

**USING PROGRAM EVALUATION TO ENSURE THE SUCCESS OF
YOUR ADVANCE PROGRAM**

ADVANCE: INSTITUTIONAL TRANSFORMATION WORKING GROUP

MAY 8, 2006

ACKNOWLEDGMENTS

Authors

Lisa M. Frehill (Team Leader)

Team Participants, in Alphabetical Order

Elena Batista

Sheila Edwards Lange

Susan Gonzalez-Baker

Cecily Jeser-Cannavale

Janet Malley

Ellen Meader

Jennifer Sheridan

Kim Sullivan

Helena Sviglin

Special thanks to *Pamela Hunt, Rebecca Zaldo, and Dina Jankowski*, all of whom provided important assistance in organizing the meetings of the working group at New Mexico State University and the University of California, Irvine.



This material is based upon work supported by the National Science Foundation under Grant No. SBE 0123690, (PI: Lisa M. Frehill). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	2
OVERVIEW	4
EVALUATION PRACTICES AT ADVANCE INSTITUTIONS	5
INTERNAL EVALUATORS	6
EXTERNAL EVALUATORS	7
ADVISORY BOARDS.....	7
BASICS OF PROGRAM EVALUATION	8
IMPACT EVALUATION	8
COST/BENEFIT ANALYSIS	11
PROCESS EVALUATION AND POLICY ANALYSIS	12
EVALUATION PLAN	13
USING EVALUATION DATA AS EVIDENCE OF PROGRAM EFFECTIVENESS	14
PRESENTING QUANTITATIVE DATA.....	15
<i>General Issues and Good Practices</i>	16
<i>Sample Presentation Graphs</i>	16
QUALITATIVE EVIDENCE OF PROGRAM EFFECTIVENESS	19
<i>Audiences</i>	21
<i>Other Considerations with Qualitative Research</i>	22
EVIDENCE OF POLICY IMPLEMENTATION → INSTITUTIONAL TRANSFORMATION.....	22
APPENDIX A: SAMPLE EVALUATION PLAN	24
APPENDIX B: SO, YOU WANT TO RUN A CLIMATE SURVEY?!	29
APPENDIX C. GLOSSARY	36
FURTHER TERMINOLOGY NOTES	38
<i>The Distinction between Evaluation and Assessment</i>	38
<i>Faculty</i>	38
<i>Logic Model</i>	39
<i>Sustainability</i>	40
TABLE 1. S&E DOCTORATE HOLDERS EMPLOYED IN ACADEMIA, BY SEX, AND DEGREE FIELD, 2003	10
TABLE 2. S&E DOCTORATE HOLDERS EMPLOYED IN ACADEMIA IN FULL-TIME JUNIOR FACULTY POSITIONS BY SEX, AND DEGREE FIELD, 2001 AND 2003	10
TABLE 3. S&E DOCTORATE HOLDERS EMPLOYED IN ACADEMIA IN FULL-TIME SENIOR FACULTY POSITIONS BY SEX, AND DEGREE FIELD, 2001 AND 2003	11
TABLE 4. TENURED AND TENURE TRACK VERSUS NON-TENURE TRACK FEMALE FACULTY	17
TABLE A1. SAMPLE EVALUATION PLAN FOR WORKSHOPS/EVENTS, DIAMONDS IN FIGURE 1 VENN DIAGRAM	25
TABLE A2. SAMPLE EVALUATION PLAN FOR INITIATIVES/PROGRAMS, TRIANGLES IN FIGURE 1 VENN DIAGRAM	28
TABLE B1. INFORMATION PROVIDED TO DEPARTMENT HEADS ABOUT CLIMATE SURVEY AT UTAH STATE	34
FIGURE 1. LEVELS OF PROGRAM EVALUATION	4
FIGURE 2. STEM FACULTY RECRUITMENT AT NMSU BEFORE AND DURING ADVANCE	10
FIGURE 3. ADVANCE PERIOD HIRING AT THE UNIVERSITY OF CALIFORNIA, IRVINE AS COMPARED TO OTHER SOUTHERN CALIFORNIA UC CAMPUSES, BEFORE AND DURING ADVANCE	11
FIGURE 4. STEM FACULTY AT NMSU, FALL 2004, BY GENDER, ETHNICITY, AND TENURE TRACK STATUS	17
FIGURE 5. FACULTY BY GENDER AT UNIVERSITY OF ALABAMA AT BIRMINGHAM, 1998-2004	18
FIGURE 6. HUNTER COLLEGE FACULTY FLUX CHART	18
FIGURE 7. RECRUITMENT OF WOMEN AT UCI BY TENURE STATUS AT TIME OF HIRING	19
FIGURE B1. FACULTY SATISFACTION AT UW-MADISON BY GENDER	32
FIGURE B2. EXAMPLE OF MAINSTREAM/NON-MAINSTREAM RESULTS—UW-MADISON	33
FIGURE B3. EXAMPLE FROM UTAH STATE DEPARTMENT LEVEL CLIMATE PRESENTATION	35

OVERVIEW

This is the second of two “toolkits” developed to provide ADVANCE: Institutional Transformation (ADVANCE: IT) awardees with guidance on program evaluation. The first toolkit focused on the numerical indicator data that are required by the National Science Foundation on an annual basis with an emphasis on developing common reporting to increase the potential for comparability of the data. Data comparability will be essential to enable a national-level evaluation of the ADVANCE: IT program. To some extent, this toolkit will provide guidance on how to make effective presentations based upon the quantitative data collection strategies discussed in the first toolkit. This toolkit will go beyond these metric “indicators” to emphasize additional strategies for documenting programmatic impact. This toolkit includes information about:

- Evaluation practices at ADVANCE institutions.
- Basics of program evaluation
- Presenting evaluation findings
- Disseminating program evaluation findings
- Glossary of evaluation terminology

Program evaluation consists of important tools that can be used to monitor and improve programs as they are being implemented (formative) and to assess the impact of programs as compared to the original goals at interim times or upon program completion (summative).

In the case of efforts such as NSF’s ADVANCE: IT Program, you can expect to use evaluation at multiple levels. The Venn diagram in Figure 1 illustrates the levels at which evaluation can occur.

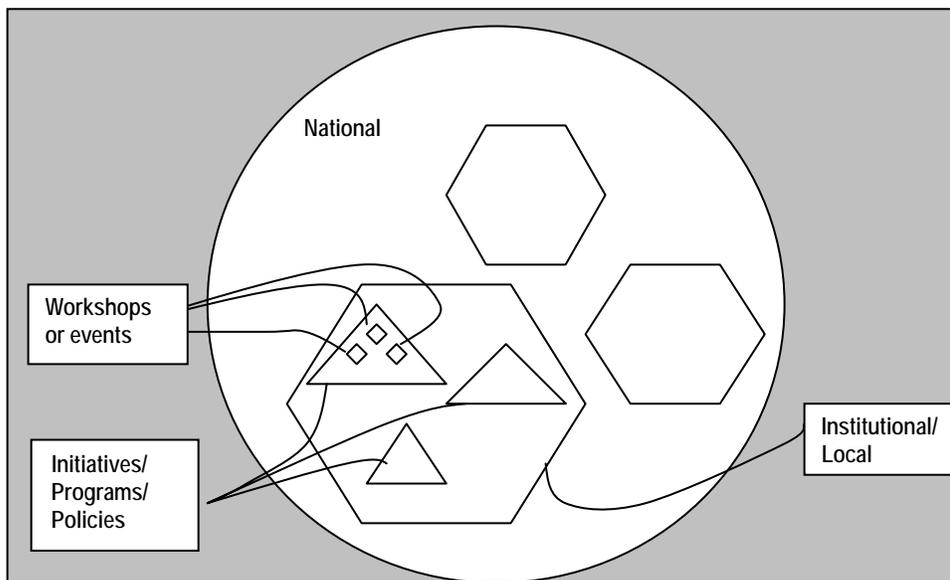


Figure 1. Levels of Program Evaluation

- National level: in the Venn diagram, the circle represents the national level. An evaluation question at this level might be: to what extent have ADVANCE: IT awardees as a group been able to realize gains in STEM faculty women’s recruitment, retention and advancement?

- Institutional/Local level: the hexagons in the diagram represent each individual ADVANCE: IT institution. An evaluation question at this level might be: to what extent has your ADVANCE: IT award transformed your institution?
- Initiative/Program/Policy level: the triangles in the Venn diagram represent different initiatives/programs/policies on which your ADVANCE: IT award is working. An evaluation question at this level might be: has your hiring intervention increased the rate of women's hiring?
- Programmatic elements: the diamonds in the Venn diagram represent specific workshops, events, or other individual-level interventions at your institution that are sponsored by the ADVANCE: IT program. Possible evaluation questions at this level might be: was the workshop effective in reaching the intended audience? Was research funding provided by the ADVANCE: IT award to an individual faculty member effective in increasing productivity?

This document focuses on what happens at the institutional/local level, rather than national and international levels. National level evaluation plans are best accomplished by a number of institutions working collaboratively.

Summative and formative evaluation strategies need to be woven throughout your local program. This toolkit provides guidance to program personnel about how to conduct local level evaluations while establishing guidelines that may permit a national-level evaluation of similar programming or initiatives of the entire NSF-funded ADVANCE: IT program. To some extent, at the national level, this effort could result in elaboration of “best practices” for institutions of various types given the differentiation of ADVANCE: IT awardees.

The ecological and individualistic fallacies are two common pitfalls with social data. The ecological fallacy is like “racial profiling,” the case where data about groups is applied to individuals. On the other hand, the individualistic fallacy occurs when data about individuals (e.g., one's own experiences) are generalized to apply to an entire group.

A common misconception about program evaluation is: who is the subject of evaluation? Oftentimes program personnel and faculty take evaluation results personally.

In program evaluation, information about the level of job performance, educational achievement, or health may well be gathered. However, the purpose is not to diagnose individuals. Instead, the purpose is to learn how well a program is helping people improve on those variables. (Posavac & Carey 1997: 10-11)

The purpose of program evaluation at ADVANCE institutions is to examine how well the program has impacted recruitment, retention, and advancement of women at the institutional level¹, represented by the “hexagons” on the Venn diagram.

EVALUATION PRACTICES AT ADVANCE INSTITUTIONS

We did a quick purposive sample survey of existing ADVANCE programs' (rounds one and two) evaluation practices (nine institutions² were asked about their practices). An analyst from each

¹ In some cases, women's advancement includes leaving your institution for opportunities elsewhere. It is useful to keep track of these gains.

institution completed a matrix that we requested with information about the use of internal evaluators, external evaluators, and advisory boards. We asked about the logistics, costs, dissemination strategies, structure, advantages, and disadvantages of each approach.

In our survey of evaluation practices—which is not comprehensive of all ADVANCE institutions—most institutions used multiple forms of evaluation because of the relative strengths and weaknesses associated with each model. For example, while internal evaluators can provide more immediate feedback, their internal status may lead to missing negative points of view. External evaluators can compensate for this drawback by virtue of their externality but logistics associated with their visit can be difficult and their reports often require more preparation time. Finally, while an external board of diverse members might provide a rapid report with findings representing multiple points of view, the logistics for their visit can be even more challenging and the brevity of their visit can be a constraint. Therefore, ADVANCE institutions should carefully craft an overall evaluation strategy that uses multiple forms of evaluation to improve program impact and sustainability.

Internal Evaluators

Internal evaluators are people who work for the institution and implement various evaluation and institutional research strategies such as those required to report the NSF indicators, workshop evaluations, sometimes climate surveys, and sometimes other kinds of in-house research efforts. In many cases, 0.5-1.0 FTE professional staff plus a 20-hour per week graduate assistant are involved in this effort, with the amount of professional FTE dependent on the scope of tasks. An internal evaluator will need adequate space, a computer and printer, and the ability to maintain a high level of data security and confidentiality. According to a few ADVANCE institutions where an internal evaluator was used, there were several benefits:

“The backbone of our ADVANCE program – evaluation results are used to enhance programs, to assist in institutionalizing programs, to publish results, etc.”
(University of Wisconsin).

“Provides important feedback at the end of each educational activity. This is helpful to promptly introduce modifications to future activities.” (University of Puerto Rico, Humacao).

“Knowledge of program, personnel and campus norms makes it easier to draw meaning from data collected.” (University of Washington).

There were a few drawbacks to using internal evaluators, often related to their status within the program. First, there was a concern that participants may be reluctant to share opinions, especially any negative perspectives about the program with program personnel. On a related note, when evaluation personnel do learn of negative findings, especially if they relate to their co-workers, there is a need for the program to have a climate that allows for constructive criticism. Also, program participants may feel “over subjectified” by being constantly asked for feedback. For the most part, however, those institutions that used internal personnel for evaluation found this to be a wise investment of resources.

