This chapter discusses issues and challenges encountered in using large-scale data sets to study educational experiences and subsequent outcomes for American Indian and Alaska Native (AI/AN) students. In this chapter, we argue that the linguistic and cultural diversity of Native peoples, coupled with the legal and political ways in which education has been used to acculturate and assimilate them, makes it imperative that a full range of social, cultural, and demographic variables be collected and that these data be analyzed, using a theory of change that emphasizes the strengths of AI/AN students and their communities rather than their perceived deficits. We envision this theory of change espousing the use of data to critically unpack and respond to the role of language and culture in shaping pathways to success in education and beyond for AI/AN students. We conclude with recommendations for constructing and analyzing large-scale data sets to better represent the diversity of cultures and experiences among AI/AN students.

Use of Large-Scale Data Sets to Study Educational Pathways of American Indian and Alaska Native Students

Susan C. Faircloth, Cynthia M. Alcantar, Frances K. Stage

Today nearly 3 million people in the United States, slightly less than 1% of the total population, identify as American Indian or Alaska Native (AI/AN) (Norris, Vines, & Hoeffel, 2012). Although often grouped together for ease of reporting and analyzing quantitative data, it is important to recognize that American Indians and Alaska Natives are in fact two distinct cultural groups, each encompassing its own unique cultures, histories, and languages (Caldwell et al., 2005). While combining American Indians and Alaska Natives provides useful data, it can also serve to mask important social and economic differences that significantly impact the overall well-being and life outcomes of individuals from these two groups.

As researchers, we are particularly interested in the ways in which data are collected and analyzed to describe and explain the educational experiences and subsequent academic outcomes for AI/AN students from
prekindergarten through postgraduate work (PK–20), institutionally as well as nationally. We situate this discussion within the larger domain of the social and institutional characteristics and the aims of elementary, secondary, and postsecondary education. While some practitioners and researchers (e.g., Tinto, 1987) use models of college success that emphasize students’ individual integration into college, both socially and academically, this notion has been challenged by other scholars who study the success of Native1 students in higher education (e.g., Pavel & Inglebret, 2007; Shotton, Lowe, & Waterman, 2013; Tierney, 1992). These scholars argue that rather than expecting AI/AN students to leave their cultures behind once they have entered institutions of higher education, these institutions must provide culturally relevant experiences for these students, thereby enabling them to be both academically and culturally successful.

This potential tension between students’ Indigenous cultures and the cultures of the institutions of higher education in which they are enrolled raises a number of questions regarding how institutions measure student success when implementing culturally relevant practices and pedagogies aimed at academically and socially engaging students. Other questions include: What are the common challenges to examining student success for AI/AN students? How can we examine the educational pathways of AI/ANs in order to identify their strengths and to determine areas in need of support?

In response to these and other questions, this chapter makes the case for the use of a critical quantitative approach to the use of large-scale national data sets to study the educational pathways and subsequent outcomes of AI/AN students. Research on academic success and postsecondary attainment is typically focused on either the kindergarten to grade 12 or postsecondary education system rather than connecting the PK–20 educational pathways of students. However, we recognize that education does not occur in a vacuum, and the educational conditions students encounter in elementary and secondary school serve to shape their aspirations, expectations, and eventual success along the pathway to college. In response, this chapter considers ways in which both PK–12 and postsecondary data sets together may be utilized to better understand the educational conditions and subsequent academic outcomes for Native students, as well as the challenges of using these data to study a culturally distinct, diverse, and historically marginalized student group.

We begin by reviewing the educational contexts in which AI/AN students attend school, followed by an overview of some of the PK–12 and postsecondary national data sets that include sufficient numbers of Native students, schools, and staff, primarily through the use of oversampling, to allow for quantitative analysis. We then discuss some of the limitations of these data and ways in which a critical quantitative approach may help to overcome these limitations. We conclude with recommendations for future research and implications for practice.
State of Education for American Indians and Alaska Natives

Any meaningful examination of the educational conditions and subsequent academic outcomes for American Indians and Alaska Natives requires, at minimum, a basic understanding of their sociopolitical history as it relates to education. Historically, education has been used as a means to civilize and Christianize Indigenous peoples worldwide (Deyhle & Swisher, 1997), resulting in the loss of Indigenous languages and, some would argue, cultures of many Indigenous groups, including those in the United States (Faircloth, 2009). One of the things that is unique about the education of Indigenous peoples in the United States is that many of these peoples’ ancestors ceded lands to the U.S. government in exchange for the provision of health, education, and basic well-being, provisions that have, more often than not, gone lacking.

