

## EPILOGUE

### Once Upon a Time: A Tale of Excellence in Assessment Richard DuFour

The distinguished authors contributing to this book offer extraordinary insight and expertise in the exploration of what represents "excellence in assessment." My contribution is not designed to examine the technical aspects of assessment, but rather to illustrate the potential power of effective assessment practices through the oldest teaching vehicle known to man—a story. Richard Axelrod once wrote: "Universities come to know about things through studies, organizations come to know about things through reports, and people come to know about things through stories" (p. 112). Good stories teach us. They convey not only how something should be done, but more importantly, why it should be done. They communicate priorities and clarify what is significant, valued, and appreciated.

The following story offers a model of what I consider to be excellence in assessment in a school setting. The protagonist of the story is a high-school teacher, but the message applies equally and with little revision to middle and elementary schools as well.

After 10 years as a high-school social studies teacher, Peter Miller was convinced that kids were kids and schools were schools. So when his wife suggested they move across the country to be closer to her family, he willingly agreed. He applied at several schools and was offered an interview at Russell Burnette High School.

The interview process at Burnette intrigued Peter. At every stage of the process, the selection committee stressed that the school had created a collaborative culture in which teachers worked together to help all students learn. Teacher teams had created a "guaranteed and viable curriculum" that specified the knowledge, skills, and dispositions all students were to acquire in each course. Peter was asked to review the "Essential Learnings" established by the U.S. history team and was struck by the fact that the curriculum stressed only 10 key concepts each semester, rather than the long list of discrete facts he had been expected to teach at his former school.

The selection committee also gave Peter copies of the curriculum pacing guide, the common assessment calendar, examples of preassessments for several units, examples of common assessments, and the rubrics for evaluating student essays and term papers—all of which had been created by the U.S. history team. The committee asked Peter to critique each document and to express his concerns as well as suggestions for improvement. Peter was impressed by the active role the other history teachers played in the interview process, and equally impressed that he had been required to spend a day at the school teaching prior to being offered the job. He gladly accepted the offer to join the staff at Burnette and looked forward to establishing himself in his new school. He had none of the trepidation and self-doubt that had characterized his first year as a teacher. He was a veteran who knew how schools worked.

## Once Upon a Time: A Tale of Excellence in Assessment

The school year at Burnette began with 3 full days for teachers to work prior to students arriving on campus. Peter was delighted; he would have plenty of time to get his room ready and to prepare his first unit. His enthusiasm diminished when he learned that mornings were reserved for teachers to meet in their collaborative teams. Peter had had little use for the faculty and department meetings in his previous school, and he quietly resented that team meetings at Burnette would intrude upon his personal time at such a busy point in the school year.

The first U.S. history team meeting, however, was nothing like the meetings at Peter's old school. Each member of the team studied the school's results from both the state assessment and the national ACT exam in social studies. The team also reviewed an analysis of the very strong correlation between results on their common assessments with the results of the high-stakes state and national exams.

"We know we are on the right track:" Ambrose, the team leader, observed. "If we can help every student be successful on our ongoing common assessments, we can be very confident they will be successful on state and national assessments as well. We can continue to assess students in other concepts we deem important, but we have an obligation to help our students be successful on the high-stakes tests they must take." The team devoted the remainder of the morning meeting to identifying the areas where students had experienced difficulty on the two external exams and brainstorming instructional, curricular, and assessment strategies to address those areas.

The second U.S. history team meeting was devoted to reviewing the results of the common assessment students had taken at the end of the first unit in the previous school year. The team had analyzed the results at the end of that unit, and now they reviewed their findings and their ideas for addressing the concepts and skills where students had performed least well. Peter was puzzled. The results looked quite good to him.

"I think you should congratulate yourselves:" Peter told his teammates. "Why are you taking time to review this exam? I don't see any evidence here of serious problems in student learning."

"It's just what we do here:" his teammate Miriam explained. "We are always looking to get better, and even on a test where students did well, there's always a concept or a few items where they do least well. If our team can identify effective strategies for addressing those areas, we can become even more effective and help more students achieve at higher levels every year."

