Executive Summary

Teachers participating in the standards-aligned classroom (SAC) training program were administered a survey designed to measure attitudes related to the implementation of standards-aligned instruction and assessment in their classrooms. Posttest scores showed that at the end of the SAC program, participating teachers reported being: 1) more familiar with the Illinois learning standards, 2) more likely to understand how to use the standards in planning classroom instruction and assessment, 3) more confident in their ability to use the standards, 4) more aware of the benefits of the standards, 5) more likely to actually use the standards to plan classroom instruction, and 6) more likely to believe a standards-aligned instructional approach has a positive impact on student motivation and participation. All of the above increases were statistically significant ($p < .001$).

Effect sizes were also computed for each increase. The effect sizes generated in the current program ranged from .48 (teachers became more familiar with the standards) to 1.04 (teachers became more confident in using the standards). Given that effect sizes of .20 are generally considered significant for educational training programs, the SAC program would appear to be particularly effective.

Interestingly, the impact of the program was consistent across teacher groups. Newer teachers were just as likely as teachers with 20 years of experience to undergo a positive shift in their attitudes towards standards-aligned instruction. The same was true in terms of grade level. Elementary teachers were just as likely to experience a positive shift in attitudes as were middle or high school teachers.

In addition, teacher portfolios developed by teachers with higher survey scores (indicating greater acceptance and endorsement of a standards-aligned approach) were found to be significantly better than those produced by lower scoring teachers for six out of seven key features associated with quality instruction and assessment material.

Finally, interviews with higher and lower scoring participants indicated that participants with higher survey scores were more likely to be members of teams that: 1) met more frequently, 2) shared leadership duties across team members, 3) volunteered to participate in the program, and 4) included team meetings at which an administrator attended regularly.
# TABLE OF CONTENTS

## INTRODUCTION

Introduction.......................................................................................................................................... 1
Organization of the Report................................................................................................................... 2
Acknowledgements.............................................................................................................................. 2

## SECTION I: Survey Development and Analysis

Developing the Survey......................................................................................................................... 5
Scoring the Survey ............................................................................................................................... 7
Analysis of Survey Data ...................................................................................................................... 9
  Who Responded? ........................................................................................................................... 9
  How Did Survey Scores Change?.................................................................................................. 9
  How Large Were the Survey Score Changes?...............................................................................11
  Did Grade Level Affect the Degree of Change? .......................................................................13
  Did Years of Teaching Experience Affect the Degree of Change? ...........................................15

## SECTION II: Scoring the Portfolios

Portfolio Scoring Summary .................................................................................................................19

## SECTION III: Summary of Interview Responses

Interview Summary: Closed-Format Questions ..................................................................................23
Interview Summary: Open-Ended Questions ......................................................................................25
  Question 16 ....................................................................................................................................25
  Question 23 ....................................................................................................................................27

## SECTION IV: Conclusions and Recommendations

Conclusions..........................................................................................................................................31
Recommendations..............................................................................................................................32
Introduction

The Illinois learning standards are specific statements that define what all Illinois students should know and be able to do as a result of their public schooling. The purpose of the standards is to provide clear targets for student learning. Furthermore, Illinois state law requires that student progress towards the achievement targets defined by the state’s learning standards be assessed on an annual basis.

While a great deal of attention has been directed towards the content and structure of the statewide assessments used to assess these standards, their existence has raised a number of related issues. It is clear, for example, that for students to do well on the assessments, the content of the standards needs to be communicated to administrators and teachers involved with classroom instruction. It is especially important that teachers be given training in how to make standards-based instruction a reality in their classroom. Only when teachers understand the standards and can incorporate standards-based instruction in their daily lessons will students begin to experience the achievement advantage provided by the learning standards.

The Standards-Aligned Classroom (SAC) initiative was designed to give teachers the kind of training mentioned above. It focused on giving teachers a way to incorporate standards-based instruction in both lesson planning and classroom assessment. Teachers who participated in the program were provided an initial two-day workshop that taught them a specific approach to bringing the standards to their students. This training was administered near the start of the 2000-2001 school year. Following the initial training, teachers participated in a local learning team. The learning teams met regularly throughout the remainder of the school year to discuss challenges and to share experiences related to implementing the training model. In addition, a specially trained coach was assigned to each learning team. The role of the coach was to support and direct the efforts of the learning team.

To determine the degree to which participation in the training and learning teams impacted teachers’ attitudes, beliefs, and actions, a formal evaluation of the training program was conducted. The evaluation had three main components. The primary component involved the construction of a survey form designed to assess key teacher attitudes towards the standards and towards standards-aligned instruction. The survey was administered either before or shortly after the initial training workshop and again near
the end of the program. The pretest/posttest design allowed for the direct assessment of attitude change as a result of participation in
the program.

