Why is Data on Disability so Hard to Collect and Understand?

Brianna Blaser  
DO-IT Center  
University of Washington  
Seattle, WA, USA  
blaser@uw.edu

Richard E. Ladner  
Paul G. Allen School of Computer Science and Engineering  
University of Washington  
Seattle, WA USA  
ladner@cs.washington.edu

Abstract—This report explores why data on disability is so hard to collect and understand. There is a reluctance to collect data by the broadening participation in computing (BPC) community even though disability is a recognized demographic in broadening participation. As a result, much less is known about the participation of people with disabilities in computing education and careers than some other groups. The reasons for this reluctance are multi-faceted. Data about disability can vary significantly depending on how you ask the question, making it difficult to understand what the data is actually indicating. Questions about functional limitations may overinflate the numbers of some populations with disabilities while undercounting other groups. This report will help the BPC community move towards collecting and reporting data on disability. Advice is given on how to ask about disability status and resources are provided to find existing data sources about disability.

Keywords—disability, data, survey questions

I. INTRODUCTION

Those of us who work on broadening participation in computing (BPC) are to a great degree data driven. We track the percentage of women and underrepresented minorities who are enrolled in and graduate from computer science programs, and who take high school computer science courses in order to see if our interventions are actually working. For disability, however, it is a different story. It seems that there is reluctance in our community to tracking the participation of students with disabilities even though approximately 15% of K-12 students and 11% of college students have a disability.

To drive this home, we recently examined the “2019 State of Computer Science Education” report, which reviews how states are doing in implementing K-12 computer science education [1]. Indeed, the report is subtitled “Equity and Diversity” to emphasize how the organizations sponsoring the report—the Code.org Advocacy Coalition, the Computer Science Teachers Association, and Expanding Computing Education Pathways—view the importance of BPC. The report is full of relevant data about women, underrepresented minorities, and rural students, but there is nothing about students with disabilities. The word “disability” does not appear in the report. It’s as if people with disabilities don’t exist or matter when thinking about BPC.

Other organizations that play a central role in BPC are also not collecting or reporting data on disability. The well-known Computing Research Association (CRA) Taulbee Survey on the state of computer science departments does not collect disability data, although it does collect data on gender, race, and ethnicity [2]. The College Board does not ask Advanced Placement exam takers if they have a disability. While they do have data about disability-related accommodations requests from exam takers, they do not publish that information despite the fact that they publish information about the participation of girls and underrepresented minorities [3].

Disability is mostly absent in broadening participation research, policy, and practice [4]. Quantitative efforts rarely ask about disability and if they do, they don’t analyze the data with respect to disability. [4, 5, 6]. Edlyn Vallejo Peña states “When scholars, researchers, and editors of top-tier journals do not engage in or include scholarship on students with disabilities, even if unintentionally, they communicate that understanding these needs and interests is less important than other issues in higher education” [6, p. 38]. We find this statement particularly powerful. When the BPC community does not engage on issues related to disability, it is problematic.

One notable exception is the CRA’s Data Buddies [7]. Data Buddies collects data from students and faculty members in computing departments nationwide and shares customized department reports with each institution about their students. Among the items on the survey, is a question about disability. Departmental reports include information about the number of students with disabilities in a department and the types of disabilities they have. Departments can also request to have their students’ data broken down by disability.

The Americans with Disabilities Act defines disability as “a physical or mental impairment that substantially limits one or more major life activities, a person who has a history or record of such an impairment, or a person who is perceived by others as having such an impairment” [8]. The social model of disability views disability as the “limit or loss of opportunities to take part in community life because of physical and social
barriers” [9, pp. 103]. This is a heterogeneous group including 
individuals who are blind or visually impaired, who are deaf or 
hard of hearing, who are autistic or have learning disabilities, 
who have mobility-related disabilities, who have mental health 
disabilities, and more.

