A Method for Evaluating Advanced Technological Education (ATE) Projects and Centers

Developed by Wayne W. Welch; University of Minnesota (ret.)
July 2014

The Advanced Technological Education (ATE) program funded by the National Science Foundation (NSF) supports projects and centers designed to improve the technological skills of America's workforce. All are required to carry out evaluation activities (National Science Foundation, 2011). The evaluation information is used to improve project and center effectiveness and to provide accountability evidence to NSF.

This document presents a new method to help ATE team leaders and others address their evaluation needs. It is a survey that consists of 32 statements made by ATE PIs. The statements addressed the strengths and limitation of their grant work. Survey respondents are asked to determine if these statements apply to their situation. They respond by agreeing or disagreeing with the statements on a five-point Likert scale. PIs and evaluators can use the survey to gather evaluation information about ATE projects or centers.[1]

This document consists of three sections: Part I is a copy of the survey; Part II includes information for scoring and interpreting the findings; and Part III is a description of survey development together with validity and reliability data.

The survey adds to the group set of procedures that might be used to evaluate the ATE program. It can be administered to PIs, staff members, faculty, students, national review committees, or project evaluators. Administration information is found on the first page of Part I and ways to score and interpret findings is found in Part II of this document.

Author Notes:

---

[1] In the remainder of this document, the term project will be used to mean both projects and centers.

Part I. ATE Evaluation Survey

Leaders of ATE projects and centers were asked to write about the effectiveness of their projects or centers. We selected a few of these statements and wonder if they are consistent with your ATE experiences.

Please read each statement and indicate the extent to which you Agree or Disagree using the following response scale. (Circle the response option at the right of the page that best represents your opinion.)

AA	*Strongly Agree with the statement*
A	*Agree with the statement*
U	*Uncertain whether I agree or disagree*
D	*Disagree with the statement*
DD	*Strongly Disagree with the statement*
NA	*Not applicable*

The survey will take about 12 minutes. Instructions for scoring and interpreting your results are found in Part 2.

	Circle the option that best represents your situation.
1. "Our administration has supported our ATE efforts."	AA A U D DD NA
2. "Persistence of students recruited through our project is lower than the college's average persistence."	AA A U D DD NA
3. "Our NSF grant has given us the confidence to seek and obtain funding from other sources."	AA A U D DD NA
4. "We have at least one industry partner who is committed to support some of our work after NSF funding has ended."	AA A U D DD NA
5. "Our NSF/ATE grant has had little long-term impact on our college."	AA A U D DD NA
6. "Student interest in technology careers has increased because of our ATE grant."	AA A U D DD NA
7. "Businesses in our area have benefited from having a more qualified pool of job candidates from which to choose."	AA A U D DD NA
8. "The ATE grant helped us to establish relationships with professionals from four-year colleges that will continue in the future."	AA A U D DD NA

9. "Our faculty has improved their teaching style because of their involvement in our ATE grant."	**AA**	**A**	**U**	**D**	**DD**	**NA**
10. "Our ATE grant experience has caused our administration to encourage other faculty to seek external funding to address workforce needs."	**AA**	**A**	**U**	**D**	**DD**	**NA**
11. "Because of our grant, local industries are more willing to ask us to provide technological education for their workforce."	**AA**	**A**	**U**	**D**	**DD**	**NA**
12. "It is doubtful that the relationships we have established with our various partners will continue after our ATE grant has ended."	**AA**	**A**	**U**	**D**	**DD**	**NA**
13. "Changes made in our technological education program will keep going after our current grant ends."	**AA**	**A**	**U**	**D**	**DD**	**NA**
14. "We have data-based evidence that our program improved the workforce skills of our graduates."	**AA**	**A**	**U**	**D**	**DD**	**NA**
15. "The ATE grant has increased our sense of worth by being a part of this national effort."	**AA**	**A**	**U**	**D**	**DD**	**NA**
16. "Alumni of our project are ambassadors to the larger technical community and tell potential students about the value of technology education."	**AA**	**A**	**U**	**D**	**DD**	**NA**
17. "Our project/center was instrumental in helping our local educators and employers understand the need for further science technology, engineering, and mathematics education."	**AA**	**A**	**U**	**D**	**DD**	**NA**
18. "We have been able to make available an educational program that was previously not offered as an option for people in our area."	**AA**	**A**	**U**	**D**	**DD**	**NA**
19. "We have little evidence that our professional development program has improved the teaching skills of our instructors."	**AA**	**A**	**U**	**D**	**DD**	**NA**
20. "Our faculty has become a part of a collegial network that will continue to share program information, workforce trends, and cutting-edge instructional technologies."	**AA**	**A**	**U**	**D**	**DD**	**NA**
21. "Most of the programs and activities started during our ATE grant will come to an end when our NSF funding stops."	**AA**	**A**	**U**	**D**	**DD**	**NA**
22. "More K-12 teachers are integrating engineering and technology content into their classrooms because of the in-service workshops we have offered."	**AA**	**A**	**U**	**D**	**DD**	**NA**
23. "We have evidence that the career awareness materials we distributed have influenced the career choices of potential students."	**AA**	**A**	**U**	**D**	**DD**	**NA**