² The following institutions were asked: University of Alabama at Birmingham, Georgia Tech, Hunter College, New Mexico State University, University of Michigan, University of Puerto Rico, Humacao, Utah State University, University of Washington, and University of Wisconsin.

External Evaluators

External evaluators are individuals or firms hired by the ADVANCE program or the institution to give a “third party” (i.e., a greater appearance of objectivity) view of the program. Depending upon the program’s needs, this individual or firm conducts climate studies, interviews, focus groups, or other data collection strategies and then prepares a formal report. In several cases, ADVANCE institutions included this formal report with their annual NSF report. Often this report was posted to the website for the program and was used to brief administrators about the program.

Preparations for an external evaluator can be intense. The extent of these preparations varied greatly across institutions, with some preparing binders of materials (akin to those assembled for the NSF site visit teams in the third year) and others providing only weblinks to the evaluator. Some institutions brought in an evaluator only once, while others brought in an evaluator each year. One institution brought in two different evaluators, with both visiting separately in the first year. The most time-consuming aspect was the extensive interview and focus group arrangements—sometimes involving meals—with university faculty and key administrators for the evaluator’s visit. After the visit all of these participants need to be thanked for meeting with the evaluator. Preparations also involve making travel arrangements and generating the contract for the evaluator. For example, one institution indicated that these preparations took the equivalent of one week for 1 FTE staff person.

There were some key benefits to bringing in external evaluators, beyond the issue of objectivity:

“The integration of the four evaluations: program activities, process evaluation, policy and effectiveness of the Program (results and impact) will provide a comprehensive view of the whole program.” (University of Puerto Rico, Humacao)

“Having experienced evaluators interview University faculty and administrators about the program provided data we probably could not have gathered ourselves. It allowed us to better gauge community response and pointed to areas where we needed to increase or modify our efforts.” (Hunter)

“An experienced external person was in a better position to hold administrators’ feet to the fire on institutional transformation and program sustainability. The external evaluator was in a better position to critique than vulnerable faculty and staff within the institution.” (New Mexico State University)

The main drawbacks to the use of an external evaluator were the set-up time and the time-lag associated with obtaining a final report. Costs were not noted as problematic: some institutions spent only a few thousand dollars on travel and consultant fees while others spent as much as \$30,000 per year for travel and consultant fees. Because NSF limits expenditures for consultant fees, you should expect to reimburse for all preparation and report-writing days in addition to visit days and you may find that there are consultants that are too expensive to engage.

Advisory Boards

Programs reported working with both internal and external advisory boards. Obviously, the logistics associated with these two types of boards will differ. In most cases, internal boards provided a mechanism to bring together faculty and administrators from different units and disciplines to review data and strategize the program. Internal boards are generally involved with oversight and monitoring of the program rather than serving a formal evaluative role.

In contrast, external advisory boards serve as a group of external evaluators. Often these are panels of 4-8 people who are invited to the subject institution for a day or two with some periodicity—e.g., annually. The composition of boards (i.e., backgrounds of individuals involved) greatly varied across the institutions that use the boards.

During the board members' visit, key program personnel usually present reports about the program, which was an important dimension along which institutions varied. That is, at some institutions, faculty participants prepared posters presenting their research for the board to review and met with the board in their labs to discuss research along with program personnel presentations. At other institutions, each of the PIs or key personnel on the program makes a PowerPoint presentation to the board and answers questions. Therefore, staff time to prepare presentations, posters, and make travel and honoraria arrangements also varies.

Oftentimes external boards include people who are very busy and in high demand. Therefore, scheduling can be difficult and necessitates long lead time. Depending upon which of your own campus administrators will be involved, you may find it quite difficult to find a time when everyone is available, so you should have enough external board members to account for the possibility that 1-2 may not be available. Indeed, a couple of institutions reported that they originally planned for 2-day board meetings but found this infeasible in practice, settling, instead, for one intense day.

Boards provide a written report to your program, sometimes in the form of a letter, which may be prepared in draft form during the visit and then edited shortly after the visit. The board's recommendations represent the outcome of deliberations by an often diverse set of members, therefore, they can provide multiple points of view, debating the merits of approaches prior to making recommendations. In most cases, ADVANCE program personnel reported that these boards were often important in considering the issues associated with sustaining ADVANCE Program components after the end of the NSF funding period. Other advantages included:

“Creates ties with important individuals/organizations, provides an outside point of view, appearance of some measure of external oversight, can be used to advance program (that is, can use external advisors to say the things you might not be able to say internally.)” (University of Wisconsin)

“Members are well-recognized and can speak with authority to the administration and university community. They provide support and help plan for sustainability of the program.” (University of Puerto Rico, Humacao)

As previously discussed, the principle drawback to the advisory board model is that the planning and preparation are intense and complicated. In addition, because of the brevity of the visit, the feedback may not be as insightful as that which could be provided if members had more time to evaluate and understand the program.

BASICS OF PROGRAM EVALUATION

As with any complex social analysis, there is much value in using different methodologies for evaluation. Because the processes and outcomes that ADVANCE Programs attempt to impact are long-term, short-term indicators of success (milestones) are important to determine whether or not an initiative is on its way to successfully transforming the institution.

Impact Evaluation

The most common kind of evaluation with which you might be familiar is impact evaluation. In impact evaluation, one seeks to answer the research question:

Did the program (intervention) have the desired effect?

Or, more generally:

What were the effects of the intervention?

Impact evaluation can occur at all program levels (as shown in the Venn diagram in Figure 1). For example, at the workshop or event level (the small diamonds in the figure) you may be interested in the short-term with whether attendees were satisfied with these individual events. Feedback provided by attendees can then be used to improve these workshops. But at some point you may wonder whether those who attended the workshop experienced longer-term effects, such as improved retention or advancement opportunities at the institution, which begs the question:

Improved relative to what?

Short-term impacts are often easier to measure than long-term outcomes, which is a challenge for ADVANCE Programs that seek long-term outcomes but with funding that is rather short-term within the scheme of the time involved in bringing about institutional change or the framework of academic careers. For example, faculty may not be formally reviewed for tenure until 5-8 years after initial hire, which means that faculty recruited during your grant are unlikely to reach the tenure milestone during the 5-year grant period.

Comparing data about your institution for a period prior to the ADVANCE award to data on similar processes during the award is one strategy to document impact. Three-year periods are the most common bases of comparison. Figure 2 shows an example of this kind of comparison.

While the results shown in Figure 2 appear impressive, it could be the case that NMSU's increased hiring of women faculty was part of a general, national trend and not really the result of the specific ADVANCE interventions at that institution. Therefore, it is imperative to determine whether the observed effects are the result of the program or due to some external force.

In order to make a case that the changes are a result of the program, then, it is necessary to identify a relevant group of institutions to which to compare the hiring at the ADVANCE institution. Most institutions have a set of other academic institutions to which comparisons are routinely made known as "peers." Typically the office of institutional research and planning can provide you with the names of the "peer" institutions. In some cases, institutions are embedded within a larger system of institutions such as the case with Hunter College (which is part of the City University of New York), the University of California, etc. In these cases, comparing the outcomes at one site to those for the system as a whole or to selected campuses can be useful to demonstrate program impact.

Unfortunately, data from one's peers are not always readily available. At the national level, neither the National Center for Education Statistics' IPEDS nor the NSF's WebCASPAR databases provide information about faculty by gender. The NSF publishes limited data about faculty in its biennial publication "Science and Engineering Indicators," which has downloadable EXCEL format tables on the web. The only table with information about faculty by gender in the most recent (2006) edition was Table 5-28. Table 1 summarizes these data, the most recent source of national-level data on science and engineering faculty. Table 1 totals include post-docs, part-time and full-time faculty and people with PhDs in other positions in academia. Table 2 provides the number of full-

time junior faculty (assistant professors and instructors) while Table 3 provides the total number of full-time senior faculty (associate and full professors) by discipline and percent female for 2001 and 2003, data which may be more useful for ADVANCE comparisons. As is customary, there is a timelag on data from these sources: as shown here, the most recent data we could access in 2006 was associated with faculty in 2003, which means these data will be of limited short-run utility.

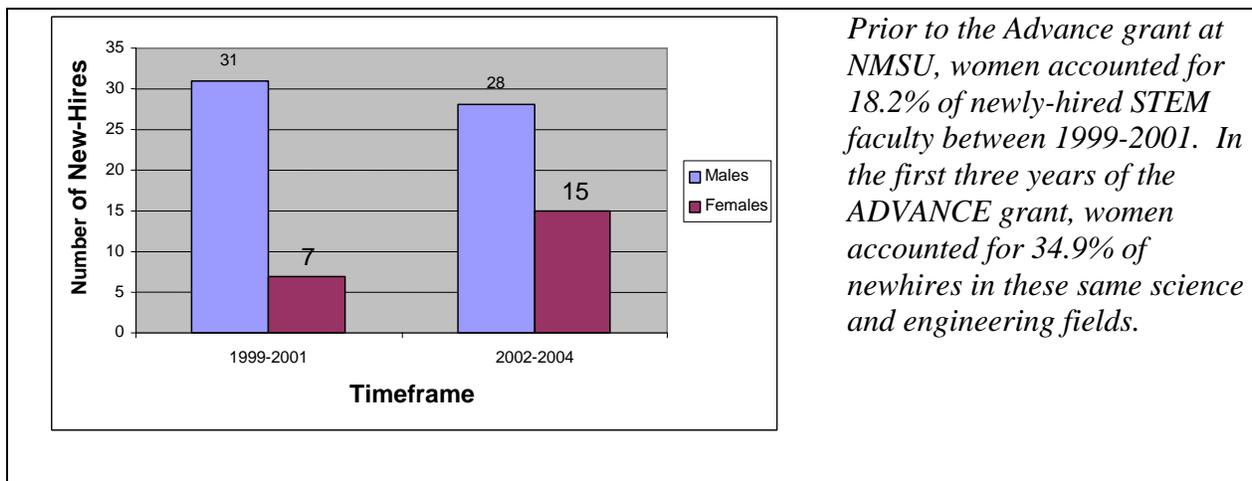


Figure 2. STEM Faculty Recruitment at NMSU Before and During ADVANCE

(Numbers in 1,000)	Both Sexes	Males	Females
Science	230.0	154.5	75.5
Physical sciences	31.5	27.1	4.5
Mathematics	15.4	12.7	2.7
Computer sciences	5.2	4.3	1.0
Earth, atmospheric, and ocean sciences	8.8	7.0	1.8
Life sciences	89.2	55.9	33.3
Psychology	31.3	15.4	15.9
Social sciences	48.6	32.2	16.4
Engineering	28.3	25.3	3.0

Table 1. S&E doctorate holders employed in academia, by sex, and degree field, 2003

(Numbers in 1,000)	2001		2003		Change 2003-2001
	Total	% female	Total	% female	
All fields	50.5	38.4%	52.7	40.6%	2.2%
Science	45.4	41.2%	47.2	43.0%	1.8%
Physical sciences	5.0	22.0%	5.5	23.6%	1.6%
Mathematics	2.3	26.1%	2.8	32.1%	6.1%
Computer sciences	0.9	22.2%	1.3	23.1%	0.9%
Earth, atmospheric, and ocean sciences	1.7	29.4%	1.8	27.8%	-1.6%
Life sciences	19.2	42.2%	20.1	44.8%	2.6%
Psychology	6.6	59.1%	6.1	62.3%	3.2%
Social sciences	9.6	42.7%	9.5	49.5%	6.8%
Engineering	5.1	15.7%	5.6	17.9%	2.2%

Table 2. S&E doctorate holders employed in academia in full-time junior faculty positions by sex, and degree field, 2001 and 2003

(Numbers in 1,000)	2001		2003		Change 2003-2001
	Total	% Female	Total	% Female	
All fields	136.9	20.9%	141.4	22.8%	2.0%
Science	120.2	23.1%	124.2	25.1%	2.0%
Physical sciences	16.5	9.1%	17.0	10.0%	0.9%
Mathematics	10.2	10.8%	10.2	11.8%	1.0%
Computer sciences	2.4	16.7%	2.9	17.2%	0.6%
Earth, atmospheric, and ocean sciences	4.2	14.3%	4.3	18.6%	4.3%
Life sciences	41.2	26.2%	43.1	28.8%	2.6%
Psychology	15.9	35.2%	15.7	40.1%	4.9%
Social sciences	29.6	26.0%	31.0	27.1%	1.1%
Engineering	16.8	5.4%	17.2	6.4%	1.0%

Table 3. S&E doctorate holders employed in academia in full-time senior faculty positions by sex, and degree field, 2001 and 2003

If your institution is part of a larger system, such as the University of California at Irvine, then you may be able to locate useful data via the system office. The University of California, Office of the President (UCOP) maintains online data about faculty, which are easily accessible, many of which are downloadable as EXCEL files. Using these data, the University of California at Irvine's ADVANCE period hiring of women was compared to that at four other southern California UC campuses, as shown in Figure 3.

