An Action Research Approach to the Design, Development and Evaluation of an Interactive E-Learning Tutorial in a Cognitive Domain

M R (Ruth) de Villiers
University of South Africa, Pretoria, South Africa

dvillmr@unisa.ac.za

Executive Summary

(This work represents an update to research published at the 2006 Informing Science and Information Technology Education Conference. De Villiers, M.R. (2006). Multi-Method Evaluations: Case Studies of an Interactive Tutorial and Practice System. Available at http://proceedings.informingscience.org/InSITE2006/ProcDeVi225.pdf)

The teaching and learning of a complex section in *Theoretical Computer Science 1* in a distance-education context at the University of South Africa (UNISA) has been enhanced by a supplementary e-learning application called *Relations,* which interactively teaches mathematical skills in a cognitive domain. It has tutorial and practice functionality in a classic computer-aided instruction (CAI) style and offers considerable learner control. A participative action research approach was used to design, develop, evaluate, and refine the application over a longitudinal period. In this process the application was formatively and summatively evaluated by different methods – questionnaire surveys, interviews, heuristic evaluation and a post-test. This article explains the purpose, structure, and operation of *Relations* and notes how the various evaluation methods resulted in iterative refinements to its functionality, learning content, and usability. The findings lead to reflection.

Conventional computer-aided instruction and learning (CAI/L) has a role to play in the milieu of e-learning. CAI can present efficient instruction, motivate and engage learners, challenge them with meaningful exercises, and can support effective learning. The students requested more such tutorial and practice environments. *Relations*' greatest strength is its excellent diagnostic feedback, attested to by learners and expert evaluators alike. Courseware authoring systems have powerful facilities that can be used to judge the learner-input and provide appropriate, detailed, tailor-made feedback. This can be done in web-based learning (WBL) too, using specialized web-programming languages, but it is more complex. A further obstacle to the use of WBL at UNISA is that many UNISA students still lack broadband Internet access.

The designer and developers of *Relations* used technology, not for its own sake, but rather to motivate and to illustrate concepts in ways that enhance cognition. Technology should be the medium and not the message. The blue-water recreational theme was well received by the majority of learners and expert evaluators, who acknowledged its role in providing brief interludes of diversion and relaxation in a demanding cognitive domain.

The concepts of usability and interaction design from the discipline of human-

computer interaction (HCI) are receiving increasing attention in the development of commercial and corporate software. It is equally important to produce usable applications in educational contexts where the users are not professionals in the workplace, but learners who must first be able to use and interact with a system before they can even commence learning. Many learners approach e-learning after exposure to commercial software. As far as possible, learning applications should use operations and keypresses that support the HCI principles of predictability and consistency with familiar systems. In certain respects *Relations* falls short in this respect, but it was found to be easy to learn and use. It adheres to the fundamental principle of internal consistency, where a system's own internal operations are characterized by predictability and visibility.

This study offers lessons for new CAI environments. The teaching approaches in *Relations* present multi-perspectives on a concept and multiple representations of content. These methods were successful and could be supplemented by multi-modal presentations in contexts where they would add value. There was a degree of learner-control, permitting choices about sections, sequence, repetition, and whether to omit theoretical segments and head straight for practice. The hyperlinks incorporated into *Relations V2.2* were a success; these hotspots provide direct access to definitions, elaborations and cross references.

Concerns were raised regarding some aspects of control and navigation. Where pedagogically appropriate, linear sequences should be supplemented with hyperlinked access. Exit and re-entry facilities are required. User-control should be provided over the level of difficulty, so that users can choose particular exercises. This would facilitate different ways of usage at different times. The control structures deactivate certain buttons in particular situations. Although this was designed for appropriate contexts, some users objected. Ways of implementing such tactics must be carefully chosen, as designers exercise discernment in deciding when to provide the facilities that users request and when to overrule for pedagogical reasons.

The multi-method evaluation of *Relations* had positive results. The complementary use of four evaluation techniques and triangulated data provided a synergic framework.

Heuristic evaluations by experts were valuable sources of critique and suggestions. Two evaluators who are leading subject matter experts tested the feedback to its limits, revealing flaws and compromises of mathematical rigour. The experts moved beyond the given criteria, identifying omissions and suggesting further features. The discernment, insight, and long-standing expertise of the heuristic evaluators accurately pinpointed the strengths and weaknesses of *Relations*, as well as subtle mathematical inaccuracies. For overall evaluation, the team of evaluators should include experts with subject matter skills and those with usability expertise.

The students' questionnaire was particularly useful with regard to educational aspects. It confirmed the value of the detailed feedback, showing that the time spent painstakingly developing it was worthwhile. The 2004 questionnaire survey elicited particularly useful information and identified certain errors that were corrected even before *Relations* was presented to the heuristic evaluators. The inclusion of open-ended questions for qualitative responses served to elaborate responses, to enrich the information obtained, and to motivate the Likert options selected. In many cases, students used these to express praise and appreciation of *Relations*. Unanticipated aspects that emerge can be probed further in interviews or subsequent questionnaires.

The interviews among end-users (students) explored further avenues, following up on usability problems. The interviews probed for explicit identification of the features that fostered learning for each individual and the features they enjoyed most. UNISA's distance-teaching context makes contact difficult and the sample of volunteers used in this study represented a fairly homogenous group of young fulltime students, typical of the shift in learner profile. However, contact with a stereotypical group is informative in its own right. To optimize on interviews, the structure was flexible. The semi-structured interview was based on a set of common core questions, with follow-up questions prompted by the participants' responses.

The post-tests gave quantitative measures that could be statistically analysed. However, scores in tests remain debatable instruments for measuring the effectiveness of an artefact or other intervention in improving learning. Tests frequently show no significant difference between methods of instruction, even when other evaluation methods generate favourable reports of the intervention. When a test, then, does indicate improved academic performance, it would appear that the method under investigation indeed enhances learning. This was the case in the present evaluation.

The participative action research process of designing, developing, evaluating, and refining the e-learning tutorial *Relations* has taught the designer and development team a great deal – not just about the application being studied, but also lessons and generic principles for the design of e-learning applications as well as lessons about the rich complementary roles of different evaluation methods and techniques.

Keywords: Action research, cognitive domain, computer-aided instruction, e-learning, human-computer interaction, learner control, theoretical computer science, triangulated evaluation methods, usability evaluation

Introduction

This article describes and discusses *Relations,* an interactive e-learning application that teaches mathematical skills in a cognitive domain. *Relations* is a supplementary computer-aided lesson that supports learning in a complex section of *Theoretical Computer Science 1* at the University of South Africa (UNISA). *Relations* is highly interactive, with tutorial and practice functionality and a considerable degree of learner control. In an iterative action research approach, development was interspersed with evaluations by different techniques: a questionnaire survey, interviews, heuristic evaluation, and a post-test. The findings with regard to functionality, usability, and contribution to learning were fed back into corrections, refinements and extensions. The article introduces *Relations* and its context of use. Action research, which was the development approach and underlying research framework, is outlined. The underlying theory, the content, and the operation of *Relations* are explained and illustrated. Findings of the four separate evaluation studies are mentioned. Conclusions are drawn about *Relations*, which lead to generic reflection on the design of computer-aided instruction and e-learning applications, as well as on the roles of the various evaluation techniques.