Today, the majority (approximately 93%) of AI/AN students attend public schools while the remainder attend schools operated or funded by the Bureau of Indian Education and tribes, and a smaller percentage attend private schools (DeVoe & Darling-Churchill, 2008). Regardless of where they are educated, many AI/AN students experience cultural incongruence between home and school (Chrisjohn, Towson, & Peters, 1988; Lomawaima, 1995; Powers, Potthoff, Bearinger, & Resnick, 2003; Suina, 2001), making it difficult for them to reconcile their Indigenous language and culture with the Westernized academic culture of schools. Unfortunately, many of those who do succeed in school and go on to graduate are still unable to attend college due to either a real or perceived lack of access to necessary financial aid and other resources (e.g., Tierney, Sallee, & Venegas, 2007).

The lack of financial resources is due in large part to high rates of poverty among AI/AN persons. For example, nearly one third (31%) attend high-poverty schools (Ross et al., 2012), and 40% of American Indian children live at or below the poverty level (Children’s Defense Fund, 2014). Poverty is especially concentrated in homes with female heads of household, with 53% of American Indian children who reside in homes with female heads of household also living in poverty. This is an important point given the relationship between poverty and academic performance and the fact that students from lower socioeconomic statuses tend to perform more poorly in school than do their peers from higher socioeconomic statuses (e.g., Duncan, Yeung, Brooks-Gunn, & Smith, 1998; Sirin, 2005).

Although many AI/AN students do graduate from high school and go on to college, a disproportionate number do not pursue postsecondary education. Data indicate only 18% of all American Indian children have parents with a bachelor’s degree or higher compared to 59% of Asians, 44% of Whites, 20% of Blacks, and 16% of Latinos (Ross et al., 2012). These findings are important given the relationship between degree attainment and the potential to earn higher wages, with those graduating from college...
often earning more than those who do not (Faircloth & Tippeconnic, 2010). What is equally troubling is that more than half of those AI/AN students who go to college end up not graduating (Knapp, Kelly-Reid, & Ginder, 2012). This raises questions about the factors that account for student retention and graduation, as well as those for student aspirations to attend college.

Data from the 2009 National Indian Education Study (Mead, Grigg, Moran, & Kuang, 2010) indicate that 57% of American Indian eighth graders who participated in this study aspired to attend college; however, the 2011 National Indian Education Study found that 63% of AI/AN eighth graders who were surveyed had not met with their high school counselor to discuss their classes or their plans following high school (National Center for Education Statistics, 2012). This suggests a disconnect between AI/AN students’ aspirations and their preparation for college. This failure to seek college counseling may also help to explain Ross et al.’s (2012) finding that only 24% of American Indian males and 33% of American Indian females between the ages of 18 and 24 are enrolled in postsecondary education, at the undergraduate or graduate level. This is the lowest rate of college attendance of all racial/ethnic groups, with 43% of White males and 51% of White females, 31% of Black males and 42% of Black females, and 26% of Latino males and 35% of Latina females attending college. An analysis of Integrated Postsecondary Education Data System (IPEDS) data also showed that American Indian students earned less than 0.8% of all associate’s and bachelor’s degrees in the 2008–2009 academic year (Stage, Lundy-Wagner, & John, 2013). While 0.8% may not seem so different from the population percentage of 1%, it suggests a 20% lower achievement rate within the population. In other words, while AI/AN students represent 1% of the population, they earn only 0.8%, or 80% of their share, of the college degrees earned in the United States. In contrast, if Whites, who make up approximately 73% of the population, earned 0.2% less than their 73% representation, the difference would be almost meaningless.

**Critical Approach to the Study of American Indian and Alaska Native Students**

As mentioned earlier in this chapter, the historical, societal, and political marginalization of American Indians and Alaska Natives has rendered them, in many cases, invisible to those within the academic arena (Shotton et al., 2013). Critical quantitative approaches are needed to reverse this trend. As Stage (2007) has argued, “When models do not accurately reflect a given population’s experiences, the task is to pose alternatives to those models. Rather than focus on explanation, or fairness, [the modeling] focuses on equity concerns that can often be highlighted through analysis of large data sets” (p. 9). The Tinto (1987) model is a good case in point. This model forms the basis of much current research on college student success. It
was originally developed in the 1970s from studies of students enrolled in elite colleges. More recently, scholars have challenged aspects of the model and searched for culturally relevant variables for particular student groups (Brayboy, Fann, Castagno, & Solyom, 2012; Harper & Hurtado, 2007; Hurtado & Ponjuan, 2005). Unfortunately, many of these variables are absent from existing large-scale data sets used to examine the educational conditions and subsequent outcomes of students in PK–20.