The second component of the program involved the evaluation of teacher-produced materials. Portfolios of teacher-developed
lesson plans and related assessment material were collected from a sample of high, average, and low scoring teachers. Materials in
these portfolios were assessed to determine the degree to which the SAC survey scores were associated with the actual quality of
classroom material prepared by the participants.

The third component focused on interviewing a subset of program participants to determine specific factors that facilitated or
hindered the implementation of the training model. Again, teachers were selected to participate in the interview segment of the
evaluation based on their survey scores. Interviews were conducted with groups of high scoring, average scoring, and low scoring
participants.

Organizational of the Report

The remainder of the report is presented in four sections. The first section provides details on how the survey items were
developed, how the total and scale scores were derived, and outcomes of key analyses related to the survey data. The second section
presents findings related to the collection and scoring of the teacher portfolios. The third section provides details and content analyses
of interviews conducted with samples of high, average, and lower scoring participants, and the final section presents conclusions and
recommendations for improving the SAC program.

Acknowledgements

The Professional Development Design Team would like to express a special measure of appreciation to the participants and
coaches, and to SAC coordinators from the Illinois Regional Offices of Education and Intermediate Service Centers. In addition, we
would like to thank the Illinois State of Board of Education for the guidance and fiscal support that made this initiative possible.
SECTION I:

Survey Development and Analysis
Developing the Survey

After meeting with the directors of the SAC training project, the evaluation team constructed a 42-item questionnaire designed to assess teacher attitudes towards the learning standards and the instructional techniques to be included in the standards-aligned training program. Survey questions covered such topics as how familiar the teachers were with the Illinois learning standards, how useful the teachers found the standards, and how helpful the standards were in guiding instruction and assessment.

Each question on the survey was phrased as a statement to which respondents indicated their level of agreement on a 5-point Likert-type scale. For example, one survey item read, “I regularly use the Illinois learning standards to help plan classroom instruction.” The teachers were asked to indicate their level of agreement with this statement by checking one box from a series of five boxes. The boxes were labeled “Strongly Disagree,” “Disagree,” “Uncertain,” “Agree,” and “Strongly Agree.” A copy of the full survey form is provided in Appendix A.
Scoring the Survey

SAC participants were administered the survey at the time of their initial two-day training (September, October, or November 2000) and again following the assessment fair (April or May 2001). Data collected during the first (n = 785) and second (n = 800) administrations were analyzed separately to determine the internal structure of the survey form and to establish the statistical characteristics of the scores.

To determine the structure of the form, a statistical technique called “factor analysis” was employed. Factor analysis groups items together based on similarities in the way that participants respond to the individual items. Items that group together can be thought of as measuring the same idea or attitude and forming a unique scale. Factor analysis of the posttest data revealed that the items on the SAC survey form could be grouped into six uniquely measurable scales.

After examining the content of the items within each grouping, a label was assigned to identify the main theme. The labels assigned to each scale, along with one or two representative items, are provided on the following page. The number in parentheses after the scale name indicates the reliability of the scale. Reliabilities for all scales are acceptably high.

Each participant received a score on each scale as well as a total score for the entire survey. The total raw score for each scale was divided by the number of items in the scale. This placed all scale scores back on the same 1 (strongly disagree) to 5 (strongly agree) scale used on the survey form. Survey scores can be interpreted by remembering that the closer a scale score is to a score of 5, the more a participant indicated higher levels of agreement with the content of each item in the scale.
Table 1: Scale Labels and Sample Items.

Scale 1 – Familiarity with the Illinois Learning Standards (.86):

“I have spent time reading the Illinois Learning Standards that cover the subject area(s) I teach.”
“I have discussed the Illinois Learning Standards with other teachers at my school.”

Scale 2 – Usefulness of the Learning Standards in Planning Instruction (.92):

“The Illinois Learning Standards help me to identify specific knowledge that a student must master.”
“The Illinois Learning Standards help me to plan problem-solving tasks for my students.”

Scale 3 – Confidence in Using the Standards (.95):

“I feel confident that I can create classroom lessons that are aligned with the standards.”
“I am good at translating the achievement targets identified by the standards into classroom lessons.”

Scale 4 – Benefits of Standards-Aligned Instruction (.95):

“Preparing standards-aligned lessons has a positive impact on my teaching.”
“Creating standards-aligned lessons makes me a more effective teacher.”

Scale 5 – Frequency of Use (.90):

“I regularly use the Illinois Learning Standards to help plan classroom instruction.”

