

Classifying Higher Education Institutions: Lessons from the Carnegie Classification

Clasificación de las instituciones de educación superior: lecciones de la Clasificación Carnegie

Alexander C. McCormick

Center for Postsecondary Research, Indiana University Bloomington, USA

Abstract

A broad overview of the purposes, issues, and challenges associated with classifying institutions of higher education is presented, drawing on the experience of The Carnegie Foundation for the Advancement of Teaching's widely used classification of colleges and universities in the United States. The Carnegie Classification serves as a convenient example to illustrate general issues and challenges related to classifying institutions of higher education. The presentation is organized around three fundamental questions that should be considered in the design of a classification system: (a) Why classify? (b) How to classify? and (c) What secondary uses and consequences can be anticipated? The article concludes with a short list of recommendations that may help inform a nascent classification effort.

Keywords: universities, classification, taxonomy, typology, policy, ranking

Resumen

En este artículo se presenta una amplia sinopsis de propósitos, problemas y desafíos asociados a la clasificación de instituciones de educación superior, con base en la experiencia de la clasificación de instituciones de educación superior y universidades por parte de The Carnegie Foundation for the Advancement of Teaching, muy en boga en Estados Unidos. La Clasificación Carnegie sirve como ejemplo para ilustrar aspectos y desafíos generales relacionados con la clasificación de instituciones de educación superior. La presentación se organiza en torno a tres preguntas fundamentales que debiesen considerarse en el diseño de un sistema de clasificación: (a) ¿por qué hacer una clasificación?, (b) ¿cómo hacerla? y (c) ¿qué usos secundarios y consecuencias pueden anticiparse? El artículo concluye con una breve lista de recomendaciones que podrían servir de información en cualquier nueva iniciativa de clasificación.

Palabras clave: universidades, clasificación, taxonomía, tipología, política, ranking

Post to:

Alexander C. McCormick
Center for Postsecondary Research, Indiana University Bloomington, USA
Indiana University Center for Postsecondary Research
Address: 1900 E. 10th Street, Bloomington, IN 47406, USA
Email: amcc@indiana.edu

The author served as director of the Carnegie Foundation's classification program from 1999 through 2007, during which time he led a major overhaul of the classification system that took effect in 2005. The views expressed are those of the author and do not represent official positions of The Carnegie Foundation for the Advancement of Teaching.

© 2013 PEL, <http://www.pensamientoeducativo.org> - <http://www.pel.cl>

ISSN: 0719-0409 DDI: 203.262, Santiago, Chile
doi:10.7764/PEL.50.1.2013.6

This article provides a broad overview of the purposes, issues, and challenges associated with classifying institutions of higher education, drawing on the experience of The Carnegie Foundation for the Advancement of Teaching's widely used classification of colleges and universities in the United States. Readers should, of course, exercise caution in making judgments about the applicability of the Carnegie Classification experience to other national contexts. My purpose is to use the Carnegie Classification as a convenient example to illustrate general issues and challenges related to classifying institutions of higher education. While it is useful to consider these issues and challenges when contemplating the creation of a classification system, by no means should this article be interpreted as endorsing the adoption in other countries of a system designed for use in the United States. Indeed, as will become evident later on, even *within* the United States there are serious concerns about contemporary uses of the classification, which was developed in the context of specific analytic purposes and yet has been appropriated for a far wider set of purposes.

The presentation is organized around three fundamental questions that should be considered in the design of a classification system: (a) Why classify? (b) How to classify? and (c) What secondary uses and consequences can be anticipated? In elaborating on each question, I offer examples and cautions drawn from the history of the Carnegie Classification. I conclude with a short list of recommendations that may help inform a nascent classification effort.

Why classify?

The “why” question goes to purposes: What is the problem a classification is supposed to solve? What can be done with a classification that cannot be done now, or what will be done better with a classification?