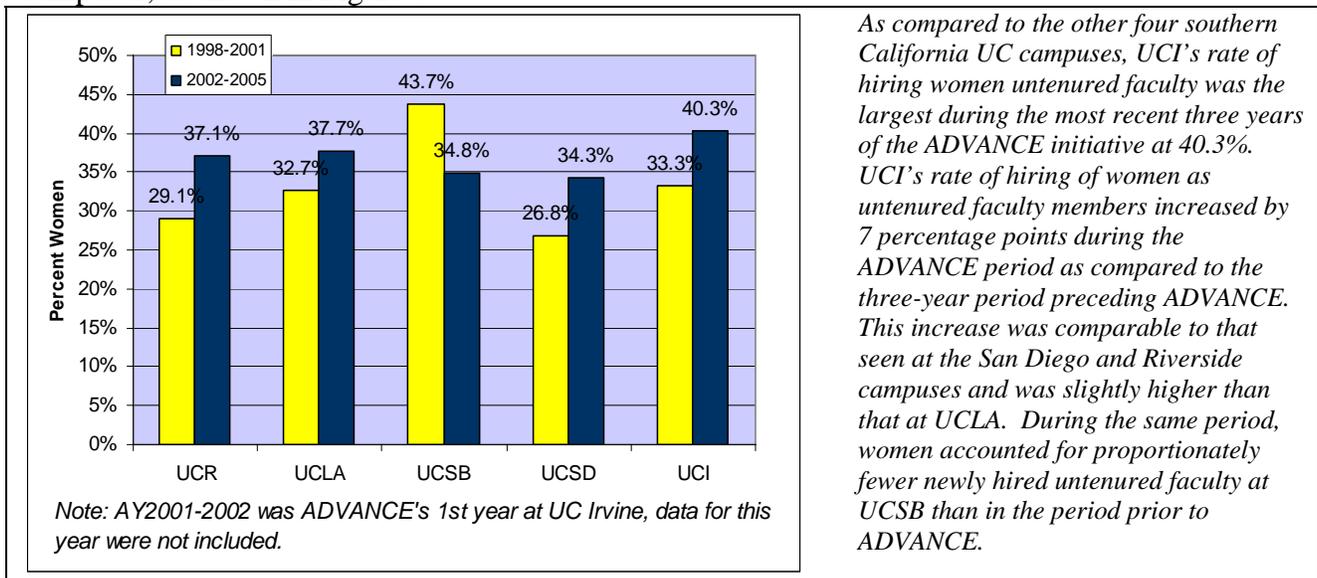


Figure 3. ADVANCE Period Hiring at the University of California, Irvine as Compared to Other Southern California UC Campuses, Before and During ADVANCE.

Cost/Benefit Analysis

Another type of evaluation poses the question:

Was the impact of the program worth the expenditure of funds?

Once an intervention's impact has been documented, determining the cost associated with that impact and then deciding whether the intervention was worth the expenditure is a next step in the evaluation process. Obviously, caution must be exercised with this kind of analysis because there

are many benefits that may not be expressed as dollar values and there is an implicit value judgment about “worth” that requires caution.

Cost/benefit analyses do not, necessarily involve straight dollar-values, however. For example, the University of Washington ADVANCE program focused attention on improving department chairs’ leadership and management skills. An outside expert was hired, who provided one-on-one coaching to chairs as well as workshops for groups of chairs. While the consultant’s services were rated as excellent, the cost associated with these services combined with the small number of chairs who took advantage of the consultant (and that many of these chairs were a self-selected group) led the program to decide to drop this aspect of the program and to reallocate funds into other activities with a wider potential reach.

Process Evaluation and Policy Analysis

Many ADVANCE institutions are attempting to improve policies and processes at their institutions. “Family Friendly” policy implementation and process improvement are viewed as beneficial for all faculty, but especially so for female faculty who shoulder a heavier burden for the care of young children or elderly parents than their male counterparts or who are less likely to have access to the informal information networks that enable them to navigate complex university processes. To some extent, these programmatic initiatives have the benefit of appearing “gender neutral” but are hypothesized to have a greater beneficial impact on women.

Process evaluation can be performed on the university’s processes (e.g., the recruitment and advancement processes) or it can be performed on the actual ADVANCE initiative itself. In either case, the key issues are whether the program is reaching the target population and the extent to which the program as it is being implemented is consistent with the original design. For example, if you had stated that your ADVANCE program would reach out to both junior and senior women faculty on your campus and then you found that only junior women attended your events, then you would not be compliant with your design and you would need to be able to explain why senior women were not participating in the program.

Flow charts of university processes are a useful analytical tool to determine ways to improve the processes within the university. Careful analysis of such processes can reveal bottlenecks. For example in recruitment processes there may be steps at which the process slows, which could impact your hiring of women faculty, since oftentimes women are high-demand candidates who are recruited by those who can complete the search process earlier than others. This analysis could also reveal leverage points for accountability. For example, the hiring process may be mostly decentralized with little accountability for diversity goals. Central administrations often suggest that they have little control over the hiring decisions made by departments. But, if, for example, departments must obtain start-up funds from a central administrative office, then these start-up funds are the leverage point for the central administration to incorporate accountability into the search process. Departments that wish to access these funds could be required to have all search committee members attend diversity training or could be required to document good faith efforts to recruit women and/or minority candidates.

Policy analysis is similar to process analysis in that one would be interested in examining the extent to which different people “used” a policy and the extent to which those who used the policy had positive outcomes as well as any negative outcomes. For example, your institution might implement a dual career hiring policy in order to recruit and retain women science and engineering faculty. To understand the effect of the policy, you would need to maintain data about the requests

made by faculty to “use” the policy to hire their partners as well as the outcomes of these requests. Also, just as you would be concerned about target population issues with process analysis, you would want to determine the extent to which the target population used these policies and what barriers might exist to their use.

For example, many universities have implemented “stop the clock” policies that permit faculty members to delay tenure decisions for various life events. The rationale behind “stop the clock” policies is that women faculty members’ biological clocks overlap their tenure clocks, making it difficult for women to start families while on the tenure track. Some universities found that women were reluctant to use these policies for fear that their colleagues would not view them as serious scholars. Princeton University made headlines in 2005 when that institution made “stop the clock” mandatory, requiring those who wished to NOT stop the clock to express this negative preference.

Evaluation Plan

According to the Venn diagram laid out in Figure 1, evaluation occurs at many levels. At the institutional level, evaluation plans can be developed for workshops and events, as well as initiative and programs. For each program (triangles in the Venn diagram) or programmatic element (diamonds in the Venn diagram) the evaluation plan outlines the relevant goals, a list of groups involved, and a description of the data to be collected to assess outcomes. Data collected to assess outcomes should be a reflection of the program goals and hypotheses about how each of the particular programmatic elements “works.” Just as program goals should be clearly specified at various levels, so too should evaluation occur at multiple program levels ranging from the individual workshop up to an assessment of whether the ADVANCE Program at the national level has had the desired impact upon women’s recruitment, retention and advancement into academic STEM and on institutional transformation.

When developing an evaluation plan, the target audience for results should be considered in order to include important scientific design principles of reliability and validity and the standards to which different audiences are likely to hold us accountable for these principles. A single mistake in how data are collected or analyzed can cost the program credibility with the audience and cause the audience to focus on the mistake rather than the evaluation results.

An efficient evaluation plan can be developed by creating a table as seen in Appendix A. This method allows you to understand evaluation at each level within your ADVANCE: IT program. The first column is “Workshops/Events” or “Initiative/Program” which are the purpose of your program. The NSF ADVANCE: IT program seeks to increase women’s recruitment, retention, and advancement in STEM to bring about institutional transformation. Initiatives might be thought of as overarching goals (e.g., retaining women in STEM) within which various local programs are established to bring about this goal. The second column then states the goals of that initiative/program. The first two columns will be similar between ADVANCE: IT grants as we all have the same overarching goals of recruitment, retention, advancement, and institutional transformation. The third column, “Involved Groups”, then determines which groups are involved in order to successfully reach your goals. The fourth column is “Program Elements” which are the workshops, programs, and components of your grant that work to accomplish the goals of your initiative/program. Each institution’s programmatic elements related to the four ADVANCE efforts (recruitment, retention, advancement and institutional transformation) will vary as will the assessment of outcomes of the individual programmatic elements’ impact related to the ultimate goals. At this level of analysis, these assessments form the basis for a determination of the overall

ADVANCE program's impact at a particular university. The final column then describes each method that you would use to measure the outcome of the initiative/program. Each initiative/program will have measures and program elements to achieve the goals.

Therefore, each ADVANCE site features a distinctive mix of workshops, events, initiatives and programs intended to achieve the program's goals in ways that are maximally efficient and effective within the context of that site's organizational structure and culture. Although workshops, events, initiatives and programs may be evaluated separately, results from each evaluation are considered in sum for how they contribute to a broader ADVANCE effort. For example, an ADVANCE site that conducts an annual workshop series on issues relevant to tenure-track faculty would benefit from an ongoing evaluation of each individual workshop so that incremental improvements can be made and so that ADVACNE personnel can ensure that the target audience is receiving these interventions. An ADVANCE site with a grant program that allows faculty to apply for funds to bring in more women speakers should evaluate the results following from each individual sub-award. When each element is evaluated systematically, it can contribute to a holistic assessment of the entire effort.

Each element in the ADVANCE effort has a fairly immediate (short-term) outcome it seeks to achieve. For example, a workshop seeks to teach a particular topic; a grant seeks to underwrite a specific activity; a mentee participating in a mentoring initiative should experience some positive career advice attributable to that mentoring relationship. Therefore, the evaluator's task is to devise an appropriate strategy for capturing these short-term outcomes and then relaying those results back to the program leadership in order to improve future elements (workshops, grant decisions, etc.).

These sorts of "formative evaluation"-driven revisions are an important part of what contributes to the highest likelihood that final summative evaluations will reflect substantive improvements.

USING EVALUATION DATA AS EVIDENCE OF PROGRAM EFFECTIVENESS

Your ADVANCE program will collect reams of data. Each year you will report the required NSF indicators, you will have an external evaluation report, and you will likely have additional institutional research projects associated with your program. Many ADVANCE institutions have produced climate studies. Others have completed exit interviews or interviews of job candidates who turned jobs down at their institutions. Still others have conducted various research projects on topics of interest to local ADVACNE participants. This section will provide you with tools so you can marshal this wealth of information to make a convincing case about program effectiveness.

In general it is important to:

- Identify the audience (e.g., Upper administrators, Regents, the public, alumni, staff, faculty members in STEM, non-STEM faculty, students, potential donors, etc.) and use an analysis and presentation mode that is familiar to the audience.
- Determine what issues are of most interest to the audience so that you can properly frame your presentation.
- Determine the amount of time you will have with the audience.
- Have a credible presenter.
- Have multiple reviewers of your presentations prior to any public display to catch small but potentially problematic errors.

Skeptics will pick up on any small omission or discrepancy and make your life difficult. This can erode confidence in your subsequent presentations. These can also result in tangential discussions that take away from the real message that you are interested in conveying.

Quantitative data, especially when it is strong and presented well in graphs, can often be quite convincing for many audiences. But keep in mind that numbers may not tell the whole story and that qualitative information can often add flesh to the skeletal structure of a quantitative presentation. In addition, the strongest indicators of program success may lie in quantitative data that will not be available until after the completion of your program (e.g., promotion, tenure, and advancement rates). Therefore, it will be important to show a *trajectory* of change and evidence that this trajectory will continue in the future.

The key issues with which we are concerned in this section are:

- Which data are most appropriate for the audience and the issue under discussion?
- How can the data be used to document or demonstrate the issue?
- How can the data be used to strengthen the argument for institutional diversity?
- How can you use these data to make change?

In this section we include a number of cautionary notes about presenting information to various audiences. For example, the Working Group members who prepared the “NSF 12” as tables for their institutions have found that reviewing all of these data could be overwhelming to faculty, department chairs, deans, and other administrators. You might recall the year that Ross Perot ran for president, paying for his own television time, during which he reviewed pie charts, bar charts and line charts until he was the brunt of jokes by every comedian in the country. Always ask yourself, do I have as many charts as Ross Perot? If the answer is “Yes,” then it’s time to simplify!

Instead, Working Group members suggest that you focus on only a couple of the indicators at any one time. The same applies to often-lengthy climate study results. In these cases, the data tables and graphs and climate survey results can form the basis for regular, periodic communications via, for example, a newsletter or presentations made to various faculty groups on campus.

