



PROJECT MUSE®

Indigenous Data Collection: Addressing Limitations in Native American Samples

Jameson D. Lopez

Journal of College Student Development, Volume 61, Number 6, November-December 2020, pp. 750-764 (Article)

Published by Johns Hopkins University Press

DOI: <https://doi.org/10.1353/csd.2020.0073>



➔ *For additional information about this article*

<https://muse.jhu.edu/article/774244>

Indigenous Data Collection: Addressing Limitations in Native American Samples

Jameson D. Lopez

This article outlines Indigenous data collection as a methodology that expands a sampling frame by using tribal Indigenous cultural knowledge, such as creation stories like those of the Cocopah and Quechan nations, to address limitations found in the methodological quality of Native American survey data in currently available datasets. Indigenous quantitative methodology and methodological limitations in existing government and institutional datasets are outlined (Walter & Andersen, 2013), and validities described by Shadish, Cook, and Campbell (2002) and the Standards for Educational and Psychological Testing (AERA et al., 2014) are explored. The conclusion presents the advantages of using Indigenous data collection to provide a sampling frame that highlights the use of Indigenous quantitative methodology.

According to Quechan (pronounced Kwat'san) oral history, the creator formed us Quechan, Maricopa, Cocopah, and Kumeyaay people from mud at the beginning of life. Eventually, the creator took us from the sacred mountain, 'Avii Kwa'amée, to teach the proper way to live in the world. The creator sent us down from this mountain, and our tribes went their separate ways. We lived our separate lives through the centuries, sometimes fighting each other and at other times fighting the cavalry together. In current times, we work together to maintain our traditions through our songs, languages, creation stories, and cremation ceremonies. Although we are certainly different nations, we come from the same place, and this

presents an opportunity to create a framework for how researchers collect and analyze data from these and other Indigenous communities.

Creation stories are meaningful to Indigenous people because they contain answers to questions we have in life. For example, creation stories like those of the Haudenosaunee may incorporate instructions for appropriate approaches to the relationship between creation and living things. Creation stories like those of the Anishinaabe may emphasize fundamentals of kinship, unity, and reverence (McGregor, 2004). The Muscogee creation story gives instruction regarding the relationship with plants, animals, and other people (Fixico, 2003). Creation stories like those of the Quechan and Cocopah help Indigenous people understand life cycles from birth to death and mourning (Alvarez de Williams, 1974; Bryant, 2013). Overall, creation stories encompass explanations of life's journeys, so it is instinctive to use creation stories as a framework to expand sample sizes in order to address limitations found in quantitative research that excludes many Indigenous students.

BACKGROUND

The majority of higher education research lacks datasets that are adequate for examining Indigenous populations through quantitative research. Additionally, government data are limited due to small samples and a lack of culturally relevant variables, which further

Jameson D. Lopez is an assistant professor of higher education at the University of Arizona.

constrain research with Native American data (Lopez & Marley, 2018). Other datasets maintained by institutions tend to have data with small samples of Native Americans, and they also lack relevant variables, resulting in data that are inconsistent, irrelevant, of poor quality, and produced and used within an environment of mistrust (Rainie, Schultz, Briggs, Riggs, & Palmanteer-Holder, 2017). Data can amplify Indigenous voices; however, with missing, inaccurate, or irrelevant data, these voices are invisible while stereotypes and misconceptions prevail. Such invisibility may contribute to social and academic marginalization. Reyes and Shotton (2018) described how misconceptions that Indigenous students encounter on campus lead to feelings of invisibility that may contribute to the lack of cultural support to counteract historical assimilation policies. Indigenous people are the First People of this country with treaty obligations from the United States, yet they are regarded as “other” or outright excluded because researchers are unable to figure out a method that is suitable for them to conduct appropriate analyses. It is extremely disheartening that since the inception of quantitative research, it is possible to see, in almost every education journal, this asterisk phenomenon of Indigenous students being invisible in the data (Shotton, Lowe, & Waterman, 2013).

Within the last year, data produced in some of the research published in top-tier journals in higher education (e.g., *Review of Higher Education*, *Journal of Higher Education*, *Research in Higher Education* and *Journal of College Student Development*) often included White, Black, Hispanic, and Asian subjects but erased Natives by categorizing Indigenous students as “other” (see Christensen & Harris, 2019; Creusere, Zhao, Bond Huie, & Troutman, 2019; DesJardins, Toutkoushian, Hossler, & Chen, 2019; Radimer & Rowan-

Kenyon, 2019). Some may think invisibility is not an issue, but when Indigenous women have a higher likelihood of facing violence than any other ethnic group, they need to be included in the conversation (Rosay, 2016). The matter of visibility becomes salient when data used by Christensen and Harris (2019) are examined closely; these authors compared only White women, Latinas, and Black women in a study measuring bystander readiness to help. The study focused on determining a person’s level of awareness, ability, and readiness to intervene in situations where there is a risk of sexual violence. The number of students that identified as American Indian / Alaska Native were purposefully not reported because of the small sample size; however, when 84% of Indigenous women face violence (Bettinger-Lopez, 2016) then visibility matters. When there are not enough data to pass policies, and when samples sizes are not large enough to find statistical significance, then visibility matters. When media and society constantly romanticize Natives, particularly in a time when people need to recognize that Native people are still here, resilient and strong in the modern world, then visibility matters. The constant invisibility of Indigenous people, for the sake of convenient or simplified statistical analyses, has been a tiring conversation since the days of the Merriam Report (1928), which even then called for better quantitative data.

Rainie et al. (2017) recommended that tribal nations use community-based, tribal nation–driven data collection to inform policy decisions and resource allocations, yet not enough tribes are engaging in data collection. For tribal nations to develop their own data collection is vital because federal data are not representative of Native Americans in the US and are likely to exclude smaller Native American tribes. Excluding tribes is detrimental because they are sovereign nations regardless of population size; a large tribe

cannot represent the majority of tribes just because they are larger. Because tribes vary in their use of sovereignty, governments and education policies are influenced differently.

