

# Quality Improvement in Aviation Education: A Framework for Programmatic Assessment and Quality Improvement

William R. Caldwell

Southern Illinois University Carbondale

## ABSTRACT

There is a growing concern in higher education for a system of assessment and program quality improvement. This paper introduces a program evaluation system that may be used for evaluating higher education programs in a manner that provides an understandable quantitative quality metric. It provides background on the movement of higher education into the realm of quality management of educational processes, the national quality movement in public education, and efforts by the International Standards Organization and its affiliate organizations to establish international standards for education. Finally, it describes a seven-step assessment and quality improvement process. These steps reflect structure found in the Kirkpatrick Model of Program Evaluation and guidelines provided by the Central Missouri State University Quality Improvement Plan.

## INTRODUCTION

The purpose of this paper is to discuss and propose a program for collegiate aviation programs that can act as a framework assessment and quality improvement. This proposal suggests what needs to be done to improve the programmatic quality of academic and technology aviation courses through the application of Central Missouri State University's (CMSU) quality program guidelines to an existing programmatic evaluation concept developed by Kirkpatrick (Phillips, 1997) called the Kirkpatrick Four-Level Approach. This paper discusses the Kirkpatrick model, the CMSU Quality Improvement Program, and the Aviation Quality Improvement Program which combines them into a useful tool for assessing and controlling aviation education programs. It suggests how to develop a program with moderate effort on the part of the department and its faculty to provide a control mechanism (the last function of management according to Bateman and Snell (2004)) that may be used to manage an academic program and improve the quality of its graduates.

Since 1983 and the publication of *A Nation at Risk* (National Commission on Excellence in Education U.S., 1980), the public school system has been struggling with the concept and requirement of measuring (assessing) student accomplishment against a set of desired learning outcomes. From the outset, there has been

resistance to evaluating students against fixed outcomes. Teachers believe that their classrooms are being invaded by people and agencies that have no understanding of the pedagogical challenges faced daily by teachers and that their traditional freedom to interpret the curriculum in their own way is being threatened by this intrusion.

A quick review of *Phi Delta Kappan - The Professional Journal for Education* for the last decade will provide ample examples of the basis of the statements above. An example would be the comments made by Professor Emeritus Maurice Holt (2002) when he wrote: "The curriculum straitjacket is the price extracted for believing that education is about assessed performance on specified content" (p.1). Holt continues with this theme by suggesting that "Commitment to standards-led school reform means creating a system of schools geared solely to the product—test results—and not to the process of creating educative experiences" (p.3).

About a decade ago, this movement began to spread into higher education, and today the budding concept of improving the quality of the educational experience is beginning to bloom in the ivy halls of higher education institutions. Ten years ago at Central Missouri State University, the North Central Association of Colleges and Schools (NCACS) indicated that the university assessment system was in need of strengthening. This concern stems from the NCACS's Academic Quality Improvement Program (North Central Association of Colleges

and Schools, 2004) that begins with a complete assessment of member school performance. Thus, for the past 10 years there has been an accelerating effort at many universities to develop and institute a meaningful system of quality improvement and assessment of student achievement.

Surely, this movement in education has been a spin-off from our national enchantment with a management movement developed during the last quarter of the 20<sup>th</sup> Century called Total Quality Management (TQM). In multiple forms, quality management has caught on in businesses around the world. This growing interest has spawned the need for quality standards and created an international clearing house of standards for quality improvement in various industries. Since 1996, various members of the International Organization for Standardization (ISO) have published proposed or final quality standards for education. For example, the *International Workshop Agreement (IWA) 2:2003* provides guidelines for the application of ISO 9001:2000 in educational organizations providing educational products. This standard basically applies the procedures of quality improvement used by industry for almost two decades to education. The goal is simple: improve the processes used to educate students, ultimately improving student learning (International Organization for Standardization, 2003). This document was preceded by a proposed base document created by the *American Society of Quality's ASQ Z1.XX: Guidelines on the Application of ISO-9001:2000 to Knowledge Work and Lifelong Learning* (Pivec, Schoening, & Sinitza, 2001) and by the American National Standards Institute (1996) in its document, *ANSI/ASQC Z1-11-1996: Guidelines for the Application of ANSI/ISO/ASQC Q9001 or Q9002 to Education and Training Institutions*.

