USING EVIDENCE IN THE

WSCUC ACCREDITATION PROCESS:

A GUIDE FOR INSTITUTIONS

Second Edition

(Updated for the *2013 Handbook of Accreditation*)

June 2015

ACKNOWLEDGEMENTS

The WASC Senior College & University Commission expresses its gratitude to Peter Ewell, from the National Center for Higher Education Management (NCHEMS), for authoring the First Edition of the Evidence Guide and for drafting many components of this Second Edition.

TABLE OF CONTENTS

INTRODUCTION

Purpose of this Guide

This *Guide* is designed to assist institutions in thinking about how to assemble and use evidence in WSCUC accreditation processes considering the accreditation standards and requirements articulated in the *2013 Handbook of Accreditation*. A central objective is to develop a common understanding throughout the region that the fundamental basis of WSCUC accreditation is concrete, verifiable evidence that an institution meets the WSCUC Core Commitments to student learning and success; quality and improvement; and institutional integrity, sustainability, and accountability.

How this Guide is Organized

This *Guide* has two main sections:

1) An Overview of Principles and Properties of Good Evidence: This section discusses, in general, principles and properties of good evidence required to support an effective accreditation culture of evidence. It also addresses evidence of student learning specifically, as well as potential pitfalls to avoid.
2) Resources: This section lists and annotates a number of useful sources on evidence and its use in higher education evaluation and decision-making.

SECTION 1

AN OVERVIEW OF PRINCIPLES AND PROPERTIES OF GOOD EVIDENCE

What is Evidence?

At the most fundamental level, "evidence" constitutes the substance of what is advanced to support a claim that something is true. There are at least five important characteristics of evidence that differentiate it from just "information," "data," or "facts." In essence, evidence also includes the careful analysis of information, data, and facts.

1. Evidence is *intentional* and *purposive*; it is advanced to address deliberately-posed questions that are important to institutions and their stakeholders.
2. Evidence entails *interpretation* and *reflection*; it does not "speak for itself." This means that sound evidence involves more than simply presenting a body of data or listing the facts. Instead, it implies that the party who advances the evidence has thought about what it means and can interpret it appropriately to support a conclusion. Indeed, for purposes of accreditation, as much emphasis should be placed upon what an institution *makes* of the information that it advances—and how it is using the conclusions that it has drawn to improve itself—as on the information itself.
3. Good evidence is *integrated* and *holistic*; it does not consist merely of a list of unrelated facts. Individual pieces of data should thus never be advanced as evidence on their own. Rather they take on meaning in the overall context in which they are presented. This means that individual pieces of evidence should mutually reinforce one another—gaining strength from the fact that information of quite different kinds and drawn from

diverse sources point in a similar direction. It also implies that judgments need to be made about any body of evidence as a whole.

4. Evidence can be based on both *quantitative* and *qualitative* information. Certainly, where available and appropriate, quantitative data will be powerful, and it is expected that much of the information that an institution advances in support of its claims for capacity and educational effectiveness will be in numeric form. But it is important for institutions to avoid automatic assumptions that quantitative measurement is what is wanted.
5. Evidence can be both *direct* and *indirect*; it does not always require obtrusive data-gathering using specially-designed instruments. (See the Glossary in the *2013 Handbook of Accreditation* for more information about direct and indirect assessment).

Evidence for Accreditation

For the purpose of accreditation, what should evidence be about? Traditionally in a self-study, institutions have used data largely to describe who they *are,* typically including enrollment counts, program inventories, faculty numbers and credentials, financial resources, space inventories, and the like. These will surely be useful in future accreditation reviews—both to orient visiting team members to the institution and to provide some indicators of capacity. The kinds of evidence advanced in the WSCUC accreditation process however, as reflected in the *2013 Handbook*, ought instead to concentrate largely on results, what each institution *does* and *how well* it does relative to its goals and standards of performance, rather than on specific structures and methods for their accomplishment.

In the realm of students, for example, the information presented should go beyond how many there are to focus instead on such things as retention/graduation rates for different kinds of students and how both aggregated and disaggregated results match institutional expectations and goals. More importantly, in the realm of student learning, institutions should cite more than just a list of assessment activities and selected performance results (such as licensure pass rates) to identify to what extent key institutional learning objectives and performance standards are being achieved. For faculty, for example, in addition to their credentials, emphasis should be given to the effectiveness of support that the institution provides in developing their scholarship of teaching or in moving toward more learning-centered instructional approaches. In the case of finances and facilities, in turn, the object of interest should be not just their extent or sufficiency, but also how effectively they are renewed and how they are deliberately deployed in support of teaching and learning.