Furthermore, the lack of high-quality quantitative education data prevents researchers from engaging in rigorous quantitative investigations of Indigenous education, which contributes to a limited literature base. Lopez and Marley (2018) indicated that these challenges demonstrate that extensive limitations exist that make valid inferences about Indigenous populations difficult. Therefore, datasets must be constructed to appropriately test quantitative education research with Native American participants (Lopez, 2018).

Four of the major challenges related to Native American samples and data are (a) small sample sizes, (b) self-identification of Native American identity, (c) generalizability, and (d) lack of relevant variables. First, when researchers include American Indian / Alaska Native samples, one of the most consistent problems in those datasets is that the small sample sizes result in biased findings and inaccuracies (Aud et al., 2013). Second, misclassification of self-identifying American Indian / Alaska Native participants results in data that do not always reflect how Indigenous people identify nor do they correlate with federal government definitions (Walter & Andersen, 2013). Data that contain self-identification of Native American heritage is in the American Community Survey, U.S. Census, and all of the federally managed postsecondary datasets (Rodriguez-Lonebear, 2016). One of the difficulties is that participants who identify as Indigenous in one year may not identify as Indigenous in later years (Feir, & Hancock, 2016). Third, researchers who use federal datasets must depend on small and nonrepresentative samples, which are problematic because

findings produced from these data are believed to be generalizable (Wine, Bryan, & Siegel, 2014). Generalizability is a major limitation. There are 574 federally recognized American Indian / Alaska Native tribes and numerous state-recognized tribes within the boundaries of the US (Indian Affairs, 2020). Unfortunately, most of the data available from government and institutional datasets do not capture the nuanced differences that exist among these tribes (Lopez & Marley, 2018). Fourth, due to all the limitations in these datasets, there is a lack of relevant variables that measure the experiences of Indigenous students; therefore, to capture American Indian / Alaska Native students' experiences, there is a need to collect accurate survey data suitable for quantitative analyses of their behaviors, attitudes, and outcomes.

AUTHOR POSITIONALITY

I am an Indigenous quantitative researcher, and I recognize that quantitative paradigms, such as postpositivism, do not align or entirely blend with my beliefs. Nonetheless, there are parallels between Indigenous and postpositivist worldviews, such as distrust of limited notions of rigorous research, the relationship between researcher epistemology/ontology and research questions, and educational challenges rooted in political influence on education research and policy (Fitzpatrick, Sanders, & Worthen, 2011; Richardson, 2015). Due to my views, I most often examine research through Indigenous methodologies that privilege community voice.

DEFINITIONS

I interchangeably use *American Indian / Alaska Native*, *Native American*, and *Indigenous* to describe the First Nations people living within the borders of the United States. Although I do understand that each of these terms

describing Indigenous people is derived from different places outside of tribal definitions, it would be difficult to refer to each individual tribe throughout this writing. Additionally, I refer to *government and institutional datasets* to describe datasets managed by the federal government, state governments, and colleges and universities.

INDIGENOUS QUANTITATIVE METHODOLOGY

At the core of Indigenous methodologies is how researchers use Indigenous positionality to conduct research with and in tribal communities and to privilege the tribal community's voice(s) to support the community (see Battiste, 2011; Louis, 2007; Windchief, Polacek, Munson, Ulrich, & Cummins, 2017). Wilson (2008) stated that scholars conducting research in Indigenous communities must engage in relational building because it "is an important aspect of ethical Indigenous research" (p. 40), which requires researchers to form mutually respectful relationships within the communities where they are conducting research. Indigenous methodology requires scholars to use a culturally sensitive statistical research lens and to reflect on our own identities and on our journeys to the research (Kovach, 2010).

Walter and Andersen (2013) characterized Indigenous quantitative methodology as "methodologies within which the practices and the processes of the research are conceived and framed through an Indigenous standpoint" (p. 83). There are two main underpinnings of Indigenous quantitative methodology: (a) creating statistical data from an Indigenous lens that privileges Native American voices, rejects dominant mainstream value systems, and refuses deficit approaches as a starting point in research; and (b) challenging statistical practices within Indigenous nations by exposing the viewpoint from which tradi-

tional quantitative research has operated in Indigenous communities (Snowshoe, Crooks, Tremblay, Craig, & Hinson, 2015; Walter & Andersen, 2013). External scholars need to help Indigenous nations with research and support to build the capacity of tribes in a respectful manner. Snowshoe et al. (2015) suggested that researchers (a) engage with the complex authority structures of tribes, (b) recognize the complexity of the traditional elder engagement process, (c) utilize culturally competent partners as mediators of the tribal partnership process, (d) take an Indigenous approach to research design that works in the community, (e) anticipate a longer timeframe for the community engagement process, (f) select culturally appropriate data collection methods, and (g) commit significant time and resources to scale development. Scholars' first step in this process is to recognize their own beliefs.

Researchers who do not recognize their personal beliefs during the research process magnify the influence of biases. Nonetheless, traditional quantitative researchers often feel uncomfortable with the subjectivity this recognition imposes. Despite the attempts that quantitative scholars make to increase objectivity and scientific validity, research is not without prejudice. The scales and constructs quantitative scholars measure are a product of what the researchers or funders view as essential (Rodriguez-Lonebear, 2016; Walter & Andersen, 2013; Zuberi & Bonilla-Silva, 2008). Past research has links to the subjugation of Native Americans and to federal policies, such as extermination and assimilation (see Echo-Hawk, 2010), that cause some Indigenous people to reject research, especially when it is disconnected from Indigenous communities (Rodriguez-Lonebear, 2016; Walter & Andersen, 2013; Zuberi & Bonilla-Silva, 2008). In more recent data analysis, mainstream educational researchers ignore Indigenous youth and

adults because of low sample sizes that make “statistically significant” inferences difficult (Shotton, Lowe, & Waterman, 2013; Sumida Huaman, Martin, & Chosa, 2016; Walter & Andersen, 2013). Additionally, the dominant society’s understanding of Native Americans constrains data and continues to reinforce biased narratives imposed by dominant cultures.

