Data-Driven Decision-Making

Sophie J. Barbu, Karen McDonald, Lisceth Brazil-Cruz, Lisa Sullivan,
and Linda F. Bisson

Abstract Addressing barriers to inclusion requires understanding the nature of the problem at the institutional level. Data collection and assessment are both crucial for this aim. In this chapter, we describe two important classes of data: (1) data on diversity that define the potential nature of the issues at stake and the need for change, and (2) data on assessing the usefulness of new programs, processes, and policies in creating a more diverse institution. Both sets of data are important for effective decision-making. At the same time, data analyses can be challenging because issues of equity and inclusion are complex and determining the basis of comparison or the “ideal” diversity target can be difficult. Nevertheless, data gathering and analysis are critical to assess progress and to provide a basis for both accountability and efficacy. Moreover, the ability to document that a problem indeed exists will help justify the need for change and, ideally, spur corrective action.

S. J. Barbu

UCD NSF ADVANCE Program, University of California, Davis, Davis, CA 95616, USA

e-mail: sjbarbu@ucdavis.edu

K. McDonald

Department of Chemical Engineering, University of California, Davis, Davis, CA 95616, USA

e-mail: kamcdonald@ucdavis.edu

L. Brazil-Cruz

UC Davis ADVANCE Program, University of California, Davis, Davis, CA 95616, USA

Institutional Effectiveness, Woodland Community College, 2300 E. Gibson Road, Woodland, CA 95776, USA

L. Brazil-Cruz

e-mail: lizcruz@ucdavis.edu; lbrazil@yccd.edu

L. Sullivan

School of Education, University of California, Davis, Davis, CA 95616, USA

e-mail: lhsullivan@ucdavis.edu

L. F. Bisson (✉)

Department of Viticulture and Enology, University of California, Davis, Davis, CA 95616, USA

e-mail: lfbisson@ucdavis.edu

Keywords Diversity • Data • Data collection • Data assessment • Program evaluation • Outcomes

1 The Nature of Data Needed

Assessing barriers to inclusion in academia requires tallying numbers and advancement rates by group, but also using surveys and other vehicles to measure the sense of belonging, if any, felt by faculty from underrepresented social identity categories. The first step in assessing current levels of inclusion is to measure various aspects of diversity and inclusion at the institutional level. For ADVANCE grant awardees, the National Science Foundation requests data for faculty on gender that address the following four questions:

1. What is the distribution of science and engineering faculty by gender, rank, and department?
2. What are the outcomes of institutional processes of recruitment and advancement for women and for men?
3. What is the gender distribution of science and engineering faculty holding leadership positions in the institution?
4. What is the allocation of resources for science and engineering faculty by gender at the institution?

Expanding these questions to other underrepresented identities requires accurate measurement of current information on these identities. Ideal or target diversity is often linked to the diversity of the general population in the region served by the institution. Underrepresented minorities, or URM, are identities with lower percentages in an analyzed population as compared to the general population. The expectation is that, if no barriers to inclusion existed, the percentages of identities across all sectors of society would be the same. Collecting data within a subgroup can also reveal systemic bias if identities are well represented at lower levels but underrepresented in leadership positions. Further, equity of salary and resource allocation can uncover bias if certain identity categories tend to have greater or lesser access to resources. However, several issues with the collection of such data can affect the data's validity.

2 Issues in Data Collection, Management, and Assessment

Ideally, data are collected with a specific question in mind, but in reality the nature of the available data often drives which questions can be asked. At the institutional level, databases may be designed with a specific intent that does not necessarily align with the goal of measuring aspects of diversity and inclusion. Consequently, the limitations of the data need to be understood and considered in a study's design. Other issues with data include the unintentional introduction of bias, such as in choice

of categories or in the nature of the inquiries made. For example, in the specific case of measuring diversity and inclusion, diversity is only representable by the categories actually made available. Categories are often thought to reflect identities, but those identities may in fact derive from societal constructs and may not represent the breadth of diversity that exists within a given community. In addition, it is often difficult to interpret the meaning or significance of differences found among different identities without knowing the specific context or the metadata. For example, start-up packages for a new faculty member are difficult to compare directly without knowing the specific situation of the new hire, such as availability and access to existing research equipment, versus the need to invest in a new area. Differences in the number of questions that women or URM faculty candidates are posed during their job interview could represent either higher interest and enthusiasm for the presentation or a questioning of the candidates' expertise.