The purpose of this paper is to explore reasons why we are not 
collecting high quality data on disability and offer suggestions 
on how the BPC community can begin doing so, either by 
collecting that data directly or by employing various national 
data sources. We believe that all of us who are working in BPC 
do care and consider people with disabilities to be a group that 
is disadvantaged and worthy of receiving interventions to 
increase their participation. And indeed, there is evidence that 
people with disabilities are underrepresented in employment 
and in STEM education. The U.S. Department of Labor reports 
that in 2018 only 30.4% of persons aged 16 to 64 who have a 
disability are employed as compared to 74.0% for those without 
a disability [10]. The 2015-16 Civil Rights Data Collection 
STEM Course Taking Report (in their Figure 10) shows that 
students with disabilities are underrepresented in STEM 
courses taught in high schools around the nation [11].

We can think about disability data as existing on two levels: (1) 
large, publicly-available existing data collections and (2) data 
collected within a research project, department, or organization. 
Many existing data collections, such as the Survey of Earned 
Doctorates and the Census Bureau’s American Community 
Survey, do collect information about disability. Beyond those 
datasets, those of us working in BPC should be empowered to 
collect information regarding disability within our projects, 
departments, or institutions and consider whether individuals 
with disabilities are successfully participating. For a particular 
project or institution, data can be obtained in a consistent 
manner to track progress at an institution or on a particular 
intervention.

In the following sections we cover reasons why disability status 
data is not collected, problems with existing data collections, 
operationalization of disability, and best practices in collecting 
disability status data.

II. WHY DISABILITY STATUS DATA IS NOT COLLECTED

Through our involvement in AccessComputing, one of the 
National Science Foundation’s BPC Alliances, we have asked 
our colleagues in the BPC community why they aren’t 
collecting data on disability. Answers vary, but include the 
following:

- We don’t know who to ask for the data: the school, the 
teacher, or the individual.
- Data on disability is considered to be confidential, so 
we don’t ask.
- We don’t know what to ask. There seems to be no 
standard definition of disability.
- Disability data is unreliable because people with a 
disability might not want to disclose due to stigma.
- Our institution doesn’t systematically collect data 
about disability.
- Our institutional data isn’t accurate because we know 
people don’t disclose to disability services.
- The information is difficult to access when obtaining 
data from institutions.

These are not sufficient reasons to ignore disability altogether 
in our broadening participation efforts.

The technical notes of the National Science Foundation report 
*Women, Minorities, and Persons with Disabilities in Science 
and Engineering* echoes many of these issues, including an 
inconsistent definition of disability, lack of institutional 
records, difficulties with self-reported data, and issues with 
changing the questions asked on a variety of surveys [12].

III. PROBLEMS WITH DATA COLLECTIONS: A CASE STUDY

The United Nations, governmental agencies, and educational 
institutions at all levels collect and report on disability. In the 
U.S., the Census Bureau, the National Center for Education 
Statistics (NCES), the National Center for Science and 
Engineering Statistics (NCSES), the Civil Rights Data 
Collection (CRDC) and the Bureau of Labor Statistics (BLS) 
all collect data on disability, although not always in the same 
way. In these data collections and elsewhere, the incidence rates 
of disability will vary depending upon how disability is defined 
and how questions related to disability are asked [13]. The 
World Bank has reports on problems with the consistency of 
disability data internationally [14].

Even though the word “disability” has been accepted 
nomenclature since at least 1990 when the Americans with 
Disabilities Act was passed, there is an international trend to 
not use the word “disability” itself in survey questions, but to 
ask instead about functional limitations. The consequences of 
moving from one nomenclature to another are apparent in the 
data from the Survey of Earned Doctorates (SED) prepared by 
NCSES. This survey has been given to all students graduating 
with a doctorate in the US every year since 1957 and reports are 
released annually. Figure 1 shows 14 years of data, tracking the 
percentage of doctorates in all fields who have a disability.

As you can see, the percentage of doctorates with disabilities 
rose from 1.5% in 2004 to 7.2% in 2017, a remarkable increase 
over such a short period. If you look more closely, there are two 
steps in increase; first in 2010 and again in 2012. Furthermore, 
the percentage was a fairly steady around 1.5% between 2004 
and 2009, jumped in 2010, and rose steadily between 2012 and 
2017. How can this chart be explained? The principal
The explanation is that the question about disability changed twice, first in 2010 and then again in 2012.