Walter and Andersen (2013) stated that statistical analyses “speak a ‘truth’ about the communities on which they shine their statistical light” (p. 9). For example, we do not measure success based on how White people can play rez ball: first, because most White people would not know what rez ball is (maybe you are wondering what rez ball is right now), and second, White people would not know the rules (or lack of rules). The dominant narratives often speak of high rates of suicide, diabetes, alcoholism, and drug abuse to characterize Native American populations; therefore, Indigenous researchers advocate for tribal nations to gather our own data to change the deficit research approach, which mostly originates with non-Native researchers.

By collecting our own data, Indigenous researchers and practitioners can ask questions that are culturally relevant to Indigenous communities, such as those from the Scale of Native Americans Giving Back (Lopez, 2020), which includes the following items:

- I help organize community events (e.g., Indian Days, Pow Wows, community dinners).
- If possible, I always try to buy from tribal businesses.
- I pray for my tribal community.
- I try to visit my tribal homeland as much as possible.
- I participated in community gatherings before coming to college (e.g., Indian Days).
- I plan on using my education to help my tribe.

Many Indigenous people still have confidence that quantitative research and analysis methods can support Indigenous nations, as long as researchers using those methods in Indigenous environments understand how quantitative methods have historically harmed and overlooked our communities (Rodriguez-Lonebear, 2016; Shotton, Lowe, & Waterman, 2013). The confidence of Indigenous nations increases when Indigenous scholars conduct the research with tribal communities. Furthermore, Indigenous quantitative methodology is not possible without making tribal community research partnerships central to the research design (Snowshoe et al., 2015). Research partnerships with tribal communities are integral because they decrease the likelihood of potentially harmful quantitative approaches to research within Indigenous nations. I assume that most traditional quantitative researchers adopt postpositivist viewpoints and accept established quantitative forms of validity and reliability; therefore, I further frame this conversation using the works *Experimental and Quasi-Experimental Designs for Generalized Causal Inference* (Shadish, Cook, & Campbell, 2002) and *Standards for Educational and Psychological Testing* (American Education Research Association [AERA et al.], 2014).

I use both of these approaches to illustrate that Indigenous data collection methodology can meet the statistical requirements of data collection in governmental and institutional datasets. The dominant society, or those in power, are still likely to question the legitimacy or validity of data so collected, which is why the Indigenous data collection framework outlines how to create stronger validity claims. Additionally, Wilson (2008) mentioned that criticizing the research of other scholars does not fit within an Indigenous framework (Windchief et al., 2017); however, traditional quantitative research situated in Western paradigms allows scholars to find fault, ques-

tion, argue, challenge, and critique the research of other scholars. While there is a need for both frameworks, I cannot fully argue for using a Western paradigm because arguing “on behalf of Indigenous nationhood within the dominant Western paradigm is self-defeating” (Grande, 2004, p. 57). I need to be able to criticize the Western paradigm using Western techniques that demonstrate the validity of Indigenous quantitative methodology to traditional quantitative researchers.

RESEARCH VALIDITIES

Shadish, Cook, and Campbell (2002) described four types of research validity: internal, external, statistical-conclusion, and construct. Researchers commonly use these four validities in educational research to design and evaluate the quality of quantitative studies. For more in-depth analyses of these concepts, see Lopez and Marley (2018).

Internal Validity

Internal validity relates to a researcher’s ability to make a causal claim about two variables. Causal claims are strongest when three characteristics are present: the cause can be manipulated to create an effect; there is a temporal order for the cause and effect; and alternative influences for an effect are accounted for (Antonakis, Bendahan, Jacquart, & Lalive, 2010; Höfler, 2005; Murnane & Willett, 2010; Shadish et al., 2002). Random assignment of participants to treatment (intervention) and control groups, matching participants with similar characteristics, and using statistical controls are ways that researchers strengthen causal claims. Random assignment is the preferred method, followed by matching, and then statistical controls.

External Validity

External validity relates to the ability to use

findings from a smaller sample to generalize to a larger population. The best approach to generalizing conclusions to a larger population is to randomly sample participants from that larger population. Random sampling is a technique that allows everyone an equal chance of selection to participate in a study (Teddle & Yu, 2007). The random sampling technique creates an unbiased selection of participants, which increases the generalizability of the findings; however, random sampling is rarely possible, because lists of entire populations are often inaccessible.

Statistical-Conclusion Validity

Statistical-conclusion validity is the ability to produce appropriate conclusions from a statistical analysis. To strengthen statistical-conclusion validity, Shadish et al. (2002) recommended maximizing power (correctly rejecting the null hypothesis) by considering the strength of a relationship or effect-size measures (e.g., Pearson’s correlation) and testing assumptions. Statistical power is the probability of detecting a relationship when one exists (Cohen, 1992). Researchers ideally want more statistical power, because it reduces the likelihood of committing Type II errors, failing to reject a hypothesis that should be rejected. To increase power, researchers will increase the sample size.

Construct Validity

Construct validity is the degree to which theoretical and empirical evidence validate the interpretations of test scores for the prepared uses of tests and instruments (AERA et al., 2014) and the combination of any evidences that relate to the meaning of participant scores (Downing, 2003; Messick, 1995). Construct validity includes measurement-validity evidence. According to the *Standards* there are five sources of measurement-validity evidence; however, the focus here will be

on evidence that is based on test content (AERA et. al., 2014). This is established by analyzing relationships between the content and the construct that the content attempts to measure (AERA et. al., 2014). Researchers occasionally establish evidence based on test content by developing questions and utilizing an expert committee to generate feedback on those questions.