Classifications of higher education institutions can serve a range of purposes. From a research standpoint, they can offer fresh insights into the structure and function of a nation's higher education system, for example by facilitating investigation into the flows of inputs and outputs. What types or sectors are best supported and least well supported by government? Which institutional types serve which types of students? Which institutional types produce the greatest social benefit with regard to knowledge production, talent development, national and regional economic development, training individuals to serve important social institutions such as schools and government, the formation of active and informed citizens, and other important purposes? Is there an imbalance with regard to the mix of types relative to social needs and national priorities? From a rational decision-making perspective, the answers to these questions can help guide the allocation of scarce resources. From a political perspective, however, a classification might be seen as merely providing political cover for controversial or contested decisions under the guise of scientific objectivity.

The Carnegie Classification was created with a specific purpose in mind. The Carnegie Commission on Higher Education was formed in 1967 by the Carnegie Foundation for the Advancement of Teaching with the aim of studying and making recommendations for U.S. higher education in the final decades of the 20th century. The Commission judged the simplistic typologies then in use—such as simple differentiation by control and degree level—to be ill-suited to its needs. The Commission thus set about developing a classification system that would better inform its research needs. The new classification was first used in a 1971 Commission publication titled *New Students and New Places: Policies for the Future Growth and Development of Higher Education*. There were no institutional listings. Rather, the new classification merely supplied the analytical categories for the report's aggregate analysis of the U.S. higher education system and its projected growth. It was not until two years later that it published the detailed list of institutions by category for use by “individuals and organizations that are engaged in research on higher education” (Carnegie Commission on Higher Education, 1973, p. vi). This history calls attention to two important but largely forgotten details of the Carnegie Classification's history: (a) the classification was created, and later disseminated, as a *research* tool; and (b) the design of the classification reflected the particular research needs and interests of the organization that created it. The latter fact helps to explain, for example, why doctorate-granting universities were much more finely differentiated (four categories for 173 such institutions in the 1973 edition) than were two-year colleges (a single category for 1,063 institutions).¹

¹ For a fuller treatment of the history and development of the Carnegie Classification, see McCormick, 2001.

Although the Carnegie Classification continues to be used in research, it has evolved into a general-purpose taxonomy quite divorced from its original purposes and applied to a wide variety of uses for which it was not designed. I will return to this important point in the discussion of secondary uses and consequences. A key lesson to be derived from this history is that a classification's uses and broader understandings of its purposes can diverge appreciably from how it was originally conceived. While not necessarily a bad thing, this should serve as a cautionary tale for the prospective classifier, and it calls attention to the importance of anticipating unintended uses and effects.

How to classify?

Institutions of higher education are inherently complex organizations performing a wide range of functions. Yet the task of classification is one of simplification—of reducing cognitive burden by identifying the most salient differentiating characteristics, given particular purposes. A well-designed classification system provides categories of roughly similar entities, such that any member of a given category resembles the other members more than it does any member of a different category. That is, the categories are relatively homogeneous, and they are also distinct from one another. Here the classifier faces a trade-off. McCormick and Zhao (2005) articulate it as balancing precision and parsimony:

As categories are defined more precisely, the number of categories increases, as does homogeneity within them, while the size of the group within each category declines. Favoring parsimony yields more manageable and more easily comprehended classifications made up of fewer categories but with more members and more variation within the categories (pp. 53-54).

Increasing the number of categories improves homogeneity within categories, but as the number of categories increases, so does the complexity of the classification, undermining the original goal of simplification. When the Carnegie Foundation expanded its classification system in 2005, it arguably sacrificed simplicity so as to acknowledge both (a) the evolution of the classification into a general-purpose taxonomy, and (b) a belief that reducing the complexity of U.S. higher education to a single set of categories and associated labels represented excessive and potentially pernicious simplification. (The new classifications were designed to enable users to simplify a given classification by aggregating related categories, but it is unclear whether this new flexibility has been recognized or embraced by many users).