24. "The professional development program(s) we developed is/are used at other sites."	AA A U D DD NA
25. "Our ATE grant has helped us produce more science and engineering technicians than we would have done without the grant."	AA A U D DD NA
26. "The grant provided the catalyst to establish and/or strengthen collaborations with business and industry partners."	AA A U D DD NA
27. "The new curriculum was created through the regular institutional approval process so all the new classes are in the regular college catalogue and are part of approved college degrees."	AA A U D DD NA
28. "The grant has permitted us to develop educational materials that otherwise would not be available."	AA A U D DD NA
29. "Very few of the graduates prepared under our grant are employed as technicians."	AA A U D DD NA
30. "The materials we have developed are seldom used by other colleges for technician preparation programs."	AA A U D DD NA
31. "The teaching methods adapted by faculty as part of our ATE project will continue to be used after the grant ends."	AA A U D DD NA
32. "The grant has enhanced our reputation as a regional leader in advanced technology education."	AA A U D DD NA

Part II. Scoring the Survey and Interpreting Results.

Scoring the Survey

The first step after the scale is administered is to compute a raw score. The response options to the Likert-type statements are Strongly Agree, Agree, Uncertain, Disagree, Strongly Disagree, and Not Applicable.

Values of 5 to 0 are assigned to the responses of positively stated items, for example, an Agree response to the statement, "We have data-based evidence that our program improved the workforce skills of our graduates" would be assigned 4 points. The scoring scheme for these items is:

Strongly Agree	5
Agree	4
Uncertain	3
Disagree	2
Strongly Disagree	1
Not Applicable	0

Items numbered 1,3,4,6,7,8,9,10,11,13,14,15,16,17,18,20,22,23,24,25,26,27,28,31,32 are scored using this point assignment.

Items that are negatively stated would be coded 0 to 5. An example is the statement, "Most of the programs and activities started during our ATE grant will come to an end when our NSF funding stops." An agree response would be assigned 2 points. In this case, points are awarded using the following.

Strongly Agree	1
Agree	2
Uncertain	3
Disagree	4
Strongly Disagree	5
Not Applicable	0

The following items are scored using this scheme, 2,5,12,19,21,29,30.

Once the points are assigned to each item, there are two ways that a scale score can be computed. The first method, called the Total Evaluation Score, is to sum the raw scores for the 32 items. Items marked Not Applicable (NA) are coded as zero.[2] The highest possible score is 160 (5 x 32). The lowest score depends on the number of NA responses. If none, the lowest possible score would be 32.

---

[2] A discussion of how to score a survey when NA is a response option is found in (Welch W. W., 2012a).

Another way to score the survey is to compute a Mean Evaluation Score. This is the average response to those items deemed applicable to a grant. You exclude items that are marked NA. For example, suppose a respondent marked NA to 5 statements. The respondent's score would be the average response to the 27 (32 – 5) items that were applicable.

You sum the raw scores for those items that are applicable and then divide that total by the number of applicable items. Using the example above, if the sum across the 27 items was 94, the ATE Mean evaluation score would be 3.48 (94 divided by 27). The highest possible score is 5.00 and the lowest would be 1.0.

In general, one would compute a total score to determine how well a project is doing in addressing all elements of the content domain. However, if a project has a narrow focus, for example, only developing materials, it would be more appropriate to compute a mean evaluation score. This would give you some idea of how well your project is doing based only on those things you promised to do.

Interpreting the Results

One way to interpret the scores is to compare your results with others who have taken the same survey. In 2010, the survey was administered to 212 ATE team leaders.[3] Their results can help you interpret your score. Table 1 provides the percentile ranks for the Total Evaluation scores and Table 2 provides percentile ranks for Mean Evaluation scores.

---

[3] The survey was sent to ATE PIs and they are the ones that usually responded. In some cases, a project manager completed the survey.