Background

Perspective on e-Learning

There are various perspectives on what is meant by 'e-learning'. Some definitions equate it only with the use of the Internet in instruction and learning, but others (CEDEFOP, 2002; Wesson & Cowley, 2003) are more inclusive. CEDEFOP's (2002, p.5) definition describes it as 'learning that is supported by information and communication technologies [which]…may encompass multiple formats and hybrid methodologies, in particular, the use of software, Internet, CD-ROM, online learning, or any other electronic or interactive media.' Such approaches include the Internet and Web-based learning (WBL), multimedia CD-ROM, online instruction, educational software/courseware, and traditional computer-aided instruction (CAI) and computer-assisted learning (CAL). The present author holds this approach and views e-learning as a broad range of learning technologies encompassing the various roles for technology given above, also including learning management systems (LMSs), as well as learners using computers as tools and for communication (De Villiers, 2005b).

Introducing 'Relations'

Relations is an interactive multimedia application about mathematical relations. Mathematical relations are a sub-area within discrete mathematics, a cognitive domain that demands analytical skills from the learners. It has tutorial functionality in the sense of a computer-based tutorial as described by Alessi & Trollip, (2001) and gives the learner extensive practice opportunities. The *Relations* CAI lesson is not a Web-based learning environment; rather it is a stand-alone CD-based system, which is not Windows-driven. It was custom-built in the School of Computing at UNISA as optional supplementary study material for the first-level module *Theoretical Computer Science 1,* which teaches discrete mathematics relevant to the foundations of theoretical computing and databases. Version 1 (V1) of the lesson was originally produced in 1994 and distributed on stiffie diskettes. It underwent learner evaluation and expert evaluation, was positively received, and was used for eight years. By 2003 it was time for a revamp. The general approach and subject matter were re-used, but the content was extended and the usability enhanced. HCI aspects, such as fonts, backgrounds, screen displays, colours, colour-coding, and interaction design, were improved and general quality was enhanced. It was developed using the multimedia authoring system, *Quest 7.0*, a tool for the creation of computer-based training courses. *Relations'* navigation is mainly linear in the traditional style of a tutorial, rather than hierarchical. However, there is learner control over which sections to do, the sequence of sections, and whether to omit theory and do exercises only. During the 2003 re-development process it was continuously formatively evaluated by lecturers and multimedia developers in the School of Computing. Following this redesign (V2.1), it was released in 2004 and comprehensively evaluated by both end-users (learners) and experts, leading to further refinements and evaluations in 2005 (V2.2), including hyperlinks to definitions and elaborations.

Context-of-use and Target Learners

UNISA is a distance-teaching institution with students worldwide. It is one of the world's mega-universities. Most students live in South Africa, but there is a focused outreach programme to the rest of Africa. Until recently most students were employed fulltime and studied part-time. The average age was mid-30s. The present learner profile is changing and many are young, fulltime students, mainly from formerly disadvantaged groups. Some of the young students study, not at home or in the workplace, but in UNISA's regional learning centres. Communication is primarily by conventional mail, because some have limited technological access. Computing students, by definition, must have access to PCs, but do not always have Internet access. For this reason, the redesigned version was implemented, not as WBL, but once again as a CAI lesson. The redesign was distributed on a CD, whereas V1 was produced on a stiffie diskette.

Between 800 and 1000 students register annually for *Theoretical Computer Science 1*, of whom about half take the module through to completion and write the exam. The mandatory study material comprises a prescribed book, a study guide, and regular tutorial letters. Assignment questions and examinations involve precise and rigorous written mathematical proofs and procedures. Detailed solutions to the assignments are subsequently provided, as well as answers for the previous year's examination questions. The CAI lesson *Relations* is optional study material that can be purchased at a reasonable price and is available for free use in UNISA's computer labs throughout South Africa. Between 30% and 40% of the students use *Relations*.

Learning Theory Underlying Relations

The learning theory ethos (De Villiers, 2005b; Reeves & Reeves, 1997) underlying an educational software application or a web-learning environment should be explicit in its design. The present researcher is the designer of *Relations* and was the module leader of *Theoretical Computer Science 1* for ten years. She is aware of the domain complexities and problems experienced by students. Her preferred pedagogical position has a social constructivist ethos. However, due to the rigid nature of its subject matter, *Relations* has fundamentally objectivist and instructivist tendencies. It has a classic CAI tutorial structure, alternating teaching and practice segments, and adopts a behaviourist stance in its stimulus-response-reinforcement approach. Yet it also embodies eclecticism, combining paradigms where appropriate.

It is operated mainly by learner-control rather than by program control. Although branching paths exist, learners choose these options themselves and are not placed on a particular route as a result of performance in a pre-test. Opportunities were taken to motivate and engage learners, to support true cognition, and to anchor the learning process in authentic contexts. Constructivist multi-perspectives are employed to portray concepts by means of text, figures, graphics, limited audio, evolving animations, and many examples. Some of the interactive exercises require synthesis skills as learners generate their own relations.

Relations does not include learning management or recording of progress, since it is intended for learning gain and not for learning measurement. The feedback provides integrated assessment, and the optional test is a capstone. For innovation and engagement, a blue water theme, comprising informal 'recreational' screen displays, is used after complex sections to help students take a break. The water analogy also offers a 'Dive' option, allowing users to omit the theory in a section and dive straight into the examples. This is useful for the high performers and is also a useful form of revision just before writing an examination. However, where a learner does choose to do the theory, they are required to work right through it.

Research and Development Process of Relations

An action research (AR) approach (De Villiers, 2005a) was used to develop *Relations*. Action research emanated in the 1940s from action-based social psychology by Kurt Lewin of the University of Michigan, who contended that complex social events cannot be investigated under laboratory conditions. AR encompasses a variety of research techniques and methods and generates both action outcomes and research outcomes. Its participative, practitioner-researcher approach lends itself to educational research, where developing interventions/products are investigated over several cycles. Commencing with the identification of a problem or situation that calls for action, AR functions as a change agent, and is (Baskerville & Wood-Harper, 1996; Cohen, Manion, & Morrison, 2000; Dick, 1999; Du Poy & Gitlin, 1998):

- *Cyclic*: iterative steps recur longitudinally, generating information for further action. Zuber-Skerrit (1992) terms its repetitive processes: *plan, act, observe,* and *reflect*.
- *Participative*: clients, end users, and researcher collaborate as co-researchers. AR is often used when practitioner-researchers examine their own work, giving the researcher a central role in both the research and the development processes.
- *Qualitative*: it operates more by verbal aspects than by numbers, although quantitative methods are also used and acceptable.
- *Reflective*: critical reflection on the process and outcomes is vital to each cycle, and is used in designing subsequent steps and events.
- *Responsive*: it reacts and adapts flexibly to the findings from each previous cycle.

An AR model, as shown in Figure 1, was used in the iterative research, development, and formative evaluation of *Relations.* The series of cycles was characterized by *central, in-depth involvement of the researcher-designer* (also the module leader) and participative input by peers and students. The diagram shows that each cycle within the series is characterized by the processes of plan, act, observe-and-evaluate, and reflect, derived from Zuber-Skerrit (1992).