Defining an institution

In approaching the task of classification, the first operational challenge one confronts is how to define a classification entity. What counts as a university for the purposes of classification? This judgment involves both simple eligibility decisions (e.g., accreditation, degree offerings, or governmental authorization and recognition) and more challenging decisions with regard to the geographic range of a university's operations. Decision must be made with regard to satellite campuses, including operations in other countries, as well as online offerings. Are all of these disparate operations considered part of the university for classification purposes? Should their attributes—students, faculty, finances, etc.—figure in classification decisions? In the case of the Carnegie Classification, a simplifying decision was made to adopt the existing institutional definitions established in U.S. government data systems (initially the identifiers created by the Federal Interagency Committee on Education, and later the U.S. Department of Education's Integrated Postsecondary Education Data System [IPEDS]). While neither system was perfect, their centrality in the U.S. data infrastructure facilitated the use of existing national data in constructing the classification, as well as the ability of researchers to use the classifications in combination with other data.

Choosing the methodology

Another critical decision involves the classification methodology. Scholars of typology and classification draw a distinction between idiographic and nomothetic techniques (see Bailey, 1994). Idiographic techniques use *a priori* criteria for assigning entities to categories (as in the case of the Carnegie Classification), while nomothetic methods take an inductive approach, using statistical techniques such as cluster analysis to group entities according to underlying patterns in the data (for an example in higher education, see Brint, Riddle, and Hanneman [2006]). Both approaches involve judgments on the part

of the classification researcher, with the most important ones involving the selection of the underlying attributes that will be used to create the classification. Idiographic approaches have the advantage of being relatively straightforward and easily understood, but at the cost of creating groupings that may make intuitive sense (e.g., reproducing conventional understandings), but that lack much empirical justification. Nomothetic methods, on the other hand, capture empirical patterns of similarity and dissimilarity revealed in the data, but at the risk of opacity: there is typically no simple way to explain why a particular entity is assigned to a given category. Another shortcoming is the requirement to use the same underlying data to classify all entities, which may lead to excessive simplification in the selection of attributes that will inform the classification. This shortcoming could be ameliorated by partitioning the universe of entities into subsets (for example, public, private not-for-profit, and private for-profit institutions; or institutions that only offer undergraduate education, those that only offer graduate education, and those that do both), and using different groups of variables used to categorize each of the subsets. This amounts to a hybrid approach, in that the *a priori* partitioning follows an idiographic logic, and nomothetic methods are subsequently applied within the subsets.

Nomothetic approaches may be more suited to pure research, while idiographic approaches may be more desirable in an applied setting if comprehensibility and transparency of classification decisions are particularly important. Because this paper focuses on the experience of the Carnegie Classification, the discussion below focuses on the idiographic approach. However, the broader points about selection of characteristics (variables) to be used for classification apply equally to nomothetic classification methods.

Selecting data elements

Specifying the purposes of a classification is the first step in identifying the specific data elements that will be used to assign institutions to categories, because it suggests something about the information needed to assess similarities and differences. In describing the Carnegie Commission's first classification, Clark Kerr wrote that the Commission "sought to identify categories of colleges and universities that would be relatively homogeneous with respect to the functions of the institutions as well as with respect to characteristics of students and faculty members" (Carnegie Commission on Higher Education, 1973, p. v). The focus on institutional "function" is why the Carnegie Classification is typically thought of as differentiating U.S. colleges and universities with respect to institutional mission. When the classification was expanded from the single classification to a set of parallel classifications in 2005, the Carnegie Foundation specified three questions that would undergird the new classifications: What is taught, to whom, and in what setting (McCormick, 2005). The new classifications, then, paid particular attention to the teaching function, responding in part to criticism that Carnegie's traditional classification paid insufficient attention to teaching, and indeed that it prioritized research over teaching.²

In considering data sources, it may be useful to think of three broad domains: inputs, processes, and outputs. Inputs are the raw materials that the university operates with, such as the physical plant and facilities (buildings, library holdings, etc.); financial resources in the form of government support, corporate and foundation support, revenue from student fees, and charitable contributions; and human resources (including students, faculty, administrative staff, and support staff). Processes are the operations by which inputs are transformed into outputs: instruction, research activity, training of staff, economic development and outreach activities, and so on. Outputs are the products: graduates; and artifacts of research and creative activity such as scholarly articles, books, reports, patents, and creative works in the fine and performing arts. Outputs may also include more diffuse and less easily measured or attributed consequences of university activity, such as employment, overall health of the local or regional economy, public health, civic participation, and more.

