Assessment & Evaluation in Higher Education, Vol. 25, No. 2, 2000

A Model for Computer-based Assessment: the catherine wheel principle

STAN ZAKRZEWSKI & CHRISTINE STEVEN, *University of Luton, Luton, UK*

ABSTRACT *The success of computer-based assessment systems is based on a structured approach to design and implementation together with a model that generates efficient and effective standards and procedures. This paper proposes a model that utilises a step-wise approach to assessment design and implementation within which the management and assessment of operational, technical, pedagogic and financial risks are made explicit. It is the strategies for risk elimination that form the basis for the standards and procedures adopted.*

Introduction

An interactive computer-based assessment system at the University of Luton centred on 'Question Designer' for Windows and originally piloted in 1993 continues to expand. In the academic year 1998–99 over 10,000 students sat summative and 3000 formative assessments. The system has evolved through a pilot phase (Pritchett & Zakrzewski, 1996) to a de-centralised department- based system to what is now a centralised university-wide system (Zakrzewski & Bull, 1998).

The pilot was launched with 150 comparative psychology students. The examination duration was 1 hour and the question types predominantly multiple choice. Objective test delivery grew in the psychology department over the next few years and its administrative staff were trained to convert examinations to 'Question Mark' format. Other departments began to take an interest; Politics and Public Policy, Biology and Computing. Each year more and more departments were using their local facilities to deliver objective tests. Formative assessments began to grow out of the summative system and in January 1998 the decision was made to centralise the provision.

The Learning Resources Centre (LRC) houses some 200 workstations in a central complex. A Modular Credit Scheme supports the student programmes of study. At the end of each semester (in January and June) the central complex is used for two weeks to facilitate computer-based summative end-of-module examinations. These examina-

ISSN 0260-2938 (print)/ISSN 1469-297X (online)/00/020201-15

tions are timetabled by the Exams Office who are also responsible for assigning academic invigilators to the system with technical support being provided by Computer Services. The system is fully integrated with the traditional examination procedures. The Centre is also reserved for phase tests (summative examinations taken during the semester) eight times a year. Formative examinations are taken on an open access basis in the Centre throughout the year. In the Academic year 1998–99 modules in Accounting, Economics, Leisure, Travel and Tourism, Design, Material Sciences, Biology, Computing, Mathematics, Social Sciences, Midwifery, Psychology, Media, Languages, Law and History used computer-based assessment as part of their assessment portfolio.

Software Engineering Principles

Today Software Engineering is recognised as a legitimate discipline that advocates the use of software process models, software engineering methods and software tools which have been adopted successfully across a wide and diverse number of industry applications (Basili, 1991). It has been learnt that where individuals and companies develop their software in an unstructured way the quality of the software suffers. It is the belief of the authors that many of the principles used for the construction of large software systems can be successfully applied to the use of that software. To solve real problems in any setting a team must develop a strategy that encompasses process models, methods and tools. One process model that is reputed by its originator to be a Meta Model and therefore of universal use is the Spiral Model (Boehm, 1988). The major breakthrough made by Boehm was in the introduction of a phase that considers the risks that may occur when developing software. He advocates the listing of the possible 'top ten' risk items in the production of software, followed by a systematic approach to the minimisation or elimination of the effects of these risks (Boehm, 1989). Pressman (1997) has proposed a variation on the spiral model by dividing it into six framework activities, also known as task regions, the original proposed by Boehm has four quadrants.

The computer-based assessment model or 'catherine wheel' put forward in this paper is based on the original idea as proposed by Boehm together with the modifications described by Pressman. Consideration is given to risk analysis and management techniques advocated by Robert Charette (Charette, 1989) and the risk assessment and control techniques as described by Barry Boehm (Boehm, 1989).

The Characteristics of a Computer-based Assessment Model

A computer-based assessment system is not just a piece of software enabling the design and delivery of objective tests. It is a complete system in which management, academic staff, support staff and students work together to achieve the system's aims and objectives. It requires human resources, physical resources, finance and quality documentation to succeed. It is 'open' in nature in the sense that it interacts with other internal institutional systems as well as external.

A computer-based assessment model must therefore depict its 'system' characteristic. A planned structured approach to the design and implementation of computer-based assessment systems must be adopted that will reflect the natural evolution of the system in its lifetime. The model must always deliver a complete system; from a pilot to a fully integrated, centralised university-wide system.