### Large-Scale National Education Data Sets

The federal government has developed a collection of hundreds of national data sets, 32 of which are education data sets managed by the National Center for Education Statistics (NCES). Researchers use these national data sets to make inferences about populations and institutions, and policy makers have used these results to inform policy and practice. Thus it is imperative that all racial and ethnic groups be represented within these data sets. Unfortunately, because of the small sample size of AI/AN students in these data sets, studies often (a) exclude AI/AN students, (b) group them as “Other” or as students who identify as multiracial, or (c) fail to report any outcomes for this group after analysis because low numbers resulted in nonsignificant findings and/or low effect sizes.

Equally important is the fact that in the recent past, no national-level data were required to be collected for AI/AN students. For example, prior to 2003, the annual report commissioned by the U.S. Department of Education, *The Condition of Education*, did not report on AI/AN students. In 2003, the U.S. Office of Management and Budget (OMB), which “is responsible for the standards that govern the categories used to collect and present federal data on race and ethnicity” (Kena et al., 2014, p. vi), revised its guidelines to include the reporting of at least five racial categories including AI/ANs. Unfortunately, the lack of data prior to 2003 continues to limit longitudinal analysis of the educational pathways of AI/AN students.

In spite of this failure to systematically include AI/AN students in the bulk of the large-scale education data sets, a limited number of data sets have been constructed that do include AI/AN children, youth, and adults. In the following section, we describe some of these data sets and selected analyses that have been conducted involving AI/AN specific data. The focus here is on PK–20 education for American Indians and Alaska Natives.

**PK–12 National Data Sets.** Four national data sets have included sufficient numbers of AI/AN children, youth, and/or adults in PK–12 to allow for disaggregation and analysis of data. These data sets are discussed in brief next.

*Early Childhood Longitudinal Study, Birth Cohort (ECLS-B).* The ECLS-B contains an oversample of American Indian and Alaska Native children from 2001 to 2006 (oversampling involves creating a sample of respondents that exceeds their actual representation in the population; see
Table 1.1. Early Childhood Longitudinal Study Data Set, by Percentage Race/Ethnicity

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total (N)</td>
<td>10,700</td>
<td>21,400</td>
<td>18,200</td>
</tr>
<tr>
<td>White</td>
<td>41.12%</td>
<td>55.08%</td>
<td>46.65%</td>
</tr>
<tr>
<td>American Indian/Alaska Native</td>
<td><strong>2.80%</strong></td>
<td><strong>0.59%</strong></td>
<td><strong>2.09%</strong></td>
</tr>
<tr>
<td>Asian</td>
<td>11.21%</td>
<td>1.33%</td>
<td>7.51%</td>
</tr>
<tr>
<td>Native Hawaiian/Other Pacific Islander</td>
<td>0.47%</td>
<td>0.45%*</td>
<td>1.23%</td>
</tr>
<tr>
<td>Black/African American</td>
<td>15.89%</td>
<td>3.66%</td>
<td>17.71%</td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td>20.56%</td>
<td>4.10%</td>
<td>21.02%</td>
</tr>
<tr>
<td>Other/More than one race</td>
<td>~7.94%</td>
<td>0.71%</td>
<td>3.30%</td>
</tr>
</tbody>
</table>


*Only Pacific Islanders. Data obtained from personal communication with project officer, October 1, 2014.

Table 1.1). This study tracks children from birth through entry into school and by collecting data on individual health, development, care, and education of children from birth through kindergarten entry. Using these data, Halle et al. (2009) found differences in cognitive development, as well as health disparities, between AI/AN children and their non-Native peers, as early as 9 months of age.

Early Childhood Longitudinal Study, Kindergarten Cohort (ECLS-K). The ECLS-K tracks students from kindergarten through eighth grade, in both public and private schools. The ECLS-K includes data from teachers, parents/families, and schools (Table 1.1). There are currently two iterations of this study, one begun in 1998–1999 and another begun in 2010–2011. Although AI/AN students were not oversampled in these data sets, they are included in sufficient numbers to allow for statistical analysis. For example, using data from the ECLS-K (1998–1999), Hibel, Faircloth, and Farkas (2008) found that AI/AN students’ referral and placement into special education could be predicted in kindergarten, based in large part on these children’s scores on standardized achievement tests in reading and math. This analysis indicated that those students who performed poorly on the selected achievement tests were more likely to receive special education services in the early grades. Also drawing on the ECLS-K (1998–1999), Marks and Coll (2007) utilized latent growth modeling to examine similarities and differences in the academic preparedness and related characteristics of kindergarten students from different racial/ethnic groups. They found that AI/AN students who came to school academically prepared did much better in school than did their peers who are less ready to learn.
National Assessment of Educational Progress (NAEP). The NAEP is the largest nationally representative data set in the United States. It includes data from assessments in math, reading, science, writing, the arts, civics, economics, geography, and U.S. history for students in grades 4, 8, and 12. In 2014, the NAEP also assessed students in technology and engineering literacy. Utilizing data from the 2007 administration of the NAEP, López, Heilig, and Schram (2013) found that culturally relevant curriculum and teaching practices had a potentially negative effect on AI/AN students’ achievement on standardized measures of reading and math in grades 4 and 8. In another study, Fischer and Stoddard (2013) used NAEP data supplements to examine the academic achievement of American Indian students. They found that family and school characteristics were strongly related to differences in achievement between AI/AN students and their peers. This study also yielded a number of interesting findings related to students’ identification as AI/AN and schools’ identification of these students’ race/ethnicity. In many cases, these two identities differed.