2.1 *Qualitative Versus Quantitative Approaches*

Both qualitative and quantitative approaches provide us with important information. However, the types of information these methodologies yield are different and therefore lead to different discoveries.

Quantitative research is concerned with unveiling facts about a particular phenomenon. Through the use of numbers, this method assumes a fixed reality, which can be measured. Therefore, this method allows us to answer the “*what*” questions. This research method is often used to test a theory in order to support it or reject it. Further, this allows for researchers to control for extraneous variables. Methodologically, data are collected through the use of instruments that help measure items, such as surveys. These data are then analyzed through statistical comparisons (Fig. 1).

Qualitative research, by contrast, is concerned with understanding social phenomena from the perspective of the research participant. It aims to understand the social reality and lived experience of participants by focusing on the “*how*” and “*why*” questions of a particular context. Qualitative research is an umbrella term covering a range of methods, though all involve an interpretive, naturalistic approach to their subject matter (Denzin & Lincoln, 2000). Data may be collected through participant observation, interviews, or focus groups, among other methods. The data may then be coded and analyzed thematically. Most important, data reflect the perspective of the participant as documented and narrated by the researcher. Qualitative researchers often do not hide or deny their own identities, social commitments, or perspectives, and may even consider them resources enabling the research process (Fig. 2).

Fig. 1 Quantitative assessment of data

Fig. 2 Qualitative assessment of data

2.2 Data Collection Issues

The questions asked and the type of data gathered will affect how data are interpreted. For example, if a primary data variable is race and if racial data are collected, then conclusions will be viewed through a racial lens and race will be the most important factor driving the analysis. Ease of measurement and accuracy of measurement are also factors in data collection design. Race is defined in the U.S. largely on the basis of physical features such as skin color and therefore is assumed to be readily determined but in fact has low value in defining groups with shared values, experiences, and beliefs. Ethnicity is used to define groups with a common culture or nationality and can reflect shared values and traits. However, ethnicity is actually fluid, as well; subcultures, such as those in rural or urban areas, may affect how ethnicity is expressed. Therefore, the process of categorization of data collected is highly problematic, as it tends to create artificial groupings that may not accurately reflect differences among individuals and may reinforce stereotypes. Equally challenging is defining the ideal forms “diversity” should take or the end goal of inclusion. Several other issues can have an impact on data interpretation.

2.2.1 Categorization

A primary issue in measuring diversity is the process of categorization. Data on identity are based on categories used by the U.S. government, as defined by the United States Census Bureau. Individuals are asked to self-identify by selecting from a list of identity categories but individuals who select a given category may only marginally self-identify with it. The process generally provides only one to three options, typically in terms of race, gender, or age, which may or may not be central to self-identification from the individuals’ perspective.

Data are often collected on race from a list of options; depending on the questionnaire or form, multiple races can sometimes be reported by an individual. “Race” has long been established to be socially constructed. There is no biological basis for characterization by race as it is largely defined on the basis of skin color, hair texture, or facial features. Sex and gender are largely reported as binary, men/male or women/female. Likewise, typically only two classifications of ethnicity are listed: Hispanic/Latino or not Hispanic/Latino. These categorizations, of course, do not represent the spectrum of identities that exist in society. Some faculty also expressed concern with categorization by surname. Individuals with Hispanic-sounding surnames, for instance, were included in the “Hispanic” category whether they identified ethnically with that category or not (see the Chapter, [‘Assessing Institutionalized Bias’](#)).

Aggregation of individuals in categories can obscure discrimination of subgroups. For example, “Hispanic” is defined by the U.S. Census Bureau as “a person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin regardless of race.” The definition states that “Hispanics can be of any

Fig. 3 The challenge of category selection. Illustration by Meghan Crebbin-Coates

race, any ancestry, any ethnicity.” This would include individuals from Europe and Spain. In 2016, 17.8% of the U.S. population was Hispanic/Latino but only 3.6% of engineering faculty nationally identified as Hispanic (Arellano et al., 2018). A deeper analysis of the national data showed that of the roughly 600 faculty members in this category, only 48 were born in the United States. Thus, barriers to inclusion for native-born Hispanics may be stronger than those for foreign-born members of the same designation. As a result, aggregation of data into a single broad category can obscure the accurate assessment of bias.