Computer-based assessment is a 'risky' activity. Academic staff concern themselves with the pedagogic value of using objective tests to assess their students, especially the

testing of higher order cognitive skills. Technical support staff will need to be convinced of the technical integrity and security of the system. Administrative staff and managers will concentrate on the reliability and integration of the operational procedures. Senior managers will need to balance the financial risks with the benefits to the institution in the short to medium term. A successful model for computer-based assessment must therefore attempt to eliminate as many risks as possible, reduce the likelihood of any of those remaining occurring, and should they occur, minimise their impact. It must generate detailed procedures, based on risk elimination or reduction, that when implemented will deliver a quality system. Finally, the model must be generic in nature, applicable equally to closed network systems, WEB-based assessment, Optical Mark Reading (OMR) systems or Optical Character Reading (OCR) systems.

The Catherine Wheel

Figure 1 depicts the 'catherine wheel'. The 'catherine wheel' has five segments: Planning, Risk Analysis and Management, Assessment Design, Evolutionary Development and Evaluation. These segments are revisited in sequence, at each stage of evolution, from a module pilot to departmental implementation and finally a fully integrated university-wide system.

Starting at the centre of the spiral, a pilot for a single module, usually at level 1 of an undergraduate programme, is planned and a risk analysis performed based on the pilot aims and objectives. An objective test is then designed, written and the pilot implemented. The pilot is evaluated and a decision is made at this stage whether to expand to a departmental system or not. The evaluation will therefore involve a pilot review and a feasibility study for further growth.

In the second stage of the spiral the departmental assessment system is planned and a risk analysis undertaken based on the pilot evaluation. The assessments may now test a higher range of cognitive skills utilising a variety of question types. The implementation will involve additional personnel, physical resources and documentation which will result in a more detailed evaluation from which a university-wide system may emerge.

The third stage of the spiral sees a significant system change as the move is made from what can be viewed as a number of decentralised departmental systems to a centralised, fully integrated university-wide system. This time planning and risk analysis are based on departmental evaluation. The quality of assessment design will continue to be enhanced with evaluation ongoing.

Risk analysis and management is conducted prior to objective test design and implementation at each cycle of the spiral and is at the heart of the 'catherine wheel principle'.

Catherine Wheel Principle

> Risk analysis and management generates strategies for risk elimination or reduction that in turn generates the detailed procedures necessary to implement successful computer-based assessment systems. The detailed procedures that characterise the system play a pivotal role in its success and therefore must be derived from the elimination or reduction of risks.