Principles of Good Evidence

Evidence supports a specific question in the context of a given community of judgment; therefore, it is important to make clear the principles of evidence that are most compelling in the accreditation review process. Five principles of evidence communicate this intent. Like any principles, these are intended to provide general guidance and should thus be applied creatively and flexibly. Indeed, several of them involve making hard choices about such matters as the level of detail to be provided, how much reflective commentary to include, and how much documentation is sufficient. Collectively, though, they frame an overall approach to using evidence in the accreditation review process.

Principles (and examples):

1. *Relevant*. Any evidence advanced must be related to the question being investigated. While this principle may seem obvious, institutions sometimes produce reams of statistics in the course of an evaluation that are only marginally related to the questions they are trying to answer. Implied by this principle, moreover, is *validity*—the extent to which the evidence advanced is capable of faithfully and fully representing the underlying concept of interest. Equally implied is the need to explain coherently exactly what any information advanced is *supposed to be evidence of*, and why it was chosen over other potential sources of information. In practical terms, this means that institutions need to select carefully the kinds of evidence that they advance in the light of specific WSCUC Standards or questions of importance to institutions themselves. Finally, present the evidence and set forth a clear rationale for why it is related to the Standard's intent.

   Example: In relation to Criterion for Review 2.6, University Y provides catalog copy indicating the specific course and credit requirements needed to earn a degree.

   Commentary: The primary intent of the criterion is that the institution be able to demonstrate that its graduates have met established and recognized standards for achievement. Relevant evidence that speaks to this point might include a) assessed results of samples of student writing that show that graduates have reached the levels of writing expected by faculty, b) curricular features such as capstone courses or performances that require students to demonstrate what they have learned in various courses, c) examples of common grading criteria or rubrics in particular fields or

departments, or d) benchmark comparisons with other institutions that indicate comparable curricular features or levels of student attainment.

2. *Verifiable*. Any evidence advanced must also allow its validity to be readily checked by others. Partly this is a matter of whether the process of assembling it is replicable, and if repeating it would likely obtain a similar result. This property, of course, corresponds directly to the concept of *reliability* in measurement. Partly, though, verifiability is also a matter of *documentation*—whether sufficient information is available to enable a reviewer (or any third party) to independently corroborate what was found.

   Example: In relation to Criterion for Review 2.7, Institution Y states that employers often express satisfaction with the match between abilities of the institution's graduates and their own needs.

   Commentary: The evidence presented could be strengthened in two ways, both involving simply reporting additional details. First, specific numbers and percentages could be cited in support of this conclusion, suggesting systematic attention to the question posed. Second, the particular methods used to collect such information such as surveys or focus group interviews could be described and could be made available to a visiting team for inspection.

   Example: In relation to Criterion for Review 2.12, Institution Z presents a description of its advising policies, together with a) results of a recent survey by the Institutional Research Office showing an overall 87% satisfaction rate with advising (with several sub-population breakdowns), and b) results of a random audit of 25 student records showing that its policies are actually being carried out.

Commentary: The second of these two evidence-gathering approaches could be easily replicated by the team on site through its own audit procedure, and documentation for both could be made readily available for further inspection or analysis. The evidence presented is in principle highly verifiable even if no further investigations are undertaken to determine its veracity.

3. *Representative*. Any evidence advanced must be *typical* of an underlying situation or condition, and not be an isolated case. If statistics are presented based on a sample, evidence of the degree to which the sample is representative of the overall population ought to be provided. Further, it is helpful to present such statistics *over time* (three to five years) to check for variation and to make any underlying trends apparent. If the evidence provided is qualitative—for instance in the form of case examples or documents—multiple instances should be given or additional data shown to indicate how typical the cases presented really are. Sampling procedures can save considerable energy and can allow much more scope for in-depth analysis and interpretation than trying to collect data about all cases. But in both sampling and reporting, care must be taken to ensure that what is claimed is typical.