Table 1. ATE Total Evaluation Score Percentile Ranks

Total Score	Percentile Rank	Total Score	Percentile Rank
158	100	118	46
156	99	116	40
154	97	114	38
152	96	112	37
150	93	110	35
148	92	108	34
146	89	106	30
144	84	104	26
142	83	102	23
140	80	100	22
138	78	98	20
136	76	96	19
134	73	94	16
132	70	92	14
130	68	90	12
128	67	88	11
126	63	86	10
124	58	84	9
122	53	≤ 82	≤ 8
120	48		

These tables enable you to compare your ATE evaluation scale score with the performance of other ATE grantees. For example, an awardee with a Total Evaluation score of 124 can see that 58 percent of other respondents scored lower.

Another way to use the survey would be to obtain ratings of your project from the faculty members involved in the project. You ask them to respond to the statements. You then compute a mean response for the group. In addition, you might also ask your business and industry colleagues to take the survey. You could then compute the mean of that group and compare those results with the responses of your faculty group.

An external evaluator may use the ATE Evaluation Survey to obtain ratings from a broader group than just the PI. He/she would administer it to a random sample of students, business/industry representatives, faculty members, and members of the National Visiting Committee. In this manner, a broader view of the success or limitations of the project could be obtained.

Table 2. ATE Mean Evaluation Score Percentile Ranks

Mean Score	Percentile Rank		Mean Score	Percentile Rank
5.00	100		4.15	56
4.95	99		4.10	52
4.90	99		4.05	49
4.85	97		4.00	46
4.80	96		3.95	41
4.75	95		3.90	36
4.70	93		3.85	27
4.65	90		3.80	24
4.60	88		3.75	20
4.55	84		3.70	18
4.50	82		3.65	13
4.45	78		3.60	12
4.40	76		3.55	10
4.35	75		3.50	9
4.30	70		3.45	7
4.25	67		3.40	7
4.20	60		≤ 3.3	≤ 5

Part III. Survey Development and Psychometric Properties

The survey was developed following well accepted guidelines recommended in the literature. See for example Borg & Gall (1983) or Worthen & Sanders (1987). Its purpose is to help grantees and others to gather evaluation information about ATE projects. The survey is based on statements made by PIs and other ATE stakeholders about the merit or worth of their ATE work. These statements sample the domain of content generated from the goals and activities of the program. There are both outcome (product) and activity (process) goals. Product goals are the expected outcomes of the program while process goals are the activities to be implemented by the program. The product goals are outlined in the congressional act authorizing the program (U. S. Senate, 1992). The process or activity goals are described in several National Science Foundation (NSF) program solicitations, for example, the one published by NSF in 2007 (National Science Foundation, 2007).

The domain of content for the survey is outlined below. It was developed using ATE grantees and an expert advisory committee. (See Welch, 2011a for details on this process.) It is limited to those projects and centers that directly implement program activities. The norm group does not include grants to conduct the annual PI meetings, nor the several research and evaluation projects supported by the Foundation.

An effective measuring instrument must be based on the domain of content of the construct one is trying to measure. The domain of content for this ATE Evaluation Survey consists of the

expected outcomes of the ATE program as outlined in the Congressional act that authorized the program. These are shown in Section I of the table below. These are sometimes called product goals.

Process goals are activities that ATE grantees are expected to implement. To evaluate these goals, one needs to ask three key questions. Were the activities actually implemented? Did someone judge these activities as meritorious? What changes occurred because of implementing these activities? These elements are found in Section II of Table 3.

Table 3. Elements of the domain of content for the ATE program

I. Products (outcome goals)
    A. More technicians will be produced
    B. The skills of technicians will be improved.
    C. There will be better STEM preparation of technicians and technician educators

II. Process (activities)
    A. The promised activities were implemented. The activities were to:
        1. Establish partnerships between academic institutions and employers.
        2. Develop curriculum and materials.
        3. Provide professional development for college faculty and secondary school teachers.
        4. Establish career pathways from secondary schools to two-year colleges and from two-year colleges to four-year institutions.

    B. The implemented activities were viewed as meritorious by those knowledgeable about the program

    C. As a result of these activities, changes occurred in the following elements
        1. People
            a. Faculty
            b. Students
            c. Administrators
            d. ATE PIs/Staff
        2. Program
            a. Curriculum
            b. Instruction
            c. Educational materials
        3. Organizations
            a. Colleges
            b. Schools
            c. Business/Industry
            d. Communities

The survey consists of a set of statements about projects and centers made by a group of 45 ATE PIs and other stakeholders.[4] More than 90 statements were obtained. After review by the author, three ATE grantees, and three test development experts, 53 were selected for an initial impact and sustainability survey (Welch W. W., 2011). This survey was administered to 212 ATE project leaders.