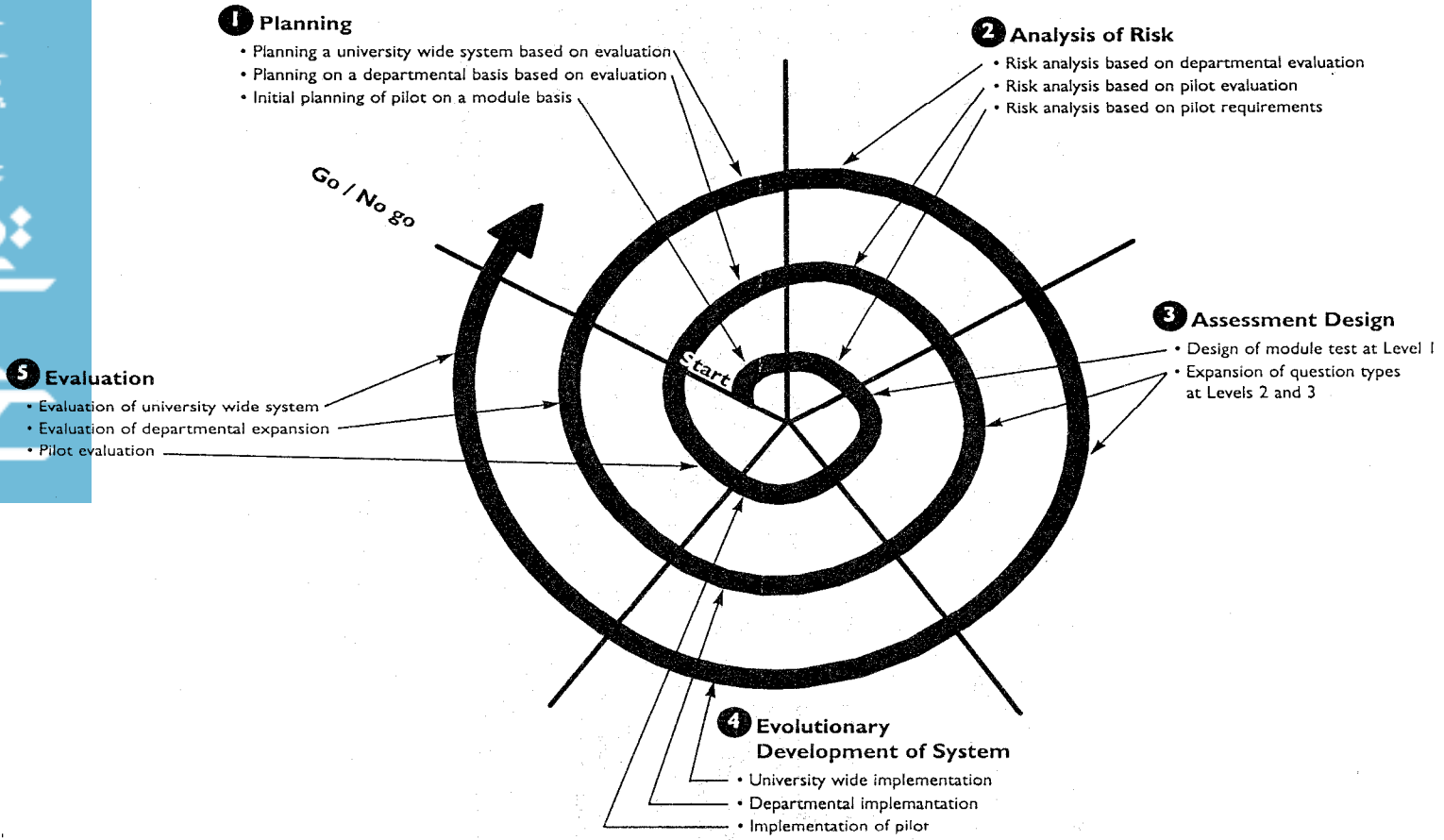

FIG. 1. Computer-based assessment—summative assessment.

Go/No Go

The decision to embark on a pilot may rest on the outcome of an initial demonstration of the software to management, academic and support staff. The objectives of a demonstration are to highlight the potential of the software and to come to a decision regarding the instigation of a pilot. The final decision regarding the running of a pilot will revolve around two factors; first the willingness of at least one member of academic staff to participate fully in a pilot and second substantial support from senior management. Thereafter Go/No go decisions will be made as a result of the evaluation of the previous stages depicted in the model.

Planning Each Stage

The scope, constraints, aims and objectives of each stage of development must be determined. Communications between management academic and support staff can break down and therefore it is essential to establish the direction of the stages and the framework under which they will operate. Personnel and their roles and responsibilities must be identified at each stage. Personnel will include management, academic staff, support staff, administrative staff, external examiners or verifiers, experts in the field and the students themselves. Attention must also be given to integration. The computer-based assessment system cannot be viewed in isolation. It must be integrated into the course design and with the existing institutional assessment procedures. Course documentation and quality assurance procedures will need to be updated to reflect the changes in assessment. Above all the planning of each stage should be based on the evaluation of the previous stage.

Risk Analysis and Management at Each Stage

Risk analysis and Management is conducted at each stage of development of the CBA system. The risk analysis described in this paper refers to university-wide implementation of summative end-of-module examinations in a closed network at the University of Luton.

The first phase analyses the potential risks. Table 1 highlights risks identified using brainstorming techniques. These risks are broken down into pedagogic, operational technical, and financial risks.

Table 2 establishes how much of the system is likely to be affected if a risk event occurred in terms of whether it is an individual student through to the entire university. Who is affected relates to all staff and students involved in the system. The combination of the 'how much' and the 'who' produces the consequence to the system if the risk event took place expressed on a scale from 0 to 1. The values for the consequence to the system in Table 2 are taken from the consequence matrix depicted in Table 3, moderately adjusted for the system at the University of Luton where:

$$\text{Consequence} = (\text{Who}_{\text{Weight}} * \text{How Much}_{\text{Weight}})/(\text{Who}_{\max} * \text{HowMuch}_{\max})$$

and

$$\text{Who}_{\max} = 5$$

$$\text{HowMuch}_{\max} = 6$$

The likelihood of the risk event occurring, if strategies are not put in place to eliminate or minimise the risk, is also stated mathematically in Table 2 on a scale