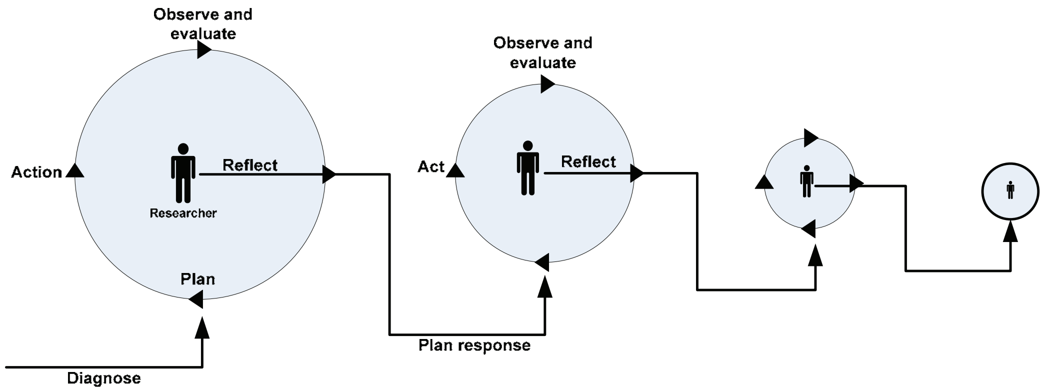

Figure 1: Action research model (De Villiers, 2005a)

Both formative and summative evaluations were conducted. During design and development *formative evaluation* (Preece, Rogers, & Sharp, 2002) was done by the development team, which comprised the course leader (who designed *Relations*), computer programmers, other lecturers, and a graphical designer. All the processes and mathematical exercises were comprehensively tested during development, both when the original V1 was produced in 1994 and again with V2 in 2003. Each set of evaluations led to reflection and responses in the form of corrections and improvements. The programmer has mathematics expertise as well as programming skills, which added value to her contribution. There was a strong designer-developer partnership throughout.

Further cycles entailed post-release *summative evaluations* (Preece et al, 2002) undertaken in 2004 and 2005 to examine, inspect, and improve the finished product. Such evaluations can also be termed *effectiveness evaluations* and *impact evaluations* (Reeves & Hedberg, 2003). The findings of these evaluations were used to add refinements, improve usability and incorporate some new minor features, making them into '*formative- summative evaluations'*. These multi-method evaluations are described in a subsequent section, while the section that follows immediately describes, discusses, and illustrates the learning content and the approach of *Relations*.

Learning Content and General Approach of Relations

Relations comprises three sub-lessons: *Background information*, *Properties of relations*, and *Special kinds of relations*. Each has its own learning-content subsections and learners use non-sequential navigation to select:

- which to do,
- in any sequence,
- how many times to tackle each – sections may be done as often as the learner chooses, and
- whether to repeat individual screens.

There is also a fourth section, an optional test, called *Test Yourself*. Figure 2 shows the structure of *Relations* and the navigation options.

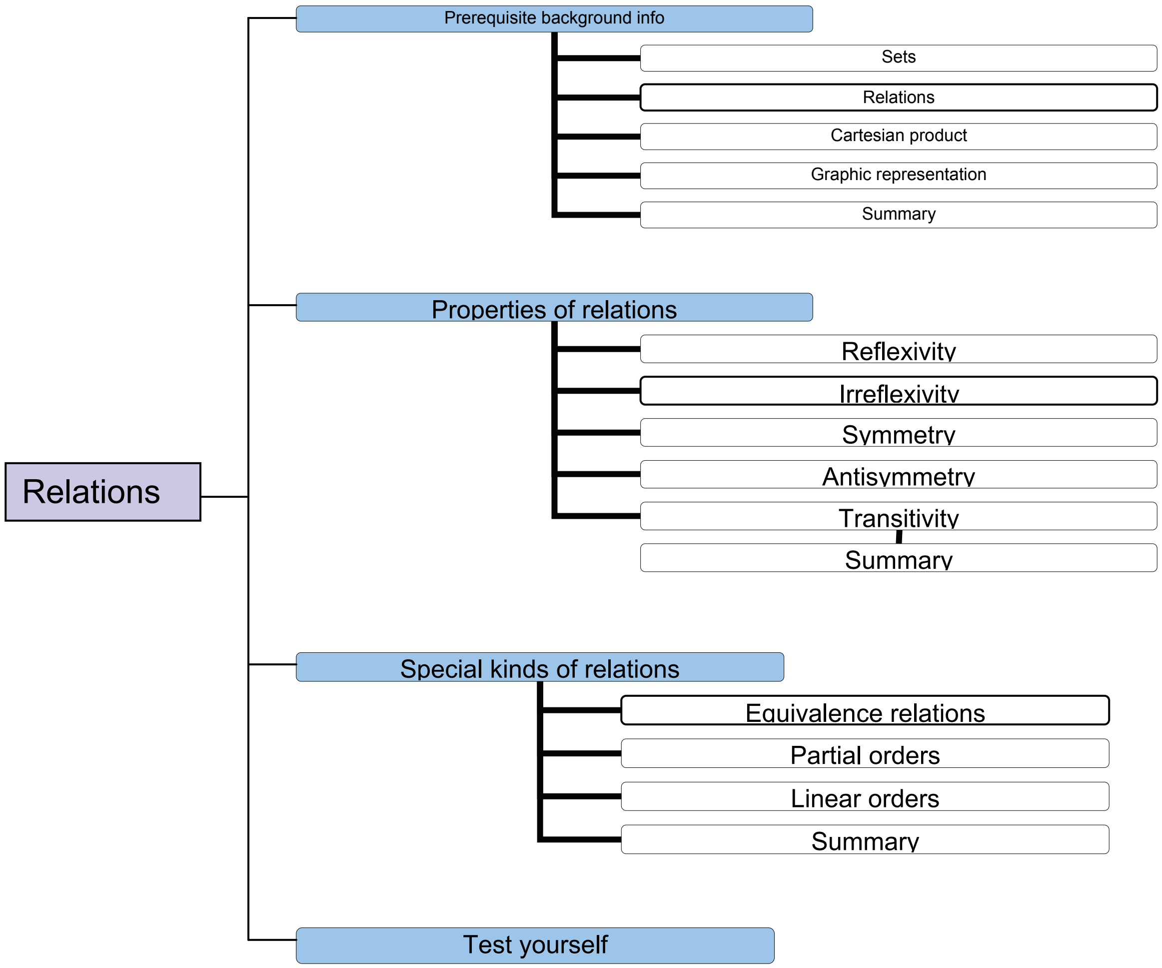

Figure 2: Structure of *Relations*

Within each section definitions, theoretical concepts, and examples are presented, illustrated by graphics, pictures, and animations that show step-by-step unfolding of a definition or process. Tutorial segments are interspersed with exercise-and-practice segments. Each major section ends with a revision summary. There are no multiple-choice questions, but a few Yes/No questions. Most of the questions are Fill-in-the-blank. Some questions are basic, but most entail a composite answer or entry of a series of mathematical characters and notations. Some of the exercises involve synthesizing a relation to meet a particular set of conditions. Due to the range of possible errors, there is considerable variation in students' input, yet *Relations* provides them with detailed feedback, and second attempts must be made after wrong answers. This diagnostic feedback with explanations is one of the lesson's greatest strengths, implemented mainly by pattern matching but also by some elementary artificial intelligence techniques. *Relations* teaches computational concepts from discrete mathematics, i.e. rigid concepts with no scope for personal interpretation. The rules of this 'well-structured domain' (De Villiers, 2005b; Jonassen, 1999) are tightly defined and each problem has a single solution or a finite set of answers that would judged as correct. As already mentioned, *Relations* makes provision for different stages of learning, giving the option to omit theory and 'dive' straight to sets of exercises or revision screens. Its purpose is learning support; it is non-threatening and does not grade learners or record progress. The optional test covers all aspects, but the score is not stored. Where appropriate, background information is presented at the beginning of a section to relate the new knowledge to prior learning. Several of the features described are illustrated in the following figures.