Some of these—especially inputs—are routinely tracked and documented by institutions or governmental agencies, and they may be the most readily available. Many processes and outputs are more difficult to measure. Some may not be measured at all, and others may not be measured in a way that permits their attribution to a particular institution (for example, if a region is served by two, three, or four universities, how should each institution's contribution to economic development or civic vitality be assessed?).

² Critics saw this as particularly ironic, given that the classification was created and maintained by an organization whose name conveys a dedication to "the advancement of teaching."

Although it is beyond the scope of this paper, each of the categories of information examined above can be measured in many different ways, and these choices must reflect the purpose of the classification. Consider some simple examples. Few would disagree that students represent an important input in higher education. But what *about* students should be taken into account? Should a classification differentiate with respect to the various degrees or qualifications that they might be pursuing, or is it sufficient to simply include the total number of students? Should student characteristics, such as their academic ability or their demographic characteristics (gender, race/ethnicity, and social class), be taken into account? If a university's contribution to educational opportunity and equity is deemed to be important for classification purposes, the composition of the student body (as well as that of those who graduate) would be an important criterion. With regard to the outputs of scholarly activity, should all outputs count equally, or should impact in the form of journal quality or citation counts be taken into account?

In addition to inputs, processes, and outputs, certain structural features may be relevant to a given classification, such as public or private control, religious affiliation, presence of a medical school, and so on.

The discussion thus far has been limited to objective measurements. But more subjective information is also potentially available to the classifier. One such example is information bearing on institutional identity, such as mission statements or perceptions of identity expressed by institutional constituents (students, faculty, leaders, trustees, alumni, etc.). Unless perceptions of identity are central to a classification's purpose, however, they should be used with caution because they may be at odds with actual behavior. One aspect of the Carnegie Classification that is seen by many as a strength is its reliance on objective measures such as degree conferrals. In this regard, the Carnegie Classification has always differentiated institutions with regard to their *enacted* mission—that revealed in their behavior—as opposed to their *espoused* mission.

More broadly, it may be important to consider whether a preference for objective measures introduces a systematic bias into the classification. Some important aspects of university activity are simply less amenable to measurement (for example, the development of so-called “soft skills” that are increasingly seen as important outcomes of education). What are the consequences of overlooking that which cannot be reliably measured? This goes to the use and interpretation of the classification and will be taken up in the discussion of secondary consequences below.

The wide variety of measures that *could* be included in a classification again emphasizes the importance of articulating the purposes of classification, and how clearly stated purposes can help to frame the selection of specific attributes to include in a classification.

Limitations of chosen data elements

Having selected a set of potential data elements, it is important to investigate their limitations and the impact of these limitations on the validity and utility of the classification. Differential availability of information on inputs, processes, and outputs can have important consequences. Because some attributes are more easily measured, care must be taken that their use does not bias the results. Consider an example from the Carnegie Classification: for several years, the Carnegie Foundation estimated an institution's research activity using a National Science Foundation data collection that reported on the fiscal obligations of selected federal agencies to colleges and universities. This had the advantage of objectivity—the data were reported by the agencies rather than the institutions themselves—but it also had consequences for the types of research activity that were detected. The limitation to federal expenditures biased the measure in favor of the sciences (especially life sciences). Regarding the bias in favor of the life sciences, a common objection was that the coveted “Research I” category was difficult to attain for institutions that lacked a medical school. Another limitation of the exclusive reliance on federal support was that research funding in fields where research is largely underwritten by private foundations was not measured. To the extent that institutions differ in their research emphasis, this measure favored some over others. But it also had the effect of communicating, quite literally, that only certain types of research “count.”³

³ For other more arcane difficulties related to this measure, see McCormick (2001).