### **THE KIRKPATRICK PROGRAM EVALUATION MODEL**

Before describing the Aviation Quality Improvement Program (AvQIP) the basis of its design should be discussed. The framework used to build this evaluation system was the Kirkpatrick Program Evaluation Model. This is

one of several models that have been used successfully to measure the effectiveness of training programs.

The Kirkpatrick model was selected from five models facilitating this process. The other models included the Kaufman Five-Level model, the CIRO Approach, the CIPP and the Phillips' Five-Level model. Kirkpatrick's model was chosen for its adaptability to the higher education process. (Phillips, 1997).

According to Nickols (2000), Donald Kirkpatrick set forth his four-level approach to the evaluation of training in a series of articles appearing in the journal of what was then known as the American Society of Training Directors. The first of these four seminal articles was published in November of 1959. The remaining three articles were published in the succeeding three months, with the fourth and final article appearing in February of 1960. These articles can be found in *Evaluating Training Programs*, a collection of articles compiled by Kirkpatrick from the pages of the American Society & Training and Development (ASTD) Journal and published by ASTD in 1975. The phases of the Kirkpatrick program evaluation model are defined below.

Level 1: Reactions. This phase is an assessment of how well the students liked a particular training program. Reactions are typically measured at the end of training. They may also be measured during the training, even if only informally in terms of the instructor's perceptions (Nickols, 2000). This level of program evaluation is common to universities and usually called an "end-of-course" evaluation. Phillips (1997) states that this level asks the question: "Were the participants pleased with the program" (p.39)? This definition suggests the validity of the reaction level is questionable because of the subjective nature of the response; i.e., students who believe they have done well in a course will tend to rate it higher than those who believe they did not do well.

Level 2: Learning. This phase is characterized by what the student learned while in the course. It measures what the student has learned – "What principles, facts, and techniques were understood and absorbed by the conferees?" (Nickols, 2000, p.1) This formative

assessment is made throughout the course via various means and instrumentalities such as examinations, quizzes, project work, etc. Usually, this assessment requires an entry diagnostic assessment of knowledge so that subsequent assessments clearly identify what was learned (Nickols, 2000). Phillips (1997) agrees with this definition stating it answers the question, "What did the participants learn in the program" (p.39).

Level 3: Behavior. This level deals with changes in behavior on the job or in other situations where the new knowledge can be applied. Nickols (2000) believes that any evaluation of change in on-the-job behavior must occur in the workplace itself:

It should be kept in mind, however, that behavior changes are acquired in training and they then transfer (or don't transfer) to the work place. It is deemed useful, therefore, to assess behavior changes at the end of training and in the workplace. Indeed, the origins of human performance technology can be traced to early investigations of disparities between behavior changes realized in training and those realized on the job. (p.5)

However in educational institutions, applying the model to the workplace is problematic, since students usually have not entered the workplace at this point in their lives. To accommodate this element of the model, the AvQIP had to incorporate a supervisor/instructor/peer evaluation system that applies to a workplace if one is available as well as to classroom application performance.

Phillips (1997) states that this level answers the question, "Did the participants change their behavior based on what was learned" (p.39). His perspective of Kirkpatrick's intent appears to be more useful than Nickols' in the context of evaluating post-secondary aviation education programs. Consequently, Phillip's concept of this level is applied to the AvQIP.

Level 4: Results. According to Nickols (2000), Kirkpatrick did not define this element of his framework. Instead, he relied on a range of examples to make clear his meaning such as: "Reduction of costs; reduction of turnover and absenteeism; reduction of grievances; increase in

quality and quantity of production; or improved morale which, it is hoped, will lead to some of the previously stated results" (p.5).