   Example: In relation to Criterion for Review 3.3, Institution Z describes its Faculty Fellows Program together with an annotated example of a particular chemistry professor's project on using classroom assessment techniques to improve her instruction.

   Commentary: The use of a particular case is appropriate and compelling because it can demonstrate in depth the kind of scholarship of teaching that individual faculty

members are engaging in and that the University is attempting to foster. But the evidence would be strengthened if data were also presented on how many faculty have participated in such programs, the distribution of participation across disciplines/departments and/or different kinds of faculty (e.g., age, rank, demographics). A simple chart summarizing the numbers and kinds of development projects that faculty have undertaken through this program (e.g., classroom research, course portfolios, etc.) could also be effective.

Example: In relation to Criterion for Review 2.7, Institution X provides a detailed account of the recent re-accreditation of its Nursing Program by the National League of Nursing (NLN) as an example of its efforts to improve program currency and effectiveness using evidence of student learning, and as an illustration of the way it involves external stakeholders in such reviews.

Commentary: Because of the detailed requirements for effectiveness reporting required by the specialized accreditation in the health professions, the case of Nursing would probably provide an excellent example of program evaluation at any institution. But to claim that it is representative, the university would be well advised to provide information on how many other programs have undergone such processes. The case for institutional commitment would also be strengthened if an additional example were chosen from among those departments that do not undergo professional accreditation. Another potential issue raised by this example is the match between the requirements of a specialized accreditor and the institution's *own* mission and educational goals.

While specific elements of nursing practice are important, so are the abilities and values that the institution seeks to instill in *all* of its graduates.

4. *Cumulative*. Evidence gains credibility as additional sources or methods for generating it are employed. Conclusions are more believable when they can be independently corroborated from quite different sources. In evaluation, using multiple methods – *triangulation* – helps guard against the inevitable flaws associated with any one approach. The same principle applies to qualitative evidence whose "weight" is enhanced both as new cases or testimony is added and when such additions are drawn from different sources. In advancing this principle, WSCUC does not mean to suggest that each and every statement advanced by an institution needs to be backed by information drawn from multiple sources. But it does suggest that the entire body of evidence should be mutually reinforcing when presented to address a particular Standard or to address an issue or question of importance to the institution.

   Example: As part of its institutional report, Institution W provides several in-depth case studies of areas that it wishes to improve. These include oral communication across the curriculum, technological literacy, and the integration of learning communities into first-year courses. Each of these case studies involves syllabus analysis (including a look at the content and difficulty of the assignments given to students), survey results comparing faculty and student perceptions of actual classroom practices, and (for the first two cases) results of a rubric-based analysis of the strengths and weaknesses of representative samples of actual student work. For oral communication, moreover, a scoring system devised by the National Communications Association is employed to

examine selected student presentations and the scoring verified by an external reviewer at another college. In the case of learning communities, information about student re-enrollment and ultimate graduation rates is also supplied.

Commentary: The evidence provided is drawn from each of the main classes of effectiveness information available to any institution. These include a) statistics drawn from existing records systems and analyzed to create appropriate indicators of performance (e.g., retention/graduation rates, syllabus analysis and examination of actual student assignments); b) self-report data on perceptions and behaviors drawn from surveys, focus groups, or interviews; and c) direct examination of student performance using, where appropriate, recognized or externally validated assessment procedures.

5. *Actionable*. Evidence should provide institutions with good information about taking actions for improvement. This entails that both the analysis and presentation of evidence need to be appropriately *disaggregated* to reveal underlying patterns of strength and weakness, or to uncover specific opportunities for intervention and improvement. It also requires that the evidence provided has been *reflectively analyzed and interpreted* to reveal its specific implications for the institution.

   Example: In presenting evidence of the support it provides for student learning, Institution Z notes that it has established explicit targets for first-year retention rates and for six-year program completion rates, provides a table indicating the actual rates attained over the past three years, and flags whether or not the established target was met.

Commentary: Establishing targets is useful, but actionability would be greatly strengthened if additional analysis was undertaken to break down these results to the level of individual schools and departments and by student demographic and behavioral groups. Further disaggregation of these data might reveal even more opportunities for action. For example, which kinds of students seem to be dropping out and when? Can these events be associated with any particular courses or other experiences? Are there populations or schools that appear to have exemplary rates, and what might explain them? And how might any best practices so uncovered be used for further improvement?