Scale 6 – Effect of Student Participation (.94):

“Students are more motivated to learn when they are involved in the creation of classroom assessments.”
Analysis of Survey Data

**Who Responded?** At the time the survey form was first administered during the Fall of 2000, there were 930 registered participants in the SAC training program. Of these, 895 were primarily classroom teachers and constituted the target sample for the survey. Pretest data were collected from 785 of these teachers (approximately 88%). Of those teachers with pretest data, 674 also provided a completed posttest survey form (approximately 86%). This matched group included teachers from 43 of the state’s 45 Regional Offices of Education. Approximately 57% were elementary school teachers, 23% were middle school teachers, and 17% were high school teachers (about 3% of the matched group did not respond or reported teaching across more than one level). The average number of years of teaching experience was 12.08, with a range from 0 years to more than 20 years.

**How Did Survey Scores Change?** The difference between pretest and posttest scores for all teachers in the matched group was computed for both the total survey score and for each scale. The average score change from pretest to posttest was investigated using a paired-sample *t*-test for the difference between means. The results indicated all score changes were strong and positive. Specifically, after participating in the SAC training and team project, teachers reported being 1) more familiar with the Illinois learning standards, 2) more likely to understand how to use the standards in planning classroom instruction and assessment, 3) more confident in their ability to use the standards, 4) more aware of the benefits of the standards, 5) more likely to use the standards to plan classroom instruction, and 6) more likely to believe a standards-aligned instructional approach had a positive impact on student motivation and participation. The average score changes are graphed in Figure 1.
Figure 1: A Comparison of Survey Scores - Pre-Test vs. Post-Test
How Large Were the Survey Score Changes? The increase in the average survey score was substantial from both a statistical and practical standpoint. Statistically, the growth in the total and scale scores from pre- to posttest was significant at the p<.001 level (see Appendix B for a summary of the t-test results).

Another way to assess the degree of change is to compute an “effect size” for each change. Effect sizes represent the degree of change in terms of the standard deviation of the variable involved. Traditionally, effect sizes of .20 or greater have been considered substantial for training programs in education. Effect sizes obtained with the SAC program ranged from .48 for the familiarity scale to 1.04 for the confidence scale. Figure 2 presents a graph of the effect sizes computed for the total score and for each scale score.
Figure 2: Effect Sizes of Pre-Test to Post-Test Score Changes
**Did Grade Level Affect the Degree of Change?** Next, to investigate whether the principal grade level at which a teacher taught was associated with changes in pre- to posttest scores, participants in the matched sample were assigned to one of three grade-level categories. Teachers who taught at kindergarten through grade 5 were placed in the first group (elementary level), teachers from grades 6 to 8 were placed in the second group (middle school level), and teachers from grades 9 to 12 were placed in the third group (high school level).

A multivariate analysis of variance was conducted using the total pre- and posttest scores as the dependent variables, and the grade-level and experience-level (see page 14) groups as the independent variables. A similar set of analyses was conducted using the pre- to posttest scores for each scale as dependent variables. No significant effect for grade-level grouping was found for either the total score or for any scale score.

These results indicate there were no significant differences in the effectiveness of the SAC project for any of the grade-level groups. Although the groups may have started out at different levels of agreement on a particular scale, high school teachers were just as likely to experience a positive shift in attitudes as were elementary or middle school teachers. Figure 3 summarizes the pretest to posttest score changes for the total survey score for each of the three grade-level groups.
Figure 3: Total Scores by Teaching Level

Grades K to 5
Grades 6 to 8
Grade Level Group
Grades 9 to 12

Average Total Score

Pre-Test
Post-Test
**Did Years of Teaching Experience Affect the Degree of Change?** To investigate whether the number of years of teaching experience a teacher had affected changes in pre- to posttest scores, participants in the matched sample were assigned to one of five experience categories. The first group included teachers with 0 to 4 years experience. The second group included teachers with 5 to 9 years of experience. The third group included teachers with 10 to 14 years experience, the fourth group included teachers with 15 to 19 years of teaching experience, and the fifth group included teachers with 20 or more years of experience.

A multivariate analysis of variance was conducted using the total pre- and posttest scores as the dependent variables, and the five experience groups and the grade-level groups as the independent variables. A similar set of analyses was conducted using the pre- and posttest scores for each scale as dependent variables. No significant effect for experience-level grouping was found for either the total score or any scale score.

These results indicate there were no significant differences in the effectiveness of the SAC project for any of the five experience-level groups. Although the groups may have started out at different levels of agreement on a particular scale, newer teachers were just as likely to experience a positive shift in attitudes as were the more experienced teachers. Figure 4 summarizes the pretest to posttest score changes for the total survey score for each of the five groups.
Figure 4: Total Score Changes by Years of Experience

![Bar chart showing total score changes by years of teaching experience. The x-axis represents years of teaching experience, ranging from 0 to 4 years to 20 or more years. The y-axis represents average total score, ranging from 0 to 5. The chart compares pre-test and post-test scores for each age group. The bars for each year group are color-coded to differentiate between pre-test (blue) and post-test (red).]
SECTION II:

Scoring the Portfolios
Portfolio Scoring Summary

Another portion of the evaluation focused on 203 of the teachers who participated in the initiative and accepted an option to receive professional development course credit for their participation. Teachers in this group were required to create a “best-example” portfolio documenting an assessment designed by the teacher around the Illinois standards. A purposive sample of 48 teachers was selected with a goal of including 16 teachers from each of the higher, average, and lower survey score groups. Of the 48 teachers contacted, we received portfolios from 22 for a response rate of 46%. Of these, 12 were returned from teachers in the higher scoring group, 2 were returned by teachers in the average scoring group, and 8 were returned from teachers in the lower scoring group. For the purpose of analysis, scores from the teachers in the high scoring group were compared to scores from teachers in a combined average and low scoring group.