SUMMARY

Although there will be limitations to any study, many of the limitations found in current government and institutional datasets could be resolved through an Indigenous data collection framework. Ideally, researchers could resolve these limitations within the government and institution; however, because of financial and logistical constraints, there are no existing datasets at the national and institutional levels that mitigate these limitations.

SAMPLING FRAMEWORK

Tribes collecting their own data will create new ways to support tribal education policies and funding for their respective members, because the data available may not be representative of members from smaller tribes. There are 574 federally recognized tribes and additional state-recognized tribes that do not have federal recognition (Indian Affairs, 2020). Citizenship numbers in these tribes can range from fewer than 100 to more than 100,000. Because some tribes are so large (like the Navajo with close to 300,000), it is likely more members of those nations will be in a sample, whereas members from smaller nations (like my own Quechan with close to 4,000) may not have any members in the sample. Thus, the likelihood that the results are representative across tribal nations is significantly decreased. The creation stories, cultures, lands, kinship, economic development, and lives of these vastly

different tribes are likewise underrepresented if not dismissed entirely. Even if tribes were to collect their own data, smaller tribes will have difficulty finding enough suitable participants for most quantitative studies in higher education.

Data collection from tribes working together may result in better services for their respective communities and students than depending on national datasets with limited sampling of Native American populations being used to solve issues. Researchers working with multiple tribes to collect data that are representative of the members can provide the sampling frame for their data by taking into account Indigenous knowledges such as creation stories. This method of data collection would address the limitations of small samples from smaller tribes.

INDIGENOUS DATA COLLECTION

The manner in which the federal government imposes its standards for tribal citizenship, tribal reservations, and boundaries does not align with tribal definitions of recognition and boundaries. For example, the Tohono O'odham tribe in Arizona lives on lands that span the border established between the US and Mexico after the Mexican–American War of 1846–1848. Following World War II, border towns filled with thousands of people of varied citizenships, races, and identities; US and Mexican officials were compelled to enforce the border (Cadava, 2011). Today Tohono O'odham tribal members living in Mexico have the same rights as those in the US, despite their Mexican citizenship. Other tribes, such as the Pai Pai of Mexico, are related to the Quechan through creation stories, language, song, and ceremony, but they do not share federal recognition or rights, as the Tohono O'odham do, despite being related to

tribes in the US. The lack of recognition and the policies around defining who is or is not Indigenous serve to further settler-colonial aims of erasure. When data misrepresent or exclude Indigenous populations, the lack of information contributes to these aims.

Neighboring tribes often share similar cultural bonds, like creation stories, so that combining these communities makes it more feasible to produce suitable samples to collect relevant data for statistical analysis. Additionally, the tribes along the Colorado River that belong to the Yuman linguistic family share similar creation stories, songs, language, and ceremonies. Likewise, the Chickasaw and Choctaw Nations are two independent tribal nations today who share a similar creation story, tribal ceremonies, and languages. It is beneficial for these tribes to collect data jointly and examine constructs influencing education outcomes for their tribes. Tribal leaders play an important role in promoting or impeding the education of their students. Collecting their own data would support their capacity to inform individual tribal decisions related to education policies and academic achievement. Institutions are often limited by what they can control while tribes do not share those limitations. In the next portion of this section, I expand on how collecting data by tribes can address some limitations found in larger datasets, beginning with descriptions of two similar tribes and how data collection with these tribes can address these limitations.

Quechan Tribe. According to our traditions, we were created and have occupied our land base since time immemorial. In English *Quechan* means “those who descended.” This is a shortened version of *Xaam Kwat’san*, meaning “those who descended by means of water.” As water is a focal point of our life and culture, it is important to know the reverence that our tribe has had for water from the beginning, as described in our creation story. Our story

is normally told over four days; therefore, I refrain from telling it here. It is very deeply rooted in the cremation ceremony that is one of the cornerstones of our culture today. During this ceremony a wake is held with traditional songs; it ends with the cremation of the deceased. It is a cultural tradition that our tribe has been able to preserve since European colonization.

In 1884, the Quechan reservation in Fort Yuma, California and Arizona, was established along both sides of the Colorado River near Yuma, Arizona. Currently, approximately 3,800 citizens originate from the tribe, and 2,200 of those citizens reside on the reservation lands. The current land base stretches for about 44,000 square acres in an area that includes land in Arizona, California, and Mexico. Most of its economic development comes from two casinos, agricultural lands predominantly leased to nontribal farmers, and hotels that cater to winter visitors. Despite this economic development, the poverty rate for our reservation is around 37%, twice that of the Arizona state as a whole, and almost half the children are considered to be below the poverty threshold (Arizona Rural Policy Institute, 2010a).

One of the elements that connects us to other tribes is our language. The Quechan language belongs to the Yuman language family, which has three major branches: River, Pai, and Delta-California. The Quechan language belongs to the River branch of the Yuman languages. Our language and creation story tie us to several tribes in surrounding areas, especially to the Cocopah, who live across the Colorado River and speak the Delta-California dialect of the Yuman language family.

Cocopah Tribe. The Cocopah people called themselves *Xawil Kunywavaei*, “those who live on the river.” The traditional home of the Cocopah is near the Colorado River delta,

where many members currently reside, and traditionally includes northwestern Mexico. Their creation story, like ours, involves supernatural beings living under the waters who emerged to create the world. Similar to our creation story, theirs currently serves as a cornerstone in their culture through their cremation ceremony. The Cocopah cremate their dead, including their possessions, and relatives often cut off their hair in mourning. This is all similar to the Quechan ritual, which is rooted in our creation story and life along the Colorado River. Today the Cocopah reservation is located on about 1,700 acres of land near Somerton, Arizona.