Example: In relation to Criterion for Review 3.1, Institution X provides statistical data on the overall composition of its faculty by discipline, age, diversity, and tenure status—together with a brief interpretive commentary that emphasizes the fact that upcoming faculty retirements will likely significantly alter its ability to staff specific disciplines in areas of high-anticipated future student demand. It also notes that, while providing a significant staffing challenge, this situation also offers an important opportunity to systematically address its diversity goals. The institution accompanies this brief commentary with a note indicating that these statistics are currently being examined by a special joint task force made up of associate Deans and representatives of the Faculty Senate to help determine a coordinated recruitment strategy. *Commentary*: The evidence provided is not only presented in enough detail to reveal its implications, but specific conclusions are also noted and actions being taken in response are described.

The presentation is thus informative and would provide a visiting team appropriate guidance about how to probe further.

Principles of Effective Evidence of Student Learning

One of the most difficult and widely discussed venues for evidence is that provided in the assessment of student learning. Here, four principles of evidence are applicable across a wide range of institutional settings and methods:

1. *Evidence of student learning should cover knowledge and skills taught throughout the curriculum.* Unless a course is designed as an integrative capstone whose coverage is comprehensive, evidence provided to demonstrate student learning should not be limited to data or information from a single course or sub-field of the discipline. The unit of analysis for evaluation for the student is the cumulative experience and level of learning of the student at the time of graduation. For programs, the cumulative effect and learning results that students achieve in an ongoing way by the completion of the program is relevant. Student learning should also be correlated to institutional goals in ways beyond using a specific course relevant to the ability level or domain of knowledge.
2. *Evidence of student learning should involve multiple judgments of student performance.* More than one person should evaluate evidence of student learning. Many techniques are available for engaging multiple reviews and reviewers such as portfolio analyses, reviews of student work products drawn from throughout the curriculum and follow-up

studies. Faculty should engage with the data to make recommended adjustments that will improve student learning results.

3. *Evidence of student learning should provide information on multiple dimensions of student performance.* In essence, this principle suggests that assessment results in more than a single summative judgment of adequacy. Information should instead be collected on a number of discrete dimensions of performance, and should be aggregated across students to provide evidence of the overall strengths and weaknesses of graduates in a program or at the institutional level. A single grade or certification of mastery is thus insufficient to meet this principle, even though it may be the result of a multi-dimensional grading process.
4. *Evidence of student learning should involve more than surveys or self-reports of competence and growth by students.* One of the first steps many institutions undertake when they begin assessment is to survey students or graduates about satisfaction and perceived growth and development. Surveys asking students to rate their own strengths and weaknesses and/or areas of growth, though helpful, are inadequate as stand-alone assessments of learning outcomes because they are indirect measures. More and different types of evidence are expected in providing evidence of student learning, including the results of the direct assessment of student learning products.

Avoiding the Pitfalls of Evidence

When using evidence in the context of WSCUC accreditation, institutions need to take care to avoid a number of potential pitfalls, including:

- *Trying to measure and report on everything.* In an evaluative situation like accreditation, it is easy to be misled into thinking that "more evidence is better." Instead, institutions should think carefully about the evidence they present and to ensure its relevance and quality. A structured and well-explained presentation, anchored on a succinct body of well-documented and reflected-upon evidence, will be far more convincing than simply a "data dump."
- *Trying to be too "precise."* Good evidence does not always have to be as precise as methodologically possible. Rather, it should be as precise as *necessary*, given the problem at hand, or the question to be answered.
- *Trying to wrap it up.* Reflecting on evidence is a process that is never really done. As a result, institutions need not always draw summative conclusions from the evidence they present to WSCUC as part of the accreditation process. Sometimes reviewing evidence does provide "answers" and suggests particular actions that might be taken—and, indeed, the Commission wants to encourage institutions to act on evidence wherever possible. But sometimes reflection yields more precise *questions* and suggests new lines of investigation that might be undertaken. This, too, is a positive outcome, and it should not be shunned. In fact, the iterative nature of the process of collecting evidence about performance and of raising questions for further inquiry is one of the hallmarks of what WSCUC means by a "culture of evidence."

