indianapolis, IN; 2014:16.

36. Lofquist D. Same-sex couple households. 2011;(September):1-4. https://www.census.gov/prod/2011pubs/acsbr10-03.pdf.

37. Office of the Spokesman. Consular report on birth abroad certificate improvements. 2012. http://www.state.gov/r/pa/prs/ps/2010/12/153636.htm.

38. Bendery J. Hilary Clinton fumed about changing passports for same-sex parents in 2011. *Huffington Post*. http://www.huffingtonpost.com/entry/hillary-clinton-same-sex-passports_us_560c68e6e4b076812700bf06. Published October 1, 2015. Accessed January 1, 2015.

39. Walpole M. Socioeconomic status and college: How SES affects college experiences and outcomes. *Rev High Educ*. 2003;27(1):45-73. doi:10.1353/rhe.2003.0044.

40. Smith KM, McCann CW. *The Validity of Students' Self-Reported Family Incomes*.; 1998. http://files.eric.ed.gov/fulltext/ED422830.pdf.

41. Donaldson K, Lichtenstein G, Sheppard S. Socioeconomic status and the undergraduate engineering experience: Preliminary findings from four American universities. In: *ASEE Annual Conference and Exposition*. Vol Pittsburgh, PA; 2008.

42. Entwistle DR, Astone NM. Some practical guidelines for measuring youth's race/ethnicity and socioeconomic status. *Child Dev*. 1994;65(6):1521-1540.

43. Hauser RM. Measuring socioeconomic status in studies of child development. *Child Dev*. 1994;65(6):1541-1545.

44. Fidelity. Time to break open the piggy bank: Parents expect kids to pay for more than one-third of college costs. http://www.fidelity.com/inside-fidelity/individual-investing/time-to-break-open.

45. Pew Research Center. *Most Say Government Policies since Recession Have Done Little to Help Middle*

Class, Poor.; 2015. http://www.people-press.org/2015/03/04/most-say-government-policies-since-recession-have-done-little-to-help-middle-class-poor/.

46. United States Census Bureau. *Selected Measures of Household Income Dispersion: 1967 to 2014*.; 2015. https://www.census.gov/hhes/www/income/data/historical/inequality/Table IE-1.pdf.

47. Conron KJ, Mimiaga M, Landers S. *A Health Profile of Massachusetts Adults by Sexual Orientation Identity: Findings from the 2001-2006 Massachusetts Behavioral Risk Factor Surveillance System Surveys. Report Prepared for the Massachusetts Department of Public Health*.; 2008.

48. Case P, Austin SB, Hunter DJ, Willett WC, Malspeis S. Disclosure of sexual orientation and behavior in the nurses' health study II: Results from a pilot study. *J Homosex*. 2006;51(1):13-31.

49. Trytten D a., Lowe AW, Walden SE. "Asians are Good at Math. What an Awful Stereotype ": The Model Minority Stereotype 's Impact on Asian American Engineering Students. *J Eng Educ*. 2012;101(3):439-468.

50. Chubin D, Donaldson K, Olds B, Fleming L. Educating Generation Net—Can U.S. Engineering Woo and Win the Competition for Talent? *J Eng Educ*. 2008;97(3):245-257.

51. McNair L, Newswander C, Boden D, Borrego M. Student and Faculty Identities in Self-Managed Teams. *J Eng Educ*. 2011;100(2):374-396.

52. Education E. Investigating the Teaching Concerns. *J Eng Educ*. 2007;(October).

53. Daly SR, Mosyjowski E a., Seifert CM. Teaching creativity in engineering courses. *J Eng Educ*. 2014;103(3):417-449. doi:10.1002/jee.20048.

54. Borrego M, Foster MJ, Froyd JE. Systematic literature reviews in engineering education and other developing interdisciplinary fields. *J Eng Educ*. 2014;103(1):45-76. doi:10.1002/jee.20038.

55. National Science Foundation. *Women, Minorities, and Persons with Disabilities in Science and Engineering*. Washington, D.C.; 2013. http://www.nsf.gov/statistics/wmpd/2013/pdf/nsf13304_digest.pdf. Accessed April 3, 2016.

56. U.S. Department of Education. National Center for Education Statistics. Students with disabilities. *Dig Educ Stat*. 2015. https://nces.ed.gov/fastfacts/display.asp?id=64.

57. Raue K, Lewis L. *Students with Disabilities at Degree-Granting Postsecondary Institutions*. Washington, D.C.; 2011. http://nces.ed.gov/pubs2011/2011018.pdf.

58. United States Census Bureau. American community survey. 2016. https://www.census.gov/people/disability/methodology/acs.html. Accessed January 22, 2016.

59. U.S. Department of Education. National Center for Education Statistics. The nation's report card. 2009. https://nces.ed.gov/nationsreportcard/pdf/studies/2009453_5.pdf. Accessed January 3, 2016.

60. National Survey of Student Engagement. Getting to know the updated NSSE. 2013. http://nsse.indiana.edu/nsse-update/. Accessed January 22, 2016.

61. Madriaga M. Enduring disablism: Students with dyslexia and their pathways into UK higher education and beyond. *Disabil Soc*. 2007;22(4):410.

62. Davis LJ. Why is disability missing from the discourse on diversity? *Chron High Educ*. 2011:4. http://chronicle.com/article/Why-Is-Disability-Missing-From/129088/. Accessed January 22, 2016.

63. Heshusius L. More than the [merely] rational: Imagining inquiry for ability diversity. *Disabil Cult Educ*. 2002;1(1):95-118.

64. Tabachnick B, Fidell L. *Using Multivariate Statistics*. Vol 6th ed. Upper Saddle River, NJ: Pearson; 1996. doi:10.1037/022267.

65. Anyon, J Social class and the hidden curriculum of work, *The Journal of Education*, vol. 162, no. 1, pp. 1–11, 1980.

66. Merrill P. Black and white: Why capitalization matters. *Columbia J Rev*. 2015. http://www.cjr.org/analysis/language_corner_1.php. Accessed March 4, 2016.