Portfolios were scored using a rubric adapted from the Arter and Busick (2001, p. 5 & Appendix B) workbook used by most teachers. Rubric scores were assigned to each of seven scorable features of the teacher’s portfolio. These features are:

- Standard 1: Are the assessment targets clear and appropriate?
- Standard 2a: Are users and uses clear?
- Standard 2b: Is communication of results effective?
- Standard 2c: Are students involved?
- Standard 3: Is the assessment method appropriate?
- Standard 4: Is achievement adequately sampled?
- Standard 5: Has bias been minimized?

Scores were assigned in each of these seven areas on a three-point scale described in Appendix B of the Arter and Busick (2001) workbook. The descriptions of the score points have been modified to more accurately reflect the criteria applied. The final score points on the rubric were:

- 3 = Clear and convincing evidence,
- 2 = Some evidence presented, and
- 1 = Little or no evidence.
Figure 5 on the following page compares the average rubric scores assigned to portfolios created by teachers in the higher scoring group to portfolios created by teachers in the combined average/lower scoring group. The figure also contains a comparison of scores on each individual rubric dimension for these two groups.

From this figure it is clear that, on average, teachers were better able to identify clear and appropriate targets for assessment (Standard 1) and identify the users and uses of the assessment (Standard 2a). Teachers were moderately successful at specifying how results would be communicated, involving students in the assessment process, creating an appropriate assessment method, and adequately sampling indicators of achievement (Standards 2b, 2c, 3, and 4). None of the teachers were successful at describing how to minimize bias in their assessments (Standard 5).

With respect to group differences, it is clear that teachers who indicated a higher level of satisfaction with the program on the posttest questionnaire were more likely to have shown evidence of success on characteristics related to Standards 1 through 4 than were teachers who indicated lower levels of satisfaction. Note that the difference between average portfolio scores for these two groups is statistically significant with a large effect size, \( t(20) = 3.83, p = .001, d = 1.64 \). This is remarkable given the small sample size and the small range of score options (i.e., 3 points on the rubric).
Figure 5: Portfolio Scoring by Survey Group
SECTION III:

Summary of Interview Responses
**Interview Summary: Closed-Format Questions**

From the sample of individuals who participated in both the pretest and the posttest, a purposive sample of 45 teachers was selected. The sample was balanced to include an even number of teachers from the top scoring one-third, middle scoring one-third, and lowest scoring one-third of the total survey score frequency distribution. Our goal was to collect data from 10 individuals in each group (for a total sample of 30 interviews), so the sample of 45 constitutes an overdraw of 33% in order to compensate for non-responses due to non-contacts and refusals. The goal of 10 interviewees per group was achieved.

Appendix D contains a copy of the interview protocol. A graduate student who is a former public school teacher with expertise in the area of teacher professional development conducted the interviews by telephone. Participants were paid $25 for their time. Responses to each interview were recorded on audio tape (with the consent of the participating teacher), and notes summarizing each answer were entered into a database.

Analysis of the closed-format interview questions produced the following results:

1) there are few differences between the three groups with respect to content specialization, grade level, or years of teaching experience;
2) interviewees with high questionnaire scores were more likely to have been volunteers than were those with medium or low questionnaire scores;
3) interviewees with higher scores tended to meet with their groups more frequently than did interviewees with medium and low questionnaire scores; and
4) the perceived quality of the SAC program was high, regardless of the interviewee’s overall survey score.

Table 2 on the following page provides the response patterns related to findings 2 through 4 listed above.
Table 2: Response Patterns to Selected Closed-Format Interview Items.

Question: How did you become involved in the [SAC] program?

<table>
<thead>
<tr>
<th>Survey Score Group</th>
<th>High</th>
<th>Average</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volunteered</td>
<td>5</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Asked by an</td>
<td>4</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Assigned by</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>administrator</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Question: How frequently did your team meet?

<table>
<thead>
<tr>
<th>Survey Score Group</th>
<th>High</th>
<th>Average</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>About once a week</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>About once every</td>
<td>3</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>two weeks</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>About once per</td>
<td>2</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>month</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Question: How [does] this program rank in comparison to your other professional development experiences?