In 1917, the Cocopah reservation was established in low-lying desert, bounded by the Colorado River, close to Yuma, Arizona. There are an estimated 1,100 tribal citizens, of which 800 reside on the reservation. Most of the economic development comes from their casino, conference center, hotel, speedway, family fun center, RV resort, golf courses, and some agriculture similar to the Quechan's. The poverty rates are comparable to the Quechan, with about 32% of family households below the poverty threshold (Arizona Rural Policy Institute, 2010b). However, there is a much higher rate of poverty among children, with an estimated 80% of Cocopah children living below the poverty threshold.

RESOLVING SURVEY DATA LIMITATIONS

Aggregating the populations from tribes with similar cultural roots for the collection of data is a part of Indigenous methodology: increasing the sample size for more robust analysis addresses the four common limitations found in federal datasets. The first limitation addressed is internal validity. By aggregating data from two similar tribes, researchers have an opportunity to statistically control for constructs that are normally excluded from

research on Indigenous populations. Often, including more relevant constructs would be impractical because the diverse samples of tribal members would make it nearly impossible to control for constructs such as culture. The second limitation that Indigenous data collection addresses is external validity. Often researchers do not have access to a defined population of people; however, tribes have population information. The third limitation addressed by combining the tribes into a single sample, is statistical-conclusion validity. Especially with education data, individual tribes do not have enough data points to find meaningful relationships, but increasing the sample size reduces the likelihood of Type II errors. Fourth, because there are cultural similarities between tribes, data should be collected from tribes with similar creation stories to ensure measurement validity. This allows researchers to develop measures that are able to capture more nuanced aspects of tribal culture in relation to education and to establish stronger construct validity. It should be noted that Indigenous data collection techniques address some types of validity more intentionally than others.

Internal Validity

A weakness in government and institutional datasets is that they lack variables, such as cultural identity, that are relevant to Native American education, resulting in exclusion of the variables from meaningful analysis. The best method to strengthen internal validity is to randomly assign participants to treatment and control groups, giving all participants an equal chance of assignment to intervention and control conditions; however, in survey research, random assignment is not always possible; other techniques such as statistical controls must be used. Statistical controls do not always work well in governmental and institutional survey datasets because these studies exclude

integral culturally relevant variables in Native American education. When researchers are unable to randomly assign or use statistical controls, it diminishes causal claims because there are other possible explanations for the relationships that are found.

By using Indigenous quantitative methodology, which centers on the voices of community, researchers have a different sampling frame to collect. Two tribes sharing similar creation stories often have comparable characteristics, as evidenced by the description of the Cocopah and Quechan. Having a dataset made up of tribes with comparable demographic, economic, and social characteristics gives researchers a unique opportunity to examine education among Native American students because of the ability to control for a plethora of variables. For example, if researchers want to examine the influence of school attendance on academic success, while controlling for cultural variables among Native Americans, they can use tribally specific cultural items, such as “I participate in our tribe’s cremation ceremony” that control for students’ cultural identity to create a robust set of items that measure the influence of school attendance on academic achievement. Essentially, researchers are able to statistically control for other explanations, which strengthens the internal validity of a study.

External Validity

One of the main weaknesses of government and institutional datasets is the small subsamples of Indigenous students that are not disaggregated by tribe. The lack of representation weakens the external validity of these datasets. The problematic issue of self-identification of students by American Indian / Alaska Native is another challenge. External validity concerns itself with the ability to generalize claims; however, this is nearly impossible because of the limitations of these datasets, the lack of representation from all tribes, and the

problematic self-identification measurement of Indigenous people in current national and institutional datasets.

By using Indigenous data collection, researchers are able to strengthen external validity in ways that these government and institutional datasets are unable to match. First, by combining two tribes, with the possibility of combining even more, researchers are able to randomly sample in a way that strengthens external validity. Because most tribes keep fairly stringent enrollment records and addresses, due to federal requirements and tribal benefits, there is a well-defined population from which to collect a sample. When researchers collaborate with a tribe on sampling, they are engaging the community, which is one of the underpinnings of Indigenous quantitative methods (Snowshoe et al., 2015). The other challenge that Indigenous data collection addresses more thoroughly than government and institutional datasets do is the never-ending argument about Native American identity. Researchers who use government and institutional datasets know that the students in those datasets who self-identified as American Indian / Alaska Native could very well not be. Self-identification can make it difficult to find meaningful relationships between variables. By using Indigenous data collection, the tribes can determine who is American Indian / Alaska Native based on their knowing who is a member of their tribe. So, if there are meaningful relationships between variables that relate to American Indian / Alaska Native education, they may be more likely to be revealed.

Statistical-Conclusion Validity

Another limitation of government and institutional datasets is small sample sizes. Researchers generally want larger samples because they increase the likelihood of finding relationships between variables. This, in

turn, increases the statistical power of a study by finding small relationships between variables. Smaller sample sizes are problematic because they make identifying larger relationships between variables less likely. The small samples of Indigenous people found in many government and institutional datasets often create insufficient power, leading to limitations in statistical conclusions. Underpowered research as a common result of a small sample size is more likely to result in a Type II error or failure to reject the null hypothesis when the null hypothesis is false.

Increasing sample sizes by using Indigenous data collection allows researchers to create a larger sampling frame for studies. With increased sample sizes, researchers are more likely to find smaller meaningful relationships that would otherwise be undetectable because the number of participants is low. Tribes, jointly, provide a large enough sample to increase the statistical-conclusion validity of a study and reduce the likelihood of Type II error, while also maintaining the voices of the community from the creation stories that are integral to the development of Indigenous quantitative methodologies.

Measurement Validity

Measurement is also a difficult feat when using most current government and institutional datasets. The predominant measurement flaw in these datasets is that they do not include variables to measure constructs important to Indigenous education, such as those variables measured by items related to culture. The lack of relevant variables creates limitations in evidence that is based on test content, whereas the measures may only be proxies for what they are intended to measure. Some researchers using government and institutional datasets use items such as “I speak a second language” that can only serve as a proxy for a cultural construct. The available variables are

possibly insufficient to measure constructs necessary to understanding Native American education. It is doubtful that government and institutional researchers creating these datasets have consulted with tribes to see how they can measure culture, but culture is important to Native American education.