Phillips (1997) on the other hand says this level asks the question: "Did the change in behavior positively affect the organization" (p.39)? Again, this definition is most useful in the post-secondary program evaluation because it does not directly tie the results to the workplace. This is important because educational institutions attempting to assess the effectiveness of their educational effort may not have the benefit of observing the student in the workplace while they are still in school. However, post-graduation surveys of student performance in the workplace should be a part of any educational programmatic evaluation. Indeed, this has been a mainstay of numerous university post-graduate program evaluation schemes.

## **THE CMSU QUALITY IMPROVEMENT PROGRAM**

The AvQIP was designed to support the CMSU Quality Improvement Program (CQIP) for Academic Departments. A quality improvement system defines a method for improving a process. In this case, the process is that of creating student learning. Hence, the purpose of the CQIP program is to improve the quality of student learning at CMSU.

But what is quality? According to Besterfield (1994) and the ANSI/ASQC Standard A3-1987, "Quality is the totality of features and characteristics of a product or service that bear on its ability to satisfy implied or stated needs" (p. 1). Besterfield continues by defining the results of a quality assurance or control system. He writes that quality assurance (a) determines the effectiveness of the quality improvement system, (b) appraises current quality, (c) determines quality problem areas, and (d) assists in correction or minimization of these problems. The AvQIP attempts to implement each of these actions.

In the CQIP model, continual process improvement requires the identification of clear programmatic objectives (student-learning outcomes) and a means of assessing the changes in student learning by measuring student

accomplishment of the objectives. It also requires the development of processes and materials that support the quest for continuing student learning improvement. It is easy to see the parallel to the quality assurance discussed by Besterfield (1994).

CQIP is a locally developed program evaluation system. It stresses three primary goals. These goals are the (a) identification and validation of student learning outcomes, (b) identification of methods to assess student achievement of these outcomes, using the results to improve student learning, and (c) implementation of student assessment that documents student progress and shows how to use this information to improve student learning.

The CQIP, however, suggests just one of several programmatic evaluation systems that have been developed over time. As noted earlier, Phillips (1997) discusses several of these, including his own. Where business profit is a factor, systems such as the Kirkpatrick Four Levels of Evaluation, Kaufman's Five Levels of Evaluation, the CIRO Approach, and Phillips Five Level Return on Investment Systems have been shown to be effective programmatic evaluation systems. In non-profit situations, the Kirkpatrick system has been shown to be effective. As the reader will recall, the Kirkpatrick system measures (a) Reaction – participant evaluation of the system, (b) Learning – what the participants learned, (c) Behavior – whether participants change their behavior based on what was learned, and (d) Results - did the change in behavior positively affect the organization (Phillips,1997). In this case, the organization(s) affected (customer) by the process is the aviation industry served by the Department, the University, and its graduates.

The goals of the AvQIP are to combine the elements discussed above into a cohesive and meaningful system that defines industry and university requirements for aviation program graduates; to establish a curriculum and supporting courses designed to bring students to this level of ability in knowledge, skills, and attitudes; to assess student ability to meet these requirements as they progress through their

university experience, at the point of graduation, and as their careers progress; and then to determine effective interventions that will improve the quality of learning as the AvQIP process ensues.

## **PREPARATION FOR AVQIP**

Attempting to develop a quality improvement program without knowing where the affected organization is attempting to go and having a management structure that supports movement toward the organization's goals are problems that should be addressed before the AvQIP can be an effective mechanism. Consequently, to prepare for AvQIP, it is recommended that the organization have a strategic plan, an operating paper that deals with the actions required by the AvQIP, as well as the AvQIP. In addition, student information guides should provide students information on the assessment processes used in the AvQIP and specifically seek their response to post-graduation surveys. Not the least of these concerns should be getting graduates of the department to help it or the related alumni association keep their address information current after graduation.