It was evident that the team did not merely consider student performance on the overall test. Members had identified the specific skills and concepts students were to learn and had established the score a student must obtain on each to be deemed proficient. This shift of emphasis from general performance to skill-by-skill analysis helped Peter to see that on one area of the test, many students failed to demonstrate proficiency. The team spent the rest of the meeting reviewing each component of the test and offering ideas for teaching and assessing the concepts and skills of the unit. Ambrose asked Peter to develop test items for the unit and to present them to the team for possible inclusion in the common assessment.

## Once Upon a Time: A Tale of Excellence in Assessment

The third team meeting was devoted to a discussion of the prerequisite knowledge students would need to be successful in the first unit and how the team would determine which students lacked that knowledge. The team had reviewed key terms and concepts recommended by the National Center for History in the Schools and the National Council of Social Studies Teachers, had selected the terms they felt were most essential to their curriculum, and then had assigned the terms to different units of instruction. The key vocabulary terms for the first unit included:

assimilation	expansionism	neutrality
autonomy	federal	protective tariff
cartography	imperialism	republicanism
colonization	inalienable	salutary neglect
constitutionalism	mercantilism	sovereign
culture	monarchy	sphere of influence
dissent	nationalism	

Miriam explained to Peter that at the start of every unit, teachers administered a brief preassessment of those terms to their students. Because there were at least two sections of history taught each hour, and because the history classrooms were next to one another, teachers were able to divide students into two different groups based on their proficiency with the vocabulary.

"So in my case, students who lacked the prerequisite vocabulary went with Frank," Miriam explained. "He introduced key terms with brief explanations, then asked students to define the term in their own words in the section of their history notebook devoted to terms and concepts. He helped them create graphic organizers and put them into pairs to review the terms."

"Meanwhile" Frank said, "students who were already proficient went with Miriam. She presented them with a high-interest article on the major themes of the unit and then led a discussion of the article."

The team then reviewed examples of some of the graphic organizers students had used in the past and discussed different strategies for presenting key vocabulary in terms students could understand. They also discussed questions they could use to stimulate discussion of the article to be presented to the second group of students.

By the end of this third day of preparation, Peter was growing uneasy. He had always enjoyed virtually unfettered autonomy in his teaching. He had been free to teach what he wanted, when he wanted, how he wanted, and to assess students in whatever manner he saw fit. Now he was part of a team that made those decisions collectively. Peter was not convinced that all this teaming and collective decision-making was in the best interests of teachers or students. On the eve of his first day of classes, he was less confident of his ability to fit in at his new school, and he questioned whether he had made the right choice in accepting the position there.

Once the students arrived, however, Peter felt far more comfortable. He discovered that he still enjoyed great autonomy in how he conducted his classroom and how he taught his content on a day-to-day basis. His concern that his instruction and daily pacing would be prescribed proved to be unfounded. His team had helped clarify what students were to

## Once Upon a Time: A Tale of Excellence in Assessment

learn in the unit, and he knew that each U.S. history teacher would present, review, and discuss those same essential learnings with students in the first few days of class. He also knew that all U.S. history students would take the team's common assessment on the same day at the end of the third week of class. In the meantime, however, Peter was free to make decisions each day regarding how to teach and how to check for student understanding.

Peter's team continued to meet for 1 hour each week. On Monday mornings, teachers reported to work 15 minutes earlier than usual, and the start of classes was delayed 30 minutes in order to create this collaborative time. Teachers were then allowed to leave 15 minutes earlier than usual on Mondays, so they were not required to work longer hours or to sacrifice personal time in order to collaborate with their colleagues.

The first few minutes of each meeting were spent debriefing members of the team on how they felt the unit was going. Members were encouraged to express any concerns. The team then turned its attention to the five-point rubric that had been created to score student responses to essay questions. Members reviewed the criteria they had established for assessing the quality of student essays and then examined different anchor essays that reflected each score. At subsequent meetings, they individually scored the same student essay and then shared their conclusions.