Using Indigenous data collection researchers can develop measures that are able to capture more nuanced aspects of tribal culture and examine the relationships to education. For example, a researcher may want to develop items that address the Cocopah and Quechan cultures to examine their influence on students’ achievement. To establish evidence based on test content, the researcher could request participation from each of their tribal historical preservation offices or other experts in tribal history (e.g., language department, culture committee, or elder committee). When the researcher develops a measure based on the tribe’s collective definition of cultural identity, the researcher could send the definition to select tribal departments for feedback, which would further provide evidence based on test content. Once the tribe collectively agrees on a definition of tribal cultural identity, the researcher could use the measure to test Quechan and Cocopah student identity through a factor analysis to see if the items measure what they intend to measure. The community and elder engagement that results from these interactions to develop evidence based on test content strengthens the use of Indigenous data collection as a method of Indigenous quantitative methodology.

LIMITATIONS OF INDIGENOUS DATA COLLECTION

One of the limitations of Indigenous data collection is that it does not address Native American identity holistically. In most governmental and institutionally managed datasets,

Native Americans are self-identified; however, it is important to remember that they may indeed not be American Indian / Alaska Native. Self-identification presents the risk of individuals claiming Indigenous heritage on a survey despite possibly having no relation to an Indigenous community (Lopez & Lucero, 2020). Non-Indigenous people often seek Indigenous status for hiring purposes, college admissions, government contracts, and other opportunities reserved for Indigenous people. For example, around 1 million more individuals self-identified as Indigenous in the 2010 U.S. Census than in the previous decade (Norris, Vines & Hoeffel, 2012). It caused a disproportionate increase in individuals with high levels of education who were not born in states with high densities of Indigenous populations. High numbers of Indigenous folks in 2000 were believed to be people committing ethnic fraud for personal gain (Lopez & Lucero, 2020).

Tribal citizenship is one of the rights that belong to tribes because they are sovereign nations; it is complex and nuanced by such factors as blood quantum, descendancy, and residency. Blood quantum, a common measure used by tribes, represents the amount of blood that traces a bloodline to a particular Indigenous nation. A tribal citizen generally holds a threshold of tribal blood, expressed as a fractional amount (e.g., 1/2). Measuring Native American identity by blood quantum is a debated topic among tribal citizens, tribal nations, and researchers because the concept of blood quantum is derived from a rule of English common law adopted by settler-colonists to differentiate between whole and half relatives in the distribution of inheritances (Spruhan, 2006). The inflow of immigrants who fled Europe to come to the US led to an expansion of this concept and extended it to define legal status for Native American ancestry. The use of blood quantum was reaffirmed in

the Indian Reorganization Act of 1934, which led tribal governments to adopt blood quantum as the measure to determine tribal citizenship (Lomayesva, 1999), a disputed topic among tribal communities, researchers, and policy makers. The problem with using blood quantum for tribal citizenship is that often Indigenous people are excluded from citizenship or are disenrolled for being below an arbitrary threshold. There are instances where the children of tribal citizens are prevented from being enrolled due to not meeting the blood quantum threshold, despite living on tribal homelands, speaking the language, participating in ceremonies, and attending community gatherings.

The continued use of a self-identified measure of American Indian / Alaska Native identity instead of Indigenous data collection can lead to ethnic fraud. Indigenous data collection is naturally mitigated, because tribes have lists of who is enrolled and can identify who is not. Oftentimes, Native Americans are not tribally enrolled because they do not meet an arbitrary blood quantum threshold. A scale to determine Native American identity would provide a more accurate measure. The scale can be developed by using Weaver's (2001) suggestion to use tribal measures to determine identity. Although this attempt at measuring American Indian / Alaska Native identity is certainly flawed, it is better than using self-identification, which tends to produce biased results due to ethnic fraud.

CONCLUSION

Although Indigenous data collection by researchers and tribes would improve the analysis of Indigenous communities, this increased labor—not necessary for the study of other ethnic groups—is not likely to be taken on by governments and institutions. While researchers have recommended over-

sampling since the early 1990s, this has yet to come to fruition (Pavel & Padilla, 1993). Indigenous people have waited for culturally relevant variables and measures of educational achievement since the Merriam Report (1928). This means that Native Americans have waited to review mediocre government-collected data for the past 90 years, and the chances of the data becoming available within the next decade are slim. Because of the need for tribes to make data-driven decisions that inform nation building, tribes must keep pushing forward while continually putting pressure on governments to continue collecting relevant data. At the very least, governments should provide resources to build the capacity of tribes to collect data. Although those in power and those with dominant perspectives will likely question the legitimacy or validity of tribal datasets, tribes will benefit from statistically sound data to inform their decisions.

As a result of limitations found in governmental and institutional datasets, a few tribal nations have started initiatives for *data sovereignty*, the right and capacity of tribes to develop processes and analytic methods to exercise influence over the collection of data by external entities (Rodriguez-Lonebear, 2016). There will still be data limitations; nonetheless, tribal nations are entitled to self-determination because they are sovereign nations. Self-determination and sovereignty relate to the right of tribes and their citizens to self-govern and maintain the trust relationship between the federal government and Native nations (Brayboy, Fann, Castagno, & Solyom, 2012; Cornell & Kalt, 2010). By exerting data sovereignty, tribes are expanding their capacity for nation building (Brayboy, Solyom, & Castagno, 2014; Reyes, 2016). Government and institutional datasets have substantial limitations that hinder their usefulness for nation building.