## **THE AVIATION QUALITY IMPROVEMENT SYSTEM (AVQIP)**

A schematic of the entire AvQIP is provided at Figure 1. The diagram depicts how the various elements of the process are related to each other and how they affect the overall quality of the aviation program. The AvQIP is composed of seven basic steps:

**Step 1 - Reaction Survey.** The reaction survey is the first element of the Kirkpatrick model. The purpose of this survey is to get the student's perspective of how the course has gone. At the end of each course, students are asked to provide their assessment of the course and the instructor using an instrument that asks questions about the quality of the course and the instruction.



follow the Mager (1984) behavioral objective format which requires a statement of the conditions of the question, the use of an action verb, and a required observable action. For example: Given an E6B computer, the student will solve a ground speed problem. After each evaluation, the instructor or department administrative personnel enter test data and learning outcomes into the Instructor's Course Assessment Spreadsheet (Table 3).

This spread sheet is designed to evaluate the difficulty, outcome effectiveness, and ability of the question to discriminate between students

who understand the material being tested and those who do not.

The background for the difficulty and discrimination values may be found in Grunlund (1998). The outcome effectiveness is a simple average for each examination question linked to its driving learning goal. The examination questions reflect the learning outcomes for each overarching goal. As was the summary of the reaction surveys, a summary of data from these spreadsheets is entered into the Department Effectiveness Summary (Table 2).

Table 3. *Instructor's Course Assessment Spreadsheet*

			Evaluation		Mid Term		Final Exam		Paper	Project
			Learning Objective #		1.1	1.2	2.1	2.2	3.1	4.1
Last	Score	Credit	Adj Score - Question #	1	2	1	2			
A	5.2	0	5.2	1	1	1	1	0.7	0.5	
B	2.2	1	3.2			1		X	0.7	0.5
C	1.2	1	2.2					X	0.7	0.5
D	4.2	1	5.2	1	1	1		X	0.7	0.5
E	4.2	1	5.2	1	1	1		X	0.7	0.5
F	4.2	1	5.2	1	1	1		X	0.7	0.5
G	5.2	0	5.2	1	1	1		1	0.7	0.5
H	5.2	0	5.2	1	1	1		1	0.7	0.5
	0		0							
J	5.4	0	5.4	1	1	1		1	0.6	0.8
K	5.4	0	5.4	1	1	1		1	0.6	0.8
L	2.4	1	3.4			1		X	0.6	0.8
M	2.4	1	3.4			1		X	0.6	0.8
N	2.4	0	2.4					1	0.6	0.8
O	5.4	0	5.4	1	1	1		1	0.6	0.8
P	5.4	0	5.4	1	1	1		1	0.6	0.8
Q	3.6	0	3.6	1	1	1		1	0.6	0.8
TOP HALF NUMBER CORRECT				6	6	7	3	5.6	4	
MIDDLE NUMBER CORRECT				0	0	0	0	0	0	
BOTTOM HALF NUMBER CORRECT				5	5	7	6	4.8	6.4	
TOTAL IN CLASS	17									
DIFFICULTY				0.75	0.75	0.88	0.38	0.70	0.50	
EFFECTIVENESS				64.71%	64.71%	82.35%	52.94%	61.18%	61.18%	
DISCRIMINATION				0.125	0.125	0	-0.375	0.1	-0.3	
MEAN SCORES	4.50		4.89							

The specific learning objectives established for each course are developed by the professor given responsibility for the course design and may be reviewed by the faculty during faculty meetings with the objective of confirming content validity of the course and with industry advisory committees for the same purpose. Courses may also be reviewed by the Aviation Accreditation Board International (AABI) teams as part of its oversight of the department's programs. In addition, the outcomes may be used by the Federal Aviation Administration (FAA) to validate courses delivered by the department.

**Step 3 - Evaluate Program-level Learning Outcomes.** The third step is to evaluate student progress in the general department learning objectives. General

department learning goals might be the following:

1. The ability to express oneself clearly and quickly in writing and speech.
2. The ability to read and comprehend literature in the student's field and have developed a reading program that will keep the student current in aviation.
3. The ability to continue one's training, education, and intellectual development when one leaves school.
4. The ability to exhibit the highest level of aviation professionalism in the student's career area.
5. The ability to solve problems in the student's aviation field.
6. The ability to work effectively as part of a team.

