

January 2014

A Replication Study of the Advanced Technological Education (ATE) Sustainability Survey

Wayne W. Welch  
University of Minnesota  
January 2014

Author Notes

Wayne W. Welch, Department of Educational Psychology, University of Minnesota (ret.)

Professor Welch is also a consultant for Rainbow Research, Inc., Minneapolis, MN

This material is based upon work supported by the National Science Foundation under Grant No. 1132099. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.”

Correspondence concerning this report should be addressed to Wayne Welch, 621 W Lake Street # 300, Minneapolis, MN 55408. Contact: [wwelch@umn.edu](mailto:wwelch@umn.edu)

Acknowledgements

This study would not have been possible without the helpful cooperation of Lori Wingate, Principal Investigator of EvaluATE and staff member, Corey Smith. This project is housed at the Evaluation Center, Western Michigan University. Thank you very much.

I also wish to thank my review panel members for their help in preparing this report; Elaine Craft, Director of the South Carolina ATE Center of Excellence, Frances Lawrenz of the University of Minnesota, and Elizabeth Teles of Teles Consulting and former co-lead for NSF's ATE program. They were most helpful in reviewing a draft of the report and providing insightful suggestions. Thanks to my wife, Jerry, for reviewing earlier versions of this report.

## A Replication Study of the Advanced Technological Education (ATE) Sustainability Survey

### Abstract

A study of the sustainability of the federally funded Advanced Technological Education (ATE) program was carried out in 2010 using a 23-item Likert-type survey. The National Science Foundation (NSF) administers this program. Based on the responses of 131 ATE team leaders, an 81.2% response rate, I concluded that ATE grantees were successful in sustaining most of the elements of their projects and center. For example, program changes were institutionalized, new collaborations are being continued, and faculty had improved their teaching methods.

A year later, several of the original sustainability items were included on an annual survey distributed by Western Michigan University. This survey gathers information about the activities and accomplishments of ATE grantees. Space and time limitations limited the number of sustainability items they could include to 7, about 30% of the original survey. The purpose of this survey was to determine if the original findings regarding sustainability were replicated when sent to a different population of ATE grantees.

The populations in the two studies were limited to active grants that were at least one year old. This reduced the study populations to 131 original sites and 144 in the replication study.

I compared the replication findings with those from the original study using a total scale score for the seven items and on an item-by-item basis. There were no significant differences between the original and replicated results. The study findings were replicated. This result enhances the confidence one could place in the original results and suggests the sustainability success could be generalized to the larger group of ATE projects and centers.

The consistent findings from the two surveys helps refute the possible effects of yea saying on the annual survey administered under the auspices of NSF. If this were occurring, one would expect to find higher scores on a survey endorsed by the Foundation threatening the validity of the survey. However, similar outcomes were obtained on both surveys which provides confidence in the validity of the responses.

The ATE program, funded by the National Science Foundation (NSF), is designed to improve the education of technicians in high-technology fields such as biotechnology, advanced manufacturing, information technology, and environmental and energy technologies. The program makes grants to support projects and centers to achieve this goal. It was mandated by Congress in 1992; its primary focus is on two-year colleges. Approximately three-fourths of the awards have been made to these institutions.

The Foundation expects grantees to plan to sustain successful outcomes of their work after NSF funding stops. I examined this issue in a report entitled “The Sustainability of the Advanced Technological Education Program” (Welch, 2011a). I concluded the ATE program had been quite successful in continuing the products and activities initiated during the ATE grant. I wrote:

Overall, the findings suggest many of the elements implemented by ATE grantees have been sustained. Program changes have been institutionalized and the collaborations the institutions have formed with industry partners have persisted. New professional development programs are in place and faculty have improved their teaching methods and formed new collegial relationships.

About two-thirds of the respondents believe other colleges have used the materials they produced; however, few have obtained revenue by selling these materials. This finding is consistent with their responses to other statements about revenue.

The grantees report that students are being employed as technicians and serving as ambassadors for technology education. Prepared students are an outcome of the grants that continues to manifest itself...at least while the students are employed as technicians. In addition, improved student recruitment methods are in place and student internships will be continued.

Certain elements of the institutions have made changes that respondents think will continue. They report a three-fourths majority of administrators are encouraging other faculty to seek grants. Reputations have been enhanced, review committees have been improved, and the colleges expect that their ATE experiences will have a long-term impact on community colleges (p. 13).