National Indian Education Study (NIES). First conducted in 2005, the NIES was commissioned by the Office of Indian Education, within the U.S. Department of Education, and the NCES. This study was initially mandated by a 2004 presidential executive order, “recogniz[ing] the unique educational and culturally related academic needs of American Indian and Alaska Native students consistent with the unique political and legal relationship of the Federal Government with tribal governments” (Executive Order No. 13336, 2004, p. 25295). The NIES includes data from two surveys—the NAEP and a separate survey aimed at better understanding the extent to which schools incorporate culturally relevant teaching and learning practices. In 2011, 5,500 AI/AN students in the fourth grade and 4,100 Native students in the eighth grade completed the reading assessment, while 5,400 fourth graders and 4,200 eighth graders completed the math assessment. Although data from the NIES is available as both public and restricted use data, the bulk of publications (e.g., Moran, Rampey, Dion, & Donahue, 2008; Rampey, Lutkus, Weiner, & Rahman, 2006; Stancavage et al., 2006) resulting from these data have been commissioned by NCES as reports of the initial findings of this study, utilizing descriptive statistics, and presentation of emergent themes, rather than higher-level statistical analysis.

Schools and Staffing Survey (SASS). The SASS collects data on the characteristics of schools and staff (i.e., teachers and school leaders) in schools across the nation. Unfortunately, after the 2007–2008 administration, the SASS discontinued the inclusion of schools operated or funded by the Bureau of Indian Affairs (BIA), where approximately 8% of all American Indian students attend schools. In the last year of inclusion of BIA schools (2008), SASS included data for approximately 9,800 public schools, 180 BIA schools, 9,800 public school principals, 180 BIA principals, 47,440 public school teachers, and 750 BIA teachers. Although the BIA schools were
included between 1990 and 2008, with the exception of Pavel and Curtin (1997) and Pavel, Skinner, Farris, Cahalan, Tippeconnic, and Stein (1998), few researchers have used these data to examine the overall characteristics of schools serving AI/AN students.

**Postsecondary Data Sets.** Table 1.2 provides percentages from data on three NCES national postsecondary data sets by race/ethnicity. We see that in the Baccalaureate and Beyond (B&b), Beginning Postsecondary Study (BPS), National Postsecondary Student Aid Study (NPSAS), and Integrated Postsecondary Education Data System (IPEDS) databases, American Indians and Alaska Natives ranged from less than 0.5% of the sample to 0.9%. These numbers are below the widely reported 1% of college students who are American Indians or Alaska Natives. These differences may seem minor, but a lower AI/AN response rate of 0.5% in the sample means that American Indians and Alaska Natives have only half the representation in the sample of their actual population proportion, which is 1%.

**Beginning Postsecondary Study (BPS).** While BPS (2004:09) oversampled for AI/AN students, response rates were relatively low (see Stage, Alcantar, & Galvan, 2013). Even with the oversampling of this population, missing data were problematic. For some variables, over 50% had missing responses. Missing data in such a small sample size force the researcher to