Finally, the selection of identity categories itself can confound assessment of the true nature of the barrier to inclusion. Underlying barriers to inclusion may be more complex than the identity categories provided (race, gender, or age), and may also involve factors, such as privilege, access, and socioeconomic standing (Fig. 3).

2.2.2 Defining “Ideal” Diversity

Another issue with data collection centers on the felt need for an “ideal” to exist, as defined in relation to diversity and inclusion efforts as the absence of discrimination. This ideal end-goal guides data-driven decision-making but poses a problem because we do not currently have a clear outcome goal, except to increase diversity and

inclusion for underrepresented minority groups. Often the term “underrepresented minority” is meant to refer to a single identity (race), so issues unique to intersectional identities (identities held by persons of “mixed race” or women of color) can be overlooked, as can the influence of additional factors besides race when race is the only metric. In other words, use of a single identity category can ignore the impact of more-complex interactions. It also assumes that lack of “seeing self”—that is, individuals of the same, specific, defining identity—is the main obstacle to achieving the desired level of diversity.

Use of overall population demographics may not be practical if there are intersecting factors such as age and race, with race percentages varying by age groups. Also, nationally aggregated data may not reflect local employee pool demographics. For example, with respect to faculty hiring, the diversity of the national pool of doctoral degree awardees rather than of the general population is often used to assess possible bias in hiring. But it may be easier to diversify an institution that is situated in a more-diverse community than to diversify one in a non-diverse region.

2.3 *Data Management Issues*

In analyzing our institution’s own data on diversity, especially with respect to the questions raised by the ADVANCE data assessment tool kit listed above, we discovered several local issues with data management that are important to address. We discuss these below.

2.3.1 *Decentralization*

At UC Davis, various administrative units have been charged with collecting demographic, hiring, and personnel data. This decentralization has led to variations in how data were collected and categorized. For example, some departments span colleges; as a result, in some cases the department is considered to be fully in one college but in other cases is considered fully in the other college. This leads to a difference in numbers of total faculty within units.

Often the specific units had their own goals in mind for data collection, so the process of collating data across units to address broader issues was challenging. For example, data on recruitment and advancement are categorized demographically, but data on start-up packages and space allocations are not. Therefore, it was tough to address question 4 above (related to resource allocation), and to weigh the potential for gender bias (or other types of bias) in resource allocation. Some data are held locally, within a college, and conventions used for data collection and categorization may differ by unit if the data are to be evaluated within the college. Units typically develop their own unique sets of questions they are trying to address, and merging these datasets can be challenging. We also discovered that personnel involved in data entry at the college level are not trained centrally to process data in a consistent

manner. Therefore, categories used for data tabulation were interpreted differently between colleges.

2.3.2 Aggregation Algorithms

Another issue with data management concerns how the data are aggregated to avoid identification of specific individuals. This can be done by grouping smaller units into larger ones, but one problem with this approach is that it can obscure local, department-specific problems. For example, gender parity varies across the STEM departments, and combining data from more-diverse larger units with less-diverse smaller ones may conceal a problem that is primarily local.

We also identified problems generated by the attempt to re-create data sets from original data. It was not clear how algorithms that aggregated data had been designed, and when individuals responsible for those algorithms leave their positions, the knowledge of aggregation procedures may be lost. In addition, datasets are inherently dynamic, so a “snapshot” taken at one point in time during the academic year can differ from another point in time. Rolling, multiyear averages are sometimes used to smooth out fluctuations, particularly in cases where small numbers would likely allow identification of specific individuals. Yet data obtained using this approach are particularly difficult to validate or confirm in retrospect.

2.3.3 Use of Terms Open to Interpretation

A final issue with data management is the use of terms open to interpretation. When combined with decentralization and lack of consistent training of data entry personnel, the same terms can be used differently. For example, the term “involuntary separation” was originally thought to mean termination for cause. With respect to faculty, this would suggest a failure to obtain tenure or to advance in rank, or perhaps the negative outcome of a misconduct charge. However, we learned that one unit used this category to document separations due to illness, when in fact that should be its own distinct category.