<table>
<thead>
<tr>
<th>Survey Score Group</th>
<th>High</th>
<th>Average</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>The best</td>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>One of the five</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>best</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Better than average</td>
<td>0</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Average</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Below average</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>One of the five</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>worst</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The worst</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Interview Summary: Open-Ended Questions

As mentioned earlier, responses to all open-ended interview questions were summarized and collected in an electronic database. The full set of responses is available upon request. However, for the current report, we have chosen to highlight the responses to two items.

A formal content analysis was conducted for survey questions 16 and 23. The content analysis was accomplished by first reading all responses to each question without knowledge of group membership of the respondent and then taking notes on the issues observed. For example, for Question 16, which asked teachers to describe their team meetings, issues identified from the analysis of teacher responses included 1) who led the team meetings, 2) what materials were used, 3) how team members interacted, and 4) who attended the meetings. Next, the issues were categorized into related themes. For Question 16, the issues identified above were grouped into two general themes—one relating to the topics, materials, and activities around which the meetings focused, and the other relating to the participants in the meetings. Finally, the responses of participants in each group relating to each theme were collated, and counts of the types of responses relating to each theme were tallied. Hypothesis tests were performed on contingency tables to determine whether trends were statistically significant and meaningfully large, where appropriate.

**Question 16.** Content analysis of Question 16 suggested three things. First, participants tended to be more satisfied when leadership of the group rotated among group members from one meeting to another, \( \chi^2(1) = 4.29, p = .04, \phi = .38 \). About 80% of the interviewees who were in the high scoring group indicated that they were members of this type of team, compared to 40% of the interviewees who were in the moderate and low scoring interview groups. Moderate and lower scoring groups were more likely to have had the coach or school principal lead most of the team meetings.

Second, teachers in higher scoring groups were more likely to have used all the available Stiggins resource materials (the text, workbook, and videos), although the trend was not statistically significant with the small sample size, \( \chi^2(1) = 4.40, p = .12 \),
\( \phi = .28 \). About 70\% of the highly satisfied teachers indicated that their group used all the materials, while only 40\% of the teachers who indicated moderate or low levels of satisfaction were in groups that used all the materials.

Third, there was a slight trend for principals to be more likely to attend meetings in highly satisfied groups than in groups with moderate or low levels of satisfaction, \( \chi^2(1) = 1.36, p = .24, \phi = .21 \).

Table 3: Content Analysis Results – Question 16

**Question 16 – Theme 1a: Type of Group Leadership**

<table>
<thead>
<tr>
<th>Survey Score Group</th>
<th>High</th>
<th>Average</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rotated across team members</td>
<td>8</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Coach or principal led most meetings</td>
<td>2</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

**Question 16 – Theme 1b: Use of Stiggins Materials**

<table>
<thead>
<tr>
<th>Survey Score Group</th>
<th>High</th>
<th>Average</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>Used all the materials</td>
<td>7</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Did not use all the materials</td>
<td>3</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

**Question 16 – Theme 1c: Principal Generally Attended Team Meetings**

<table>
<thead>
<tr>
<th>Survey Score Group</th>
<th>High</th>
<th>Average</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generally attended</td>
<td>7</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Generally did not attend</td>
<td>3</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>
**Question 23.** This question asked interviewees to identify what they thought were the most useful team activities or experiences that came from participation in the SAC project. Content analysis of Question 23 revealed three trends across the groups of higher, average, and lower scorers with respect to the types of team activities felt to be most useful. First, teachers with high or average survey scores were more likely to see hands-on activities (e.g., discussing the textbook, working through workbook activities, or preparing an assessment fair project) as the most useful activities (60% versus 30% for teachers with low levels of satisfaction), $\chi^2(1) = 2.40, p = .12, \phi = .28$. On the other hand, lower scoring participants were more likely to suggest that the opportunity to get together with other teachers in their schools was the most beneficial activity in the program (70% versus 25% for teachers with moderate or high levels of satisfaction), $\chi^2(1) = 5.63, p = .02, \phi = .43$. Finally, there was a slight trend for teachers with higher survey scores to cite team reviews of team members’ classroom assessments as the most valuable team activity (40% for highly satisfied, 20% for moderately satisfied, and 10% for low satisfaction), $\chi^2(2) = 2.61, p = .27, \phi = .29$. 
Table 4: Content Analysis Results – Question 23

**Question 23 – Theme 1a: Conducting Hands-On Activities**

<table>
<thead>
<tr>
<th>Survey Score Group</th>
<th>High</th>
<th>Average</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>More Useful</td>
<td>6</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Less Useful</td>
<td>4</td>
<td>4</td>
<td>7</td>
</tr>
</tbody>
</table>