Presenting Quantitative Data

Toolkit 1 involves in-depth discussion about collection and analysis of quantitative data by ADVANCE: IT awardees, specifically the “NSF 12” required indicators. Many ADVANCE: IT awardees conduct climate surveys. Guidelines related to collecting and reporting climate data appear in Appendix B. Here we include collective wisdom about presenting the large amount of quantitative data collected as part of the “NSF 12” for best impact with campus ADVANCE: IT stakeholders.

The University of Washington ADVANCE team realized that presenting ALL of the data at once was too overwhelming. They decided, with the advice of their site visit team, to hold quarterly “evaluation summits” for the leadership team (deans, department chairs, evaluation team, program staff, in Arts & Sciences and Engineering). By presenting just a few of the indicators at each meeting, the UW team found that people could be better engaged in discussing the underlying processes that produced the quantitative outcomes being reviewed.

General Issues and Good Practices

This section discusses how to avoid common issues that can “derail” presentations, causing an audience to take discussion in an ineffective direction while also discussing good practices to make your presentation a success. Create an effective presentation to showcase the impact of your program. Make sure each chart has appropriate notes of:

- Inclusion/exclusion criteria
- The time period of the data
- The source/quality of the data
- Any other assumptions, definitions or key analytical notes.

These notes will allow the audience to understand the quality of the data and will reduce the questions about validity. Using careful notes also shows that you are not trying to hide or manipulate data. Have multiple reviewers of your graphical and tabular data prior to any public presentation to catch small but potentially problematic errors can be especially important to avoid skeptics who will seize any opportunity to challenge your findings and assertions.

Quantitative data can be used to effectively show that ADVANCE: IT has impacted your institution. There are some techniques to use when presenting quantitative data to ensure that it has maximum impact on the audience. You should present both numbers and percents because invariably the audience wants what you don’t have! It also allows the audience to process the data in the way that makes the most sense to them. Consider using a metric and format that is familiar or well-accepted by your audience. In order to successfully portray your point, consider using a presenter that has the most credibility to your audience. This will allow more time for presenting and discussing the data instead of having to defend the credentials of the person presenting the data. Clearly state the message the graph, table, figure is addressing and label it appropriately. A color printer and PowerPoint are essential tools for crafting effective and convincing presentations. You might consider purchasing a projector and laptop so that you can always have the appropriate technology to make presentations to campus groups and at conferences, regional meetings, etc.

Sample Presentation Graphs

Looking at data at one point in time

Stratify on relevant demographics (e.g. tenure status, rank, gender, race/ethnicity, citizenship status, department, school, field, etc.) The pie charts in Figure 4, shows STEM faculty at NMSU by sex and ethnicity separately for tenured/tenure-track and non-tenure track faculty. Because ethnicity is an important factor at New Mexico State University (a Hispanic-serving institution in a state that does not have a majority of any one ethnic group) it is imperative to show the representation of both women and under represented minorities in STEM faculty. These also highlight the different composition of the tenured/tenure-track and non-tenure track faculty.

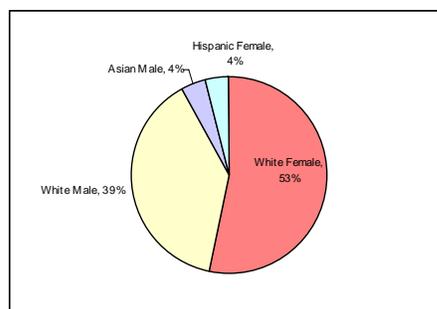
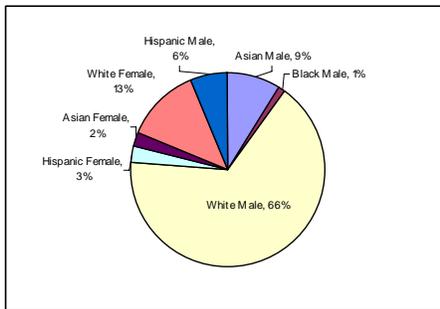


Figure 4. STEM Faculty at NMSU, Fall 2004, by Gender, Ethnicity, and Tenure Track Status

Consider using various denominators to compute percents (within gender, rank, department, tenure status, etc.) To answer the question: What percentage of women versus men are in non-tenure track positions? requires that you look at within gender track differences while answering the question: what percentage of tenure track faculty are women requires that you examine within track gender differences. In answer the first question, the denominator is “All women faculty” or “All men faculty” while in the second question the denominator is “All tenure-track faculty.” Table 4 indicates an example of both of these analyses. For example, while women accounted for 18.9% of tenure track STEM faculty in the College of Arts and Sciences, they accounted for an overwhelming majority of non-tenure track STEM faculty and 37.5% of all STEM faculty women were non-tenure track.

	Tenured and Tenure Track			Non-Tenure Track			Non-Tenure Track as % All Females
	All	Female	%Female	All	Female	% Female	
College of Arts & Sciences	106	20	18.9%	17	12	70.6%	37.5%
Astronomy	8	1	12.5%	1	1	0.0%	50.0%
Biology	19	4	21.1%	1	1	100.0%	20.0%
Chemistry and Biochemistry	19	1	5.3%	3	1	33.3%	50.0%
Computer Sciences	11	2	18.2%	2	2	100.0%	50.0%
Geological Sciences	6	2	33.3%	0	0	0.0%	0.0%
Mathematical Sciences	29	10	34.5%	9	7	77.8%	41.2%
Physics	14	0	0.0%	1	0	0.0%	0.0%

Table 4. Tenured and Tenure Track Versus Non-Tenure Track Female Faculty

Changes in Faculty Over Time

Cross-sectional changes in faculty over time can be represented in a number of ways. Figure 5 shows a graph used by the University of Alabama at Birmingham to document women’s relative numbers among the STEM faculty as compared to men between 1998 and 2004. Another representation, shown in Figure 6, is a “faculty flux” chart used by Hunter College’s ADVANCE Program. This graph is basically a cross-sectional analysis but in order to show the transitions (flows into and out of positions) longitudinal data were also used. Longitudinal data provide information about individuals at multiple points in time.

The stacked bar chart that forms the basis of both graphs in Figures 5 and 6 are fairly easy to construct in spreadsheet programs like EXCEL but then manual addition of many elements must be manually added, which increases the time required to construct these charts. On the other hand, charts like these have the benefit of summarizing quite a bit of information—data that might span many tables—into one visual image, improving the ability of audiences to make connections among the processes of recruitment, retention and advancement.

It should also be noted that as the size of the institution increases (i.e., there are more faculty), the difficulty of compiling these data also increases. Institutional datasets for successive years would likely need to be compared. To some extent manual comparisons would need to be done.

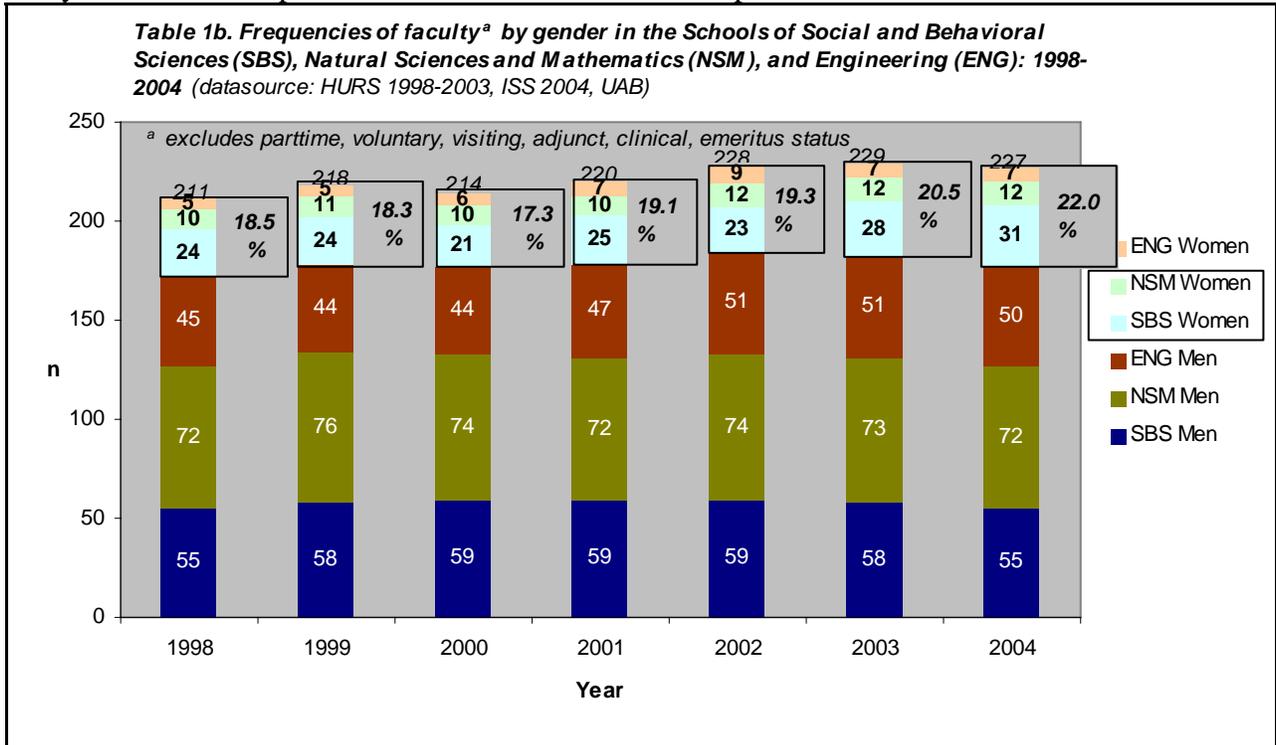


Figure 5. Faculty by Gender at University of Alabama at Birmingham, 1998-2004

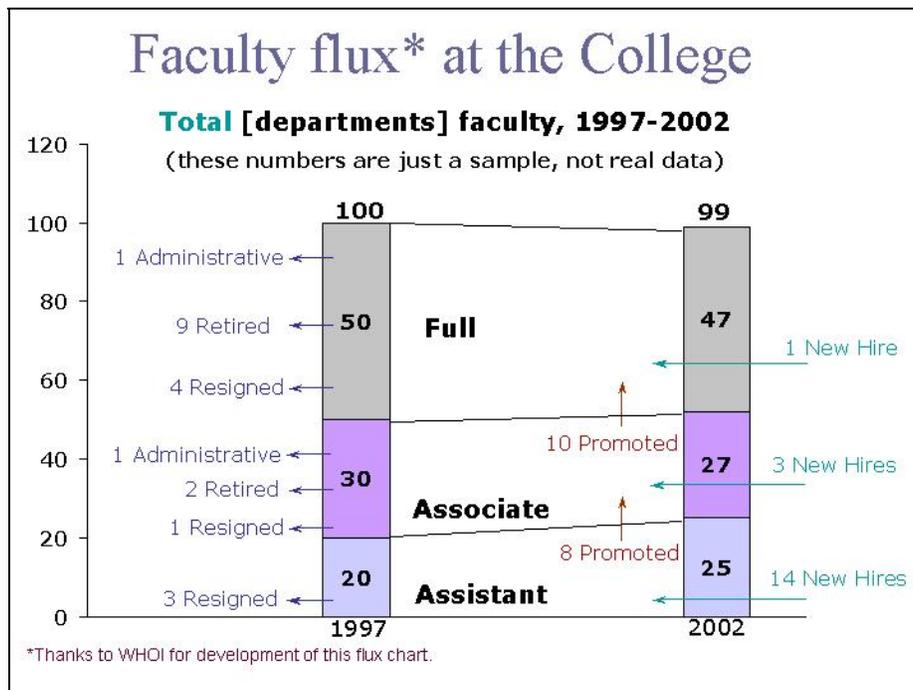


Figure 6. Hunter College Faculty Flux Chart

Institution-Specific Controls

Figure 7 shows a graph comparing women’s representation among newly hired faculty at the University of California, Irvine (UCI) versus the rest of the California system separately for untenured and tenured faculty. Such an analysis is not involved in the graph for NMSU’s recruitment results, previously shown in Figure 2 because almost all new faculty at NMSU are untenured assistant professors. UCI, however, is an institution that is attempting to increase its departments’ national rankings, which necessitates hiring more senior faculty, where, especially in the STEM fields, women are more scarce. In addition, the processes by which senior versus junior faculty are hired differ: many of the published guides on increasing faculty diversity, for example, provide excellent tips for hiring faculty members who are within a few years of completing PhDs but not so many tips for how to recruit prestigious, well-known faculty. Hence, to understand the impact of ADVANCE at UCI, it is necessary to examine data for hiring within these two categories while at NMSU, this would not make sense.