The limitations found in these govern-

mental and institutional datasets include internal, external, statistical-conclusion, and measurement validity. A study's internal validity will be decreased by using the government and institutional datasets because they often lack variables relevant to Native American education within tribal communities, thus limiting causal claims due to the presence of other possible explanations. Furthermore, accurate tribal data provide multiple avenues for constructing variables to examine alternative explanations to factors influencing tribal education, especially creating constructs related to culture. This increases measurement validity, which is a limitation in existing government and institutional datasets. Government and institutional datasets often have only a small sample of self-identified American Indian / Alaska Native students, weakening the external validity and leading to an underpowered study.

Access to accurate data about the experiences of Indigenous students presents several opportunities. First, representative data reflect the lived experiences of Indigenous students including their environmental and cultural demographics, which can change practitioners' approaches to including Native students. Second, student affairs professionals can use the data to improve the effectiveness of programming to support Indigenous students, leading to higher levels of retention. Third, access to data can help practitioners understand how to support Indigenous communities by revealing the needs of tribal communities.

Collecting data from tribes with similar creation stories gives researchers an opportunity to work with a larger dataset, providing larger samples from an identified population which will increase the power of a study, external validity, and statistical-conclusion validity. Finally, implications and recommendations derived from these data may not apply across tribes, a problem yet to be overcome by these

federal studies that remains a major barrier to researchers and policy makers who use these datasets for public policy regarding Native American education. Researchers need to re-examine how data are collected and to begin thinking of ways to collect data by using

Indigenous knowledges, such as creation stories, to give credibility to Indigenous voices.

Correspondence concerning this article should be addressed to Jameson D. Lopez at jamesondlopez@email.arizona.edu

REFERENCES

- Alvarez de Williams, A. (1974). *The Cocopah People*. Phoenix, AZ: Indian Tribal Series.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *Leadership Quarterly*, 21, 1086-1120. <https://doi.org/10.1016/j.leaqua.2010.10.010>
- Arizona Rural Policy Institute. (2010a). *Demographic analysis of the Fort Yuma-Quechan reservation using 2010 census and 2010 American community survey estimates*. Flagstaff, AZ: Arizona Rural Policy Institute, Center for Business Outreach, W. A. Franke College of Business. Northern Arizona University.
- Arizona Rural Policy Institute. (2010b). *Demographic analysis of the Cocopah Indian tribe using 2010 census and 2010 American community survey estimates*. Flagstaff, AZ: Arizona Rural Policy Institute, Center for Business Outreach, W. A. Franke College of Business. Northern Arizona University.
- Aud, S., Wilkinson-Flicker, S., Kristapovich, P., Rathbun, A., Wang, X., and Zhang, J. (2013). *The condition of education 2013* (NCES 2013-037). Washington, DC: National Center for Education Statistics.
- Battiste, M. (2011). *Reclaiming Indigenous voice and vision*. Vancouver, BC: University of British Columbia Press.
- Bettinger-Lopez, C. (2016) *Violence against Indigenous women and girls: Commitments from the North American Leaders Summit*. Retrieved from <https://obamawhitehouse.archives.gov/blog/2016/07/25/violence-against-indigenous-women-and-girls-commitments-north-american-leaders>
- Brayboy, B. M. J., Fann, A. J., Castagno, A. E., & Solyom, J. A. (2012). *Postsecondary education for American Indian and Alaska Natives: Higher education for nation building and self-determination*. San Francisco, CA: Jossey-Bass.
- Brayboy, B. M. J., Solyom, J. A., & Castagno, A. E. (2014). Looking into the hearts of Native peoples: Nation building as an institutional orientation for graduate education. *American Journal of Education*, 120(4), 575-596.
- Bryant, G. (2013). *Xiipúktan (first of all): Three views of the origins of the Quechan people*. Cambridge, England: Open Book.
- Cadava, G. L. (2011). Borderlands of modernity and abandonment: The lines within Ambos Nogales and the Tohono O'odham Nation." *The Journal of American History*, 98, 362-383.
- Christensen, M. C., & Harris, R. J. (2019). Correlates of bystander readiness to help among a diverse college student population: An intersectional perspective. *Research in Higher Education*, 60, 1195-1226.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155.
- Cornell, S. and Kalt, J. (2010). *American Indian self-determination: The political economy of a policy that works*. (HKS Faculty Research Working Paper Series RWP10-043). Cambridge, MA: John F. Kennedy School of Government, Harvard University.
- Creusere, M., Zhao, H., Bond Huie, S., & Troutman, D. R. (2019). Postsecondary education impact on intergenerational income mobility: Differences by completion status, gender, race/ethnicity, and type of major. *Journal of Higher Education*, 90, 915-939. <https://doi.org/10.1080/00221546.2019.1565882>
- DesJardins, S. L., Toutkoushian, R. K., Hossler, D., & Chen, J. (2019). Time may change me: Examining how aspirations for college evolve during high school. *Review of Higher Education*, 43, 263-294. <https://doi.org/10.1353/rhe.2019.0096>
- Downing, S. M. (2003). Validity: On the meaningful interpretation of assessment data. *Medical Education*, 37, 830-837.
- Echo-Hawk, W. R. (2010). *In the courts of the conqueror: The 10 worst Indian law cases ever decided*. Golden, CO: Fulcrum.
- Feir, D., & Hancock, R. L. (2016). Answering the call: A guide to reconciliation for quantitative social scientists. *Canadian Public Policy*, 42, 350-365.
- Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2011). *Program evaluation: Alternative approaches and practical guidelines* (4th ed.). Upper Saddle River, NJ: Pearson.
- Fixico, D. L. (2003). *The American Indian mind in a linear world*. New York, NY: Routledge.
- Grande, S. (2004). *Red pedagogy: Native American social and political thought* (1st ed.). Lanham, MD: Rowman & Littlefield.
- Höfler, M. (2005). The Bradford Hill considerations on causality: A counterfactual perspective. *Emerging Themes in Epidemiology*, 2, 11.
- Indian Affairs. (2020). *About us*. Retrieved from <https://www.bia.gov/about-us>
- Kovach, M. (2010). Conversational method in Indigenous research. *First Peoples Child & Family Review*, 5, 40-45.
- Lomayeva, F. (1999). Indian identity: Post Indian reflections. *Tulsa Law Journal*, 35, 63-72.