TABLE 1. Risks and their descriptions

Risk Number	Description
P1	Assessment method not integrated into the curriculum
P2	CBA for a particular degree programme is not acceptable to the relevant professional bodies
P3	Unsuitable or badly designed questions unacceptable to the external examiner
P4	Academic staff leaving the system because of complete test redesign each academic year
P5	Examination contains errors
P6	Students very anxious about a new assessment method
P7	Lack of interest from academic staff
P8	CBA exam unacceptable to a student
O1	Unrealistic schedules for delivery of CBA
O2	Personnel skills shortfalls (Academic)
O3	Personnel skills shortfalls (Support)
O4	Students accessing CBA paper(s) Network security
O5	Students that cannot take the exam on computer or need more time are not accommodated
O6	Computer workstations too close to each other encourage viewing adjacent screens
O7	Module size too large for the number of workstations available
O8	Students start to take the wrong exam
O9	Students arrive late for the examination
O10	Students leaving the examination early
O11	Inadequate access to workstations
O12	Unauthorized access to answer files
O13	Students NOT taking CBA affected by lack of access to computing resources during examination period
O14	Student attends correct exam but wrong session
T1	Workstation collapse during exam
T2	Server collapse during exams
T3	Answer files not being stored or updated
T4	Answer files lost once the examination is completed
T5	Network load too high
T6	A badly designed user front end
T7	An unbelievable result
T8	Computer becomes disconnected
F1	Lack of commitment from Management
F2	Unrealistic budgets for delivery of CBA
F3	CBA system not cost effective

from 0 to 1 by a combination of experience, judgement and external advice during a brainstorming process.

Then in Table 2:

Risk Severity = Consequence * Likelihood

It is also necessary to estimate when the risk event is most likely to occur: before the examination (B), during the examination (D) or post-examination (Po) as this will yield an important timing requirement that will enable the eventual operational system procedures to be ordered. Once completed, the risks are sorted on Risk Severity as shown in Table 2.

All risk events who's Risk Severity is less than a pre-determined Critical Severity Factor (0.04 is the figure suggested by the authors from experience) are then abandoned as being both very unlikely to occur and having low consequences to the system if they did.

TABLE 2. Risk analysis sorted by risk severity

RNumber	How Much	Who	Consequences	Likelihood	Risk Severity	When
O13	University	St	0.99	0.95	0.9405	D
P4	University	Ac/St	0.8	0.95	0.76	B
O6	University	St/Ac/Su	0.8	0.9	0.72	D
O2	University	Ac/M	0.8	0.75	0.6	B/D
O3	University	Su/M	0.8	0.75	0.6	B/D
P7	University	Ac	0.8	0.7	0.56	B
O11	University	St/Ac/Su	0.8	0.65	0.52	D
P3	One department	Ad/Ac	0.66	0.75	0.495	B
T6	University	Su/Ac/St	0.9	0.5	0.45	B
F1	University	M	0.6	0.7	0.42	B
O5	One exam	Ac/Su/St	0.4	0.99	0.396	D
F2	University	M/Su/Ac	0.4	0.75	0.3	B
O7	One group	St/Ac/Su	0.3	0.7	0.21	B
O1	University	Su/M	0.5	0.4	0.2	B/D
T5	University	Su	0.4	0.5	0.2	D
P1	One exam	Ac/St	0.33	0.6	0.198	B
F3	University	M/Su/Ac	0.4	0.45	0.18	B
P5	One exam	Ac/St	0.3	0.6	0.18	D
O4	University	Ac/Ad/Su/St	0.6	0.3	0.18	B/Po
O10	One student	St/Ac/Su	0.2	0.85	0.17	D
O9	One student	St/Ac/Su	0.2	0.8	0.16	D
P2	One department	Ad/Ac/St	0.6	0.25	0.15	B
T4	One group	St/Ac/Su	0.5	0.3	0.15	Po
P6	University	Ac/St	0.9	0.15	0.135	B
O8	One exam	St/Ac/Su	0.3	0.4	0.12	D
O12	One exam	St/Ac/Su	0.3	0.35	0.105	Po
T3	One group	St/Ac/Su	0.5	0.2	0.1	D
T1	One student	St/Ac/Su	0.2	0.4	0.08	D
T2	One group	St/Ac/Su	0.5	0.1	0.05	D
T7	One student	Ac/Su	0.2	0.2	0.04	Po
O14	One student	Su/St	0.17	0.15	0.0255	D
T8	One student	Su/St	0.07	0.05	0.0035	D
P8	One student	Ac/St	0.17	0.01	0.0017	B

The second phase manages the remaining risks. Columns 1, 6 and 7 are brought forward from Table 2 and sorted on when the risk is likely to occur. Strategies for eliminating the risks or at least minimising the likelihood of them occurring and minimising their impact on the system are then decided upon. The result of risk management is shown in Table 4. The adopted strategies for risk elimination or reduction are in a timed sequence from which implementation procedures and responsibilities are generated.

Let us take four examples from the University of Luton.