7. Possess the knowledge, skills, and attitudes necessary to be a success in the student's area of the aviation industry.
8. Possess the basic understanding of the leadership and managerial skills graduates will need to be an effective leader in the aviation industry.
9. The ability to successfully compete for employment in the student's aviation field.
10. The ability to do basic research, interpret and analyze the data and make useful presentations based on that research.
11. Possession of the basic knowledge, skills, and attitudes needed to be a useful participant in the student's profession, society, and country.
12. Possession of the university-desired skills of higher order thinking, communicating, interacting, managing information, and valuing.

This step of the process is accomplished by post-course evaluations of student progress in and application of the general goals established by the collegiate entity and is conducted by students and instructors using the form shown in Figure 5.

During this phase, the Program Learning Outcome Evaluation is presented to students near the end of each course. The student completes the student self-evaluation portion of the form and returns it to the course instructor for the instructor's evaluation. The form is then forwarded to department administration and its data entered into a Program Goals Evaluation Spreadsheet (Table 4).

**Step 4 - Monitor Progress.** The fourth step is to monitor student progress through the curriculum. Each full-time faculty member is assigned a list of students to monitor, mentor and advise. Freshmen are contacted at the beginning of their first semester and advised on the normal progression and course schedule for their degree program. After this initial contact, department administration advises faculty members if one of their students departs from this recommended schedule or does not perform satisfactorily during a course. This is accomplished at the end of each semester by

entering student grades into a four-year course plan and record form. This record is kept on file in the department. In addition to their advisory role, faculty members are encouraged to establish a mentoring relationship with their assigned students.

Student advisement includes the responsibility to:

1. Help the student develop his/her personal program plan (following the established four-year program whenever possible but adjusting for unique student needs such as being a transfer student or coming into the program with a FAA certificate),
2. Counsel when the student has difficulty with the plan or courses in the plan,
3. Recommend any necessary course substitutions,
4. Initiate credit by evaluation requests if appropriate, and
5. Monitor student progress.

**Step 5 - Evaluate Application of Learning.** This step complies with Level 3 of the Kirkpatrick model – measure how well student learning is applied on the job. Hence, work for university airport management, the flight operations management program, internships, or maintenance management is monitored through supervisors, and work in the classroom is monitored by course instructors. Each semester students, instructors, and departmental supervisors of aviation students complete the Program Learning Outcome Evaluation form (Figure 2) to provide an assessment of the student's ability to apply what he/she has learned during their coursework to their current or future job. These evaluation forms are collected, entered into a Program Goals Evaluation Spreadsheet (Table 4) and then filed in the student's hard copy file.

In addition, all four-year students are required to complete a capstone course which requires students to apply what they have learned during their degree program to simulated aviation problems. Students taking the capstone course complete the Program Learning Outcome Evaluation form at the end of the class just as they would for any other course.





Table 4. *Programs Goals Evaluation Spreadsheet*

Student Number	Delivery 1		Delivery 2	
	Self Evaluation	Instructor Evaluation	Self Evaluation	Instructor Evaluation
1	85	85	85	85
2	57	66	57	66
3	85	85	85	85
4	57	66	57	66
5	85	85	85	85
6	57	66	57	66
7	85	85	85	85
8	57	66	57	66
9	85	85	85	85
10	57	66	57	66
Mean	71	75.5	71	75.5
%	83.78	89.09	83.78	89.09

It is possible to reduce the manpower involved in this program through automation of the data collection and analysis programs. With time, the spreadsheets and other records required by the program can be generated automatically using electronic data collection and reporting technology to minimize the student, faculty, and staff effort required to input, analyze and report on the data by using the Scantron Par System to reduce much of the handwork discussed above.

The Scantron Company's ParSYSTEM is an integrated suite of powerful software modules that allow you to create, administer and score tests on paper, via networks or over the Internet. With ParTEST, teachers can develop multi-format tests from item banks. ParTEST Online enables test takers to take tests online or on a network. And ParSCORE completes the suite allowing educators to manage student records, analyze test outcomes and create reports. (Scantron, 2006, p. Products/ParSystem).