Many of the survey comments were value-based, that is, they addressed issues of merit or worth. For example, one statement was "Our faculty has improved their teaching style because of their involvement in our ATE grant." This statement implies change has occurred and the change can be viewed as positive or negative depending on whether the respondents agree or disagree with the item.

The 53 items were examined to determine which were value-based. I looked for statements that included a verb indicating a change in the system and this change could be judged as positive or negative. For example, "student interest increased," "collaborations with industry were established," or, "persistence of ATE students was lower."

Thirty-nine (39) statements were identified for further consideration. A Likert-type response option ranging from Strongly Agree to Strongly Disagree was attached to each statement. In addition, there was an option to respond, Not Applicable (NA). This occurred when the respondent thought the item did not apply to their specific project or center.

Two additional steps reduced the final set of survey items to 32. In a few cases, the number of NA responses was quite large. I decided to exclude items if 25% or more respondents chose the NA option. It did not seem reasonable to include statements it they did not apply to a large proportion of the respondents. Four items exceeded this level and were excluded from the survey leaving 35 statements.

I then conducted an item analysis of these items by computing the item to total score correlation. If the items are measures of the same construct, they will be correlated with one another and add to the reliability of the scale. If the correlation is low, then including the item reduces the total scale reliability.

I selected an $r \leq .30$ as my cut-off point. There is no standard value for this but if excluding an item increases the total score reliability, then the item could be deleted. My experience has shown this occurs when the item-total score correlation is below .30. Three items did not meet that standard leaving 32 items for the final version.

---

[4] Approximately 35 ATE PIs responded to our request for statements. There were about 10 others knowledgeable about the program who made statements, for example, ATE consultants and NSF program officers.

Psychometric Properties of the Scale

Content Validity

An assessment device, such as a survey, has content validity when the items are measures of the elements of the domain of content. The 32 items show a strong correspondence to the content domain as shown in Table 3. The domain list in that table includes 11 main sections. Every one of these sections is represented by at least one item on the scale. The average number of items per sub-section was 2.9 and the median was 2.0.

Recall that the domain was defined by the Congressional mandate and NSF's program solicitation. The statements on this scale are a sample of a universe of items that could be used. Other sets of items might be developed; however, this sample represents the dimensions deemed important by PIs and others knowledgeable about ATE. This and the fact that PIs and others familiar with ATE wrote these items support the claim of content validity.

Reliability

I administered the evaluation scale to a population of 261 ATE team leaders, either PIs or program managers and received 212 responses, a response rate of 81.2%. I used their responses to calculate reliability and percentile ranks.

Scale reliability is the consistency of a measure. A measure has high reliability if it produces similar results under the similar conditions. I computed the internal consistency of the evaluation scale using Cronbach's alpha and obtained a value of 0.94, well above generally accepted standards. (Note: This value is based on 89 respondents who did not mark Not Applicable as an option. SPSS calls this a list-wise deletion.)