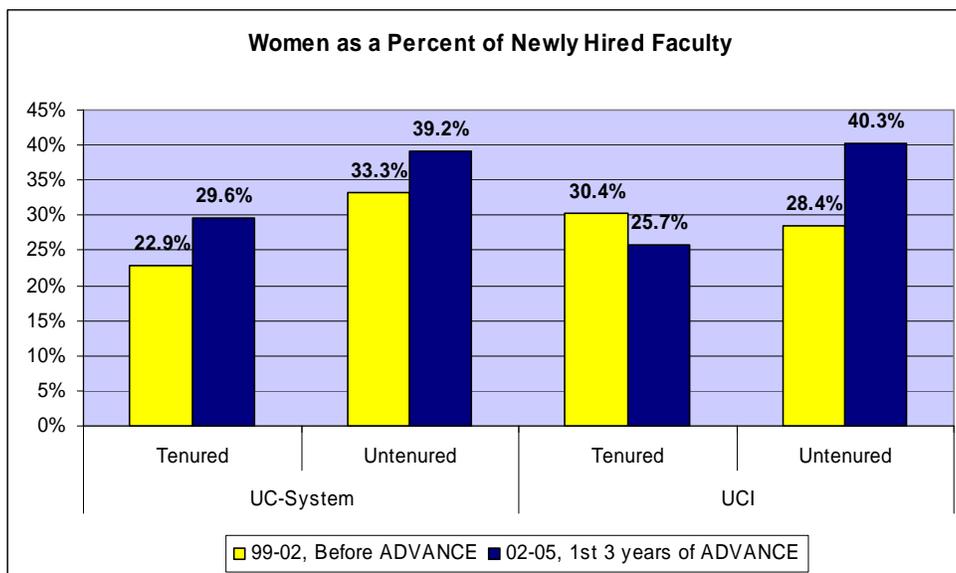


Figure 7. Recruitment of Women at UCI by Tenure Status at Time of Hiring

It is also useful to produce charts separately for each unit or different discipline groups. Some basic pointers on chart construction:

- Whatever source of data you decide to use to “count” new hires, this should be specified and consistent across time.
- Another useful chart to show recruitment success can indicate how your institution performs as compared to the “availability pools” at the national level. The University of Washington has used line charts, while NMSU has used bar charts to present these data to key decision makers (department chairs, deans, and provosts).

Qualitative Evidence of Program Effectiveness

What is qualitative research? Qualitative research is a scientific inquiry that seeks to understand how people make sense of the social world by engaging in inductive processes using one or more of the following methods:

- Interviews

- Direct observation
- Focus groups
- Workshop evaluation forms with open-ended questions
- Applications for programs that ask applicants to prepare an essay on a specific topic
- Meeting minutes, notes, websites, e-mails, and other written materials as historical documents.

Reliability and validity are key concerns with qualitative analysis, which must be well-demonstrated to ensure that the data are convincing. Because of the subjective elements of qualitative approaches, it is customary to use more than one method to understand the phenomenon of interest as a way of demonstrating the reliability of findings. For example, qualitative interviews might reveal that women feel isolated from colleagues. A climate survey might have also asked questions on this same topic. Likewise, you might also find that at meetings that women sit away from a central conference table and do not get the floor to speak as often as their male colleagues. In this case, there are three sources of information to “check” and “complement” each other in understanding how women feel isolated and to what extent.

Another way to establish the reliability of qualitative analysis is to use multiple coders. Coders are researchers who read and interpret the qualitative data using a coding scheme that has evolved from the data itself. If there are three coders, then inter-coder reliability can be established, which is simply the extent to which they agree as the material is coded.

Because qualitative methods emphasize the words and ideas expressed by participants themselves, and quotes can be used to illustrate points in the analysis in a powerful way, establishing validity with qualitative data can be more straightforward than with items on a fixed-choice questionnaire. For example, respondents to a survey might be asked to rate the level of agreement with the following statement: “I feel my colleagues value me” using a standard 4-point scale (strongly agree-strongly disagree). But when they complete the survey, they don’t know what you mean by “value me.” With a qualitative interview, they instead describe their relationship with their colleagues in a way that reflects the extent to which they feel valued and even define “value” for themselves. Therefore, in a qualitative interview you can ascertain that the concepts you are exploring are constructed by the respondents and then organized thematically by the researcher.

Qualitative methods, while rich in contextual detail, require substantial investment of time and resources. Interviews are usually transcribed and often analyzed using textual analysis software. Analysis and interpretation often require multiple coders to check each other for possible bias. The inductive analysis requires a deep understanding of the data that goes beyond simple codes. Therefore, results from qualitative studies may not be rapidly available.

So if qualitative research is so time consuming, why use it?

- Validity: as a researcher you can feel confident that these issues are important to respondents, that you are not putting words in their mouth or misunderstanding them.
- Building programmatic support—during an interview or focus group, information about the program can be provided to faculty to involve more faculty in solving problems.
- Small sample sizes: because there are so few women in STEM fields, the issues they experience are complex, which can be detected using qualitative methods.
- Interaction preference, especially in some cultures: face-to-face or verbal communication via phone may be preferred over “anonymous” questionnaires.

- Testimonials: for fundraising or seeking administrative support.
- Vignettes for training purposes.

Testimonials can be gleaned from data that have been collected using qualitative methods. Such testimonials can have a big impact in multiple contexts. Foundations, women’s organizations, administrators (advancement office) the press, etc. all appreciate testimonials as a way of putting a face on stark numbers. Washington, Wisconsin, NMSU are in the process of using these data to put a human face on the program and to demonstrate the personal impact of the ADVANCE program. Men’s program participation can be highlighted in these ways to illustrate that the program is not just for women. Even if there are few men involved (e.g., transitional support programs or other kinds of “emergency” grants), it may be important to document that these awards CAN go to men to avoid “backlash.”

E.g. The University of Wisconsin’s Life Cycle Grant program provided support to four faculty (three women and one man) in its inaugural year. Qualitative interviews with these recipients revealed powerful stories about the significance of the support to their careers at a critical time. These powerful stories provided the impetus for continued funding of the Life Cycle Grant program by the estate of William F. Vilas.

Qualitative data can also be very useful to assemble realistic case studies for training purposes. Combine elements from multiple stories to mask identities, being cautious to protect the respondents’ identities.

Presentations based on qualitative research need to emphasize the methodological strengths and be up front about the weaknesses. Sample size, how people were recruited to the research, who conducted the research how the analysis was performed need to be specified to a greater or lesser extent depending upon the audience.

Audiences:

- STEM faculty
- Non-STEM faculty (more likely to have a knowledge of qualitative research)
- Administrators

Each audience will have a different level of background with qualitative research and will have different preferences for communication style. Generally you will need to provide different levels of detail about the inquiry, results, and recommendations. With qualitative data, you will have a list of themes and the frequency with which the themes arose in the data. These themes can then be illustrated using vignettes or quotes from the data.

With STEM faculty and administrators, it is important to avoid specialized jargon. For example, “we used multiple modes of inquiry” really should be stated instead as “we used several methods to determine why . . .” STEM faculty, are often unfamiliar with qualitative research, it is important to provide enough methodological detail to demonstrate that the data are “scientific.” For this group, it is important to use the multiple methods approach to establish reliability to avoid accusations of subjectivity/bias/over interpretation. Quotes, therefore, are very important “raw data” to provide to this group.

Non-STEM faculty who may be more familiar with qualitative approaches, it may be necessary to simply specify the inter coder reliability to address possible concerns about bias. Because the

method is possibly more familiar, connecting the research to the literature might be more important and pertinent to this group. Theories and implications related to the research are also more likely to be of interest to this group.

Administrators are often pressed for time and interested in what action steps they need to take based on your research. Therefore, you often do not want to spend precious time describing the research and, instead, get to the “punch line.” State how you collected the information, your key findings, and what the administrator should do based on these findings. The communication should be clear, succinct, precise, and emphasize no more than one or two important points. It is also useful to prioritize these actions for the administrator. Which one is most important? Provide the strongest evidence to support the most important action.

Other issues in dealing with administrators:

- Administrators will also need to know the projected cost of the recommendations.
- Usually interested in the “big picture,” therefore do not overwhelm them with minor details.
- What’s in it for them? Why should they, from their own personal standpoint, take the actions that you suggest? (E.g., are there risks associated with not acting?)

Other Considerations with Qualitative Research

Confidentiality: human subjects approval is generally necessary, especially if you would like to use quotes from interviews or other materials.

Clearing the quote with the respondent is also a good practice. Do not use a quote or information that the respondents request that you not use.

It is important to collect qualitative data in various ways so that you can have many choices of quotes, especially if some are so sensitive that people would prefer you not to use them.

When these data are collected by various people, it would be useful to have one person who can systematize these data to improve analysis and permit generation of results across contexts. You may need to convince members of your leadership team that something is “data.”

E.g., when program staff meet with female candidates, these meetings are a source of data. Documenting these meetings can be used to justify the perpetuation of the program and its efforts. (They hire women after ADVANCE “intervention.”) Avoid relying on your own brain as a processor or storage device.

Tape record and transcribe interviews for best reliability and accuracy. It would be useful to manage these data, as you would quantitative data, using appropriate qualitative analysis software. This could enable construction of a larger database.

Evidence of Policy Implementation → institutional transformation

For institutional transformation to take place and our efforts sustained, policy changes and implementation must occur. The policy changes and implementation is a success of the ADVANCE:IT grant so it is important to document these changes. Measuring institutional transformation is a challenge, especially when the grant focuses on science and engineering faculty, so the most straightforward way to measure impact on the institution is to document successes in

policy changes and implementation. The question then is what is the best way to document policy changes and implementations?

The following table is an example of how to measure and document policy changes and implementations. The first column are examples of changes in policy that will occur at the end of the five year ADVANCE:IT grant and the second column are indicators or measures of those successes in policy.

Motivate—document—best practices
Data on the ADVANCE best practices

c. Major Points	Indicators
Policies and procedures are public and transparent	<ul style="list-style-type: none"> • Posted on websites • Accessed easily • Those affected are informed of policies and procedures and made sure that they are understood • No bias in evaluation processes/objective standards
Committed leadership	<ul style="list-style-type: none"> • Resources are given • Policies and procedures are created, changed • Media used to inform community • Evaluation of leaders
Inclusive of diversity in recruiting	<ul style="list-style-type: none"> • Recruitment pool reports • Increased diverse hires
Welcoming climate	<ul style="list-style-type: none"> • Climate studies show positive changes over time • Norm of success—that the institution ensures that all new faculty are given the resources (physical, emotional, social, etc.) to succeed. • All faculty feel invested in the success of a new faculty member—if a junior faculty member leaves, then it is a failure of EVERYONE. • Incorporation of retention and successful promotion and tenure outcomes into the chair evaluation. • Student enrollment—<u>grad/undergrad</u>, retention
Collaborations with o/s stakeholders	<ul style="list-style-type: none"> •
“Spirit of community,” Collaboration	<ul style="list-style-type: none"> • Increased mentoring of faculty • Opportunities for faculty to meet and learn about each other’s research • Networking structures (incl personnel that enable this)
Women feel more powerful	<ul style="list-style-type: none"> • Women in leadership positions • Tracking of women who might leave our institutions but move on to leadership positions elsewhere.