We also noted differences in data entry practices. A high-interest category is “retention” and the reasons faculty give for leaving a position, especially faculty in underrepresented identity categories. We learned that if faculty members did not select a reason from a predetermined list, the data entry personnel made a “best guess”—which could be wildly incorrect. It is important to understand the actual trends in retention failures. Individuals leave for many reasons: to accept a better offer (salary, position, or resources) at another institution; to better match the partner’s career goals or family-of-origin needs; to resolve dissatisfaction with their current position, which may require leaving academia altogether. This latter reason is more important than the others in evaluating the inclusiveness of a campus climate.

2.4 Data Assessment Issues

Analysis and interpretation of data can be truly challenging. Categorizations of individuals have changed over time, and individuals who initially responded to categories under an older system might not respond the same way with a more-expanded set of category identifications. Cohort issues may also span changes in policy or practice. Prior to the 1996 passage of California Proposition 209 (the “California Civil Rights Initiative”), several affirmative action programs were in existence at UC Davis and drove diversity goals. These programs were terminated after the proposition was passed by the state. Cohorts hired under these different sets of conditions may therefore vary in how they are “counted” with regard to measures of diversity.

An analysis of University of California faculty salary equity, made in 2014, found that faculty hired before that year had lower current total and off-scale salaries, had been appointed at lower interval steps, and had lower off-scale salaries at the time of hire. Salaries of more recent hires were higher because of greater competition for candidates and cost-of-living adjustments. Time of hire was a more important factor than gender, race, or ethnicity. There was no connection seen between advancement through the ranks and salary equity, suggesting that faculty who sought outside positions to bolster retention efforts (including garnering off-scale salary) was more a determining factor in salary inequity than in the programmatic impact of the individual. However, the campus administration considered the incidence of salary inequity minor and not associated with either gender or URM status, indicating that no systemic bias was in play, possibly due to the prior existence of salary equity review for all advancement actions.

3 Uses of Data

Diversity data are collected and rigorously examined for four principal reasons:

1. It is vital that faculty members know their own situation and how over- or under-represented they are in terms of diversity goals. It is also important to assess diversity along the entire academic path—undergraduate, graduate, and postdoctoral—leading faculty positions, so as to learn where the blockage or barriers to inclusion arise. If undergraduate majors display the same demographic percentages as the general population but graduate school enrollees do not, then a barrier may exist at the point of application or acceptance into graduate school; or there may be other, unidentified reasons or incentives for graduates to pursue other paths. In many cases, rather than doing a comparative assessment of general population demographics, a comparison to Ph.D. graduate demographics might reveal more about problems with faculty hiring. If, for example, Ph.D. graduate and junior faculty demographics are similar, then we can reasonably conclude that the issue with lack of representation occurs not

- at the point of faculty hiring but at the point of retaining diversity in graduate programs.
2. Data are important drivers first for educating the community served by the institution and then for making the case for change. Campus data demonstrated that bias was indeed operating at the point of hire for URM scholars, because the demographics of various pools of interviewees did not match the demographics of actual new hires. These data were vital to justifying the need for mandatory training in implicit bias. Comparing data across the faculty career spectrum can identify the different points at which bias may come into play. Our data suggested that, once hired, there was no disadvantage to any particular group during advancement, meaning the major barrier is indeed at the point of hire. Retention data suggest that there may be some differential loss of URM scholars. However, this seems to be a result of recruitment away from the institution by other institutions, rather than a lack of equity in offering retention packages.
 3. Data are important for assessing the effectiveness of programmatic change. Surveys of the social climate for inclusion conducted both before and after the implementation of new policies and practices can define the most successful interventions as well as the best practices for creating a more inclusive climate. Our initial COACHE survey identified mentoring as being a deficiency in creating supportive academic cultures. Several programs, such as the LAUNCH program for new faculty (described in the Chapter, '[Making Visible the Invisible: Studying Latina STEM Scholars](#)'), were initiated at the start of the ADVANCE grant. The follow-up COACHE survey at the end of the grant period showed measurable increases in satisfaction with mentoring.¹
 4. Finally, data enable more astute comparisons to programs in other institutions. Comparison institutions are, by definition, similar in research status and size of faculty. Identifying institutions and comparing self-assessments of diversity can determine those with better (or worse) track records, and, in turn, which information can be used to define and refine best practices for creating more inclusive campus climates.

4 Assessment of Program Effectiveness

Equally important to generating initial data is the measurement of the effectiveness of changes made to enhance inclusion. It is critical that such measures be designed and gathered by an outside group or evaluation team. This prevents bias in interpretation of the findings that might arise if individuals with a vested interest in the outcome were responsible for data interpretation. Also, use of external evaluation can serve to encourage participation in the review process by ensuring the anonymity of the responders.