These findings were based on a survey administered to ATE PIs and program managers in 2010. The research population consisted of all active ATE grantees that were more than one year old. The potential pool of sites was 160. Responses were received from 131 sites, an 81.9% response rate.

The findings provide evidence for the success of the ATE program and were well received by the National Science Foundation and the two-year college community. This was the first time that the Foundation has made a large number of grants to two-year colleges and there was considerable opposition from the Foundation until Congress mandated the program in 1993. An evaluation finding that the work of the project and centers is being sustained is seen as strong justification for program continuation.

During the early years, the ATE program funded projects and centers. In 2007, the Foundation added a third track to the program called Targeted Research on Technician Education. Its stated purpose was

“Employing rigorous standards of research and scholarship, project proposals should pose a research question or outline a topic of broad interest, survey previous research and scholarship on the issue, conduct original research and compile data, prepare cogent analyses, present conclusions, and describe how the results can inform practices in technician education programs.” (Section II C, paragraph 10, National Science Foundation, 2007).

Later program solicitations for the Targeted Research Track called for investigators to advance the knowledge base needed to improve technician education programs (National Science Foundation, 2011). One way to do this is to improve the effectiveness of research techniques. For example, one study designed to improve research methodology explained to researchers and practitioners how to handle nonresponse bias in large-scale surveys studies (Welch & Barlau, 2013 b).

The sustainability study mentioned above provides another opportunity to improve research methodology by employing an oft-recommended but seldom implemented procedure called replication. Replication is the *sine qua non* of scientific research and every educational research textbook calls for it. It is the process of repeating a research study with a different group of subjects using the same or similar methods. According to Borg and Gall (1983), “Results of a study are more ‘significant’ in the sense of inspiring confidence that they represent true differences or relationships...if a new study yields similar results” (p. 383). Unfortunately, educational researchers seldom do them.

A survey implemented by the Evaluation Center at Western Michigan University (UWM) provided an opportunity to replicate part of the study described above. This survey, administered online, gathers data from ATE grant recipients about the grantees and their activities, accomplishments, and impacts for each calendar year. It has been administered since 2000 with response rates above 90%. The project agreed to include seven of my sustainability items on their 2011 annual survey. This survey asks questions about products and activities carried out in 2010. I compared the findings of their survey with the findings I found for 2009.

The specific research question for this study was, “Were the sustainability findings obtained in the replicated study at WMU consistent with the results obtained in the original targeted research study?” If they are, it should give ATE stakeholders more confidence in the validity of my results and further justify the generalization of the findings to the larger ATE program.

### **Method**

The original study population consisted of 160 active projects and centers that were at least one year old at the time of the survey. A 23-item sustainability survey was mailed to this group in the spring of 2010. After three follow-ups, two mailed and one online, 131 responses were received for a response rate of 81.9%. See (Welch, 2011a) for details on the development and implementation of the survey.

The WMU survey included seven of the 23 sustainability statements that were on the original survey. They were selected to represent the dimensions of sustainability, for example, collaborations, institutional change, and professional development. This survey was administered online in the spring of 2011 to all grants that were active during the 2010 calendar year. Grants that did not respond were

sent several reminder emails from Western Michigan and one from NSF.<sup>1</sup> Eventually, 220 sites returned the survey, a response rate of 94.0%. Appendix A contains the items and the format used by WMU. Further information on the survey can be found in Wingate, Westine, & Gullickson (2011).

Among the 220 sites that responded to the WMU survey were five targeted research projects and one site that did not answer any of the questions. It also included 70 grants that were less than one year old. Because I was studying established projects in this comparison, I excluded these sites leaving 144 sites for the replication study.

The survey contained both positive and negative worded statements. Sustainability success is indicated when a respondent agrees or strongly agrees with a positively worded item or disagrees or strongly disagrees with a negatively worded statement. An example of the former is “changes made in our program will continue after current funding ends.” Agreeing with this statement provides evidence that the grant work is being continued. An example of a negatively worded statement is “Materials that we have developed are seldom used by other colleges.” In these cases, sustainability success is indicated when a respondent either disagrees or strongly disagrees with the item.

I used two measures of sustainability. The first was a scale score and the second focused on individual items. The scale score is a summary statistic of sustainability success while the second compares responses on individual items.