**Question 23 – Theme 1b: Getting Together with Other Teachers**

<table>
<thead>
<tr>
<th>Survey Score Group</th>
<th>High</th>
<th>Average</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>More useful</td>
<td>2</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Less useful</td>
<td>8</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

**Question 23 – Theme 1c: Review of Team-Produced Classroom Materials**

<table>
<thead>
<tr>
<th>Survey Score Group</th>
<th>High</th>
<th>Average</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>More useful</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Less useful</td>
<td>6</td>
<td>8</td>
<td>9</td>
</tr>
</tbody>
</table>
SECTION IV:

Conclusions and Recommendations
Conclusions

The analysis of the pre- and posttest survey data clearly showed that teachers in the SAC program became more positive towards standards-aligned instruction over the course of their participation. Specifically, by the end of the SAC program, teachers reported being: 1) more familiar with the Illinois learning standards, 2) more likely to understand how to use the standards in planning classroom instruction and assessment, 3) more confident in their ability to use the standards, 4) more aware of the benefits of the standards, 5) more likely to actually use the standards to plan classroom instruction, and 6) more likely to believe a standards-aligned instructional approach has a positive impact on student motivation and participation. These positive changes held true regardless of the grade level at which the teacher taught or the number of years of teaching experience the teacher had. The degree of change was both statistically significant and larger than changes typically obtained with educational training programs.

In addition, teacher portfolios developed by teachers with higher survey scores (indicating greater acceptance and endorsement of a standards-aligned approach) were found to be significantly better than those produced by lower scoring teachers for six out of seven key features related to quality instructional and assessment material. This suggests a relationship between the survey data and the production of real classroom materials. Teachers who rated themselves as more comfortable and knowledgeable regarding the standards tended to translate that knowledge into classroom products more effectively.

Finally, interviews with higher and lower scoring participants indicated that participants with higher survey scores were more likely to be members of teams that: 1) met more frequently, 2) shared leadership duties across team members, 3) volunteered to participate in the program, and 4) included team meetings regularly attended by an administrator. This suggests that the role of the team in mediating attitude change may be significant. Indications from the interview data support the idea that team structure and task focus during team meetings can contribute positively to overall outcomes.
Recommendations

The first year of the SAC initiative was a pilot year. While the success of the current SAC program is apparent from the data, interviews with selected participants identified a number of areas in which improvements could be made to further enhance the overall effectiveness of the program in the future. The primary recommendations that emerged from the interviews were: 1) restructure the initial training so that more information is provided to the teachers in advance of the first meeting, 2) spend more time communicating to the teachers about program structure and expectations for team meetings and products, 3) provide coaches with additional training and resources, and 4) restructure the assessment fair to give more time for in-depth presentations focusing on teacher-produced materials. A detailed list of recommendations from various participants is provided below.

1) **Restructure the two-day training session by providing a detailed overview of the direction of the program prior to the training.** Many teachers, even those who were highly satisfied with the program, indicated that they felt the two-day training session was very intense. There was so much information presented in the training session that some teachers felt overwhelmed. In addition, because it was a first-year pilot program, the information presented may not have included an adequate description of the expected outcomes of the program (particularly with respect to the products to be presented at the assessment fairs).

2) **Be sure to provide an overview of the direction the program will go during the school year.** As early as possible, preferably even before the training sessions, be sure to tell teachers about the activities that will take place during the school year, and emphasize that the assessment fair will be a culminating activity. There were a few interviewees who felt that too much emphasis was placed on preparing for the assessment fair.

3) **Provide teachers with the textbook as the primary material for the program.** Some teachers felt that the text, videos, and workbook are redundant. Some also reported that they did not find the videos very useful during meetings, although they may be useful if loaned to individuals for viewing on their own time. The workbook activities did not seem to be used much by some teachers, particularly those in the most highly satisfied groups—they spent more time reviewing their own materials. Overall, the book seemed to be the most appreciated resource.
4) **To the extent possible, match resources to the needs of the team.** Some teams complained that the coach did not have the experience needed to help their teams. In addition, there were several complaints that the assessment fair had too little (and in a few cases, too much) of an emphasis on lower elementary grades.

5) **Allow for adequate meeting time, and instruct coaches and team leaders not to rush through activities.** Teachers indicated that the program takes a great deal of time due to the detailed work involved.

6) **Emphasize reviewing teachers’ own classroom assessments during team meetings.** Overwhelmingly, teachers indicated that this was the best use of their time during the meetings. The most highly satisfied groups all seemed to spend a good deal of time engaged in reviews of team members’ materials (rather than spending a majority of their time viewing videos or working through workbook activities).

7) **Bring in Rick Stiggins again as a presenter at the assessment fair, and utilize him as a resource during the training if possible.** Several interviewees really appreciated Rick Stiggins’ presence and presentation at the assessment fair.