---

### Table 1.2. NCES National Postsecondary Data Sets, by Percentage Race/Ethnicity

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total (N)</td>
<td>17,110</td>
<td>18,640</td>
<td>110,790</td>
<td>17,330</td>
<td>34,330</td>
</tr>
<tr>
<td>White</td>
<td>72.9%</td>
<td>61.54%</td>
<td>57.88%</td>
<td>63.59%</td>
<td>82.50%</td>
</tr>
<tr>
<td>American Indian/Alaska Native</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>5.83%</td>
<td>4.71%</td>
<td>5.60%</td>
<td>12.88%</td>
<td>6.4%**</td>
</tr>
<tr>
<td>Native Hawaiian/Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pacific Islander</td>
<td>0.38%</td>
<td>0.37%</td>
<td>0.51%</td>
<td>0.55%</td>
<td>N/A***</td>
</tr>
<tr>
<td>Black/African American</td>
<td>8.56%</td>
<td>13.79%</td>
<td>16.09%</td>
<td>11.78%</td>
<td>5.50%</td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td>9.19%</td>
<td>14.90%</td>
<td>16.03%</td>
<td>8.72%</td>
<td>3.10%</td>
</tr>
<tr>
<td>Other</td>
<td>0.20%</td>
<td>1.31%</td>
<td>N/A*</td>
<td>N/A*</td>
<td>N/A*</td>
</tr>
<tr>
<td>More than one race</td>
<td>2.47%</td>
<td>2.75%</td>
<td>2.98%</td>
<td>2.04%</td>
<td>2.10%</td>
</tr>
</tbody>
</table>


**“Other” not available.**

**Asian category in NSOPF includes Pacific Islander.**

**Native Hawaiian/Pacific Islander category included in Asian category.**
identify proxy variables and hope for less missing data or use mean imputation or other techniques to save as many cases as possible for analysis. Researchers attempting to uncover unique factors related to AI/AN students’ access to and success in college found their efforts stymied (Stage, Alcantar, & Galvan, 2013). While data were collected from 448 AI/AN respondents for the Beginning Postsecondary Study data set (2.4% of the total sample), several hundred had missing data on many of the individual student response measures.

**National Postsecondary Student Aid Study (NPSAS).** NPSAS is a national longitudinal study that collects data about the ways in which students and families finance their postsecondary education. The national sample includes funded and nonfunded students from various types of institutions (private, public, two- and four-year, undergraduate, and graduate levels). NPSAS also served as the base-year data collections for the BPS and B&B at different years. The data come from students and from administrative data. NPSAS has been collecting data since the 1990s. NPSAS has a larger overall sample size than other NCES data sets and a larger AI/AN undergraduate student sample (0.91%; see Table 1.2). A report on the national education trends of AI/AN students developed by NCES demonstrated that in the 2003–2004 school year, 82% of AI/AN full-time undergraduate students received financial aid (DeVoe & Darling-Churchill, 2008).

**Integrated Postsecondary Education Data System (IPEDS).** IPEDS is an annual survey that includes all institutions that participate in the federal student financial aid programs. Institutions must report on enrollment; program completion; graduation rates; faculty and staff; finances; several student and institutional indicators, such as enrollment and graduation rates; and finances and institutional prices (U.S. Department of Education, 2010). Although this data set is designed to represent the majority of colleges and universities across the United States, there are still issues with it. One issue is that difficulties exist in obtaining accurate data for small populations—thus, it is important to examine changes in numbers over time as well as the percentage of change. A second issue associated with IPEDS is that some institutions fail to report data, necessitating the imputation of data for these institutions. Finally, a remaining issue concerns the use of self-reported racial/ethnic data—often misreported for several reasons that are discussed later.

**Examples of Critical Approaches Using Large Data Sets**

Although these national data sets have sometimes proven problematic, researchers have examined underrepresented populations using these data sets with a critical quantitative approach. Often these national data are used to inform data derived from state- and institution-level surveys and other means of data collection. It is imperative we learn from these studies in
order to inform the development, data collection, and analysis of research on AI/AN students.

Oseguera and Hwang (2014) used the Educational Longitudinal Study 2002–2006 to study college access for low-income students using a critical quantitative approach. They chose a theory that reflected experiences of low-income populations as well as particular variables that were relevant to low-income students’ college experiences. Because the sample size was already low, they used multiple techniques to impute values (Oseguera & Hwang, 2014). By using more complex imputation techniques, they were able to reduce loss of cases through missing data.

Stage, Alcantar, and Galvan (2013) used the BPS (2004:09) to study AI/AN student success in college, defined by months enrolled, degree expectations, and persistence or degree attainment. Using a critical approach, they included measures of high school academics; characteristics of college including distance from home and percentage minority enrollment; remedial course taking in college; and participation in college activities, academic support groups, and college study groups. Variables such as high school academic course taking, college distance from home, personal goals such as becoming a community leader, and engagement on campus—meeting or talking with faculty and advisors, club participation, community service, and study groups—all influenced outcomes positively. Additionally, given the relatively low graduation rate for AI/AN students, the additional measures to model persistence, total months attended, and highest degree expected proved useful.