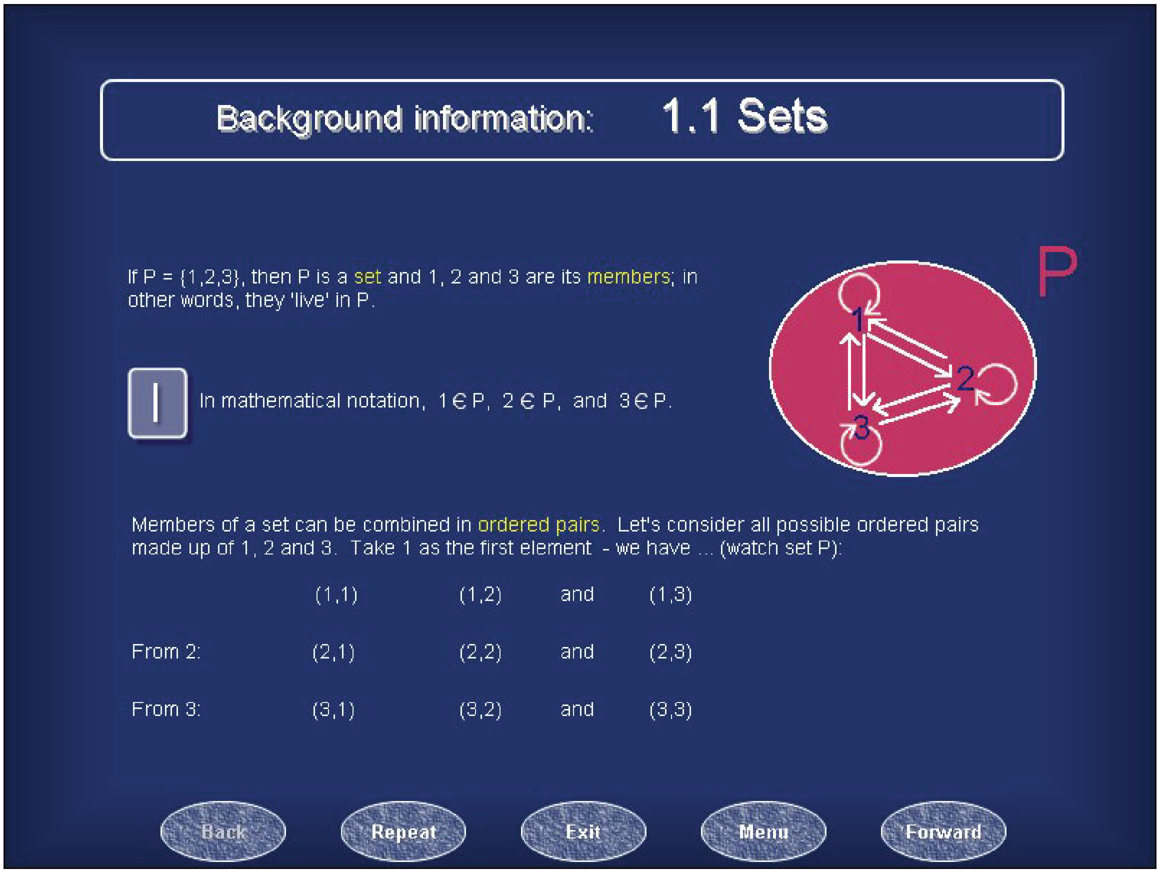

Figure 3: Background information to contextualize the new knowledge

Presentation of previous learning to contextualize the new knowledge is depicted in Figure 3. In real-world use, this screen would demonstrate animated development of the ordered pairs in the set {1,2,3}, cumulatively adding each pair as a corresponding arrow appears in the circle. Note the buttons in the bottom control bar, offering users the control options of:

- **Back** Return to previous screen;
- **Repeat** Redo current screen;
- **Exit** Leave the tutorial immediately;
- **Menu** Return to main- or submenu, from which the current screen was entered;
- **Forward** Continue to the next screen.

Figure 4 portrays a question for which a learner entered an incorrect answer, followed by detailed feedback, which develops the concept in detail. Judgement and feedback are diagnostic, customized, in many cases, to the nature of the error.

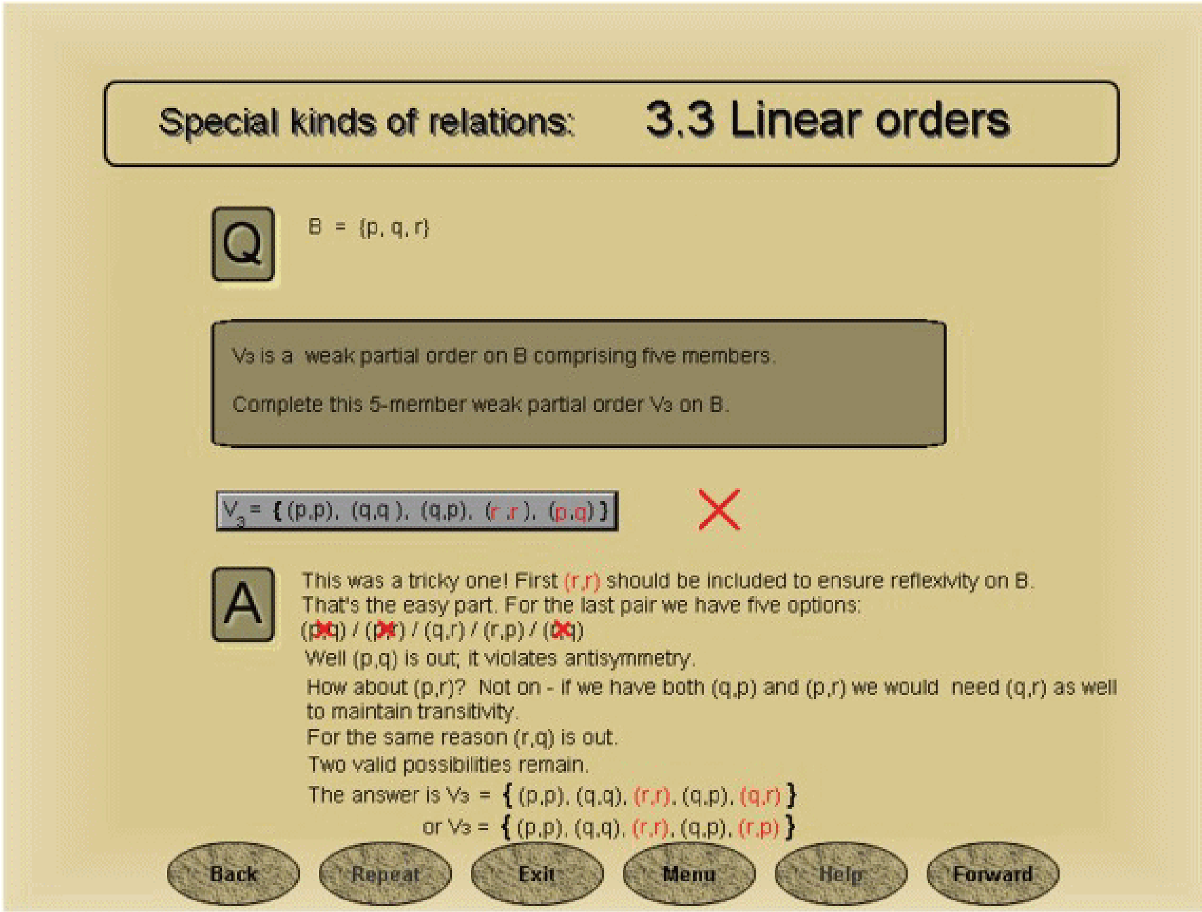

Figure 4. A question with an incorrect answer

Figure 5 demonstrates a completed colour-coded animation, along with a question about equivalence relations and cosets, where the concepts entail different styles of brackets: { }; (,); and [ ]. To answer the question, learners enter series of characters and notations, which are not assessed until they have all been entered. The boxed '**Q**' and '**A**' are symbols indicating question and answer respectively.