Distribution of Scores

Total Evaluation Scores

I calculated the distribution of the scores of the 212 ATE respondents to the survey to determine if the scale could discriminate among responders. A scale where everyone obtained the same result would not be too useful. The distribution is shown in Figure 1. Using an SPSS option, I superimposed a normal distribution curve on the distribution. As that figure shows, the distribution is comparable to a normal distribution but is negatively skewed. The scores indicate that the responders to the evaluation survey were quite positive in their assessment of their own projects and centers.
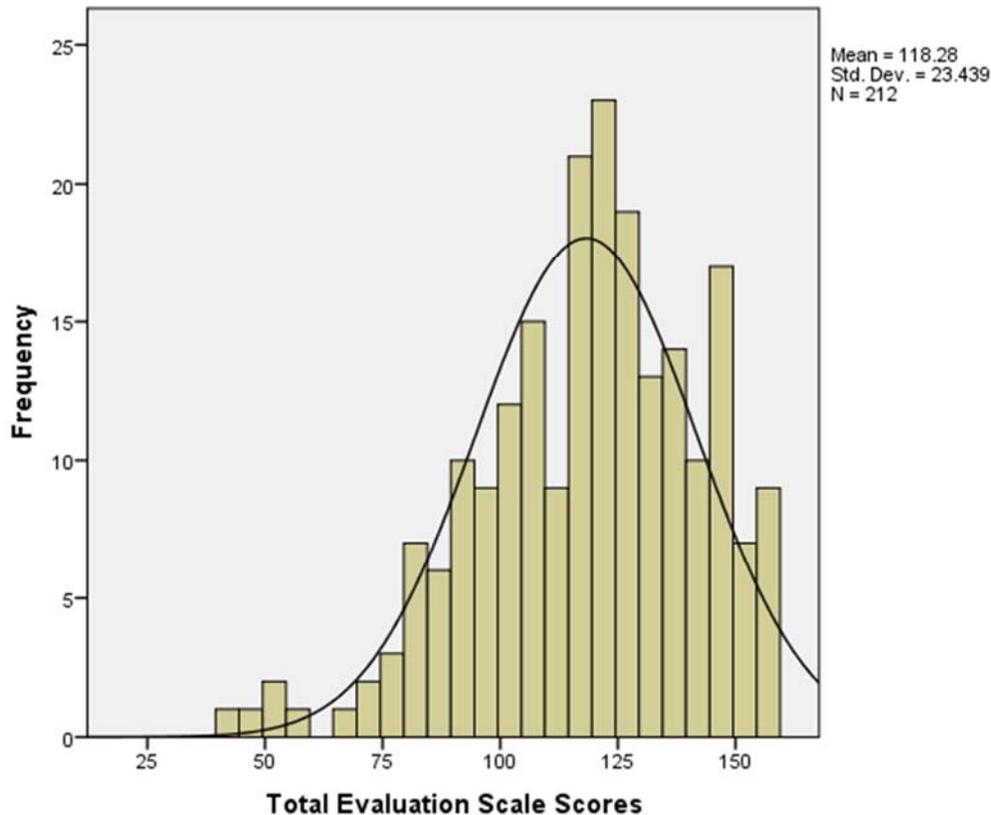
Figure 1. Frequency distribution of total ATE evaluation scale scores

We see there was variability among the respondents' total scores. The scale did differentiate among the ATE grants. The scores ranged from 42 to 158. The mean was 118.3 with a standard deviation of 23.4. The median was 121.

The scores are negatively skewed. They are shifted toward the higher end of the scale. The mean score of 118 was about one standard deviation above what one would expect if the evaluation responses were normally distributed. In this case, the mean would be 96, half way between the lowest score possible, 32, and the highest score, 160. The scores indicate that the responders to the evaluation survey were quite positive in their assessment of their own projects and centers.

Mean Evaluation Scores

I also calculated the mean scale responses of the 212 team leaders and computed the frequency distribution of these scores. This is shown in Figure 2.
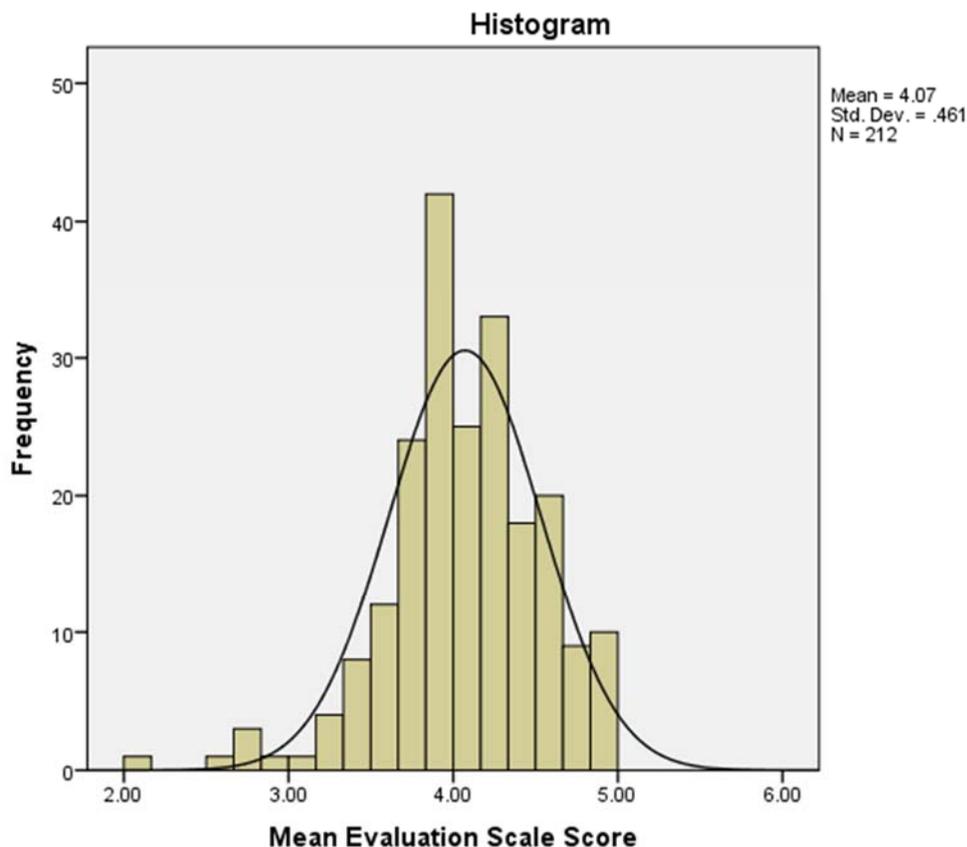