Figures 2, 3, and 4 have the questions asked between 2004 and 2009, in 2010-2011, and between 2012 and 2017, respectively.

For the years 2012-2017, respondents were designated as disabled if the degree of difficulty of completing a task was moderate, severe, or unable to do. This question gathers data about functional limitations instead of disability identity, which is what was asked about in prior years.

A closer look demonstrates the problem further. According to NCES (Table 204.30), about 0.1% of children age 3 to 21 students have a visual impairment [15]. According to the SED, in 2011, 0.6% of new doctorates were “blind/visually impaired” and in 2017, 3.1% of new doctorates had “visual limitations” from moderate to unable to do. How can 0.1% become 3.1%, more than a factor of 30, unless they are counting fundamentally different people? It should be noted that the SED is self-reported data, while NCES data is from schools who provide services through the Individuals with Disabilities Education Act (IDEA). As the IDEA students have to document their disabilities to receive services and it would be unlikely that there are people with visual impairments who are not receiving accommodations, this number should be fairly accurate.

This problem with using functional limitation questions is apparent when looking at people who are unemployed due to “chronic illness or permanent disability.” In 2017 National Survey of College Graduates (NSCG) data 700,000 scientists and engineers who are unemployed due to chronic illness or permanent disability, 178,000 (25.4%) are identified as not having a disability because of the way they answered questions related to functional limitations [16]. The ADA definition of disability considers someone to be disabled if they have a condition that impacts a major life activity, including working. By that definition, 100% of people unemployed due to chronic illness or permanent disability have a disability.

IV. UNDERSTANDING EXISTING SURVEYS ON DISABILITY

Many other surveys have moved away from the use of the term “disability” [17]. The SED and organizations such as the United Nations and the US Census Bureau’s American Community Survey (ACS) all use questions that measure functional limitations rather than those that measure disability identity in their surveys. Unfortunately, there is no agreement on the functional limitation language. For blindness, the UN and ACS use the language “serious difficulty seeing,” while SEC uses the range “moderate” to “unable to do.” The situation is so confusing that the National Federation of the Blind simply reports all the different percentages on blindness on their web page addressing data [18].

We admit that we do not fully understand the movement away from “disability” nomenclature. It could be motivated by the word “disability” having a stigma attached. Indeed, the guidelines recommended by the UN explicitly state not to use...
the word “disability’ or other negative terms and phrases” [19]. Many people with disabilities would cringe at the assertion that “disability” is a negative term. The person-first language, for example using “person with a disability” instead of “disabled person,” seems to imply that “disability” is a secondary trait that is stigmatized. Organizations of (not for) people with disabilities tend not to use person-first language because they do not believe that their disability is stigmatizing. Examples of such organizations are the National Federation of the Blind, National Association of the Deaf, and the Autism Self Advocacy Network. Increasingly, disabled people are embracing disability-first language, an indicator that “disability” is not as stigmatizing as it once was [20, 21, 22].

For a number of reasons, we do not recommend adopting the functional limitations measures. First, there is no agreement on how to ask the question. Is “moderate difficulty,” “severe difficulty,” or “serious difficulty” the standard for having a disability? What set of functional limitations do you focus on? Second, the concept of functional limitation is not one that people readily understand. It seems to apply nicely to seeing, hearing, and walking, but what about other disabilities such as learning disabilities, attention deficit, anxiety, and autism. The percentage rise from 2012 and 2017 in Figure 1 demonstrates the confusion about the question. Clearly, the understanding of this question by new doctorates changed during this period. Although the meaning of the word “disability” has changed over time, it is relatively stable and far less stigmatizing than it was in the past.

And yet, arguments for asking about functional limitations seems compelling. These disagreements are part of what make collecting data on disability difficult. Internationally, surveys that ask someone if they identify as disabled have the lowest rates of disability [17]. The UN Washington Group on Disability Statistics argues that even though it will miss some people with disabilities asking about functional limitations will identify the majority of them [19]. Furthermore, advocates for these questions argue functional limitation questions fit within a social model of disability that views disability as an interaction between environment and a person whereas a question asking about disability fits within a medical model of disability that views disability as a problem within the person [13].