A more fundamental problem with using funding as a proxy for research activity involves the assumption that dollars spent on research can stand for research itself. This has the effect of attaching greater weight to fields where research is simply more costly (consider, for example, the costs of scholarly activity in high-energy physics relative to linguistics or philosophy, and the consequence of relying on funding as the activity measure).⁴

Operational considerations

In closing the discussion of how to classify, it may be helpful to consider some specific operational challenges, drawing on the experience of the Carnegie Classification. Several of these involve responding to special cases or requests for exceptional handling.

If a classification relies largely on self-reported data (as in the case of the Carnegie Classification), there may be concerns about inaccurate or inconsistent data reporting. One approach to improving the quality of self-reported data would be to implement an audit procedure, wherein a subset of institutions are sampled and asked to supply supporting evidence for selected data elements. A practical decision rule may also be required for responding to requests by institutions to amend their officially reported data due to reporting errors revealed by the classification.⁵

An important decision to be made involves when to update a classification. Updates will be necessary for two reasons: the institutional universe will change due to openings, closures, and mergers of institutions; and institutions themselves may undergo programmatic or operational changes that bear on their classification. If the classification is only updated rarely, a policy regarding case-by-case adjustments may be needed. Case-by-case updates may keep the classifications of particular institutions current, but they may undermine the comparability of classifications as the time frame for classification is no longer consistent across institutions.

As an idiographic classification derived from objective measures, each category of the classification is defined by explicit criteria that refer to underlying measures (for example, the proportion of undergraduate majors in traditional arts and sciences fields). In most cases there will be institutions that are very close to the line that separates categories, but on opposite sides.⁶ This leads to clear violations of the previously expressed objective that any member of a given category resemble the other members more than it does any member of a different category. That is, with respect to this attribute, two institutions close to but on opposite sides of the line will resemble one another more than they will the typical member of their assigned category. If classifications are periodically updated, institutions close to a border may exhibit instability in their classification due to small fluctuations in this measure. One approach to these problems is to simply disclose cases where an institution is close to a category border, as the Carnegie Foundation does in the institution display on its Web site. Another possibility is to define a zone around a category boundary and exercise flexibility in classifying these cases (for instance, by recourse to expert judgment, or by inviting the affected institutions to assist in their proper placement).⁷

As a result of the decision to use existing objective data about institutions—enrollments, degree conferrals, measures of research activity, etc.—the Carnegie Classification is inherently retrospective, describing institutions as they were at the time of data collection. For new or rapidly changing institutions, this can lead to classifications that may not keep pace with the contemporary reality of these institutions. For example, new programs that are in operation but that have not yet produced graduates would not be reflected in data on degree conferrals, rapidly growing programs would be underemphasized by relying on

⁴ These shortcomings were addressed in the 2005 revision of the classification, when the Foundation adopted a multi-measure index of research activity that (a) incorporated other proxy measures for research activity beyond funding, (b) used a different source of research expenditures data that was not limited to federal agencies (though this necessitated a shift to institutionally reported data), and (c) disaggregated research expenditures with regard to focus on science and engineering versus other fields. For more information, refer to technical documentation for the Basic classification available at classifications.carnegiefoundation.org.

⁵ Positive responses to such requests may enhance accuracy of classifications, but at the cost of reproducibility using the official sources. In these cases, careful documentation and disclosure of these decisions is important.

⁶ This is almost inevitable when a continuous variable is used to make categorical assignments. The only exception would be if the distribution exhibited a “natural” gap, such that no cases are close to the dividing line.