For the ADVANCE grant, we engaged both external and internal evaluators. The external evaluator reviewed survey data and conducted interviews with individuals

¹ COACHE stands for the Collaborative on Academic Careers in Higher Education.

and administrators on campus to help provide an independent assessment of the effectiveness of programs aimed at increasing diversity, equity, and inclusion.

The internal evaluation team, composed of evaluators in the UC Davis School of Education, was independent of the ADVANCE group. They came on board at the very beginning of the grant to help guide the institutional change framework and to align both impact and outcome measures with grant goals. The internal evaluation team worked independently of and collaboratively with grant leadership to document implementation activities, successes, and challenges. The information gathered through the program evaluation was used to refine and adjust grant activities and to gauge outcomes.

Evaluation activities included surveys conducted of attendees at ADVANCE events to assess the impact of the event in terms of the delivery of information content as well as individuals' sense of belonging. The team also evaluated the impact and effectiveness of training in implicit bias through surveys of participants. This survey information was vital in showing the effectiveness of the program and was used to make adjustments to content. The internal evaluators also provided an independent assessment of the COACHE surveys on faculty job satisfaction. Finally, the external evaluator conducted interviews of key stakeholders and participants, particularly with respect to new programs offered to bolster a sense of belonging at the institution. Such third-party interviews are important in assessing program effectiveness.

5 Key Findings from Internal Evaluation

The internal evaluation team documented the success of new policies and practices as well as ongoing challenges. Most important was the documentation of the impact that implicit bias training had on all members serving on faculty search committees. Surveys were conducted after each Strength Through Equity and Diversity (STEAD) training. Collation of the data indicated high degrees of satisfaction with both the program and its content. The comments section invited respondents to indicate which methods or content worked well, and comments were used to refine the training. The team was also able to document increased mentorship and support of Latina faculty, greater job satisfaction, and an increased sense of belonging as some of the consequences of newly launched mentorship programs. The data on reported satisfaction with these programs justified their continued support. It is essential that such evaluations be conducted at arm's length by knowledgeable people with no vested interest in the outcome.

The evaluative assessments also revealed the following:

- In general, the lag time between new policy implementation and impact makes it challenging to measure immediate impact.
- If multiple processes are changed simultaneously, it can be difficult to know which one or ones had the greatest impact.

- It is difficult to sustain momentum for change under new leadership, if individuals feel uncertain about the new leaders' goals and priorities.
- Attempting to measure the inclusivity of the local campus culture climate can be challenging if specific colleges or departments have attracted few individuals belonging to underrepresented minority groups. Individuals from such groups might be dissatisfied with the local social climate but feel uncomfortable providing feedback.
- The impact of new policies and practices on campus climate among members of underrepresented identity categories can be difficult to determine over a short time span and may not be immediately measurable.

The surveys underscored the impact that mentoring can have on new faculty members' sense of belonging and ultimate job satisfaction, and also supported the continued use of LAUNCH committees. Such committees are composed of three to five senior faculty members, some of whom come from other departments than the new hire. These committees are an effective sounding board and a valuable source of advice to all new hires. More important, the committees represent a commitment to the success of the new hire, which in turn will bolster an authentic sense of belonging in academia.

6 Conclusions

Meaningful data are essential to making the case for institutional change, defining the changes needed, and assessing impact. Data, especially from surveys, can define the points in the recruitment and retention process at which barriers to inclusion exist. Data collection is essential for identifying the most successful interventions that enhance inclusion. In our case, by far the most successful tool was the development of STEAD training in implicit bias. Implicit bias was neither generally known as a phenomenon nor understood as having a negative effect on recruitment practices; this has changed as a result of the training implemented. Another summary observation is that some data are challenging to obtain in a usable format, so the campus is committed to greater centralization of data so they can be both more robust and more useful in decision-making.

Acknowledgements Authors thank Meghan Crebbin-Coates for drawing Fig. 3.

References

- Arellano, G. N., Jaime-Acuña, O., & Graeve, O. A. (2018). Latino engineering faculty in the United States. *MRS Bulletin*, 43, 131–133.
- Denzin, N. K., & Lincoln, Y. S. (2000). *Handbook of qualitative research* (2nd ed.). SAGE Publications.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.