1. Unsuitability of badly designed questions unacceptable to the external examiner (P3)

The design of effective objective questions is an acquired skill and many staff will need help and advice in developing these skills. A staff development session to describe CBA at the university and explore the pedagogic issues surrounding the construction and use

TABLE 3. Consequence matrix

		Who				
		Administration	Support	Management	Academics	Students
How Much		1	2	3	4	5
1 student	1	0.03	0.07	0.10	0.13	0.17
1 exam	2	0.07	0.13	0.20	0.27	0.33
1 group*	3	0.10	0.20	0.30	0.40	0.50
1 Dept	4	0.13	0.27	0.40	0.53	0.67
1 Faculty	5	0.17	0.33	0.50	0.67	0.83
University	6	0.20	0.40	0.60	0.80	1.00

Note: 1 group refers to a set of students taking examination(s) at the same time.
Values to be interpreted by the user (by judgement or statistical averaging).
Grid to be used as a GUIDE only.

of objective tests is therefore a key element of the strategy. This strategy can be translated into a formal system procedure.

- All academic staff involved in CBA will go through a staff development programme.
- The programme will be launched by the Staff Development Unit (SDU) and will form an integral part of the annual staff development programme of the university.
- The staff development programme will be incorporated into the annual staff development events diary.
- No academic member of staff will be able to undertake CBA without attending the staff development session.

2. Students anxious about a new assessment method (P6)

Introducing a new method of assessment causes student anxiety and must be viewed as a pedagogic risk. Strategies must therefore be deployed to eradicate this risk. Part of the strategy to alleviate student anxiety is to introduce sample questions on the network before the examination starts. These sample questions, perhaps 7 or 8 questions, will be available to students on an open access basis and would serve a dual purpose, first to introduce the students to the differing question types and second to enable students to get used to the technology. The sample questions would have to be accompanied with an instruction set. This strategy can be translated into a formal system procedure:

- Sample questions will be designed by academic staff.
- Four weeks before the commencement of examinations all sample questions will be mounted in the central IT suite by computer services.
- The learning technology department will provide student instructions to take the sample questions.
- Student instructions that will be made available behind the counter in the IT suite.

3. Workstation collapse during exam (T1)

In the central IT suite 5% of workstations comprise a contingency area. Should a workstation collapse during an examination, the student who is affected is either directed to the contingency area or given a hard copy of the examination. This strategy can be translated into a formal system procedure.

TABLE 4. Risk management—strategies to be adopted.

RNumber	Risk Severity	When	Possible Strategy	Strategy Adopted
O4	0.18	B/Po	A. Set policy for installation and deletion of exams from servers B. Reconcile issues and returns of any hard copies C. No shelving of hard copies in the library D. Exam installed on server immediately prior to start of exam with invigilating staff in exam area E. Protected test directories on server F. Dedicated server for examinations	″A,B,C,D,E″
O2	0.6	B/D	A. Specify roles and responsibilities of academic staff B. Design and implement a staff development programme C. Pre-Schedule key personnel	″A,B,C″
O3	0.6	B/D	A. Specify roles and responsibilities of support staff B. Staff with top talent C. Design and implement a staff development programme D. Pre-Schedule key personnel	″A,B,C,D″
P4	0.76	B	A. Ensure that past CBA papers are not available to students B. Devise question update policy with Quality Assurance division C. Amend regulations at departmental or university level	″A,B,C″
P7	0.56	B	A. Demonstration to academic staff B. Integration of CBA into University strategic plan C. Ring-fenced funding D. Integration with existing QA procedures	″A,B,C,D″
T6	0.45	B	A. Adopt national/international HCI standards B. Investigate 'add-on' software	B
F1	0.42	B	A. Demonstration to senior managers B. Integration of CBA into University strategic plan C. Ring-fenced funding D. Integration with existing QA procedures	″A,B,C,D″
F2	0.3	B	A. Breakdown costs into definitive categories B. Requirement Scrubbing C. Evolutionary development	″A,B,C″
P3	0.495	B	A. Initiate internal staff development B. Conduct literature searches on effective question design C. Invite guest speakers D. Join appropriate user group E. Support from a central Unit	″A,B,C,D,E″
P6	0.135	B	A. Introduce sample questions on network 3/4 weeks prior to exam on open access basis B. Good instruction set C. An overview of the system is given to students D. Q and A sessions	″A,B,C,D″
P2	0.15	B	A. Introduce the assessment method at a level/year that does not count towards award of degree B. Abandon the assessment method for this module and consider other programmes C. Negotiate with professional bodies	″B,C″