⁷ To some degree, this only displaces the problem: some institutions will be close to but on opposite sides of the new lines defining this zone of flexibility.

the degree record, and programs that have recently been closed would still be included in the degree record. Similarly, some institutions may experience an unusual fluctuation—a spike or dip in a normally stable characteristic—that may cause an inappropriate or misleading classification decision. Smaller institutions and smaller programs within institutions are typically more subject to such instability problems.⁸

A related problem arises when the criteria result in a classification that is at odds with an institution's professed identity. In the Carnegie Classification, this has arisen for institutions that both identify and promote themselves as liberal arts colleges, despite the presence of graduate programs of sufficient size to lead to a different classification. It has also arisen for institutions that offer a relatively specialized educational program (for example, institutions specializing in engineering), that would normally be classified among "special focus" institutions. But to the degree that these institutions *also* identify as research universities, their leadership objected to them being classified apart from other research universities, and indeed perceived it as placing them at a disadvantage by separating them from their peer institutions.

Each of the examples described above led institutional leaders to appeal a classification decision, in essence asking to be classified outside of the formal criteria. In almost every case, their objection was not about the classification *per se*, but how it is interpreted and used by other individuals and organizations. These requests for exceptional consideration must be handled with great care, respecting the underlying "truth" of the situation while avoiding the slippery slope of subjective and unwarranted reclassification decisions, as well as the perception that certain groups of institutions are entitled to special treatment.

What secondary uses and consequences can be anticipated?

As noted earlier in this article, the Carnegie Classification was developed as a research tool for the specific needs of the Carnegie Commission on Higher Education, but it quickly evolved into a general-purpose classification that was adopted for uses that sometimes bore important consequences for institutions. Part of the reach and wide adoption of the classification was no doubt due to its provenance, having been developed by an independent, legitimate, and widely respected authority on higher education. Another likely factor was face validity. The groupings made intuitive sense (they also closely paralleled the structure of the 1960 California Master Plan for Higher Education). The best-known universities were grouped together, as were the most selective liberal arts colleges (McCormick & Zhao, 2005).

In *The Rise of American Research Universities* Hugh Davis Graham and Nancy Diamond (1997), write that "the Carnegie system was designed to pull the attention of policy makers away from the nation's research institutions, and to emphasize instead the variety and social importance of the vast majority of institutions that were not research oriented" (p. 53). In its first publication to use the classification system, the Carnegie Commission on Higher Education (1971) wrote, "[we] find no need whatsoever in the foreseeable future for any more research-type universities granting the PhD" (p. 5) and its recommendations included "preserving and even increasing the diversity of institutions of higher education by type and by program [and] resisting homogenization." (p. 8). Yet historian John Thelin (2004) writes that the classification "set off a competitive rush by institutions to meet the operational criteria" to climb the hierarchy (p. 320).

Thus a particular irony of the Carnegie Classification is the extent to which it engendered (or perhaps more precisely, facilitated) institutional ambitions to "move up" the status hierarchy of U.S. higher education, particularly among the four categories of doctorate-granting institutions (which were unhelpfully labeled Research Universities I and II followed by Doctoral-granting Universities I and II, with the Research I category including the nation's most prestigious and well regarded research universities). The classification was interpreted as both signal and roadmap for ambitious institutions and their leaders. This effect persists to this day. For example, following the 2010 update of the Basic classification, North Dakota State University issued a press release proclaiming, "NDSU reaches top national rank from Carnegie

⁸ Prior to 2005, the Carnegie Classification used a three-year average to smooth out year-to-year fluctuations in some variables, but this procedure increased the classification's insensitivity to change by incorporating older data. In 2005 the Carnegie Foundation opted to use the most recent year's data for all data elements. When institutions asserted that the selected year represented an anomaly, they were offered the option to have their classification based on a three-year average as in prior versions of the classification.

Commission on Higher Education” and listing the Ivy League and prestigious public universities whose ranks it had presumably joined by virtue of its new classification.