Finally, in Chapter 2, Vaccaro, Kimball, Wells, and Ostiguy make suggestions about interpretation of data with large data sets and small sample sizes. Their study of students with disabilities shared similar challenges as research with AI/AN students. Within the population of students with disabilities are many differences based on type of disability. One suggestion was to conduct multiple operationalizations to test for robust findings across multiple analyses. Several significant results for a particular variable might suggest greater validity. Additionally, the authors suggested that researchers discuss important findings, whether significant or not. For example, a variable with a large effect size but no significant influence on the outcome may still be an important variable whose effect size should be further explored and discussed.10

Beyond the Data. As previously discussed, researchers have cited a number of problems associated with studies of American Indians and Alaska Natives relevant to the use of large-scale data sets (Kena et al., 2014; Ross et al., 2012). These problems include: (1) issues related to self-identification of race and ethnicity, (2) issues of tribal identity, (3) difficulty gaining access and entrée into AI/AN communities for research purposes, (4) the need for the development and use of culturally appropriate strategies for increasing AI/AN representation in large-scale data sets, and (5) a limited pool from
which to draw AI/AN samples of sufficient size for large-scale data sets. These issues are discussed in brief next.

1. Who is American Indian or Alaska Native? Who gets to decide? What does it mean to be classified as American Indian? A number of issues impact whether a population identified as American Indian, Alaska Native, or Native American is included in nationally collected data sets. For example, Demmert, Grissmer, and Towner (2006) assert that when the term “Native American” is used, we often have no way of knowing whether this term is used accurately, as some individuals identify as Native American based on the fact that they were born or reside in the United States. Although individuals have the right to self-identify as any racial or ethnic group they choose, it is important to acknowledge that for Native Americans, the question of identity and tribal membership is one that is culturally and politically complex, as it is rooted in principles of tribal self-determination and local tribal control (e.g., National Congress of American Indians, n.d.). While tribes have the right to determine who their tribal members are, most national surveys allow respondents to self-report to which racial/ethnic group they belong, thereby opening the door for those who are not tribally recognized to identify as American Indian. According to the Bureau of Labor Statistics, “Research on the collection of race/ethnicity data suggests that the categorization of American Indian and Alaska Native is the least stable self-identification” (Kena et al., 2014, p. viii). In other words, individuals may self-identify in these categories inconsistently. Kena et al. (2014) also point out that the way in which questions of identity are posed can also impact the way in which individuals respond, particularly if these individuals consider themselves to be of more than one racial or ethnic group. This fact must be considered when analyzing the results of surveys and other types of data where race/ethnicity is based on self-reporting.

2. The question of identity is critical in that analysis of data from large-scale data sets lends itself to clustering respondents in such a way that the heterogeneous nature of groups, such as American Indians and Alaska Natives, may be overlooked or overshadowed (Lavelle, Larsen, & Gundersen, 2009). The question of identity is important given that there are more than 600 different state and federally recognized tribes, each with its own culture and many with their own language—as more than 200 different Native languages are still spoken today. Within each of these tribes are unique cultural and linguistic traditions, practices, and characteristics that are not easy to tease out or to explain without interacting and talking with members of these groups. By combining tribal groups into one larger catchall category, we “can sometimes mask significant differences between
subgroups” (Ross et al., 2012, p. 1). In effect, we may be masking the variables that have the greatest effect or impact on student outcomes.

3. Gaining access and entrée into AI/AN communities for the purposes of research is difficult. According to Lavelle et al. (2009), issues of “mobility, geographic dispersion, and higher proportion of non-telephone households” (p. 386) make it difficult to access many Native peoples. These issues of physical access are also complicated by larger issues of social isolation and distrust of outsiders.

4. There is an ongoing need for the development and use of culturally appropriate strategies for increasing AI/AN representation in large-scale data sets. According to Lavelle et al. (2009), there are a number of ways in which to improve AI/AN participation in large-scale studies. These include:
   a. The establishment of agreements between tribes and tribal organizations and the agency or organization conducting the research, explaining what information will be collected, how it will be used, and the ways confidential information will be protected;
   b. Formal documentation from the tribe/tribal organization supporting the study;
   c. Training in cultural awareness for those administering surveys or other means of data collection; and
   d. Involving community/tribal members in the collection of data.

   Such strategies lend themselves to creating an atmosphere in which Native students feel welcome.