APPENDIX A:
SAMPLE EVALUATION PLAN

Workshops/ Events	Goal	Involved Groups	Workshop Elements	Assessment of Outcomes	Possible sources of information
Promotion & Tenure workshops	Provide faculty members with information on processes, policies	Faculty Deans Department chairs P&T committee members	Writing statement Identifying letter writers Compiling the package Do's and don'ts in process Understanding when "ready"	# of faculty attending workshops # of faculty who are promoted and tenured satisfaction with workshop elements perceptions about relevance and usefulness of information presented	
Chair training	Increase leaders awareness of gender equity Improve leaders ability to govern Improve leaders ability to manage faculty (conflict resolution, faculty development, etc.) Improve processes of implementing policies	Chairs Faculty Deans Provost EOO	Communication skills Managing difficult people Understanding policies and human resources regulations Leadership skills Self-reflection of management style Gender bias in ratings and evaluations Unconscious bias	Change in knowledge about issues Interviews with faculty about change Satisfaction with event elements Perceptions about relevance and usefulness of information presented Change in policy utilization and implementation Implementation of anti-bias processes Dissemination of information at faculty meetings	
Applying for grants	Increase the number of faculty applying for research grants Improve likelihood of faculty receiving research grants	Chairs Faculty Deans Funding agency program officers	Writing tips Information about the processes and funding agencies	#number of grant applications submitted percentage of grants funded	

Table A1. Sample Evaluation Plan for Workshops/Events, Diamonds in Figure 1 Venn Diagram

Initiative/ Program	Goal	Involved Groups	Program Elements	Assessment of Outcomes	Possible Sources of Information
Recruitment	<p>Improve awareness of national availability</p> <p>Improve transparency of search processes</p> <p>Implement policies that impact women's careers</p> <p>Increase awareness of subtle biases in search processes (e.g., gender schemas, ad placements, etc.)</p>	<p>Search com chairs & members,</p> <p>Dept chairs</p> <p>Deans</p> <p>Provost</p> <p>EEO/OEOD/ .</p> <p>Human Resources</p> <p>Faculty Senate</p> <p>Relevant allied groups</p>	<p>Search assistance</p> <p>Start-up enhancement funds</p> <p>Dual career assistance</p> <p>Family-friendly policies</p>	<p># workshops run (incl frequency)</p> <p>attendance at workshops (# of partic, rank, admin level, and breadth of disciplines attending)</p> <p>program interactions with Deans, Provost related to gender equity in recruitment</p> <p>Quality of program relationship with: EEO, HR, Faculty Senate, and other related allied groups.</p> <p>% female newhires (TK1) pools among interviewees among offers made/refused</p> <p>documentation of new procedures/processes (e.g., Dean meets with each search committee to convey diversity as an important goal)</p> <p>equitable start-up packages (TK1)</p>	
Retention	<p>Improve climate – nurturing environment (w/in depts. & across univ)</p> <p>Ensure transparent resource allocation processes</p> <p>Inclusion in dept/univ decision processes</p>	<p>Faculty members</p> <p>Dept chairs</p> <p>Deans</p> <p>Provost</p> <p>Human Resources</p>	<p>Dual career assistance</p> <p>Family-friendly policies</p> <p>Climate studies</p> <p>Resource allocation studies</p> <p>Life cycle grants</p> <p>Research and travel subawards</p> <p>Mentoring programs</p>	<p>Climate studies show that regardless of gender:</p> <ul style="list-style-type: none"> • Faculty feel they have a voice • Understand processes • Involved in decision making <p>Resource allocation studies indicated that regardless of gender:</p> <ul style="list-style-type: none"> • Salaries are equitable • Space allocation is equitable • Graduate students/post-docs' funding 	

Initiative/ Program	Goal	Involved Groups	Program Elements	Assessment of Outcomes	Possible Sources of Information
				<p>equitable</p> <p>Involvement in mentoring programs (#, sex, rank, disciplines, etc.)</p> <p>Satisfaction & perceptions of mentoring program (benefits)</p> <p>Additional grants awarded</p> <p>Publications and scholarly presentations</p>	
Advancement	<p>Prepare women for leadership opportunities</p> <p>Increase leaders' awareness of gender equity</p> <p>Increase advancement of women in academe</p>	<p>Faculty</p> <p>Dept chairs</p> <p>Deans</p> <p>Provost</p> <p>President</p> <p>Advancement office</p>	<p>Transparent process: promotion from assoc to full</p> <p>Fund-raising to enable endowed/named professorships for women</p> <p>Knowledge about national-level opportunities</p> <p>Educate chairs and deans about advancement processes and gender bias therein</p>	<p>Additional grants awarded</p> <p>Publications and scholarly presentations</p> <p>Awards & recognition (dept, campus, discipline, national, etc.)</p> <p>Leadership positions (dept, campus, discipline, national, etc.) (TK1)</p>	
Institutional transformation	<p>Facilitate and maintain committed leadership at all levels</p> <p>Develop and communicate transparent policies and procedures</p> <p>Guarantee diversity in recruiting</p> <p>Create incentives for STEM women to remain in the academic enterprise</p>	<p>Faculty</p> <p>Department chairs</p> <p>Deans</p> <p>Provosts</p> <p>Presidents</p>		<p>On-going monitoring of recruitment, retention, and advancement of STEM women within the institution's permanent evaluation/planning structure/office</p> <p>Analysis and documentation of the changes in policies and procedures across time</p> <p>Compare STEM applicant pools with availability in terms of gender and race/ethnicity</p> <p>Analysis of promotion and attrition of</p>	<p>Analyze structures and rewards that are available to STEM faculty to succeed</p> <p>Assess leadership capacity building programs or other strategies to ensure STEM women are advancing into leadership roles in and outside of the institution</p> <p>Evaluate opportunities for interaction, "mentoring," and networking structures inside and outside of the institution</p>

Initiative/ Program	Goal	Involved Groups	Program Elements	Assessment of Outcomes	Possible Sources of Information
	Develop a welcoming climate for STEM women Assist STEM women in developing leadership competencies in order to acquire academic leadership positions Generate a spirit of community and collaboration within and outside the institution			STEM women over time	

Table A2. Sample Evaluation Plan for Initiatives/Programs, Triangles in Figure 1 Venn Diagram

APPENDIX B

SO YOU WANT TO RUN A CLIMATE SURVEY?!

Climate survey data can be a useful evaluation tool. Different universities have used climate survey data:

- to determine campus needs,
- to provide baseline data before interventions occur,
- to evaluate the effect of interventions especially at the department level,
- as a measure of accountability for deans and department heads/chairs,
- to determine directions for the ADVANCE program, and
- to provide quantitative data on job satisfaction, relationship quality and isolation among faculty members, and more.

The first thing to do once you decide to do a climate survey is to find out if your campus is already taking part in a national survey of faculty such as HERI³ or Harvard's COACHE⁴ project. If so, your institution may be able to add questions specific to your institution and/or request special tabulations. Any additional surveying you do should complement existing campus surveys, not duplicate them. There is an increasing tendency to over-survey faculty, and response rates are suffering accordingly. In order to enhance the success of your survey, and all such surveys on your campus, work in collaboration rather than competition with existing faculty survey initiatives.

Design Considerations

To develop your own survey, you might start with survey instruments at the Utah State ADVANCE website⁵. Ensure that a person experienced in survey methodology designs the instrument. If you use items from existing surveys you will be able to compare your results to the previous work, and thus provide benchmark data for your campus. Some of the important considerations to debate while creating your climate survey include:

1. Timing. Try to run the climate survey in Year 1 to establish baseline data. Be especially mindful of the academic year and avoid sending your survey into the field during a time you know will be difficult (e.g., during finals week, in the middle of summer.) Approach your administration well in advance of your field date in order to establish the sample (Faculty in sciences only? Entire faculty? Off-ladder faculty?) Some sites have found that they administer their surveys as planned in their ADVANCE grants, only to find their campus's administration wanting to increase the sample to populations not surveyed by ADVANCE. Intra-campus comparisons (e.g., between science faculty and humanities faculty) can be important; try to establish the sample before going into the field, because re-surveying the faculty/staff you left out can be time consuming and possibly give unreliable results.

2. Mode. Most sites use either paper or web surveys, with advantages and disadvantages to each.

- **Paper Surveys:** More expensive, requires addresses (either home or office), requires data entry, longer to get results, better response rates, better for long surveys

³ <http://www.gseis.ucla.edu/heri/heri.html>

⁴ <http://www.gse.harvard.edu/~newscholars/>

⁵ <http://websites.usu.edu/advance/Document/index.asp?Parent=6337>

- **Web Surveys:** Less expensive, requires email addresses, dataset generated automatically, lower response rates, less-good for very long surveys

3. Confidentiality vs. Anonymity. There are advantages to running a confidential survey rather than an anonymous survey. The coding on a confidential survey allows for follow up of non-responders and results in a higher overall response rate. It also allows tracking of individual responses over time, and the linking of publicly-available data (such as productivity, grants, space, salary, etc.) to the survey responses. If you plan to maintain identities of survey respondents in order to link successive climate surveys, or surveys with other data, ***be sure to obtain specific IRB approval to do this!***

4. Survey Content. Consider mining other surveys, including those done at other ADVANCE sites, to uncover a core set of questions/items that all schools agree to include⁶. Topic areas to consider include: overall job satisfaction, intention to quit, psychological safety, respect and collegial relationships, isolation, satisfaction with the promotion and tenure process, access to and satisfaction with resources, work-family balance, satisfaction with administration, satisfaction with the hiring process, leadership opportunities and leadership development, mentoring, and sexual harassment. Also consider using your survey to evaluate programs and policies specific to your campus.

5. Survey Administration. We recommend using a reputable survey organization to administer your climate survey. On many campuses the planning and assessment office can administer the survey. You do not want your ADVANCE team to be responsible for mailing surveys, tracking responses, and coding the (hopefully) hundreds of questionnaires that are returned. A professional survey outfit can also assist you with survey design, advice about how to increase response rates (sometimes based on experience with your specific campus), obtaining IRB approval, pre-testing your instrument, and more.

6. Response Rates. In order for your data to be taken seriously by campus decision-makers, it is imperative that you do everything possible to ensure high response. As mentioned above, the highest response rates are obtained when a mail survey is used, combined with multiple mailings. Personal requests from ADVANCE PIs can have a large impact on response. Some sites have PIs/co-PIs call individual faculty non-respondents, asking them to mail their survey back (but make sure that PIs are cleared by the IRB before accessing information about individual respondents!) For others, the PIs visited department chairs, requesting the *chairs* to encourage their departmental faculty to fill out the survey. It might be a good idea to request a person in authority (Provost/President/Chancellor) to send a postcard or email to all faculty urging them to respond. (Use knowledge of your own campus to optimally time such a message. Some campuses might get better response if the Provost makes the request up front; others might find their response rates decline if faculty think the Provost has access to the data.)

7. Cover Letter. The cover letter you send with your survey (or text that you present at the beginning of your web survey) should be approved by your IRB. The content of this letter is important, not only to inform respondents of their rights and of the uses to which you intend to

⁶ The Association of American Universities Data Exchange (AAUDE) is convening a working group to determine a comment set of climate questions; watch their website for more information, <http://www.pb.uillinois.edu/AAUDE/>.

put their data, but also to assure them that you will safeguard their responses with the highest levels of confidentiality. Use the cover letter wisely, and it will help increase trust with your respondents, and therefore, your response rates.

8. Department-Level Data. All Institutional Review Boards (IRBs) will require you to ensure the confidentiality of survey responses. Consider how you will ensure data provided by faculty will remain confidential. Some faculty feel very vulnerable when answering questions about their work lives and need to be assured that there will be no way that someone will be able to identify them personally. In particular, consider whether you will request information about *department*, and how you will use this information. Results from a climate survey are almost useless if you do not have information about department, as most climate is locally experienced. Most of the time, you will need to ask the respondent to name his or her department to get this information, and this particular item makes most faculty feel vulnerable and identifiable, especially women and minorities. Decide *in advance* whether you will identify individual departmental results, to whom these results will be reported, and include this information in a cover letter to faculty. Some institutions promised to never release data at the departmental level (aggregating to the school/college or divisional level instead.) Others provided departmental-level data, but did not break it down by gender in order to maintain confidentiality. If you decide that you will release results at the department level, decide how you will keep responses from minority members of the department (e.g., women, faculty of color) confidential. We advise not reporting department identity outside of the institution.

Using Climate Data

Done well, a campus climate survey can provide rich data and important information to your ADVANCE program and to administrators and committees on your campus. Providing these data to your campus can increase the visibility and the value of your ADVANCE program.