The status-seeking effect was not limited to doctorate-granting universities, as public comprehensive (master’s granting) universities sought to break into the doctoral ranks, and the classification’s published criteria again showed exactly what would be required. The desire to move up the hierarchy was no doubt reinforced by state funding formulas that either made use of the Carnegie categories or that relied on a derivative classification model. An institution’s classification in the desired categories was also seen as important to its ability to compete for research grants and to recruit desirable faculty and senior administrators.

Anecdotal reports reached the Carnegie Foundation indicating that the classification was even affecting internal funding decisions, wherein funds were reallocated away from departments whose research funding was not included in the data source used in the classification’s published criteria in favor of departments whose research would “count” (see foregoing discussion of data limitations). This is a classic example of “gaming,” that is, a strategic response oriented to the specific measures used and not to the underlying attributes and processes those measures are intended to represent. While gaming is frequently observed (and condemned) in the context of rankings, classifications can produce similar responses if benefits accrue to certain categories, but not others.

Other consequential uses included grant programs that explicitly tied eligibility to certain classification categories, and the use of the classification by US News & World Report to define the comparison groups used in its annual rankings (McCormick, 2007). In the latter case, this could mean the difference between a relatively high ranking as a “regional” university (one of the labels used by US News for the repurposed classification), or a much lower ranking among “national” universities. In other instances it could mean the difference between being ranked among “national” liberal arts colleges or regional universities. It could even determine whether an institution is included at all, as US News excludes institutions in Carnegie’s “Special Focus Institutions” and “Tribal Colleges” categories. In these cases, the use of the classification in the annual rankings affects the institution’s visibility to prospective students, with obvious consequences for student recruitment—and potentially, survival. In a less consequential but nonetheless telling example of the classification’s wide use, some higher education organizations even based their dues structure on an institution’s Carnegie Classification.

As the Carnegie Classification came to be seen as “the” way to think about institutional mission, identity, and differentiation, the categories came to be reified as institutional characteristics unto themselves. This in turn seemed to limit recognition of the many other ways in which institutions resembled or differed from one another. Concern was also expressed that activities that were not explicitly recognized in the classification criteria were seen as lacking in importance or legitimacy. All of these factors were behind the Carnegie Foundation’s decision to adopt a multidimensional classification framework, as well as the introduction of a new “elective” (voluntary) classification with respect to community engagement.⁹

The point of this discussion is clear. The impact of a classification framework can reach far beyond its original purposes, and it would be wise to anticipate these consequences in the initial design.

⁹ Because no satisfactory national data were available for the new community engagement classification, this voluntary approach instead required an extensive process of documentation and review by a panel of experts.

Discussion and recommendations

The key lesson from the Carnegie Classification is that a classification framework can have consequences that bear little or no relationship to its formal purposes, and which may indeed even run counter to its original intent.¹⁰ And it may not be possible to get the genie back into the bottle. Although the term “Research I” has not formally been part of the Carnegie Classification for more than a decade, it is firmly established in the *lingua franca* of U.S. higher education. At my own institution, I frequently hear invocations of the abbreviation “R1” to signify the most research-intensive universities.

Classification can be valuable for research and it can also have valid policy uses, but the use of a classification in formal policy increases the risk that it will confer differential benefits, and consequently that institutional leaders will engage in strategic action to advance the interests of their institutions. Thus my aim in elaborating the limits of classification is to encourage readers to consider the risks in addition to the potential benefits of a classification system.

I conclude with a brief set of recommendations based on the experience of the Carnegie Classification.