Overall, there’s a lack of clarity when examining the definitions of disability as used in theoretical models, legal terminology, and lay usage [23]. Indeed, sociologist Barbara Altman argues: Disability is a complex social phenomenon that involves the interaction of individuals with specific limitations, possibly due to a physical impairment, with their identity, their physical and social surroundings, and the societal culture and normative systems. Disability, then, is undefinable empirically unless one reduces the focus of the definition to a specific aspect of experience [23, p. 80]. In BPC research, we need to determine what aspects of disability we are interested in when constructing questions about disability status.

V. DEFINING DISABILITY

There is a need to define disability in order to identify who should receive services or accommodations. As a result, in education and employment settings, conversations about the representation of people with disabilities and conversations about services or accommodations are necessarily intertwined. Generally, students and employees are identified as having a disability if they qualify for and receive accommodations like documents in alternative formats, extended time on tests, flexible work hours, or interpreting services. To qualify, a person must have documentation of their disability by a
recognized professional. Many individuals who have a disability that do not register with disability services for a variety of reasons.

A. PreK-12 Education

In the PreK-12 education setting there is more uniformity with regard to what constitutes a disability than in other settings because of two laws: the Individuals with Disabilities Education Act (IDEA) [24] and Section 504 of the Rehabilitation Act of 1973 [25]. Students with a documented disability may receive accommodations under either law. Under IDEA, the student will have an Individualized Education Program (IEP) that includes specifics about the students’ education goals and accommodations that a student will receive. The education goals and accommodations are designed to meet the individual needs of the student. The education goals may be different than those of other students in the same school. A student with a Section 504 Plan has the same education goals as other students, but may need some accommodations to help meet those goals. Accommodations may include extra time on exams because of a learning disability, screen reader access to computers because of blindness, or a sign language interpreter because of deafness. The student, their parents, teachers, and administrators all know if they are covered under IDEA or Section 504.

IDEA requires states to collect data about PreK-12 students with disabilities [26]. The number of students covered under IDEA and Section 504 has varied over time, but is relatively stable. One striking change has been the number of students with autism who have been identified over the past 10 years. In 2008-9 the number was 336,000 nationwide and in 2017-8 the number had almost doubled to 710,000 [14]. There is also some variation in the application of these laws by state. For example, in 2017-8, in Texas only 9.2% of students are covered under IDEA, while the percentage in Pennsylvania is 18.6% [14].

In addition to NCES, the Office for Civil Rights in the Department of Education maintains data in its Civil Rights Data Collection (CRDC) [27]. Reports are published annually on the student demographics, including disability status, of participation in STEM courses. Since 2017-2018, data on computer science courses has been collected, but has not yet been reported on.

As states implement policy related to PreK-12 computer science (CS) education, some are enacting laws that require data collection about computer science education in their state. In Washington State, HB 1577 requires data collection about CS education, including demographic data of students enrolled in CS courses starting in 2020 [28]. Specifically, the special education status of students enrolled in CS courses must be reported. We hope to see other states follow suit to collect data about all students with disabilities taking PreK-12 CS courses.

B. Postsecondary Education

Students with disabilities at the postsecondary level are guaranteed access by Section 504, but not under IDEA. The educational goals for these students are the same as for the rest of the student body. Colleges are required by law to provide accommodations for students with disabilities; however, unlike in K-12 education, students must request those accommodations from an office of disability resources for students (DRS). To receive accommodations the disability must be documented. Students who are employed as a teaching assistant, research assistant, or otherwise on campus and need accommodations for their employment may need to request those services from an office of disability resources for employees. In such situations, disabled students may have to deal with two offices.