Some of the ways ADVANCE programs have used the climate data:

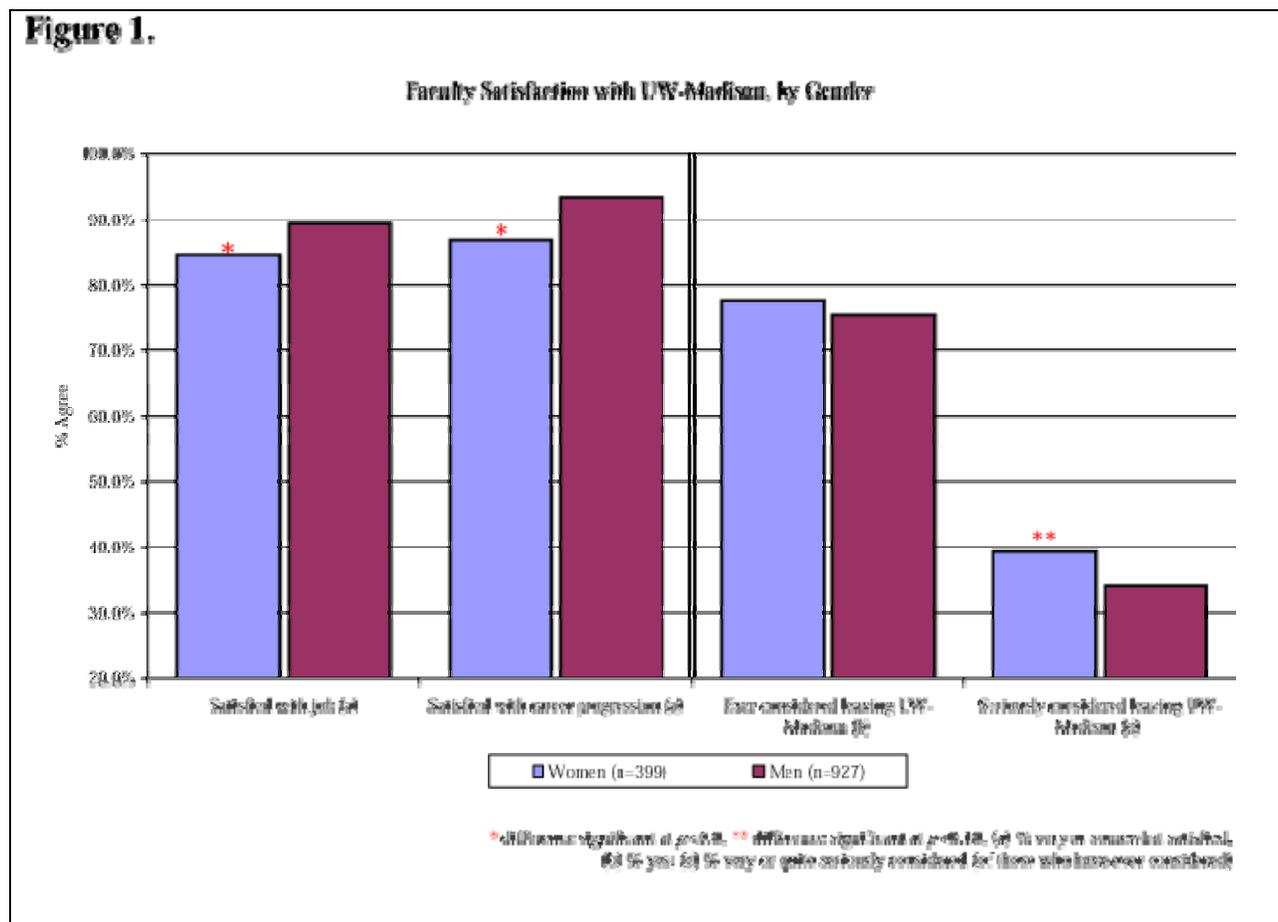
- Climate survey results can be an entree to a department to discuss climate workshops and department transformation efforts.
- Chairs/heads of high scoring departments are asked to speak at head/chair workshops on how to build a positive department climate.
- Evaluation reports on specific programs (e.g., tenure clock extensions) have been produced and presented to campus decision-makers.
- Data obtained in the first year are compared with that in a later climate survey, to assess institutional change.
- Climate results can be presented in multiple forums on campus to begin a conversation about climate.
- Results can be used for publishable papers and conference presentations.
- Results can be combined with other campus data to look at gender differences in resources, tenure rates, etc. For example, UW-Madison combined survey data on childbearing with administrative data on promotion to full professor, to look at whether women who delay childbearing until after tenure have a slower rate of promotion to full professor, compared to men.

Presenting Climate Results

When presenting data from climate surveys, do not present too many categories at once; stick to one topic, or aggregate questions to create overall indices (e.g., relationship quality or isolation.) Do not make it difficult for your audience to read the data and interpret your charts. Be prepared for questions on statistical significance, what tests were used, nuances of questions used, standard errors for data bars, questions about basic methodology, and requests for anecdotes or quotes to illustrate the quantitative results.

Example 1. Overall Results by Gender

Figure B1 shows an example of climate survey results presented to highlight gender differences in faculty satisfaction as indicated by four variables from the University of Wisconsin climate survey.



Source: <http://wiseli.engr.wisc.edu/initiatives/survey/results/faculty/satisfaction/summary.htm#Fig1>

Figure B1. Faculty Satisfaction at UW-Madison by Gender

Example 2. Highlighting Issues Other Than Gender in Reporting Climate Results

Consider emphasizing groups other than gender in the presentation (e.g., racial/ethnic, untenured vs. tenured, STEM vs. non-STEM, Mainstream vs. Non-Mainstream research, Chairs vs. faculty.) Also, when presenting data where you find major gender differences, consider presenting data where you do NOT find such differences. This helps to diffuse the “women are whiners” objection to the results.

UW-Madison supplements data on isolation and “fit” of women faculty with Figure B2, which shows how faculty who do research considered “non-mainstream” experience the same isolation and lack of “fit” that women (and minority) faculty experience. We use this to point out to chairs that it is not just gender and race/ ethnicity that might make a faculty member feel isolated in the department, and that working on ways to include ALL department members benefits all members.

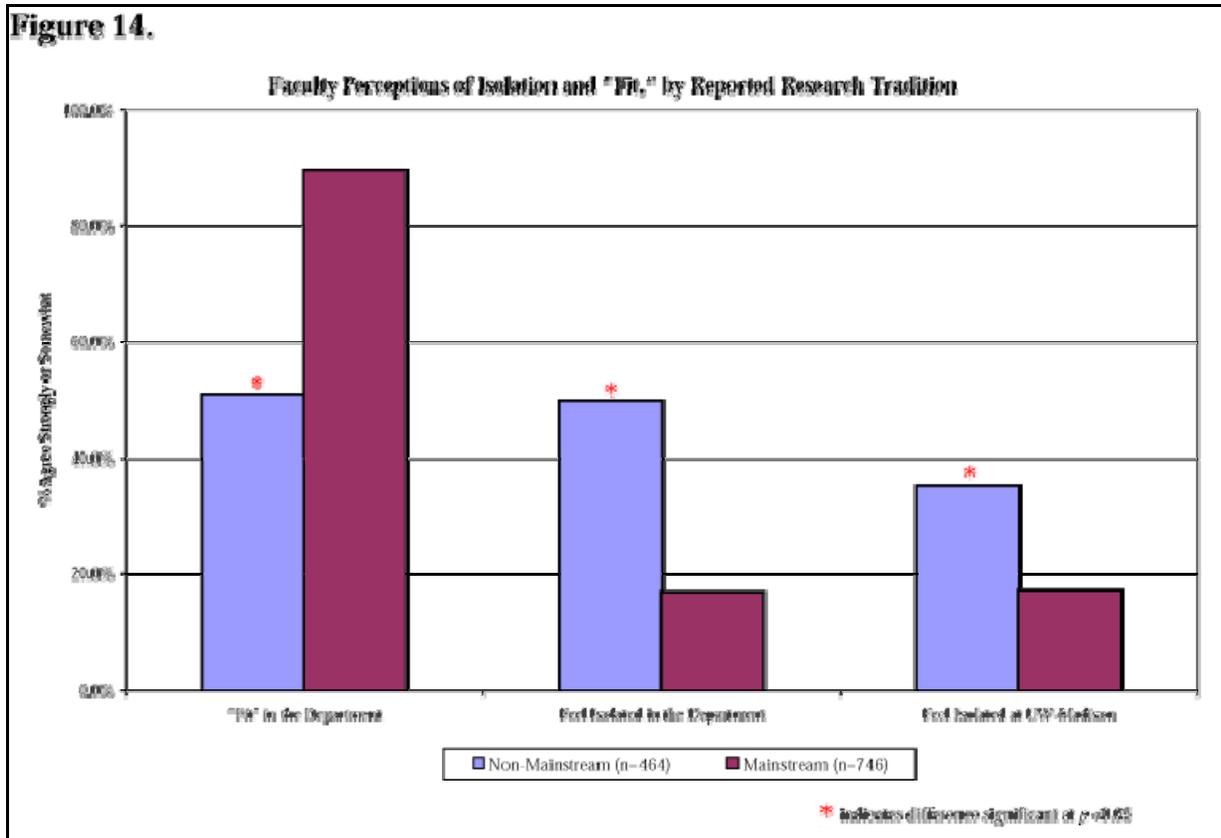


Figure B2. Example of Mainstream/Non-Mainstream Results—UW-Madison

A final thought: climate surveys are an effort that require institutionalization. Doing one survey is helpful, but establishing a systematic effort to assess climate on campus every few years is a step towards institutional change. Remember—we care about what we measure, and if you can ensure that your campus establishes a system to measure climate over time, you are on your way to changing that climate for the better. Working with other campus units to create the questionnaire, determine sample, share results, etc. can assist you in your efforts to establish a mechanism to continue periodic surveys after the ADVANCE award is over.

Example 3. Department-Level Reporting: Utah State University ADVANCE

Utah State ADVANCE held a retreat to share department-level results from their climate survey. Table B1 and Figure B3 show one graph and related text distributed at the retreat. ADVANCE

PI, Ronda Callister, combined seven survey items into one scale, which reduced the amount of data presented. A research citation related to the scale was also included and “reverse coded” items were indicated so that the scale interpretation was more straightforward.

<i>Explaining and contextualizing the climate response rate⁷</i>	A total of 563 surveys were returned for an overall response rate of 70%. In organizational research, over 50% is considered a good representative sample.
<i>Explaining the measurement of Psychological Safety</i>	Psychological safety and learning behavior in work teams. Edmondson, A. (1999) <i>Administrative Science Quarterly</i> 44(2): 350-383. alpha= .82 These questions measure the belief that the department is safe for interpersonal risk-taking. 1. If you make a mistake in this department it is often held against you (R) 2. Members of this department are able to bring up problems and tough issues 3. People in this department sometimes reject others for being different (R) 4. It is safe to take a risk in this department 5. It is difficult to ask other members of this department for help (R) 6. No one in this department would deliberately act in a way that undermines my efforts 7. In this department my unique skills and talents are valued and utilized.
<i>Why were these variables presented?</i>	The variables that show significant differences (at the .05 level) across all of the 38 departments . . .
<i>Explaining the graph labeling about departments</i>	Those departments with high scores are listed by name – all others have a confidential code name which each department head and dean should have received as they arrived at the department head retreat.
<i>Who’s NOT on the graph and why?</i>	Five departments were too small or had fewer than 8 respondents and were eliminated from this report in order to maintain the confidentiality of respondents. These responses are included in analyses when larger groups (such as faculty at each rank) are analyzed. The departments that are not included in this report are: Industrial and Technology Education; Aerospace Studies; Journalism & Communication; Landscape Architecture; Interior Design. Intensive English is included with the English department.
<i>What do the results mean and how do they relate to gender?</i>	These department climate measures are strongly related to job satisfaction and intentions to quit. Intentions to quit are generally very predictive of subsequent job turnover. We will track turnover for the next few years to see if department climate predicts turnover at Utah State. We are also in the process of interviewing faculty members in our target colleges who left Utah State in the past five years to ascertain their reasons for leaving. Analyses of the climate survey results show that women faculty report higher levels of work outcomes – higher job dissatisfaction and higher intentions to quit,-- but that the relationship between gender and work outcomes is completely explained by the quality of department climate. This suggests that while department climate seems to be important to the quality of life for both men and women, it appears that high quality department climates are even more instrumental in retaining women faculty.

Table B1. Information Provided to Department Heads about Climate Survey at Utah State

⁷ Utah State University’s ADVANCE website includes all of the graphs and information presented at their department heads retreat. We accessed this site on May 5, 2006 at: <http://websites.usu.edu/advance/FileManager/Download.asp?Parent=6308&FilePath=Dept+head+retreat+Climate+Pres+8+15+05%2Edoc>. The graph in Figure B3 has been modified slightly from the original to fit in available space. Content in grey is directly from the USU website.

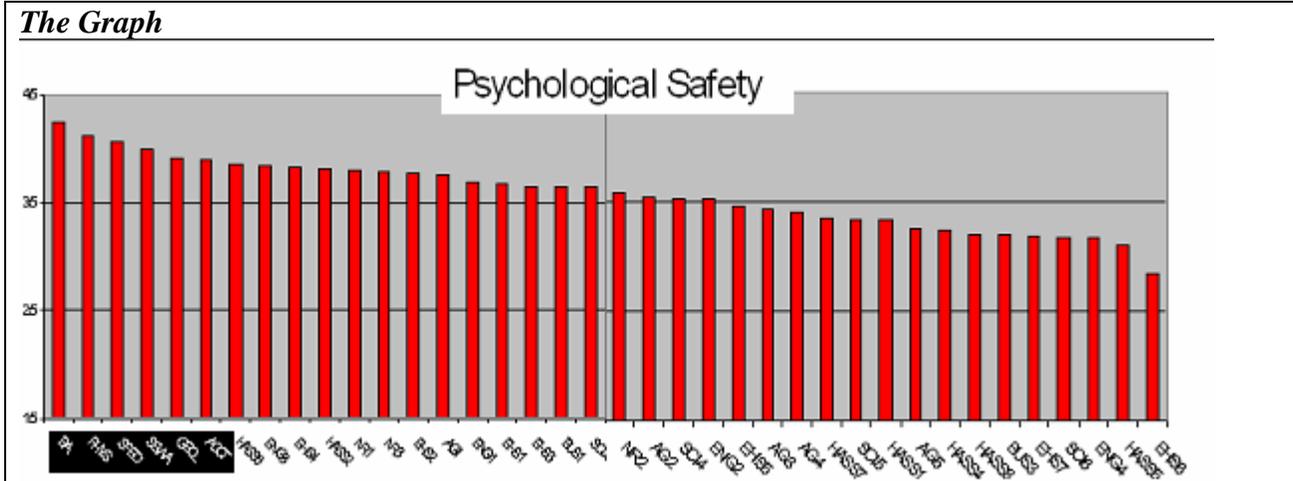


Figure B3. Example from Utah State Department Level Climate Presentation

Departments with “high scores” on the variable were identified publicly (see the black box), but those with lower scores were only identified to the actual department heads. In this way, departments could be provided with feedback about their department and see how they compared to others. USU ADVANCE team members reported that many STEM department heads questioned the metrics and measurement scales, which reflected a general skepticism with social science in general.