TABLE 4. Risk management—strategies to be adopted.—*continued*

RNumber	Risk Severity	When	Possible Strategy	Strategy Adopted
O7	0.21	B	A. Conduct workstation audit before exams are scheduled B. Ensure maintenance on workstations completed C. Adopt a 2-group policy. The first group leaves as second group enters exams area D. Design a second exam E. Invest in additional workstations	″A,B.C″
P1	0.198	B	A. Update course documentation B. Seek support from the relevant academic standards committee	″A,B″
F3	0.18	B	A. Academic Staff time analysis B. Cost/ Benefit analysis	A
O13	0.9405	D	A. A fully integrated CBA system B. Students informed well in advance of closure of computing areas for exam purposes	″A,B″
O6	0.72	D	A. Timetable different exams adjacent to each other B. Provide privacy screens C. Provide shuffling capabilities for questions and/or distractors	″A,B,C″
O5	0.396	D	A. Additional hard copies are provided B. Additional computer time is provided	″A,B″
O11	0.52	D	A. Booking policy for computer resources B. Computing resources used for no other purposes than delivering exams during exam weeks	″A,B″
T3	0.1	D	A. Understand how the a software operates B. Flag set for answer-file storage is a check-list item C. Test storage of answer files with fictitious user name 15-20 minutes prior to exam D. Check for correct number of answer files and current scores after commencement of exam E. Issue hard copy of exam F. Reschedule exam	″A,B,C,D,F″
O1	0.2	D	A. Breakdown schedule into definitive categories B. Critical Path Analysis C. Requirement Scrubbing D. Evolutionary development	″A,C,D″
T2	0.05	D	A. Reschedule exam with resit paper and design new resit paper B. Adopt a 2-server policy C. Disk- mirroring	A
T5	0.2	D	A. Load testing B. Upgrade network	″A,B″
O10	0.17	D	A. Design a policy for leaving early for one and two group sessions B. Integrate policies for leaving early into Quality Assurance procedures from traditional exams	″A,B″
O9	0.16	D	A. Design a policy for lateness for one and two group sessions B. Integrate policies for lateness into Quality Assurance procedures from traditional exams	″A,B″
O8	0.12	D	A. Use student ID as username B. Academic staff activate the correct exam from a presenting menu having seated students C. Use additional software to register students for exam	″A,B″

TABLE 4. Risk management—strategies to be adopted.—*continued*

RNumber	Risk Severity	When	Possible Strategy	Strategy Adopted
T1	0.08	D	A. Leave a certain percentage of workstations free B. Provide hard copy backups.	"A,B"
P5	0.18	D	A. Academic staff check paper for errors in text B. Academic staff check correct flagging of right answer	"A,B"
T4	0.15	Po	A. Provide answer file backup B. Reschedule exam with resit paper and design new resit paper	"A,B"
T7	0.04	Po	A. Provide software with comprehensive reporting facilities	A
O12	0.105	Po	A. Answer files collated and deleted from servers immediately after exam ends B. Protected answer file directories on server C. Dedicated server for examinations	"A,B"

- If there is a crash on an individual workstation within 15 minutes of the start of the examination the student is asked to move to a different machine and restart. The 15 minute time limit has been set to allow students to still complete the examination on computer before a second group, if any, takes the examination.
- If there is a crash on an individual workstation after 15 minutes have elapsed, the student will be accompanied to the Examinations Office where a decision will be made as to whether (a) a hard copy is given to the student, (b) the student takes the examination in another group or (c) the case is referred to the Examination Board.

4. Answer files not being stored or updated (T3)

Student answer files not being stored or updated is a technical risk and must be eliminated. The strategy to eliminate the risk involves the use of a checklist to ensure that the appropriate flag is set to store the answer files for that examination together with a formal systems test to ensure that the answer files are being stored prior to the students entering the examination arena. Once the examination has started the number of answer files are compared with the number of students sitting the examination.

This strategy may be translated into a formal system procedure:

- The learning technology department will use a checklist to ensure that answer file flags are set for each test in the test control information block.
- This checklist will be signed by the head of the learning technology department.
- The learning technology department will conduct a formal test to ensure answer files are being stored on the examination day. This test will be conducted 15 minutes prior to the students entering the IT suite.
- A formal count will be made 5 minutes after the start of the examination and the number of students sitting the examination matched with the number of answer files.
- A computer seat number and student identification number will uniquely identify and pinpoint a mismatch if it occurs and the student will be asked to restart the examination.

Assessment Design at Each Stage

The effective creation of objective tests will involve:

- staff development;
- test specifications;
- test design.

It is important for all academic staff involved in CBA to go through a staff development programme which forms an integral part of the annual staff development programme of the university. The aim of the programme is to describe the current state of the CBA system at the university and to explore pedagogic issues surrounding the construction and use of objective tests. The elements of the staff development programme will be reflected through the pedagogic risks identified during risk analysis and management. Academic staff will also need updating and therefore an annual update programme should be implemented.