Unlike PreK-12, there are no regulations requiring colleges to report on students with disabilities [26]. Typically, in their offices of institutional research, colleges track data about disability. There is likely data about whether a student received accommodations. Some colleges ask about disability in the college application, in which case that data is in a student’s record as well. Naturally, an individual student’s data is protected by the Family Educational Rights and Privacy Act (FERPA). Publishing of aggregate data, however, does not violate FERPA. Nonetheless, colleges rarely publish aggregate data about students with disabilities whereas they often do publish data about gender and minority status.

Regardless, aggregate data about disability may be available by request to the office for institutional research. Computing departments can track this data as part of the work they are doing related to BPC in their own departments. In working with partners at institutions across the country, we have found that many computing departments struggle to obtain and report on institutional data. Some institutions did not agree to provide information to faculty, while others blamed updates of databases, DRS records that did not include information about students’ majors, or paper-based systems within their DRS office. Even if you are able to successfully obtain institutional data, it will necessarily undercount students with disabilities. Some students in postsecondary settings are unable to get adequate documentation, which can be costly to procure, in order to get accommodations [29]. Other students fail to disclose because of concerns around stigma [30, 31].

In 2016, 88% of students who participated in the Data Buddies survey disclosed disability information. 8% indicated that they have a disability [32]. The overwhelming majority (68%) had cognitive disorders (attention deficits, autism, or mental health disabilities). 12% of students had learning disabilities and 12% reported sensory disabilities (vision, hearing, or speech-related disabilities). In 2017 data from Data Buddies, students with disabilities were significantly more likely to report feeling like an outsider (32%) than majority men without disabilities (17%) [33]. Women with disabilities (46%) and underrepresented minorities with disabilities (45%) were the most likely to report feeling like an outsider. Findings such as this highlight the importance of including disability in conversations about BPC.
C. Employment

Section 504 and the ADA protect the rights of people with disabilities in employment. Employers must provide reasonable accommodations to employees with disabilities. Like in the postsecondary setting, to receive accommodations disabled people need to disclose their disability to their employer and provide documentation. Some people with disabilities may decide not to officially disclose to their employer because they do not require accommodations in a work setting.

The 2017 Survey of Doctorate Recipients found that among doctorate holders employed in universities and 4-year colleges working as computer and information scientists, 8.5% have a disability [34]. This is comparable to the 9.4% of doctorate holders that have a disability across all occupations. The 2017 NSCG found that 9.1% of employed computer and information scientists have a disability compared to 10.3% of all employed scientists and engineers [16]. Note that this survey uses functional limitation questions to ask about disability status.

Recent years have seen several large tech companies release diversity reports about their workforce. Until very recently, very few of these reports contained information about disability. A 2016 article from TechCrunch interviewed seven major companies, none of whom had included information about disability in their diversity reports [35]. Many companies pointed to inclusion efforts related to disability or their work in accessibility but refrained from talking about why they weren’t reporting on disability representation in their companies. In the time since that article was written, more companies—including some interviewed for the TechCrunch article—have released disability-related data.

- Clover has reported that company-wide, 8% of their employees have a disability [36].
- Google has reported that of the 39% of global employees who self-identified, 7.5% have a disability [37].
- LinkedIn has reported that 1.6% of US employees had a disability [38].
- Slack has reported that 1.7% of their U.S. workforce identified as having a disability but notes that this statistic is based on the small number of employees who have chosen to disclose [39].

The variability between these data points to some of the difficulties with collecting this information. Different questions or different data sources could lead to wildly different data about the prevalence of disability within these companies.

Beyond data available from companies, Stack Overflow issues an annual survey of developers [40]. In 2019, they received 90,000 responses, 18.0% of respondents reported a mood or emotional disorder, anxiety disorder, concentration or memory disorder, or autism. In terms of physical disabilities, 1.5% blind or difficulty seeing, 0.8% deaf or hard of hearing, 0.3% difficulty walking or standing, and 0.3% difficulty typing. Because of the way this question is asked, it’s not clear whether all of these individuals have a disability. For example, not everyone with an anxiety or mood disorder necessarily has a disability, although there are certainly some of them may have a mental health disability.