APPENDIX C. GLOSSARY

<u>Advancement</u>	A promotion as in rank; encouragement of the progress or growth; the act of moving forward toward a goal; gradual improvement or growth or development.
<u>Assessment</u>	Assessment is the process of documenting, in measurable terms, knowledge, skills, attitudes and beliefs.
<u>Audience</u>	Refers to the target group at which an item of communication is directed. A group of people who are receptive to a medium or message.
<u>Climate survey</u>	A quantitative tool to measure the quality of work life as perceived by the faculty.
<u>Cohort</u>	A group of people who are of the same age or who entered a system at a similar moment in time.
<u>Ecological fallacy</u>	Inferences about individuals inappropriately drawn from data gathered about groups, societies or nations. (Frankfort-Nachmias and Nachmias (pg. 517)
<u>Evaluation</u>	The systematic determination of merit, worth, and significance of something or someone. (Source: Wikipedia, accessed online April 24, 2006 at http://en.wikipedia.org/wiki/Evaluation)
<u>Faculty</u>	The scholarly staff who teach at colleges or universities.
<u>Formative</u>	Evaluation activities undertaken to furnish information that will guide program improvement. (Rossi, Lipsey and Freeman, p. 426)
<u>Generalizability</u>	The extent to which the research findings can be generalized to larger populations and applied to different settings (Frankfort-Nachmias and Nachmias (pg. 518).
<u>Individualistic fallacy</u>	Inferences about groups, societies, or nations inappropriately drawn from data gathered about individuals (Frankfort-Nachmias and Nachmias (pg. 519).
<u>Inner coder reliability</u>	When coding open-ended questions, different coders may classify the same response differently (Frankfort-Nachmias and Nachmias (pg. 311).
<u>Institutional</u>	Relating to or constituting or involving an organization or mechanisms of social structures.
<u>Institutional review board (IRB)</u>	An independent body constituted of medical, scientific, and non-scientific members who ensure that the rights, safety, and well-being of animal and human subjects is protected in the conduct of research.

<u>Logic model</u>	Defines a situation and priorities as well as the inputs, outputs (i.e., activities, participants, etc.), outcomes and anticipated impacts (short, medium, and long-range) of a program as well as the assumptions and external factors associated with the plan and context. The logic model shows the chain of connections of how a program is expected to work to achieve the desired results.
<u>Program</u>	A set of structured activities.
<u>Recruitment</u>	An activity in which the organization attempts to identify and attract candidates to meet the requirements of anticipated or actual job openings.
<u>Reliability</u>	The consistency of a measuring instrument, that is, the extent to which a measuring instrument exhibits variable error (Frankfort-Nachmias and Nachmias (pg. 524).
<u>Response rate</u>	The percentage of individuals who respond to a given questionnaire (Frankfort-Nachmias and Nachmias (pg. 524)
<u>Retention</u>	The number of faculty in a cohort who return to the institution the subsequent time period.
<u>Stakeholder</u>	Individuals, groups or organizations having a significant interest in how well a program functions, for instance, those with decision-making authority over the program, funders and sponsors, administrators and personnel, and clients or intended beneficiaries. (Rossi, Lipsey, and Freeman, p 435).
<u>Summative</u>	Evaluation undertaken to render a summary judgment on certain critical aspects of a program's performance, for instance, to determine if specific goals or objectives were met. (Rossi, Lipsey and Freeman, p. 435).
<u>Sustainability</u>	A systemic concept, relating to the continuity of economic, social, institutional and environmental aspects of human society. It is intended to be a means of configuring civilization and human activity so that society, its members and its economies are able to meet their needs and express their greatest potential in the present, while preserving biodiversity and natural ecosystems, and planning and acting for the ability to maintain these ideals in a very long term. Sustainability affects every level of organization, from the local neighborhood to the entire planet. (<i>Wikipedia accessed online at http://en.wikipedia.org/wiki/Sustainability#Sustainability_and_Competitiveness April 24, 2006</i>).
<u>Target group</u>	Learners for whom the program is designed and whose needs it is particularly intended to meet.
<u>Transformation</u>	A marked change in appearance or character, especially one for the better.
<u>Validity</u>	The degree to which an instrument measures what it is supposed to measure (Frankfort-Nachmias and Nachmias (pg. 526)

Further terminology notes

A couple of notes about terminology are necessary based on the working group discussions. The term “program” can be used in many ways. In some cases, “program” can refer to the national-level, the local level, or to pieces of a local level entity. Your particular ADVANCE program may use different terminology and have different ways of organizing your effort. The point in this toolkit is that

1. You need to evaluate pretty much everything that you do.
2. There are benefits for a national-level evaluation if individual ADVANCE: IT awardees collaborate to document the impact of the ADVANCE program.

“Institutional” is another term that was debated by our working group because there are multiple levels at which “institutional” can be understood. For example, we might think of the “institution of high education” in very broad terms. By this we would refer to the entirety of organizations that enable the pursuit of higher education. “Institutional” can also refer to the specific institution at which you are located. This means that when we speak of “institutional transformation,” we need to look at how the institution has changed—which includes units that are not within the NSF-defined STEM areas—regardless of the extent to which your program engaged these various units.

The Distinction between Evaluation and Assessment

In the field of evaluation, there is some degree of disagreement in the distinctions often made between the terms 'evaluation' and 'assessment.' Some practitioners would consider these terms to be interchangeable, while others contend that evaluation is broader than assessment and involves making judgments about the merit or worth of something (an evaluand) or someone (an evaluatee). When such a distinction is made, 'assessment' is said to primarily involve characterizations – objective descriptions, while 'evaluation' is said to involve characterizations *and* appraisals – determinations of merit and/or worth. Merit involves judgments about generalized value. Worth involves judgments about instrumental value. For example, a history and a mathematics teacher may have equal merit in terms of mastery of their respective disciplines, but the math teacher may have greater worth because of the higher demand and lower supply of qualified mathematics teachers. A further degree of complexity is introduced to this argument when working in different languages, where the terms 'evaluation' and 'assessment' may be variously translated, with terms being used that convey differing connotations related to conducting characterizations and appraisals. (Source: Wikipedia, accessed online April 24, 2006 at <http://en.wikipedia.org/wiki/Evaluation>)

Faculty

Faculty can be any college or university staff who teach or conduct research. This means, as discussed in Toolkit #1, that it is essential that you define WHICH faculty you are counting and to make a careful notation of this group whenever you report results. Colleges and universities differ in terms of the ranks that are included on the tenure track and in the duties and prestige accorded to those who hold doctoral degrees but engage in activities off the tenure track (such as teaching or funded research).

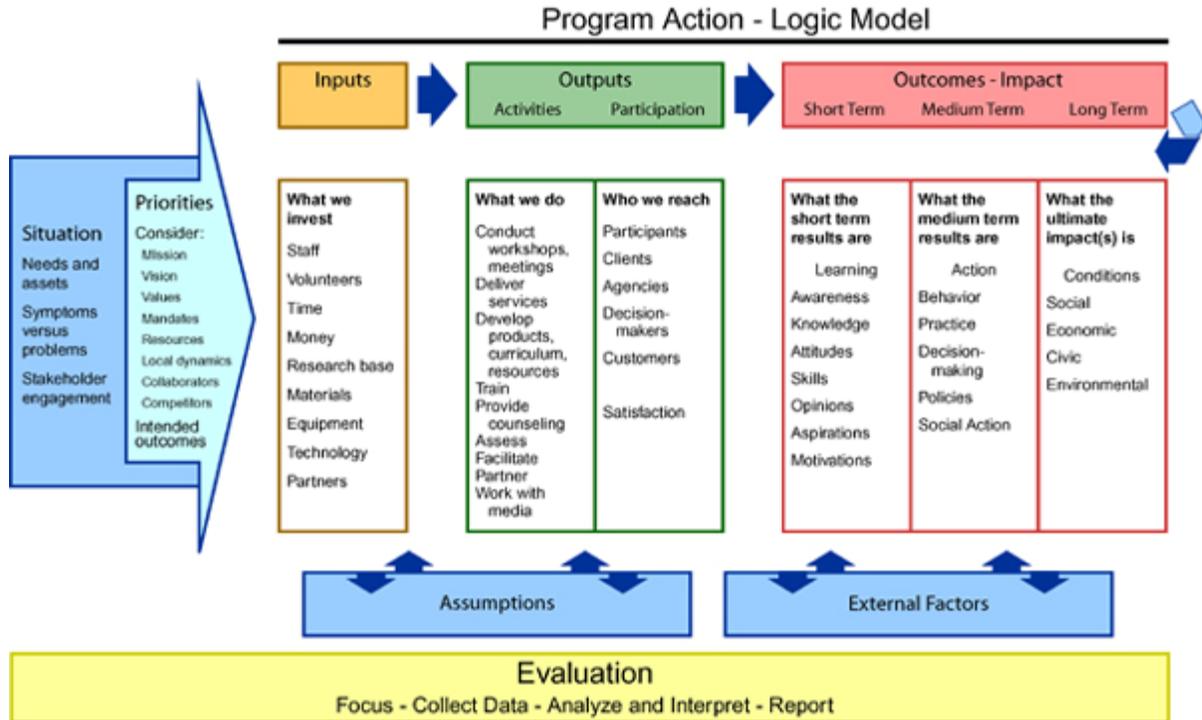
Logic Model

According to the University of Wisconsin, Extension Program Development office, the logic model displays actions that describe what the program is and will do – how investments link to results. They include 5 core components in this depiction of the program action:

1. **INPUTS:** resources, contributions, investments that go into the program
2. **OUTPUTS:** activities, services, events and products that reach people who participate or who are targeted
3. **OUTCOMES:** results or changes for individuals, groups, communities, organizations, communities, or systems
4. **Assumptions:** the beliefs we have about the program, the people involved, and the context and the way we think the program will work
5. **External Factors:** the environment in which the program exists includes a variety of external factors that interact with and influence the program action.

In UW-Extension, the logic model is used in planning, implementation, evaluation and communication. While the term “program” is often used, a logic model is equally useful for describing group work, team work, community-based collaboratives and other complex organizational processes as we seek to promote results-based performance.

University of Wisconsin Extension includes the following diagram and has a downloadable template at their website: <http://www.uwex.edu/ces/pdande/evaluation/evallogicmodel.html>.



Sustainability can be defined both qualitatively in words, and more quantitatively as a ratio. Put in qualitative terms, sustainability seeks to provide the best of all possible worlds for people and the environment both now and into the indefinite future. In the terms of the 1987 Brundtland Report, sustainable development is development "*[m]eeting the needs of the present generation without compromising the ability of future generations to meet their needs.*" This is very much like the seventh generation philosophy of the Native American Iroquois Confederacy, mandating that chiefs always consider the effects of their actions on their descendants through the seventh generation in the future. The original term was "sustainable development", a term adopted by the Agenda 21 program of the United Nations. The 1995 World Summit on Social Development further defined this term as "*the framework for our efforts to achieve a higher quality of life for all people,*" in which "*economic development, social development and environmental protection are interdependent and mutually reinforcing components.*" (Wikipedia accessed online April 24, 2006, at http://en.wikipedia.org/wiki/Sustainability#Sustainability_and_Competitiveness)

REFERENCES

Frankfort-Nachmias, Chava and David Nachmias. 2000. *Research Methods in the Social Sciences, 6th Edition*. (New York: Worth).

Posavac, Emil J. and Raymond G. Carey. 1997. *Program Evaluation: Methods and Case Studies, 5th Edition*. (Upper Saddle River, NJ: Prentice-Hall).

Rossi, Peter H., Mark W. Lipsey, and Howard E. Freeman, 2004. *Evaluation: A Systematic Approach, 7th Edition*. (Thousand Oaks, CA: Sage).