An important element in assessment design is test specification. Academic staff will need to be in a position to consider the assessment topics, their learning outcomes and from these identify levels of learning (Bloom's Taxonomy) for each topic in their assessment. The test specification will inform academic staff on the balance of their tests in terms of the number of questions committed to assessing a particular level of learning (Heard *et al.*, 1997). Having constructed the balance of the test, academic staff will then be in a position to consider how to test the learning outcomes in each topic. Different question types will need to be considered and the answer(s) to the questions will determine the question type used.

Academic staff will design their tests from the test specification ensuring reliability and validity. Questions should be not be ambiguous or trivial. They should reflect 'current thinking' in the subject discipline and test relevant knowledge and skills.

Evolutionary Development at Each Stage

Complete systems are required at evolutionary stages from pilot to departmental expansion and finally to university-wide systems. The procedures for implementation, derived from Risk Analysis and Management, will now be put into practice. These procedures will change as the system expands because the risks and/or their impact may change with expansion. Nevertheless certain commonality is inevitable. Roles and responsibilities will need to be clearly identified. The documentation on all system and quality assurance procedures must be written. A training plan for the support of the system must take place which will include both technical support and academic duties. The physical site must be prepared and formally tested which must include both backup and security measures.

The candidate for the pilot should emerge from a successful demonstration. It is important to achieve system reliability and system simplicity during pilot implementation is a key factor in achieving reliability. The trial group should be of manageable size and the test limited in its multimedia elements and suitable for first-year undergraduates.

If departmental expansion takes place additional personnel will be involved, documentation to support the system will increase and a strategy to accommodate multi-assessments will be evolved. The staff development unit will play a more active role. Questions will become more challenging as they begin to test or develop higher order cognitive skills and a user group could be formed to exchange good practice

and learn from the experiences of external consultants or guest speakers invited to the department. Administrative tasks will increase and it may be necessary to train departmental personnel in the correct procedures that facilitate integration and the production of the tests in electronic form. A senior member of staff in each department may be given the responsibility of controlling test production and delivery and managing the quality assurance processes.

Whereas the second stage can be viewed as an expansion, the third stage is a significant system change as the move is made from what can be viewed as a number of de-centralised departmental systems to a centralised university-wide system. The systems at this stage would be fully integrated with the traditional assessment system of the university. A central examinations office may be responsible for timetabling and scheduling the examinations and a quality assurance department would take responsibility for the integrity of the processes. A large centralised system may very well be supported by a research-based assessment team.

Evaluation at Each Stage

A formal evaluation of the system must take place at each stage of development at the end of each academic year. Student and staff perceptions are paramount. The outcome is a review report that not only covers quality assurance issues and recommends changes but also makes a recommendation for expansion to the next stage based on pedagogic, operational technical and financial feasibility. The review report may cover:

- suitability of the assessment method;
- the comparative evaluation of CBA achievement against traditional assessment formats;
- the review of the effectiveness of staff training support systems;
- the evaluation of CBA administration;
- staff and student comments on CBA;
- recommendations for improvements to the structure and format of CBA, staff support systems, administration and control systems;
- the CBA external comments (based on the review of documentation, sampling of CBA specifications and examinations and discussions with relevant staff);
- costs and benefits;
- feasibility study for the next stage of development.

Conclusion

The ‘catherine wheel’ is a generic model. It is applicable equally to summative or formative assessment on closed network systems, WEB-based systems, OMR or OCR systems. The model itself does not change, only the risks drawn from the model change when applied to differing assessment systems. At the University of Luton the formative assessment system grew naturally from the summative system on a closed network resulting in the complete ‘catherine wheel’ depicted in Figure 2.

Notes on Contributors

DR STAN ZAKRZEWSKI was Computer Adviser at the University of Luton before becoming Head of Learning Technology. He now manages and researches the