VI. BEST PRACTICES FOR COLLECTING DISABILITY DATA

We encourage computer science education researchers to ask about disability when they are collecting demographic data from participants. Recent work has explored best practices for asking about gender on surveys in the area of human computer interaction [41]; similar recommendations with regard to disability could move forward conversations about individuals with disabilities in computing education and employment. Using a variety of measures to inquire about disability status is problematic in that different measures will classify individuals differently [23, 42]. As a result, it’s important that the BPC community adopt consistent measures.

When working with preK-12 students, researchers can ask whether a student has an IEP or 504 Plan or explore the use of institutional records to obtain this information. Teachers will know whether students in their classes have an IEP or 504 Plan. In the context of CS education, we are interested in asking about disability status in order to determine how a disability has impacted a student’s education. Since most preK-12 students whose education is impacted by a disability should have an IEP or 504 Plan, asking about that is sufficient in identifying students with disabilities in preK-12.

When working with adults, whether looking at students in computing or teachers participating in professional development, the situation is a bit more complicated. As discussed, for a variety of reasons adults may not use accommodations in the context of postsecondary education or employment, so a similar measure wouldn’t be sufficient. We feel strongly about avoiding functional limitation questions because of the potential to overestimate some populations, namely those with vision impairments, while also underestimating other populations, including autistic individuals.

Part of the difficulty of determining who has a disability and who does not is that there isn’t a black and white line between the groups but rather a fair amount of grey area. Some individuals with mental health diagnoses may identify as disabled whereas others do not. Individuals who have age-related disabilities may hesitate to identify as disabled.

Based on our own experiences and conversations with other experts in disability, people with disabilities, and other stakeholders, we suggest using the following two questions when asking adults about their disability status:

1. Do you identify as having a disability or other chronic condition?
   a. Yes
   b. No
   c. Prefer not to disclose
2. How would you describe your disability or chronic condition?
   a. Attention deficit
   b. Autism
   c. Blind or visually impaired
   d. Deaf or hard of hearing
   e. Health-related disability
   f. Learning disability
   g. Mental health condition
   h. Mobility-related disability
   i. Speech-related disability
   j. Other (please specify)

We feel that addressing disability directly in the question is important. Asking the second question gives the research richer information for analysis. Alternatively, individuals could write-in their specific disability. In our experience, however, this approach can make it difficult to categorize disabilities that may affect individuals in multiple ways. Asking about disability-type is also less invasive than asking someone to specify their exact diagnosis.

A. Where to Find Data

National data on disability is scattered in various tables and reports some of which were mentioned earlier. Principal data sources include the following:

1. National Center for Education Statistics maintains data on PreK-12 education. [43]. NCES periodically publishes a report The Condition of Education that covers students with disabilities. They also publish a report titled the Digest of Education Statistics that has data about postsecondary students.
2. National Center for Science and Engineering Statistics maintains data on postsecondary education [44]. They administer the Survey of Earned Doctorates [45]. They also maintain an interactive tool for creating custom tables from their data [46]. Unfortunately, this tool is limited and doesn’t allow access to disability data.
3. The Department of Education Civil Rights Data Collection (CRDC) includes information about students with disabilities taking various STEM courses nationally [27]. Starting in 2017 they are collecting data on computer science courses.
4. NSF Committee on Equal Opportunities in Science and Engineering has biennial reports to Congress that have data on disability [47].
5. Bureau of Labor Statistics has data on employment including data about disabilities [48].
6. Census Bureau’s American Community Survey has general data on disability [49].
7. Centers for Disease Control and Prevention has some data and advice on how to ask about disability status [50].
8. Cornell University Disability Statistics reviews several resources of U.S. disability statistics [51].
9. Disabled World web site reviews disability statistics from a number of resources [52].

VII. Conclusion

Over the last couple of decades, the BPC and computer science education communities have spent a significant amount of time collecting data about and researching the involvement of women and underrepresented minorities in computing education. We’d like to move towards a future where people with disabilities are more substantively included in conversations about diversity in computing. To do this, we need more people to collect data about disability status. We hope this paper can start a discussion about ways to get this data and serve as a call to action to the community.

References