Figure 2. Frequency distribution of total ATE evaluation scale scores

Here, again, we see that the scale discriminates among our respondents. As was true for the total score scale, the average scores were negatively skewed. In addition, a high positive response occurred. The mean was 4.07 and the standard deviation was .46.

Concluding Remarks

This document presents a new instrument that PIs, ATE team members, NSF, and evaluators can use to gather evaluation information about ATE projects and centers. It is a 32-item Likert-type survey that uses evaluative statements made by ATE stakeholders. Respondents are asked to express their degree of agreement or disagreement with these statements. Two ways are described to score the survey and percentile ranks are presented so grantees can compare their scores with those of other ATE grantees. The psychometric properties of the scale are described which support the validity and reliability claims for the scale.

The survey can be used in several ways. It could be administered at the end of project to provide summative evaluation information to a National Visiting Committee or to NSF. It could be used to compare the responses of various groups involved in the project, that is, faculty, business/industry representatives, or students. One could also give it annually to check to see if response patterns are changing over the course of the project work. An external evaluator could use the instrument as part of the evaluation activities to gather information about the success or limitations of the project.

I think it would be important for PIs to identify which of the statements are most important to them in terms of their stated objectives and to track progress over time (e.g., annually or semi-annually) to ensure that the project is headed toward success in ways that matter most to them. If survey users are interested in obtaining results for the individual items, they can do so by contacting the author at wwelch@umn.edu.

References

Borg, W. R., & Gall, M. D. (1983). *Educational Research: An Introduction.* New York and London: Longman.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (2nd. ed.).* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Costello, A. B., & Osborne, J. W. (2005). *Best practices in exploratory factor analysis: Four recommendations for getting the most out of your analysis.* Retrieved from Practical Assessment, Research & Evaluation: http://pareonline.net/getvn.asp?v=10&n=7

National Science Foundation. (2007). *Advanced Technological Education (ATE) Program Solicitation NSF 07-530.* Retrieved from www.nsf.gov: http://www.nsf.gov/pubs/2007/nsf07530/nsf07530.htm

National Science Foundation. (2011). *Advanced Technological Education (ATE).* Retrieved from Program Soliciation NSF 11- 692: www.nsf.gov/pubs/2011/nsf11692/nsf11692.htm

Nunnally, J. (1978). *Psychometric Theory (2nd ed.).* New York: McGraw-Hill Book Company.

U. S. Senate. (1992). *Library of Congress Thomas.* Retrieved from S.1146 : Scientific and Advanced-Technology Act of 1992: http://thomas.loc.gov/cgi-bin/bdquery/z?d102:SN01146:|TOM:/bss/d102query.html|

Welch, W. W. (2011). *A study of the impact of the advanced technological education program.* Retrieved from DECA, University of Colorado: http://www.colorado.edu/ibs/decaproject/pubs/index.html

Welch, W. W. (2011a). *Research Report 2: The Impact of the Advanced Technological Education Program.* Retrieved from Evalu-ATE: http://evalu-ate.org/resources/sustainability_of_ate/

Welch, W. W. (2012a, June). *Measuring the Sustainability of the Advanced Technological Education (ATE) Program.* Retrieved from Evalu-Ate: http://evalu-ate.org/resources/sustainability_of_ate/

Welch, W. W. (2013, January). *Developing and Analyzing a Scale to Measure the Impact of the Advanced Technological Education Program.*

Worthen, B. R., & Sanders, J. R. (1987). *Educational Evaluation: Alternative Approaches and Practical Guidelines.* New York: Longman, Inc.