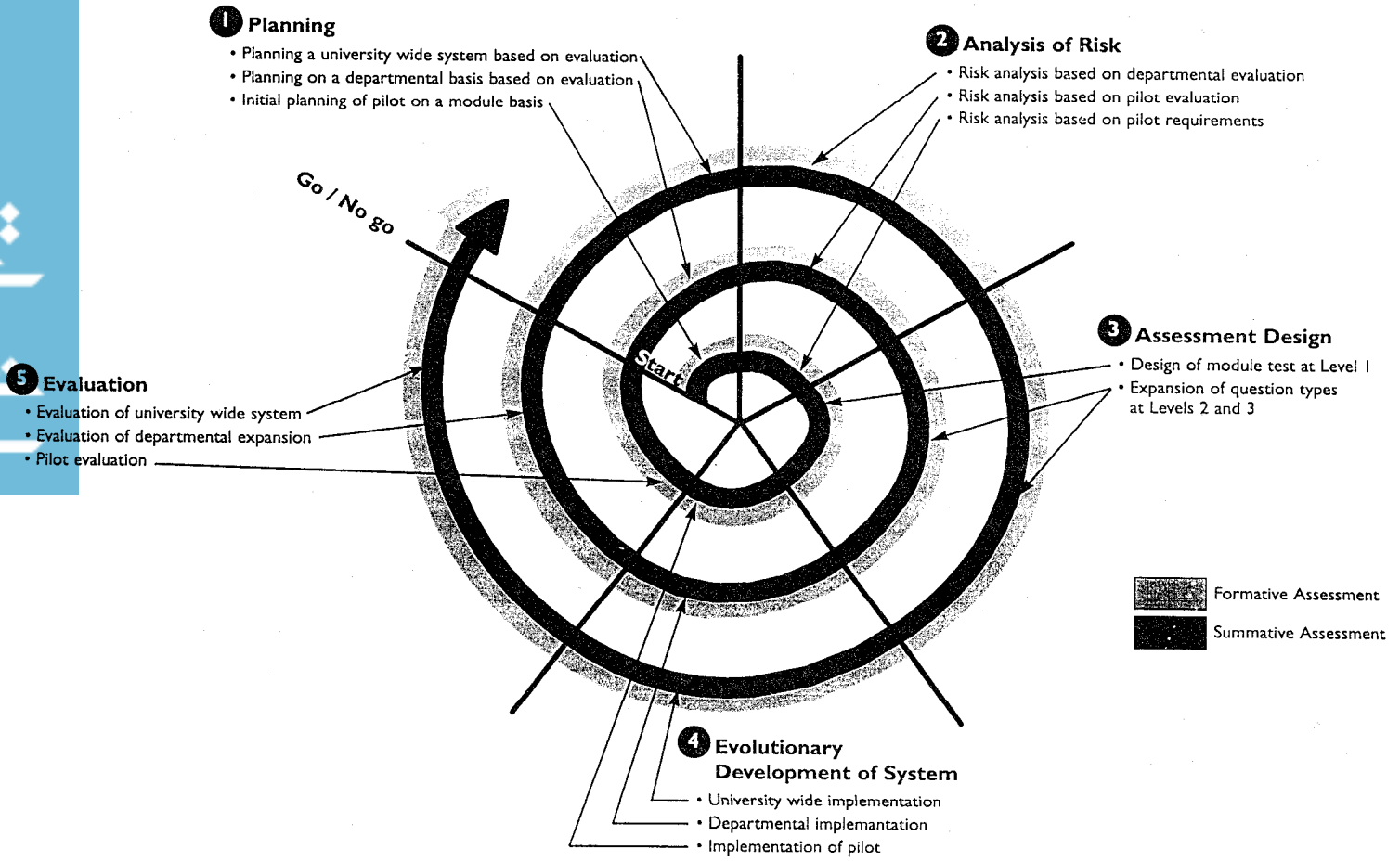

FIG. 2. Computer-based assessment model—summative and formative assessment at the University of Luton.

implementation of learning technology. *Correspondence:* Dr Stan Zakrzewski, Head of Learning Technology, University of Luton, Park Square Campus, Luton, Bedfordshire LU1 3JU, UK

CHRISTINE STEVEN was until recently a Principal Teaching Fellow in the Faculty of Science, Technology and Design. She now combines teaching computer science part-time with her research into assessment techniques.

REFERENCES

BASILI, V. R. (1991) The future engineering of software: a management perspective, *Computer*, 24(9), pp. 90–96.

BOEHM, B. W. (1988) A spiral model for software development and enhancement, *IEEE Computer*, 21(5), pp. 61–72.

BOEHM, B. W. (1989) *Risk Management Practices: The Six Basic Steps* (Washington, IEEE).

Charette, R. N. (1989) *Software Engineering Series* (New York, McGraw Hill).

HEARD, S., NICOL, J. & HEATH, S. (1997) *Setting Effective Objective Tests* (University of Aberdeen, MERTal Publications).

PRESSMAN, R. S. (1997) *Software Engineering—A Practitioner's Approach,* 4th edn European Adaptation (London, McGraw Hill).

PRITCHETT, N. & ZAKRZEWSKI, S. (1996) Interactive computer assessment of large groups: student responses, *Innovations in Education and Training International*, 33(3), pp. 242–247.

ZAKRZEWSKI, S. & BULL, J. (1998) The mass implementation and evaluation of computer-based assessments, *Assessment and Evaluation in Higher Education*, 23(2), pp. 141–152.