- Lopez, J. D. (2018). Factors influencing American Indian and Alaska Native postsecondary persistence: AI/AN Millennium Falcon persistence model. *Research in Higher Education, 59*, 792-811.
- Lopez, J. D. (2020). Examining construct validity on the scale of Native Americans giving back (SNAG). *Journal of Diversity in Higher Education*. Advance online publication. <https://doi.org/10.1037/dhe0000181>
- Lopez, J. D. & Lucero, D. D. (2020). Measuring cultural identity in Indigenous Communities. In Sumida-Huaman, E., & Martin, N. D. (Eds.), *Indigenous communities and research: Research with, by, and for local peoples*. Ontario: Canadian Scholars Press.
- Lopez, J. D., & Marley, S. C. (2018). Postsecondary research and recommendations for federal datasets with American Indians and Alaska Natives: Challenges and future directions. *Journal of American Indian Education, 57*(2), 5-34.
- Louis, R. P. (2007). Can you hear us now? Voices from the margin: Using Indigenous methodologies in geographic research. *Geographical Research, 45*, 130-139.
- McGregor, D. (2004). Coming full circle: Indigenous knowledge, environment, and our future. *American Indian Quarterly, 28*, 385-410.
- Merriam, L. (1928). *The problem of Indian administration: Report of a survey made at the request of Hubert Work, Secretary of the Interior, and submitted to him, February 21, 1928*. Baltimore, MD: Johns Hopkins Press.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*, 741.
- Murnane, R. J., & Willett, J. B. (2010). *Methods matter: Improving causal inference in educational and social science research*. New York, NY: Oxford University Press.
- Norris T., Vines P. L., & Hoeffel E. M. (2012). *The American Indian and Alaska Native population: 2010*. Retrieved from <https://www.census.gov/history/pdf/c2010br-10.pdf>
- Pavel, D. M., & Padilla, R. V. (1993). American Indian and Alaska Native postsecondary departure: An example of assessing a mainstream model using national longitudinal data. *Journal of American Indian Education, 32*(2), 1-23.
- Radimer, S., & Rowan-Kenyon, H. (2019). Undergraduate men's alcohol consumption: Masculine norms, ethnic identity, and social dominance orientation. *Journal of College Student Development, 60*, 1-16. <https://doi.org/10.1353/csd.2019.0000>
- Rainie, S. C., Schultz, J. L., Briggs, E., Riggs, P., Palmanteer-Holder, N. L. (2017). Data as a strategic resource: Self-determination, governance, and the data challenge for Indigenous nations in the United States. *International Indigenous Policy Journal, 8*, 1-29.
- Reyes, N. A. (2016). "What am I doing to be a good ancestor?" *An indigenized phenomenology of giving back among Native college graduates*. Retrieved from ProQuest Dissertations & Theses Global. (1797939292)
- Reyes, N. A. S., & Shotton, H. J. (2018). *Bringing visibility to the needs and interests of Indigenous students in higher education: Implications for research, policy, and practice*. Association for the Study of Higher Education and the National Institute for Transformation and Equity.
- Richardson, T. A. (2015). Indigenous methodologies and educational research for meaningful change: Parsing postpositivist philosophy of science and mixed methods in collaborative research settings. *Journal of American Indian Education, 54*, 33-62.
- Rodriguez-Lonebear, D. (2016). Building a data revolution in Indian country. In T. Kukutai & J. Taylor (Eds.), *Indigenous data sovereignty* (pp. 253-272). Canberra, Australia: Australian National University Press.
- Rosay, A. B. (2016) Violence against American Indian and Alaska Native women and men. *National Institute of Justice Journal, 277*, 38-45.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Belmont, CA: Wadsworth Cengage Learning.
- Shotton, H., Lowe, S. C., & Waterman, S. J. (Eds.). (2013). *Beyond the asterisk: Understanding Native students in higher education*. Sterling, VA: Stylus.
- Snowshoe, A., Crooks, C., Tremblay, P., Craig, W., & Hinson, R. (2015). Development of a cultural connectedness scale for First Nations youth. *Psychological Assessment, 27*, 249-259. <https://doi.org/10.1037/a0037867>
- Spruhan, P. (2006) A legal history of blood quantum in federal Indian law to 1935. *South Dakota Law Review, 51*, 1-50.
- Sumida Huaman, E., & Martin, N., & Chosa, C. (2016). "Stay with your words": Indigenous youth, local policy, and the work of language fortification. *Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas, 24*, 1-29.
- Teddlie, C., & Yu, F. (2007). Mixed methods sampling: A typology with examples. *Journal of Mixed Methods Research, 1*, 77-100.
- Walter, M., & Andersen, C. (2013). *Indigenous statistics: A quantitative research methodology*. Walnut Creek, CA: Left Coast Press.
- Weaver, H. N. (2001). Indigenous identity: What is it and who really has it? *The American Indian Quarterly, 25*, 240-255.
- Wilson, S. (2008). *Research is ceremony: Indigenous research methods*. Black Point, NS, Canada: Fernwood Publishing.
- Windchief, S., Polacek, C., Munson, M., Ulrich, M., & Cummins, J. D. (2017). In reciprocity: Responses to critiques of Indigenous methodologies. *Qualitative Inquiry, 24*, 532-542. <https://doi.org/10.1177/1077800417743527>
- Wine, J., Bryan, M., and Siegel, P. (2014). *2011-12 National postsecondary student aid study (NPSAS:12) data file documentation* (NCES 2014-182). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.
- Zuberi, T., & Bonilla-Silva, E. (Eds.). (2008). *White logic, white methods: Racism and methodology*. Lanham, MD: Rowman & Littlefield.