"We're willing to accept a difference of one point on the five-point scale" Ambrose explained, "but if two members present scores with a variance of more than one point, we'll discuss the variance, review our rubric and the anchor essays, and then determine an appropriate score." Peter was somewhat chagrined when he was the only team member whose score deviated from the rest of the team the first two times they practiced applying the rubric. His colleagues, however, were very supportive. They explained the thought process they used in scoring the sample and encouraged him to articulate his reasoning. The dialogue was helpful, and on the third attempt to review a sample essay, his score was consistent with his colleagues.

The ability to write a well-reasoned, persuasive essay that incorporated historical evidence was one of the essential outcomes all history students were expected to achieve. So Peter followed the lead of his teammates and taught his students the rubric to ensure they understood the criteria they should use in judging the quality of their own work. He devoted class time to reviewing the rubric with his students, providing them with sample essays from the past, and leading the class in scoring essays of different quality.

Peter had already discovered the importance of checking for student understanding on an ongoing basis. He felt he was proficient in using classroom questions and dialogues for that purpose. He directed questions to students randomly, rather than relying primarily upon volunteers. He extended wait time whenever students struggled and refused to let any student simply declare he or she did not know the answer. He would prod, rephrase, ask them to explain their thought process, and insist they clarify exactly what they did understand and exactly where they were confused. Students soon learned that a simple shrug would not suffice for Mr. Miller. They also learned that he rarely affirmed or corrected an answer immediately. Instead, he would provide more wait time and then direct a student's response to several other students for analysis and comment. He encouraged debate and insisted that students explain their thought process.

## Once Upon a Time: A Tale of Excellence in Assessment

Peter did not limit his strategies for checking student understanding to questioning during class. He would typically begin each class by directing students to write in their notes, "At the end of today's class, I will be able to . . ." and asking them to explain how that day's lesson was linked to the essential learnings of the course. At the conclusion of the class, he would pose a question, ask students to write a response in their notes, and quickly check each student's response to see if there was confusion. He frequently called upon students to identify similarities and differences between historical events and eras or to develop analogies between historical situations and contemporary events. He often presented a statement, challenged students to explain whether or not they agreed, and then used disagreements or confusion as an opportunity to clarify. He did not believe in giving homework every day, but when he did assign homework, he made a point of providing specific feedback to students. In short, Peter was confident his students were well-prepared when they took the team's first common assessment.

The assessment was in two parts. The first section included multiple choice and matching items, while the second presented an essay question. Peter presented the results from the first part of the assessment to his department chairman and received two printouts the next day. The first showed how his students had performed on each skill and concept the team had assessed, compared to the performance of all the students who completed the assessment. The second printout presented an item analysis that compared the results of his students to all students on each item on the assessment.

The night before the next team meeting, Peter's wife asked how his classes were going. "Well, I'm generally pleased," Peter told her, but on our common assessment, my students struggled with one concept—distinguishing between different forms of government. Their scores prevented our team from achieving its target for that concept." He grimaced. "I'm not looking forward to admitting that tomorrow." Privately, he hoped he would be able to avoid saying anything.

The next team meeting was a revelation to Peter. Although each teacher had received only the analysis for his or her own students compared to the total group, teachers were extremely open with their results. "My students obviously didn't get the concept of republicanism," Miriam said. "How did the rest of you teach that?" Various team members shared their strategies, then brought up the weak spots in their own students' performance.

Encouraged by their openness, Peter shared his concerns about his students' understanding of different forms of government. The team's response could not have been more positive. Frank and Miriam suggested instructional strategies. Ambrose offered a graphic organizer he had developed that had helped students use comparison and contrast to understand the concept. Skill by skill, concept by concept, the team reviewed student performance, identified whose students had excelled and whose students had struggled, and engaged in lively dialogue about strategies for teaching concepts more effectively.

The team then turned its attention to the item analysis and identified three items on the 30-item test that warranted review. The team quickly discovered that all three items assessed the same skill and that one of the items had been poorly written. They also discovered that the skill had been the last one taught in the unit. The team decided to rewrite the poorly written item and to change the pacing of the unit so members could devote more time to the skill prior to giving the next assessment.