1. Proceed with caution. Recognize that classification involves more than a set of technical challenges.
2. Think hard about the purposes of a classification system, and whether classification is the only or the best way to achieve those ends.
3. Keep it simple. Differentiate only as much as is needed to achieve the classification’s purposes. Elaborations can be added later.
4. The call for simplicity is counterbalanced by the possible advantages of explicitly acknowledging multiple dimensions of similarity and difference. Higher education is a complex social institution, and there may be some benefit in building a classification system that discourages reductionism and the reification of categories.
5. Anticipate inferences and interpretations. Don’t produce a list of winners and losers. Labels and order of presentation matter. Provide clear guidance about what the categories represent. The anticipation of inferences and interpretations pertains not only to the variables used to define the classification, but also to what’s not included.
6. Recognize that if clear benefits accrue to members of certain categories, institutions will respond to the incentive to qualify for the privileged category.

As stated above, classification can have valid policy uses, but its incorporation into formal policy in a way that advantages certain categories and disadvantages others (or indeed, any use that has that effect) risks inducing strategic responses that can undermine the classification’s purpose as a neutral and objective tool. This is a manifestation of what has come to be referred to as Campbell’s Law, after the late Donald T. Campbell (1979): “The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor” (p. 85).

Earlier in this article, in distinguishing between nomothetic and idiographic classification techniques, I suggested that “idiographic approaches may be more desirable in an applied setting if comprehensibility and transparency of classification decisions are particularly important.” But if a classification’s use and interpretation risk inducing a gaming response, then transparency of classification decisions may in fact be undesirable. In this circumstance, the more opaque nomothetic approach may prove more resistant to gaming, because it does not establish the clear and unambiguous targets that can facilitate gaming. Even if the variables used in a nomothetic classification are known, each entity’s category assignment results from the combination of all variables used to produce the classification. Thus the specific actions required to influence the result are far from obvious.

¹⁰ This is not always the case, however. Other classifications have been developed for U.S. higher education (McCormick & Zhao, 2005), but none have had the wide adoption, secondary uses, impact, and staying power of the Carnegie Classification.

The task of classifying higher education institutions may seem simple. But each stage of the classification process—design, execution, and maintenance—involves important, often difficult decisions that can have enduring consequences.

The original article was received on December 19th, 2012

The revised article was received on January 28th, 2013

The article was accepted on February 9th, 2013

References

- Bailey, K. D. (1994). *Typologies and taxonomies: An introduction to classification techniques*. Thousand Oaks, CA: Sage Publications.
- Brint, S., Riddle, M., & Hanneman, R. (2006). Reference sets, identities, and aspirations in a complex organizational field: The case of American four-year colleges and universities. *Sociology of Education*, 79, 229-252. doi: 10.1177/003804070607900303
- Campbell, D. T. (1979). Assessing the impact of planned social change. *Evaluation and Program Planning*, 2(1), 67-90. doi: 10.1016/0149-7189(79)90048-X
- Carnegie Commission on Higher Education (1971). *New students and new places: Policies for the future growth and development of American higher education*. New York: McGraw-Hill.
- Carnegie Commission on Higher Education (1973). *A classification of institutions of higher education*. New York: McGraw-Hill.
- Graham, H. D., & Diamond, N. (1997). *The rise of American research universities*. Baltimore: Johns Hopkins University Press.
- McCormick, A. C. (2001). The 2000 Carnegie Classification: Background and description. In Carnegie Foundation for the Advancement of Teaching. *The Carnegie classification of institutions of higher education, 2000 edition*. Menlo Park, CA: The Carnegie Foundation for the Advancement of Teaching, 2000.
- McCormick, A. C. (2005). *A new set of lenses for looking at colleges and universities*. Retrieved November 25, 2012 from <http://www.carnegiefoundation.org/perspectives/new-set-lenses-looking-colleges-and-universities>
- McCormick, A. C. (2007, May 10). Hidden in plain view. *Inside Higher Ed*. Retrieved November 25, 2012 from <http://www.insidehighered.com/views/2007/05/10/mccormick>
- McCormick, A. C., & Zhao, C. (2005). Rethinking and reframing the Carnegie Classification. *Change*, 37(5), 50-57.
- Thelin, J. R. (2004). *A history of American higher education*. Baltimore: Johns Hopkins University Press.