The figure indicates several correct answers, marked with a '√' in red font and one wrong response as part of a composite answer. Here no comments were given, but the correct answer was provided in red after the '**x**'.

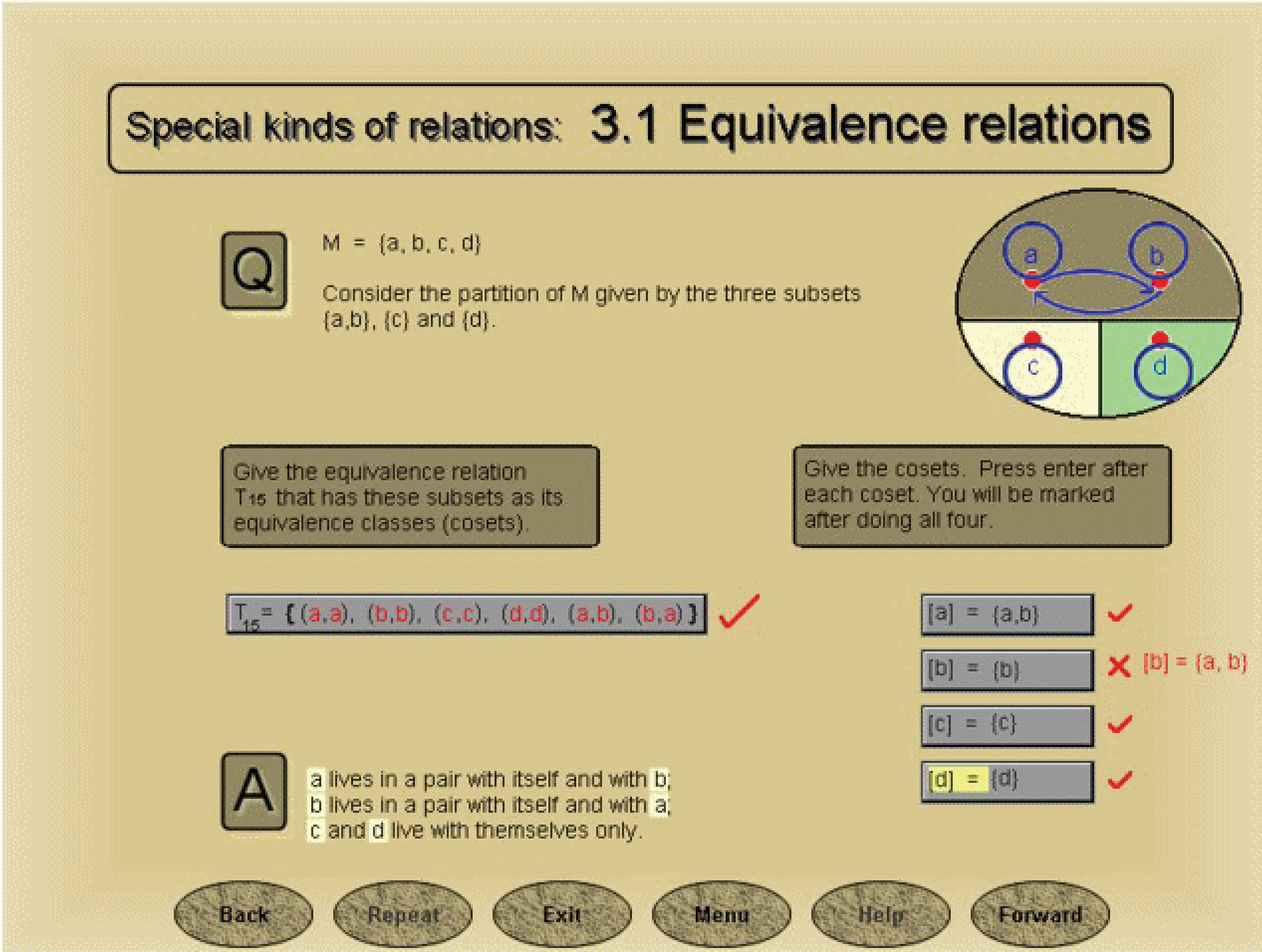

Figure 5. A completed colour-coded animation

Boxed '**Q**' and '**A**' symbols were previously mentioned. Further icons, shown in Figure 6, are '**I**', which represents important relevant information that scaffolds understanding of the concepts in hand, and '**D**', which indicates a definition. Some essential definitions appear as an integral part of the text, while others are pop-up overlay definitions of background theoretical concepts, which appear as required when learners click on a hyperlink. Both kinds of definition occur in Figure 6, the pop-up form being a feature added after the need for such emerged during evaluation.

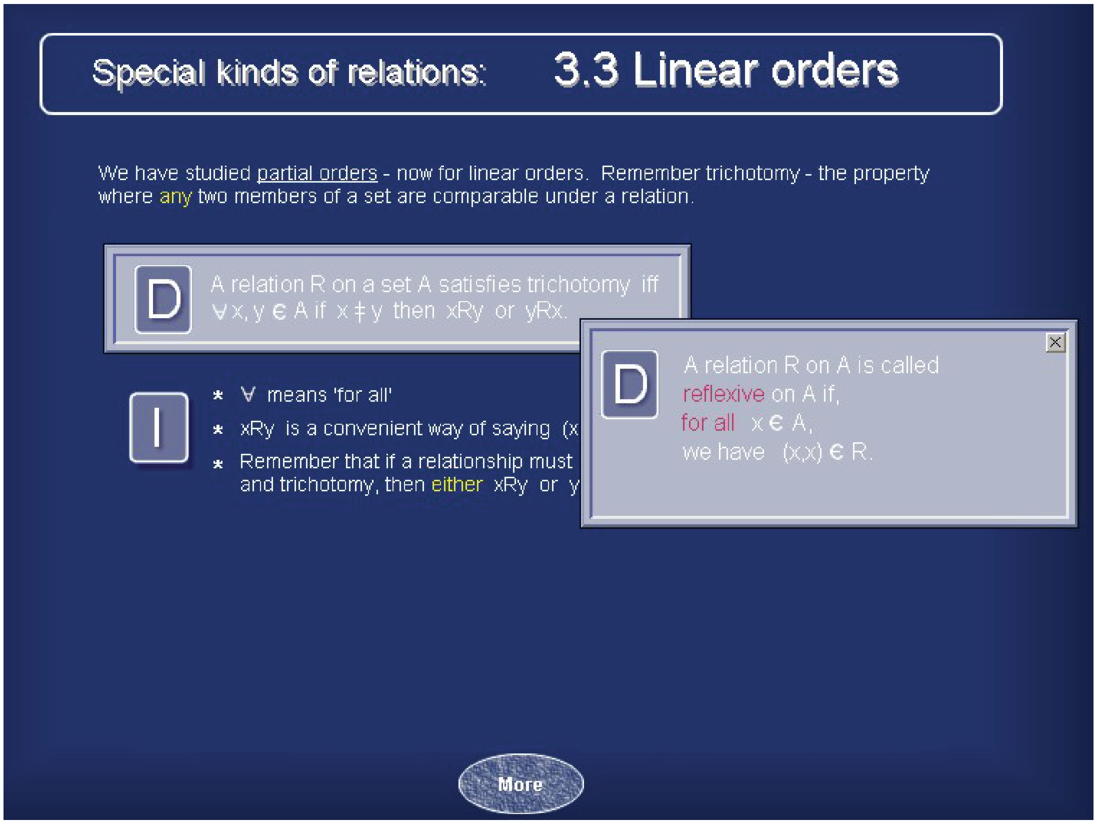

Figure 6. Illustrations of the 'I' icon and two forms of 'D'