## Once Upon a Time: A Tale of Excellence in Assessment

Following the meeting, Peter asked Miriam, "What happens if we use all these strategies and as result, student performance on that skill reaches proficiency?"

"Why, we'll celebrate our success, of course," she said. "And then we'll look for the next items where students did less well. There will always be 'the lowest 10 percent' of items on any assessment we give. We attack those items, implement improvement strategies, celebrate our success, and then look for the next items. That is the beauty of continuous improvement. You never really arrive, but there is always a lot to celebrate."

Peter was perplexed by the team's policy regarding the essay portion of the assessment. Teachers were expected to provide specific feedback to each student regarding how he or she could improve the essay according to the team rubric, but they did not assign a specific grade to the essay. Students were then required to prepare a second draft of the essay that incorporated the recommendations before they would receive a grade.

"I don't understand the rationale behind this process" Peter said. "Why not grade the first essay and average the scores?"

"Well" Frank said, "we just don't think it's reasonable to assign a grade to skills students are attempting to use for the first time. We want our kids to have the benefit of specific feedback before we grade their efforts."

"We think giving feedback tells students that we expect them to achieve a standard," Miriam chimed in, "and that we'll ask them to refine and improve their work until they reach it. Later in the year, they won't have this chance, but for now, early in the learning process, we feel it's imperative that students benefit from practice and specific feedback before we assign grades to their work."

This feedback was part of a systematic structure to ensure learning. Shortly after the team administered the common assessment, teachers were required to complete progress reports sent to counselors, advisors, and parents. Students in danger of failing were required to report to the tutoring center, where they devoted extra time to their studies and received small-group and individualized tutoring—during the school day. Burnette High had created a schedule that ensured each student had one period available each day to receive this additional time and support for learning. Upperclassmen who did not require this intervention were given the privilege of unstructured time, while freshmen and sophomores were assigned to study halls.

Two weeks later, the students who had completed this first intervention were given another opportunity to demonstrate they had learned the key concepts of the previous unit by taking another form of the assessment. If they performed well, their failing grade was dropped and replaced with the higher grade for students. Miriam explained, "We say we want them all to learn; we don't say that we want them all to learn *fast* or the *first time*. If some students have to work harder and take longer before they demonstrate proficiency, so be it. In the final analysis, if they demonstrate proficiency, we give them a grade that reflects that."

Peter was still a little skeptical. He thought that an opportunity to take a second assessment would cause students to "blow off" the first test. Afterwards, however, he had to

## Once Upon a Time: A Tale of Excellence in Assessment

admit that he was wrong. His juniors truly valued their unstructured time. They knew that poor performance on the first assessment would mean not only the loss of that privilege, but also an extra commitment of time and effort to learning what they should have learned in the first place. Peter could see no evidence that students were indifferent to the results of their first test. In fact, there was a palpable sense of academic press—a clear expectation that students must demonstrate they had actually acquired the essential knowledge and skills of the unit—that he had never experienced before.

By the end of his first month at Burnette, Peter had come to the realization that he was not in Kansas anymore—this school was very different from those in which he had worked in the past. He had never experienced practices like working in teams, developing common assessments, aligning those assessments with state and national tests, using the results from previous assessments to guide instruction, identifying prerequisite knowledge for success in the unit, regrouping and sharing students, providing students with specific feedback rather than grades, providing systematic interventions when students were unsuccessful, and allowing students additional opportunities to demonstrate proficiency.

The difference in the use of assessments was one of the most striking contrasts between Peter's past practice and his new school environment. In his former school, individual teachers had either developed their own assessments or simply used the assessments provided in the textbook and teacher's manual. There, administering a test signaled the end of a unit, and the purpose of the test was to assign grades. Students who did not do well were exhorted to do better and try harder, but they rarely received specific feedback on how to improve—and almost never were given a second chance to demonstrate their learning. Students and teachers alike understood that taking a test meant the unit was over, and the class would move forward.