8) **Make sure that coaches are trained and remain well-informed about the direction, requirements, and goals of the program.** Particularly in groups with low levels of satisfaction with the program, complaints about non-attendance of coaches at the meetings, lack of knowledge about the program on the part of the coach, and lack of direction to the team meetings were common.

9) **Restructure the assessment fair to focus on small group presentations of projects.** Some teachers did not feel that the poster sessions were as useful as they may have been. In addition, some felt that the roundtables were too crowded, and it was difficult to hear the presenters. Most teachers indicated that this was the most valuable part of the assessment fair, and they were disappointed that more options were not available and that the ones that were available did not give them the opportunity to learn more than they did about the presenting teachers’ projects.

10) **Begin the program as early in the school year as possible, preferably during the summer.**
Appendix A:
The Standards-Aligned Classroom Initiative Survey Form
Appendix B:
Table of Paired $t$-test Results
Paired Sample $t$-tests:

<table>
<thead>
<tr>
<th>Pair</th>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Standard Error</th>
<th>95% Confidence Intervals</th>
<th>t</th>
<th>df</th>
<th>Signif.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair 1</td>
<td>Total Score</td>
<td>.4953</td>
<td>.4829</td>
<td>.020</td>
<td>.4559 .5347</td>
<td>24.705</td>
<td>579</td>
<td>.000</td>
</tr>
<tr>
<td>Pair 2</td>
<td>Confidence</td>
<td>.6660</td>
<td>.6219</td>
<td>.024</td>
<td>.6175 .7146</td>
<td>26.943</td>
<td>632</td>
<td>.000</td>
</tr>
<tr>
<td>Pair 3</td>
<td>Usefulness</td>
<td>.3418</td>
<td>.5757</td>
<td>.023</td>
<td>.2972 .3864</td>
<td>15.045</td>
<td>641</td>
<td>.000</td>
</tr>
<tr>
<td>Pair 4</td>
<td>Benefits</td>
<td>.4663</td>
<td>.5615</td>
<td>.022</td>
<td>.4231 .5095</td>
<td>21.188</td>
<td>650</td>
<td>.000</td>
</tr>
<tr>
<td>Pair 5</td>
<td>Frequency</td>
<td>.6402</td>
<td>.9353</td>
<td>.036</td>
<td>.5691 .7113</td>
<td>17.677</td>
<td>666</td>
<td>.000</td>
</tr>
<tr>
<td>Pair 6</td>
<td>Participation</td>
<td>.4453</td>
<td>.7932</td>
<td>.030</td>
<td>.3850 .5056</td>
<td>14.499</td>
<td>666</td>
<td>.000</td>
</tr>
<tr>
<td>Pair 7</td>
<td>Familiarity</td>
<td>.2979</td>
<td>.8228</td>
<td>.032</td>
<td>.2349 .3609</td>
<td>9.287</td>
<td>657</td>
<td>.000</td>
</tr>
</tbody>
</table>
Appendix C:
Multivariate Analysis of Variance Results
Multivariate Analysis of Variance Results

Pretest and Posttest scores entered as dependent variables, and the five “years of experience” groups and the three “grade level” groups entered as independent variables.

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within Cells</td>
<td>62.93</td>
<td>545</td>
<td>.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>40.01</td>
<td>1</td>
<td>40.01</td>
<td>346.45</td>
<td>.000</td>
</tr>
<tr>
<td>Experience</td>
<td>.61</td>
<td>4</td>
<td>.15</td>
<td>1.32</td>
<td>.263</td>
</tr>
<tr>
<td>Grade Level</td>
<td>.12</td>
<td>2</td>
<td>.06</td>
<td>.53</td>
<td>.589</td>
</tr>
<tr>
<td>Exp. By Level by Time</td>
<td>1.65</td>
<td>8</td>
<td>.21</td>
<td>1.78</td>
<td>.078</td>
</tr>
</tbody>
</table>
Appendix D:
Interview Protocol
Follow-up Interview

(Complete the information below prior to making the call.)

(1) Name: __________________

(2) Phone Number: __________________

(3) School: __________________

(4) Coach Name: __________________

(5) Team Number (Leave Blank): __________________

(6) Coach Number (Leave Blank): __________________

(7) Area Number (Leave Blank): __________________

(8) ROE Number (Leave Blank): __________________

Introduction

Hi, this is INTERVIEWER NAME from Michigan State University. I'm calling to conduct the interview that we spoke about previously. If you recall, the interview will take about 45 minutes. Is this a convenient time for us to do that interview? (If YES, Continue. If NO, reschedule the interview.)

As I told you the last time, I'm working as an external evaluator of the Standards-Aligned Classroom Initiative. The purpose of this interview is to gather information about how your team functioned and the value of your team experience.
Your answers will be held in confidence. Nobody but you, me, and the data analyst will know your identity. If we quote anything that you say, we won't associate your name with that quote, and you won't be identified as a person that we interviewed in any reports that are written. We'd like a candid description of how you feel about your team's experiences, so nobody associated with your group or school will be told what you say to me during this interview.