Evaluations

The redesigned *Relations* was released and evaluated in 2004. In line with the reflective and responsive action research approach, this summative evaluation was used to further improve the operational product in a formative way, as findings from the 2004 student questionnaire survey, along with input from a heuristic evaluation, were used to further correct, refine, and slightly extend V2.1. The subsequent V2.2 was released in 2005 and evaluated by a further student questionnaire survey, by interviews among students, and by heuristic evaluation. In both 2004 and 2005, students' performance in a particular question in the final exam was used as a post-test. A comprehensive and detailed account of the evaluations is given in De Villiers (2006), but in the section following, pertinent extracts are given from the findings of these evaluations by four different techniques.

Questionnaire Survey among Learners

Various sets of guidelines, criteria, and principles exist for evaluating usability and effectiveness of educational software and web-based learning environments. Examples are those of Albion (1999) and Costabile, De Marsico, Lanzilotti, Plantamura, & Roselli (2005). The surveys of *Relations* were conducted using criteria of Squires and Preece (1999), who took Nielsen's (1994) classic set of usability heuristics and filtered them through the concepts of cognitive authenticity, contextual authenticity, and socio-constructivism. They converted them to eight criteria that integrate usability and learning factors for predictive evaluation of educational software by experts *prior* to selecting systems for use. Seven of their criteria (some slightly amended) were used in the present study, along with two new criteria (9 and 10 below). Based on the criteria, evaluation questions were generated for questionnaire surveys among learners *after* use:

1. *Match between designer & learner models*: The software should represent cognitive tasks in ways that foster formation of a learner model consistent with the designer model.
2. *Navigational fidelity*: This involves investigating the structure of navigation, cosmetic authenticity, and the effectiveness of the representation of the world within the system.
3. *Appropriate levels of learner control*: This relates to the balance between learner control, self-direction, customization, consistent protocols, and system responsibility.
4. *Prevention of peripheral cognitive errors and usability errors*: There is a relationship between domain complexity and error prevention. Peripheral usability-related errors should be anticipated and avoided, and novice's versions could be provided.
5. *Understandable and meaningful symbolic representation*: The representational forms and symbols used are vital. Interfaces should present low cognitive demands, and learners should not have to remember the forms of interaction. Symbols, icons, and names used for learning objects should correspond with the subject domain and be used consistently.
6. *Personally meaningful forms of learning*: There are multiple methods and representations for supporting varying learning styles. Metacognition should also be supported, and software should be used in tandem with other learner support materials.
7. *Strategies for cognitive error recognition, diagnosis and recovery*: While usability-related errors should be avoided (see item 4), cognitive errors are part of the learning process. Techniques, such as scaffolding and bridging, should be used to promote a recognition–diagnosis–recovery cycle.
8. [*Match with the curriculum*: Irrelevant to the present study, since the software was custom-designed for the curriculum.]
9. *Distinctive features*: Features should be provided that are unique to the environment under evaluation and that support the particular requirements of its content and context.
10. *Capacity of the system to engage learners*: E-learning environments should motivate learners and hold their attention.

A pilot questionnaire survey with 20 students early in 2004 (De Villiers, 2004) elicited highly positive ratings. Two errors in the learning content were identified and corrected.

In 2004 and 2005 evaluation questionnaires were included in the tutorial material, eliciting 49 returns in 2004 and 37 in 2005. The groups of respondents were samples of volunteers. In both groups the respondents comprised about 10% of the 'serious students' who persevered and wrote the final examination. In general they rated *Relations* very highly. There were some differences between the 2004 and 2005 cohorts. Despite using the improved V2.2, the 2005 group were less positive in some judgements; for example, they found the system less easy to learn and use and the subject matter more difficult. The findings were statistically analysed and indicate that responses to certain questions show significant differences between the groups and some anomalies. This is probably due to the rapidly shifting profile of the student population, in line with overall changes in the stereotypical UNISA student. For example, 80% of the 2005 group were aged 30 or younger (versus 60% in 2004), 50% were fulltime students (38% in 2004), and 39% were in full time employment (versus 56%). Most of them, 77% (51%), were from formerly disadvantaged minority groups – a distribution in line with the changing module population. It is natural that younger students with less computing experience would find it harder to relate to a new type of computing application and therefore the minor decline in some ratings is acceptable. On investigation of results in the final examinations, the average percentages obtained in 2005 exam results in all modules that comprise Computer Science 1 were 2% to 7 % lower. Across the board, the 2005 cohort were lower performers than the class of 2004.

Many of the evaluation participants were ‘on-campus’ students – young people who study full-time in the university’s regional libraries or learning centres, and have social interaction with peers.

Questions in the form of statements investigated aspects of the nine selected criteria on a Likert scale. There were 40 in total, some with open-ended sections for elaboration of responses. This paper extracts and discusses the responses to some of the criteria as examples. There were five options on the Likert scale, but the data is condensed by integrating categories. The findings of 2004 and 2005 are kept separate. The 2005 group did not identify any content errors in *Relations*, which is not surprising given the iterative investigation and refinements it had undergone.

Match between learner and designer models (Criterion 1):

There were highly positive ratings regarding the cognitive effectiveness of the lesson, i.e. how the learners understood the theory of relations. Some open-ended responses were: ‘The teaching content is absolutely brilliant.' / ‘Practicing makes it easier to learn and understand.’

Doing exercises in the *Relations* lesson helps me understand the theory of relations:					
Strongly agree or Agree	2004 2005	98% 95%	Maybe or Disagree	2004 2005	2% 5%
The elaborations and explanations help me understand:					
Strongly agree or Agree	2004 2005	98% 84%	Maybe or Disagree	2004 2005	2% 16%
Relations helps me understand a difficult section of the module, *Theoretical Computer Science 1*:					
Strongly agree or Agree	2004 2005	78% 72%	Maybe or Disagree	2004 2005	22% 28%

Navigational fidelity (Criterion 2):

The 2004 responses mainly described *Relations* as easy to use and easy to navigate. The 2005 responses were more conservative; 22% had ‘got lost’, probably due to their novice status.

I learned to use *Relations*:					
Very quickly or Quickly	2004 2005	75% 59%	Slowly	2004 2005	25% 41%
When I use Relations, I know exactly where I am, what parts I have done, and what I should still do:					
Strongly agree or Agree	2004 2005	92% 81%	Maybe or Disagree	2004 2005	8% 19%
The lesson is easy to operate:					
Strongly agree or Agree	2004 2005	94% 86%	Maybe or Disagree	2004 2005	6% 14%
I got ‘lost’:					
Strongly agree, Agree or Maybe	2004 2005	6% 22%	Disagree or Strongly disagree	2004 2005	94% 78%

Some asked for more sensitive navigation facilities: ‘I would rather use the menu to be exactly where I want to be next.’ / ‘It would be useful to navigate to any parts of the lesson, rather than only having the selection between doing the lessons and exercises, or diving straight to the exercises.’ / ‘...frustration when you’ve got to go through all the trial examples to reach the next part.’