A sustainability subscale score was created by summing the responses to the seven items. Five points were assigned to a strongly supportive response, 4 points for a supportive response, 3 for uncertain, 2 for a denial of sustainability, and 1 for a strong denial. A Not Applicable response was given 0 points.

The possible scores ranged from 7 to 35 and this turned out to be the actual range. One person received a score of 7.0 and five respondents reported scores of 35. The mean for all 275 respondents was 25.52 with a standard deviation of 5.84.

I computed the reliability of the subscale using Cronbach’s alpha and obtained a value of 0.66. That value is too low to make decisions about individual scores but should be adequate for making decisions about groups.

I also computed a sustainability score for each item by computing the percent of respondents agreeing or strongly agreeing to the positive items and those disagreeing or strongly disagreeing with the negatively worded items. I then compared the item scores for the replication group with those of the original group.

## **Findings**

The first null hypothesis tested was, “There were no statistically significant differences between the sustainability scale scores obtained during the original study and those obtained during the replication study.” I used a t-test to test this hypothesis. The findings are shown in Table 1.

---

<sup>1</sup> WMU did not record how many times the sites were contacted, but I suspect that in some cases it was 10 or more.

Table 1

*Comparison of Mean Sustainability Scores: Original vs Replication*

	Group	N	Mean	Standard Deviation	Difference t-test
Sustainability Scores	Original	131	25.26	5.99	-0.50
	Replication	144	25.76	5.70	t = .71 (p = 0.48)

These findings suggest strongly that the original results were replicated by the second study. The means differed by only .50 points; the difference was not significantly different. Another way to express the difference is to compute the effect size between the groups. A value of 0.09 was obtained using Cohen's  $d$  (Cohen, 1988). This is well below the generally accepted standards suggested by Cohen. He considered .20 to be low. Here, again, the difference is very small which means the original findings were replicated.

The second method I used compared the original and replication groups on each of the seven items. I used a Chi-squared ( $\chi^2$ ) test to determine whether there were significant ( $p \leq .05$ ) differences between the groups. I also computed the effect size for each comparison. These results are shown in Table 2.

Table 2

*Sustainability Item Scores: Original vs Replication Groups<sup>a</sup>*

Statement	Original Group (%)	Replication Group (%)	Difference (%)	Chi Square <sup>b</sup>	Effect Size
Changes made in our technological education program will keep going after our current grant ends	91.1 (123) <sup>c</sup>	93.9 (132)	- 2.8	.77	.11
It is doubtful that the relationships we established with our various partners will continue after our ATE grant has ended	90.7 (129)	88.1 (143)	+ 2.6	.48	.08
The materials we have developed are seldom used by other colleges for technician preparation programs	74.5 (110)	68.2 (129)	+ 6.3	1.16	.14
The teaching methods adapted by faculty as part of our ATE project will continue to be used after the grant ends	91.9 (124)	96.2 (131)	- 4.3	2.08	.18
Very few of the graduates prepared under our grant are employed as technicians	76.0 (96)	78.4 (107)	- 2.4	.16	.06
Our ATE grant experience has caused our administration to encourage other faculty to seek external funding to address workforce needs	82.6 (121)	78.4 (134)	+ 4.2	.74	.11
We will be able to keep our project/center going by obtaining revenue income for specific education services	38.4 (112)	44.0 (128)	- 5.6	.77	.11

a. Percent agreeing or strongly agreeing with positively worded statements or disagreeing or strongly disagreeing with negatively worded statements

b. Chi square values significant at the .05 level are shown with an \*

c. Number of respondents in each group

Item differences ranged from a – 5.6 to + 6.3 percentage points. In some cases, the original group was slightly higher; in others, the replication group was higher. None of the p values for the  $\chi^2$  tests reached the .05 level of significance.

I computed the effect size for each comparison using an effect size calculator developed by Wilson (2010). The effect sizes are presented in the last column. They are low ranging from .06 to .18. All are below the .20 standard suggested by Cohen. Thus, we have support for accepting the null hypotheses of no difference between the groups. The findings from the original study can be generalized with confidence to the larger population of ATE projects and centers.

The preceding discussion is based on the assumption that the two groups are independent of each other. That means a team member who responded to the original survey may be included in the replication study. However, the two surveys were administered one year apart and may have been answered by two different people.<sup>2</sup> Because of the possibility that there may be some dependence for those places that were in both groups, I decided to repeat the analysis only using the replication sites that were not part of the original study. There were 66 projects or centers that were in both groups. Excluding these from the replication group reduced the number of replication sites from 144 to 78.