In the meantime, the author will provide electronic copies of the spreadsheets and forms designed to support this system.

### CONCLUSION

The quality improvement system proposed in this article is based on a classic program evaluation model and techniques and processes produced by practitioners in quality

management. The process has seven primary steps that comply with Central Missouri State University Quality Improvement Plan goals and follow the challenges our students and the aviation industry will face tomorrow.

In summary, post-secondary educational institutions are feeling increasing pressure to improve the product of their institutions through student assessment and quality improvement programs that make use of outcome-based assessment data. While some will feel that this is an encroachment on the academic freedom post-secondary education has traditionally enjoyed, a more positive view would suggest that moving in this direction may be the only way for modern educators to keep up with the exponential growth in the knowledge pool at a time when the world is moving faster and faster toward an information-based-economy that demands that its workforce be able to access this knowledge pool and use it efficiently and effectively.

### RECOMMENDATIONS

This paper proposes one way to structure a teaching organization, to monitor its processes and products, and to continually improve its processes both to the betterment of the segment of industry and the economy served by the organization and its graduates. If there is a recommendation supported by this paper, it is not to emulate the system explained in this paper, but to set about producing a system that fits one's own environment while accepting the underlying rationale that the quality of the educational experience provided students today can be and must be steadily improved to meet the steps of program evaluation recommended by Kirkpatrick.

## REFERENCES

- American National Standards Institute (ANSI). (1996). *ANSI/ASQC Z1-11-1996: Guidelines for the application of ANSI/ISO/ASQC Q9001 or Q9002 to education and training institutions*. Milwaukee: American Society for Quality.
- Bateman, T. S., & Snell, S. A. (2004). *Management: The new competitive landscape* (6th ed.). New York: McGraw-Hill.
- Besterfield, D. H. (1994). *Quality control* (4th ed.). Englewood Cliffs: Prentice Hall.
- Gronlund, N. E. (1998). *Assessment of student achievement* (6th ed.). Needham Heights: Allyn & Bacon.
- Holt, M. (2002). It's time to start the slow school movement [Electronic Version]. *Phi Delta Kappan - The Professional Journal for Education*, 84, 265-271. Retrieved October 31, 2006 from [http://www.pdkintl.org/kappan/k0212\\_hol.htm](http://www.pdkintl.org/kappan/k0212_hol.htm).
- International Organization for Standardization (ISO). (2003). Quality management systems – Guidelines for the application of ISO 9001:2000 in education. Retrieved April 7, 2004, from <http://www.iso.ch/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=38866&ICS1=3&ICS2=120&ICS3=10>
- Mager, R. F. (1984). *Preparing instructional objectives*. Belmont: Lake Publishing.
- Nickols, F. W. (2000). *Evaluating training: There is no "cookbook" approach*. Retrieved April 2, 2004 from: [http://home.att.net/~nickols/evaluating\\_training.htm](http://home.att.net/~nickols/evaluating_training.htm)
- North Central Association of Colleges and Universities (NCACU). (n.d.) *Academic quality improvement program (AQIP)*, Retrieved April 2, 2004, from <http://www.aqip.org/WhatisAQIP.html>.
- National Commission on Excellence in Education U.S. (NCEE). (1983). *A nation at risk*. Department of Education, Washington, DC.
- Pivec, M., Schoening, J., & Sinitsa, K. (2001) *ASQ Z1.XX: Guidelines on the application of ISO-9001:2000 to knowledge work and lifelong learning*. Milwaukee: American Society for Quality.
- Phillips, J. J. (1997). *Handbook of training and evaluation and measurement methods* (3rd ed.). Houston: Gulf Publishing.
- Scantron, (n.d). *Parsystem*. Retrieved October 29, 2006 from <http://www.scantron.com/siteSpecific/common/categoryDisplay.aspx?categoryID=104&tree=2|29|42|104>
- Skymark, Inc. (n.d.). *Cause & effect diagram*. Retrieved April 2, 2004, from [www.skymark.com/resourses/tools/cause.asp](http://www.skymark.com/resourses/tools/cause.asp)