5. Small sample size and the limited pool from which to draw AI/AN samples are often problematic for the development and analysis of large-scale data sets. Small sample size impacts the reliability of the data and limits the ability of researchers to analyze data in ways that are meaningful for Native students (Ross et al., 2012). Small sample sizes also result in high standard errors, thus impacting findings of statistical significance. According to Kena et al. (2014), “survey data for American Indians/Alaska Natives often have somewhat higher standard errors than data for other racial/ethnic groups. Due to large standard errors, differences that seem substantial are often not statistically significant and, therefore, not cited in the text” (p. vii). As a result, asterisks or dotted lines are often placed beside the category of American Indian/Alaska Native to indicate inability to report due to small sample sizes. The practical implications of these symbols are aptly described in the book Beyond the Asterisk: Understanding Native Students in Higher Education (Shotton et al., 2013). In that book, the authors call for increased data collection and analysis specific to American Indians in higher education. In addition to the recommendations by Shotton et al. (2013), Lavelle et al. (2009) recommend the use of oversampling and increasing the overall sample size, strategies that may prove cost prohibitive for some researchers.
While the research issues just described were illustrated primarily by difficulties with large national data sets, similar issues can be found in state and single institutional-level data sets. According to Demmert et al. (2006), “research on Native American achievement has mainly been small scale, non-experimental, non-longitudinal and methodologically problematic” (p. 6). Most existing studies include sample sizes smaller than 500 and focus on individual tribes or subgroups, thus limiting the ability to make between group comparisons. While small quantitative and qualitative studies have provided useful information, more large-scale work needs to be done. These and other issues suggest that caution is warranted for institutional researchers and other scholars seeking to conduct quantitative studies of AI/AN students.

Implications for Research, Policy, and Practice

It is important that institutional researchers and scholars develop an increasing awareness of these issues and address them in research design and data collection. While many examples cited above referred to national data sets, these issues remain for any quantitative analysis of data focusing on students generally by racial and ethnic subgroups, including national, state, college system, and single institution. Issues regarding sample size for institutions or state systems can be minimized by oversampling for relatively small campus populations in order to draw meaningful analyses and conclusions. For example, rather than taking a 10% sample of the entire student body, perhaps a 20% or even greater sample should be taken for smaller populations such as AI/AN students. This issue of sample size, coupled with the need to ensure the sample is representative of the population studied, is critically important. According to Demmert et al. (2006), small sample sizes “limit the types and complexity of analysis possible in trying to understand Native American achievement” (p. 5).

Challenges researchers face in conducting studies of American Indians and Alaska Natives ultimately have broad consequences for national, state, and institutional policies and practices (Lavelle et al., 2009). Analyzing data from a critical perspective offers a way to make both scholarly and practical sense of the often confounding nature of research, particularly research seeking to draw on large numbers of American Indian students. As this chapter documents, numerous dangers exist in simply mining large-scale data sets without awareness of limitations and appropriate cautions.

The concerns just listed can be remedied by developing policies and practices to ameliorate the use of culturally misaligned data. If, as researchers, we impose our own thoughts and beliefs onto the data without giving careful consideration to alternative perspectives and approaches, a cultural mismatch will result between what we purport to measure and what is actually measured in multiple ways. This issue can be minimized by (1) ensuring that constructs being assessed are culturally relevant (see...
Recommendations for Future Research

As discussed earlier, there is a paucity of published quantitative research that focuses on AI/AN students. Additionally, undersampling of such students, given small population numbers, promises to slow progress. As a result, a number of important educational, social, and cultural questions remain unanswered. The focus of this chapter is on the need to engage in a more critical analysis of national, state, and institutional data.

As López et al. (2013) point out, the collection of information regarding American Indian students’ cultural experiences within school is important; however, studies must also assess the extent to which students, parents, teachers, and school leaders perceive education to be academically and intellectually rigorous, the extent to which students are expected to excel academically, and the quality of education students receive. This process of assessment includes measuring the extent to which students see the importance of their Native culture(s) as well as the ways in which teachers are taught to engage in culturally relevant teaching and learning practices during their teacher preparation programs. López et al. (2013) argue:

> Without these considerations, researchers who use [studies such as] the NIES to examine how culture is related to achievement may assert that the practice is associated with lower achievement, and make policy recommendations to limit AIAN [American Indian/Alaska Native] culture—returning education for AIAN students to the 1950s policies that diminished the AIAN-focused curriculum (Tippeconnic & Swisher, 1992) or even used to justify retrograde standards policies in US history and other subjects that neglect the history of AIAN students. (p. 534)

Finally, it is important to acknowledge that the complexity of the AI/AN population may not be fully understood by utilizing existing quantitative data without supplementing these data with qualitative data in the form of individual and group interviews, observations, and other forms of in-depth data collection and analysis. According to Hines (1993), the use of mixed methods may help to improve the overall quality of research with culturally diverse groups, such as American Indians and Alaska Natives. The use of data collection techniques that allow researchers to probe more deeply and to provide a more nuanced understanding of the educational
experiences and subsequent outcomes of AI/AN students is an important point given the fact that the educational experiences and subsequent outcomes for AI/AN peoples have been shaped by more than 500 years of cultural, political, and legal subjugation, played out in part by educational systems that have attempted to save the individual by killing the Indian (e.g., Adams, 1995). For many individuals and communities, this attempt to acculturate and assimilate has resulted in historic traumas that portray education as a means of oppression rather than the great equalizer it is often heralded to be (e.g., Grande, 2004).