Do you mind if I tape record the interview, so that I can later check my notes about our conversation? (If YES, begin tape recorder. If NO, take detailed notes.)

Do you have any questions before we start? (Answer any questions.)

Background Information

Before I begin asking questions about your team, would you please tell me what content area you teach, the grade level you teach, and how many years you've been a teacher?

(9) Content Area: __________________

(10) Grade Level: __________________

(11) Years Teaching: ________________

(12) Also, tell me about how you became involved in this program. ________________

Section 1: Team Activities

The first questions that I'd like to ask concerns how your team operated during the school year. In this section of the interview, I'll ask you about the structure of your meetings and the process that you followed.

First, let me ask you a few short questions about how often you met.

(13) How frequently did your team meet (Check one)?
Section 2: Team Successes

The next thing that I'd like to talk about concerns the things that your team did that you thought were particularly helpful or particularly successful. In this section of the interview, I'll ask you questions about how the Standards-Aligned Classroom Initiative compares to other professional development activities in which you've been involved and what impact your involvement in the Standards-Aligned Classroom Initiative has had on the way you plan, teach, and assess.

First, I'd like you to rank this program relative to other professional development experience you have had. Think about the professional development experiences that you've had in the last five years (or since you began teaching if you have not been a teacher for at least five years). Please tell me how this program ranks in comparison with those other professional development experiences.

(17) Was it (check one):
[ ] the best?       [ ] one of the best 5?
[ ] better than average? [ ] average?
[ ] below average?   [ ] one of the worst?
[ ] the worst?

(18) In about how many professional development programs have you participated (check one)?
[ ] about 1 to 5 programs  [ ] about 6 to 10 programs  [ ] about 11 to 15 programs
Now I'll ask you a few open-ended questions about your experiences.

(19) Overall, how valuable was your participation in this project as a professional development experience (summarize the response)?

(20) How did this program compare to other professional development experiences you've had? In other words, why did you rank it the way you did? (summarize the response)

(21) What skills or knowledge did you learn in this program that will be most valuable to you as a teacher (summarize the response)?

(22) What kind of impact, if any, do you envision your participation in this program having on your success as a teacher (summarize the response)?

(23) Of all of the things that you did during your team's work, what were the most useful activities or experiences for you (summarize the response)?

Now give me a few before and after snapshots. Tell me what you know and can do now that you did not know or could not do prior to participating in this program.

(24) Specifically, what did you know about the Illinois standards before and after participating in this program (summarize the response)?

(25) What do you now do differently in planning instruction than you did before participating in this program (summarize the response)?

(26) What do you now do differently in presenting instruction to students than you did before participating in this program (summarize the response)?

(27) What do you now do differently in the area of assessment than you did before participating in this program (summarize the response)?
Section 3: Improvements

The section of the interview concerns ways that the experience could be changed to make the program more useful to you, personally. We're asking these questions because it is likely that the Standards-Aligned Classroom Initiative will be funded for a second year, and we'd like to provide recommendations for improving the program. Specifically, I'd like you to address five aspects of the program--the training, the coach, the structure of the meetings, the materials, and the assessment fairs.

(28) Before I ask about each of these aspects of the program, let me ask, do you have any suggestions for improving the program? (If YES, then continue. If NO, then skip to the final section.)

[ ] Yes [ ] No

(29) So, tell me how the program could be improved? ((Summarize the response. Probe for clarification, and prompt the respondent to address each of the following aspects: (a) the initial two-day training session, (b) the coach, (c) the structure of the team meetings, (d) the materials used in the team meetings, and (e) the activities at the assessment fairs)

(30) Did you receive any incentives for participating in the Standards-Aligned Classroom Initiative? ((Summarize the response. Probe for amount of money, release time, etc.)

(31) Are there additional incentives that would have made participation in the program better for you? ((Summarize the response. Probe for amount of money, release time, etc.)

Conclusion

(32) You have given me useful information for evaluating the program. Do you have any questions or additional comments? (summarize the response)

(33) In our previous conversation, I mentioned to you that we will send you $25 for your time. Would you please tell me the name you would like to appear on the check and the mailing address where you would like the check to be sent? (Read the information back to the person to verify accuracy.)
We will probably mail the checks around the end of June. If you have not received a check by the first of July, please contact Ed Wolfe. Here's his phone number and email address. Let me know when you are ready.

His phone number is 517-355-8538. His email address is wolfee@msu.edu. Dr. Wolfe is also the person to contact if you have any questions or concerns about the interview.

Thanks for your time. Have a good summer.

(34) Survey Score
[ ] High Survey Scores   [ ] Medium Survey Scores   [ ] Low Survey Scores