SECTION 2

RESOURCES

There is a vast literature on evidence and its use in evaluative settings like the accrediting process. As institutions begin to marshal evidence in preparation for accreditation, the following sources may be especially beneficial:

Bers, Trudy H., with Jeffrey A. Seybert (1999). *Effective Reporting*. Tallahassee, FL: Association for Institutional Research (AIR). A brief and highly readable guide to presenting data and information in the context of institutional research. Addresses the reporting of both qualitative and quantitative information, and is especially strong on the use of graphics and the emerging possibilities of web-based reporting. A more thorough (and probably the definitive) treatment of graphics can be found in Tufte, Edward R. (1983). *The Visual Display of Quantitative Information.* Cheshire, CT: Graphics Press.

Borden, Victor M. H.; and Banta, Trudy W. (1994). *Using Performance Indicators to Guide Strategic Decision Making, New Directions for Institutional Research #82*. San Francisco: Jossey-Bass. This edited collection describes a number of approaches to constructing performance indicators in higher education settings. Particularly useful is an extensive appendix listing some 250 higher education performance indicators grouped under 22 categories of performance.

Council for Higher Education Accreditation (2000). *The Common Data Project*. Washington, DC: Council for Higher Education Accreditation (CHEA). Reviews the current data requirements of both regional and specialized accrediting agencies and proposes a common core of data for use in the accreditation process. Definitions and sources for proposed data elements are included.

Ewell, Peter T. (1989). *Enhancing Information Use in Decision Making*, *New Directions for Institutional Research #64*. San Francisco: Jossey-Bass. This is an edited collection of essays that discusses a range of techniques for using information more effectively in college and university settings. Includes analyses of lessons drawn from evaluation practice, the organizational context for information, the psychological dimensions that affect information use, and techniques for effective reporting. For additional examples of innovative reporting formats, see Kinnick, Mary K. (1985). Increasing the Use of Student Outcomes Information, in P. T. Ewell (ed), *Assessing Educational Outcomes, New Directions for Institutional Research #47*. San Francisco: Jossey-Bass, pp. 93-109.

Ewell, Peter T.; and Lisensky, Robert (1988). *Assessing Institutional Effectiveness: Re-Directing the Self-Study Process*. Washington, DC: Consortium for the Advancement of Private Higher Education (CAPHE). Based on a project involving ten colleges, provides guidance on how to identify existing data and information resources and how to organize the presentation of

evidence around strategic themes. Emphasizes the notion of institutionalizing information as a permanent strategic resource.

Jones, Dennis P. (1982). *Data and Information for Executive Decisions in Higher Education*. Boulder, CO: National Center for Higher Education Management Systems (NCHEMS). Addresses the basic properties of data and information in a higher education context, with particular emphasis on the need for information to be tailored to the characteristics of users and particular kinds of decisions. Provides a useful review of the properties of good information in a decision-making context, as well as a conceptual overview of the structure and contents of a comprehensive management database for colleges and universities.

Light, Richard J.; Singer, Judith D.; and Willett, John B. (1990). *By Design: Planning Research on Higher Education*. Provides an unusually readable and accessible approach to the basics of designing and implementing evaluation research in college and university settings, based on the first five years of experience at the Harvard Assessment Seminar. Specific topics addressed include formulating appropriate research questions, identifying target populations, choosing the right evaluative methods, and presenting results in an actionable form.

Webb, Eugene. J.; Campbell, Donald. T.; Schwartz, and Richard. D. (1999). *Unobtrusive Measures: Non-Reactive Research in the Social Sciences, Revised Edition*. Sage Classics Series, 2. New York: Sage Publications. This is the classic treatment of unobtrusive measures such as direct observations and "footprint" data, revised and updated. Still provides the best general introduction to this topic.

Whiteley, Meredith A.; Porter, John D.; and Fenske, Robert H. (1992). *The Primer for Institutional Research*. Tallahassee, FL: Association for Institutional Research (AIR). Provides a basic orientation to the principal methods and tools of institutional research in the form of a dozen essays prepared by leading practitioners. Among the topics addressed are student impact, faculty workload analysis, persistence and student tracking, diversity, cost analysis, peer comparison, and academic program review. An earlier edition covers a different set of topics and is also useful [Muffo, John A.; and McLaughlin, Gerald W. (1987)].