While taking cognizance of this, it must be stated that the strict control in certain parts deliberately forces users to work through sets of exercises for pedagogical reasons.

Appropriate learner control (Criterion 3):

Students appeared generally satisfied, but open-ended responses show that some want more control, also exit and re-entry facilities and the option to leave a section without completing it. In the 2004 group, 88% wanted to choose the level of difficulty (a facility not provided), whereas of the more tentative respondents in 2005, only 58% would appreciate that choice, the others preferring the pre-set sequences with increasing difficulty. The percentages are not given in detail.

Some requested more sophisticated customization and control: 'It would be nice to completely disable certain kinds of screens.' / 'Maybe a separate menu for exercises.'

Prevention of usability errors and peripheral cognitive errors (Criterion 4):

Domain complexity should not be simplified. Content-related errors should be permitted as part of the learning process, but situations that cause usability errors and perception errors should be avoided. Only one item is extracted from the data and reported here, namely, the issue of trying to use operations familiar from commercial systems. This occurred more with the 2004 group, probably due to their greater familiarity with other software. (This also occurred with two heuristic evaluators – see evaluation case study 3.) In general *Relations* was pleasant to use, e.g.: 'It is visually stimulating and appealing'.

I made mistakes because I used operations and keystrokes I know from another system:					
Strongly agree or Agree	2004 2005	28% 8%	Maybe or Disagree	2004 2005	72% 92%

Meaningful symbolic representation (Criterion 5):

Ratings about the symbols, icons and names in *Relations* were positive, although there were a few complaints about typing in brackets and commas. The different bracket styles are vital notations in set theory and relations and it was important that students should be able to enter the appropriate ones independently. In some cases the 2005 students were more positive than those of 2004.

The symbols and names that represent mathematical objects are used consistently in the lesson:					
Strongly agree or Agree	2004 2005	83% 94%	Maybe or Disagree	2004 2005	17% 6%
The screen layouts are easy to read:					
Strongly agree or Agree	2004 2005	85% 92%	Maybe or Disagree	2004 2005	15% 8%
There is too much information on the screens and it confuses me:					
Strongly agree or Agree	2004 2005	19% 25%	Maybe or Disagree	2004 2005	81% 75%

Personally meaningful ways of learning (Criterion 6):

Socio-constructivist learning theory holds that learners should experience personal satisfaction with the process and style of learning. The tutorial structure of *Relations* tends more to behaviourist than constructivist but, where possible, personal learning styles were supported in the design.

When a concept was taught or illustrated in more than one way, it helped me understand it:			
Strongly agree or Agree	2004 90% 2005 91%	Maybe or Disagree	2004 10% 2005 9%
The ways that *Relations* helps me understand are through (participants could choose more than one option):			
Diagrams 2004 69% 2005 74%	Real life pictures 2004 17% 2005 24%	Animations 2004 32% 2005 38%	Practicing 2004 69% 2005 65%
I enjoy approaching studies collaboratively, i.e. working with a fellow-student:			
Strongly agree or Agree	2004 42% 2005 40%	Maybe or Disagree	2004 58% 2005 60%

The multiple representation approach was used frequently. Concepts were defined and then illustrated with diagrams or animations. This was appreciated: 'It is easier to understand, since a picture is formed in one's head.' There were spontaneous comments about how the examples and diagrams helped them to grasp concepts and concretize the theory, i.e. inductive learning. Collaborative work is part of the social constructivism trend and although it is complex to implement in distance learning, it is occurring naturally with the students who work in learning centres or labs. Several said they had used *Relations* with fellow-students: 'We can discuss, argue …'. Some had used it along with the textbook or with their own summaries.

Cognitive error recognition, diagnosis and recovery (Criterion 7):

This relates to domain complexities, misconceptions and pitfalls, and the adequacy of the feedback in helping learners to correct them. The redesign of the lesson had concentrated on HCI features and enhanced control. An aspect not changed was the original feedback of the 1994 version, which paid meticulous attention to potential errors. Apart from some requests for 'harder exercises', the feedback obtained high praise: 'It accurately guessed what I did wrong and pointed me in right direction.' / 'I could easily pinpoint my mistake the next time.' / 'It explains clearly where I went wrong. Those hints really help.' / 'The incorrect part was pointed out and they give another chance to try.' / ' The feedback uses definitions and helps me to remember definitions and theory.' / 'This … helped me answer the assignment questions correctly.'

I made mistakes in doing the exercises:			
Got most wrong 2004 0% 2005 8%	Got some wrong 2004 42% 2005 31%	Got a few wrong 2004 46% 2005 42%	Hardly any mistakes 2004 12% 2005 19%
The feedback (system responses) to my incorrect answers helped me:			
Strongly agree or Agree	2004 92% 2005 90%	Maybe or Disagree	2004 8% 2005 10%
When I got an answer wrong, I was able to get it right on the next try.			
Always 2004 25% 2005 36%	Nearly always 2004 56% 2005 44%	Sometimes 2004 15% 2005 17%	Not often 2004 4% 2005 3%

General: distinctive features; motivation and engagement (Criteria 9 and 10)

After complex sections, an innovative water-based recreational theme emerges in the form of displays showing animated water activities – a pool, ducks, boardsailing – accompanied by audio. Figure 7 below the next set of ratings shows *Relations'* concluding screen. In line with Criterion 10 of motivation and engagement (Hodges, 2004; Keller & Suzuki, 1988), these features were designed with dual means of motivation (Alessi & Trollip, 2001), namely, intrinsic motivation – by content that engages learners, and extrinsic motivation – by moments of informality and humour. Responding to the first question below, over 70% chose either the first or second option or both, indicating appreciation of the recreational theme. A very low percentage was 'irritated'.

The recreational screens with the water theme (participants could choose more than one option):			
Are fun 2004 43% 2005 44%	Give me a break 2004 43% 2005 45%	Make no difference 2004 26% 2005 25%	Are irritating 2004 6% 2005 3%
The shortcut 'Dive' screens to avoid redoing theory are useful.			
Strongly agree or Agree	2004 84% 2005 70%	Maybe or Disagree	2004 16% 2005 30%

Open responses were mainly positive: 'They were amusing.' / 'It was FUN.' / 'The fun screens attracted me – they make the program friendly' / 'A bit of a surprise! I liked the sounds, they help me relax' / 'It's too stressful to do the whole CD in one go – the screens with water help me take a break.' But: 'When I was tired, the breaks were useful, otherwise they wasted time.' / 'Fun screens slow you down when seriously revising.' A 19-year old criticised the animations: 'Purchasing this CD was a good investment. It helped me grasp concepts much easier. Thanks, guys, for a terrific job. But the animation is quite poor. I know you aren't Hollywood studios, but you are capable of something more exciting. Take it as a challenge'. Another 19-year old said: 'Well done, you guys, it's great! It beats reading the book as it goes straight to the point'. So *us guys* get bouquets and brickbats ... Are any Disney animators looking for instructional design posts in Africa?

Students appreciated the 'dive' option to bypass theory and go straight to exercises. 'If you know the theory, you can just go for the exercises.' / 'It's just great not re-doing the theory after the first time.' / 'It saves time.' / 'It will be useful to dive when revising for exams.' Referring to the neutral character that symbolizes the dive, 'Pictures of real babes diving would be nice!'