I repeated the analysis described above for these two groups. First, I compared the means using a t-test and then compared the item responses using the crosstabs procedure.

The mean comparison is shown in Table 3.

Table 3

*Sustainability Scale Scores: Original vs Non-Duplicated Replication*

	Group	N	Mean	Standard Deviation	Difference and t-test
Sustainability Scores	Original	131	25.26	5.99	-0.14
	Non-duplicated Replication	78	25.40	6.07	t = .16 (p = 0.87)

The results are similar to the earlier findings suggesting that I can accept the null hypotheses of no differences between the groups when I exclude the duplicated projects and centers. This result is consistent with the earlier finding and supports the idea of generalizing the results to most of the ATE grants.

The findings for the Chi Square analysis are shown in Table 4. I set the probability level set at .05 and reported the effect size of the comparisons.

---

<sup>2</sup> I could not check this because I did not have the names of those that responded to the WMU survey. Also, that survey is set up so that different people may respond to different parts.



Table 4

*Sustainability Item Scores: Original vs Non-Duplicate Replication Groups<sup>a</sup>*

Statement	Original Group (%)	Non-Dup Replication Group (%)	Difference (%)	Chi Square <sup>b</sup>	Effect Size
Changes made in our technological education program will keep going after our current grant ends	91.1 (123) <sup>c</sup>	93.2 (73)	- 2.1	.27	.07
It is doubtful that the relationships we established with our various partners will continue after our ATE grant has ended	90.7 (129)	90.9 (77)	- .02	.00	.01
The materials we have developed are seldom used by other colleges for technician preparation programs	74.5 (110)	69.1 (68)	+ 5.4	.62	.12
The teaching methods adapted by faculty as part of our ATE project will continue to be used after the grant ends	91.9 (124)	97.3 (73)	- 5.4	2.28	.22
Very few of the graduates prepared under our grant are employed as technicians	76.0 (96)	78.7 (61)	-2.7	.15	.06
Our ATE grant experience has caused our administration to encourage other faculty to seek external funding to address workforce needs	82.6 (121)	76.7 (73)	+ 5.9	1.02	.14
We will be able to keep our project/center going by obtaining revenue income for specific education services	38.4 (112)	43.9 (66)	- 5.5	.53	.11

a. Percent agreeing or strongly agreeing with positively worded statements or disagreeing or strongly disagreeing with negatively worded statements

b. Chi square values significant at the .05 level are shown with an \*

c. Number of respondents in each group

None of the Chi-square comparisons reached the .05 level of significance. The comparison effect sizes for the two comparisons (See Table 2 and Table 4) are similar and in the same direction. One statement about the teaching methodologies had an effect size of .22. A similar result was found for the item when all of the replicated sizes were included (E.S. = .18). Neither of the chi-square values for the two comparisons were significantly different ( $p \leq .05$ ). Even though the effect sizes for this statement were somewhat higher than the other comparisons, they are still considered small.

In summary, the replicated findings are consistent with those of the original study. This means the null hypothesis could be accepted and the sustainability findings from the original study were replicated by

the WMU results. They provide support for the notion of generalizing the original study findings to the larger ATE program.

### **Concluding Remarks**

The original study included 23 items and the replication study only had 7 items. They were a representative sample of the domain of content, not a random sample. It would have been better to replicate all the items but time and space limitations made this impossible.

I do have one piece of validity evidence for the 7-item scale that addresses this problem. It differentiated between projects and centers in the same way that the 23-item scale did. The original study found the following. “The centers ( $M = 84.83$ ,  $SD = 11.67$ ) reported significantly higher levels of sustainability than did the projects ( $M = 71.89$ ,  $SD = 17.37$ ),  $t(91) = 5.78$ ,  $p = .00$ , two-tailed,  $d = .79$ .” (Welch, Measuring the sustainability of the advanced technological education (ATE) program, 2012, p. 13)

When I repeated this comparison using the 7-item sub-scale, I found that centers scored significantly higher ( $M = 28.59$ ,  $SD = 4.70$ ) than projects ( $M = 25.02$ ,  $SD = 5.75$ ),  $t = 3.09$ ,  $p = .00$ , two-tailed. The effect size was .64. The results of the 7-item subscale were consistent with the results found for the 23-item scale. The subscale behaved in the same way as did the longer sustainability scale.