At Burnette, however, assessments were used to determine if students needed assistance in acquiring prerequisite skills prior to teaching each unit, to inform individual teachers of the strengths and weaknesses in their instruction, to help teams identify areas of concern in the curriculum, to identify students who needed additional time and support for learning, and to give students additional opportunities to demonstrate that they had learned. Assessment seemed to represent the most critical component of the collaborative culture that characterized the school, and the way teachers used assessments sent students a clear message that they were required, rather than invited, to learn.

By the end of his first semester, Peter considered Burnette's practices so powerful and practical that he questioned why he and his colleagues had not implemented them in his former schools. If certain background knowledge was an essential prerequisite for success in a unit, it just made sense to identify students who did not have that knowledge and to intervene on their behalf at the outset of the unit. If all the teachers of a course were expected to teach the same concept, it was certainly more efficient to work collaboratively in planning the unit, gathering materials, and developing assessments than to work in isolation and duplicate each other's efforts.

"In my old school," Peter told the U.S. history team one day, "what students learned, the rigor of their assessments, and the criteria used to judge the quality of their work depended on who their teacher was. We each worked in isolation. Here, however, our approach is so much more equitable. Students have access to the same curriculum and

## Once Upon a Time: A Tale of Excellence in Assessment

assessments of equal rigor, and we judge their work according to the same standard. At Burnette, I know all students are receiving the best education possible, in every classroom.”

Peter had come to recognize the power of assessments in the service of learning—for students and teachers alike. The common assessments provided him with timely feedback on the success of his students in meeting an agreed-upon standard, on a valid assessment, in comparison to other similar students attempting to achieve the same standard. For the first time in his career, he was able to identify areas of strengths and weaknesses in his teaching and to use that insight in his dialogue with teammates to improve his instruction. Assessments had become a powerful tool in informing his practice.

More importantly, however, Peter had discovered the potential of assessments to enhance the learning of his students. By administering common assessments at the end of each unit, the members of his team were able to identify students who needed additional time and support for learning. Burnette’s systematic intervention process required those students to continue to work on acquiring the essential skills in the tutoring center. Tutors then used the assessments to identify the specific skills and concepts a student had been unable to master and to provide precise instruction and feedback in a small group setting to assist the student. It certainly made sense to Peter to use assessments not only to point out that a student had not learned, but also to provide the student with the specific feedback and information to improve upon the learning. Because assessments in his former school had been regarded as the conclusion of a unit rather than a critical element in the learning process, poor performance on an assessment sometimes had a devastating effect on a student’s motivation. A series of bad test scores early in the semester could doom a student to a failing grade. Burnette’s practice of allowing additional opportunities to demonstrate learning never deprived students of hope.

It had been a semester of growth for Peter. He had reexamined not only his practices, but also some of his fundamental assumptions. Not all schools were alike. Some school cultures and structures are far more effective in helping students learn. Even a veteran teacher like himself could learn to approach his profession from a new perspective. He had made the right choice in coming to Burnette, because this school had taken the assertion, “All Kids Can Learn” and added an even bolder proposition: School can be a place where even the adults could learn.

Burnette High School’s assessment story is not a utopian ideal. It takes place in real schools with real teachers and real students. It does not require a windfall of new resources. It does, however, require something even more rare—the willingness to change the fundamental assumptions and practices that have characterized public education for decades.

Attention to this vital area must be a cornerstone of any school improvement effort. Schools simply cannot meet the challenges they face unless educators unleash the potential of effective assessment. To limit the use of this powerful instrument to ranking, sorting, and selecting students is analogous to using a computer as paperweight. When done well, however, assessment can help build a collaborative culture, monitor the learning of each student on a timely basis, provide information essential to an effective system of academic intervention, inform the practice of individual teachers and teams, provide feedback to students on their progress in meeting standards, motivate students by demonstrating next



## Once Upon a Time: A Tale of Excellence in Assessment

steps in their learning, fuel continuous improvement processes—and serve as the driving engine for transforming a school.

### **Reference**

Axelrod, R. (2002). *Terms of engagement: Changing the way we change organizations*. San Francisco: Berrett-Koehler.