Critical quantitative measures, coupled with qualitative data, have the potential to assist educators and other education researchers in understanding issues such as these three: (1) the role language and culture play in shaping AI/AN students’ precollege educational experiences and subsequent academic and life outcomes; (2) the disconnect between early aspirations to attend college and historically low rates of college entry, persistence, and completion; and (3) the meaning of academic success for American Indian students, their homes, schools, and communities (i.e., what does it mean to be successful to Native peoples? How does this relate to the practice of schooling at the PK–20 level?).

**Need for a New Critical Theory of Change**

Critical quantitative approaches to researching the educational conditions and subsequent academic outcomes of American Indian and Alaska Natives are required if we are to fully understand the educational conditions and subsequent academic outcomes of Native students. As we work to develop these approaches, we must reflect on the work of Native scholars, such as Tuck (2009), who caution against the use of what she describes as “damage-centered research” (p. 409)—research that focuses solely on what is wrong with Native peoples rather than on their strengths as well as their needs. This shift requires researchers, particularly those who engage in education research using large-scale databases, to adopt a new theory of change, one that we characterize as utilizing data to critically unpack and respond to the role of language and culture in shaping pathways to success in education and beyond. As Tuck (2009) so aptly argues, this process involves moving away from theories of change that focus on what is wrong, missing, or absent among Native students, their schools, and communities—in other words, deficit frameworks—and instead moving toward research that seeks to identify what is right with these students, their schools and communities, and what can be done to improve the culture, climate, and conditions of academic institutions in order to better enable these students to flourish both academically and culturally. Doing this requires asking hard questions about “who participates in the research, who poses the questions, how data are gathered, and who conducts the analysis” (Tuck, 2009, p. 423). All are
questions that can and should be addressed through a critical quantitative approach.

**Conclusion**

This chapter discusses issues and challenges encountered in analyzing data, both large-scale data and smaller, institutional-level data, to study the educational experiences and outcomes for AI/AN students. Lessons from analysis of national data sets are extended to inform institutional researchers and other scholars who aim to study issues related to the educational experiences and subsequent outcomes of American Indian college students, as well as other issues at the institutional, state, and national levels. These lessons are also applicable to student affairs practitioners who are interested in improving college access and educational attainment among AI/AN students. Issues discussed include the dangers of using ill-suited or irrelevant models that do not fully capture or take into consideration the full linguistic and cultural diversity of AI/AN students, the lack of salient measures, and small sample sizes. The authors conclude with recommendations for constructing and analyzing large-scale data sets to better represent the diversity of cultures and experiences found among AI/AN students.

**Notes**

1. The term “Native American” is often used interchangeably with the term “American Indian/Alaska Native.”
2. For additional information, see http://nces.ed.gov/ecls/index.asp
3. For additional information, see http://nces.ed.gov/nationsreportcard/about/
4. For additional information regarding the sample composition of the NIES, see http://nces.ed.gov/nationsreportcard/nies/faq.aspx#participated
5. For additional information regarding the SASS, see http://nces.ed.gov/surveys/sass/
7. For additional information regarding the BPS, see http://nces.ed.gov/surveys/bps/
8. For additional information regarding the NPSAS, see http://nces.ed.gov/surveys/npsas/
9. For additional information on IPEDS, see http://nces.ed.gov/ipeds/datacenter/
10. Measures and results should be examined whether significant or nonsignificant. Research on testing and assessments have pointed out the importance of examining the consequential validity of measures and results (Messick, 1979). Consequential validity is the social consequences of using certain measures and results. One area to be particularly cautious with is the relationships among sample size, significance level, and effect size. Large sample sizes can produce statistically significant results with low and meaningless or irrelevant effect sizes. Conversely, very small samples can produce large effect sizes but be statistically nonsignificant. The potential impact of small sample size is particularly important for studies on the educational outcomes of AI/AN students. Often these studies have small sample sizes and render nonsignificant results, yet the results are useful for policy and practice. Large effect sizes should be discussed as possible factors to be explored in future research.
References


---

**SUSAN C. FAIRCLOTH** is an associate professor of education in the Department of Leadership, Policy, and Adult and Higher Education at North Carolina State University.

**CYNTHIA M. ALCANTAR** is a research associate for the Institute for Immigration, Globalization, and Education and a doctoral student at the University of California, Los Angeles.

**FRANCES K. STAGE** is a professor of higher and postsecondary education in the Department of Administration, Leadership, and Technology at New York University.