Although it was unintended, this study also provided some validity evidence for the annual survey administered by Western Michigan University. A concern expressed by an early advisory committee was the impact of possible yea-saying. This might occur when a survey has close ties with the National Science Foundation. The Foundation requires projects and centers to respond to the survey every year and program officers are enlisted to contact the site if the survey is not returned. This might encourage team leaders to respond in a manner that portrays their grant in the best way possible.

One goal of the original study was to reduce yea-saying tendencies. This was accomplished by asking respondents to agree or disagree with statements made by their peers rather than to those made by NSF or an organization that is viewed as being closely aligned with the Foundation. Another tactic was to have a researcher implement the study who was largely unknown to ATE PIs. In addition, the survey was mailed out from a neutral site, the University of Colorado in Denver.

Several of the items were stated negatively which forces respondents to think more about their replies instead of answering several questions that all portrayed their site in a positive manner. There were also items that respondents gave low positive ratings suggesting that the PIs were willing to state they did not do certain things. Some examples were the items about obtaining revenue. Their average supportive response was less than 40%.

The responses to the items administered by a group with strong ties to NSF were quite similar to the original study, one that was designed to reduce response bias. This provides some evidence that the WMU survey does not suffer from yea-saying tendencies. This helps support the validity claims of the annual survey.

Finally, this study adds to the limited literature on replicating educational research findings. A method for doing this was described and the results were positive. I presented evidence that supports generalizing the sustainability findings to the larger ATE program. I hope it also provides encouragement to other educational researchers to consider replicating their research findings.

#### References

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd. ed.)*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- National Science Foundation. (2007). *Advanced Technological Education (ATE)*. Retrieved from Program Solicitation NSF 07-530: <http://www.nsf.gov/pubs/2007/nsf07530/nsf07530.htm>
- National Science Foundation. (2011). *Advanced Technological Education (ATE)*. Retrieved from Program Solicitation NSF 11- 692: [www.nsf.gov/pubs/2011/nsf11692/nsf11692.htm](http://www.nsf.gov/pubs/2011/nsf11692/nsf11692.htm)
- Welch, W. W. (2012). *Measuring the sustainability of the advanced technological education (ATE) program*. Retrieved from Evalu-ATE: <http://www.colorado.edu/ibs/decaproject/pubs/>
- Welch, W. W., & Barlau, A. B. (2013 b). *Addressing Survey Nonresponse Issues: Implications for ATE Principal Investigators, Evaluators, and Researchers*. Retrieved from DECA, University of Colorado: <http://www.colorado.edu/ibs/decaproject/pubs/index.html>
- Wilson, D. B. (2010, August 11). *Meta-analysis macros for SAS, SPSS, and Stata*. Retrieved from Practical meta-analysis effect size calculator: <http://mason.gmu.edu/~dwilsonb/ma.html>
- Wingate, L., Westine, C., & Gullickson, A. (2011, July). *ERIC*. Retrieved from Advanced Technological Education Survey 2011 Fact Sheet: <http://files.eric.ed.gov/fulltext/ED541557.pdf>

## Appendix A

This is a copy of the page from the WMU survey, which contained the sustainability items. The small circles are bitmap images. The respondents can use their cursor to mark their responses.

**The following statements have been made by other ATE PIs about the sustainability of their project or center or aspects of it. Please indicate the extent to which you Agree or Disagree with each one.**

	Statement	Strongly Agree	Agree	Uncertain	Disagree	Strongly Disagree	Not Applicable
a.	Changes made in our technological education program will keep going after our current grant ends.	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b.	It is doubtful that the relationships we have established with our various partners will continue after our ATE grant has ended.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
c.	The materials we have developed are seldom used by other colleges for technician preparation programs.	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d.	Changes made in our technological education program will keep going after our current grant ends.	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e.	The teaching methods adapted by faculty as part of our ATE project will continue to be used after the grant ends.	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f.	Very few of the graduates prepared under our grant are employed as technicians.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
g.	Our ATE grant experience has caused our administration to encourage other faculty to seek external funding to address workforce needs.	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
h.	We will be able to keep our project/center going by obtaining revenue income for specific education services.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
i.	Student interest in technology has increased because of our ATE grant.	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Note that items a. and d. are identical. I think this was a typo. In addition, WMU added item i. as a general measure of ATE success. This item was not on the original survey.