Figure 7. Concluding screen – an animated boardsailing race

Finally, 'This lesson encourages you'. / 'These exercises are much more fun … compared to the study guide.' / 'More of the same, please!' / 'Keep it up for giving us such a lesson on computer. I think you can do it for other chapters too.' / 'I would like CAI lessons for all modules'.

Query Method: Interviews with Students

Semi-structured interviews (Preece et al, 2002) were held with five volunteers in 2005 to investigate aspects not explicitly addressed in the questionnaire. The subjects were fulltime students in their first or second year of study, with average age 21. One had a prior diploma qualification in electronics.

They had all experienced a synergy from the mixed model of textbook and CAI lesson. Added value came from the detailed explanations, animated diagrams showing development of concepts, and the diagnosis of errors so that they could be corrected. A quote: 'The computer is your teacher and you don't have to make an appointment'. All five found *Relations* easy to learn and use. In line with the motivation and engagement criterion, it engaged them and they experienced flow (Csikszentmihalyi, 1990): 'I was just enjoying it…' / 'I <u>want</u> to do this'.

The relative value of the theoretical teaching segments and the practical exercises was investigated to see if *Relations* fitted students' personal learning styles. It appears to be on target in addressing more than one style, because the respondents were divided equally in their preferences. Asked which they *liked doing most*, two chose the theoretical teaching, saying respectively: 'the explanations are different from the book' and 'because of those diagrams', and two chose the exercises: 'with that second chance to get it right'. The fifth enjoyed both equally. Regarding *learning value*, they did not all make the same choice as they had made for *liking*. One chose 'the theoretical definitions, because they come with examples'; two selected exercises, because: 'you learn more when you <u>do</u> things' and 'when you just read, you get lazy'. The other two were adamant about both together: 'the teaching to know and the exercises to prove you know' and to 'show how far you have learned from the teaching'.

Collaborative work was addressed. This style was previously impractical for UNISA, since the student population consisted of classic distance-learners working in isolation. But with the advent of fulltime students, working on-campus or at learning-centres, it is feasible. Four liked working with fellow students. Two had used *Relations* jointly with a friend, and two preferred to do the lesson alone and discuss issues with peers afterwards.

When asked about *Relations'* special features, all five spontaneously mentioned the diagrams. The more experienced postgraduate student felt, 'The most important thing about *Relations* is its usability'. Another stressed the different ways of using it at different times – for exam revision, he would use the Dive facility to skip theory and go direct to practice.

Post-examination investigation of the five subjects indicated above-average performance. Of the 400 students who wrote the exam in 2005, only 52% actually passed. But these five all passed, obtaining between 50% and 65%. The average mark for the exam question on mathematical relations was 57.5%, but the interview group averaged 65% for that question, showing their diligence and enhancing their credibility.

Inspection Method: Heuristic Evaluation by Experts

Heuristic evaluation (HE) (Dix, Finlay, Abowd & Beale, 2004; Nielsen, 1994; Preece et al, 2002; Squires, 1997) is an inspection technique developed by Jakob Nielsen and Rolf Molich. Expert evaluators examine a prototype or an operational system, using a set of guidelines or usability principles called *heuristics*. HEs of *Relations* were conducted by six evaluators, peer academics who are subject matter experts. Five are involved in teaching various levels of Theoretical Computer Science and the sixth is a young postgraduate student. In the participative action research approach, two had been involved in the initial 1994 evaluation. They mainly used the same criteria as those in the students' questionnaire, but as subject matter experts, their emphasis was on learning content, rigour, usability, and navigation. They worked through the lesson comprehensively, simulating students' activities. Four of them evaluated V2.1 after minor improvements from the 2004 student evaluation had been made. Two subtle mathematical shortcomings were identified; certain additional functionality and minor modifications to the control structure were suggested. These improvements were made, and V2.2 was released in 2005 and evaluated by the other two experts. No real usability problems were identified in V2.1 or V2.2. Despite the era of hyperlinked web-based learning environments, there was general consensus that the control and usage patterns of *Relations* – based on traditional CAI architecture – operate effectively. The data from the six assessments was integrated under the given criteria as headers (de Villiers, 2006). Notable points are mentioned below.

Match between designer and learner models

The expert evaluators unanimously supported the educational value of *Relations*, all agreeing that the exercises, judgement-and-feedback, elaborations, and explanations in the CAI lesson would help students learn the subject matter.

Navigational fidelity

They all learned to use *Relations* quickly and found it easy to operate. Navigation was not complex; five of the six knew their exact location within the lesson at any time. Screen layouts were easy to read, and one evaluator commented particularly on the simplicity, aesthetics, and visibility of the uncluttered screens compared to other CAI lessons she had encountered.

One evaluator suggested improvements to *Read-Me* on the CD, which were incorporated.

Learner control

The level of learner control – in terms of sequence and section control – was appreciated, although the young evaluator would have liked even more control, expressing concern about places where the *Back* button is explicitly deactivated. He viewed this as 'system control'. Most of the evaluators felt that users should be able to choose the level of difficulty of exercises, rather than working through each entire set. Another problem is the lack of exit and re-entry facilities in the sets of exercises. While noting these concerns, the modifications were not made due to the major architectural changes. The points are noted for future efforts.

Usability errors and problems

The experts did not find *Relations*' interface and interaction facilities difficult to use. Two made initial errors by using operations and keystrokes familiar from commercial software, but which were invalid in *Relations*, e.g. double clicking on an object when a button in the control area should be clicked. This problem, also encountered in the student survey, indicates non-adherence to Dix et al's (2004, p. 260) generalizability subprinciple of the 'principles to support usability'.

Symbols, icons, and names

No problems were identified. The evaluators agreed that *Relations* does not make cognitive demands on memory. There is adequate visibility; it operates on recognition rather than recall; the meanings of all symbols and icons are clear. No evaluator got 'stuck'.

Personally meaningful ways of learning

The experts appreciated the multi-perspective approach of demonstrating concepts. They felt that the diagrams and animated developments communicated messages effectively and that the extensive practice reinforced concepts. One of the evaluators suggested that more detailed information and definitions could be obtained by clicking on hotspots. This feature was implemented in V2.2 and was appreciated by its evaluators, as well as by V2.1 evaluators who saw it.

Recognizing cognitive mistakes and recovering

The experts made very few cognitive mistakes! When they did, the feedback helped them get it right on the next try. However, they used their mathematical skills to deliberately challenge the system with novel or obscure responses, resulting in some refinements to diagnostic feedback.

Special features

Four of the six agreed that the recreational screens were fun; one found them irritating and another felt they made no difference.

Post-test

The role of post-tests in measuring learning is debatable, but they are a traditional quantitative measure of the effectiveness of a learning system/intervention. In an educational context they are also appropriate instruments to assess usability, since without ease of use, functionality is obscured. *Relations* includes a revision test, but due to its role as a non-threatening system, designed for learning gain, scores are not recorded, so this test is not a post-test. As a suitable 'post-test', the participants' performance in the final examination was compared to the performance of the rest of the students – see Table 1. One of the exam questions related directly to the subject matter on mathematical relations, as taught in *Relations*. The scores for this question are given in the central row of the table, and the last row indicates (as a percentage) how many students passed that question, irrespective of the mark they got for it.

Table 1: Examination performance of survey participants

Criterion	**2004**		**2005**	
	Participants in the questionnaire survey (study group)	**Other students (control)**	**Participants in the questionnaire survey (study group)**	**Other students (control)**
Average percentage in final end-of-year examination	58%	54%	53%	50.5%.
Average score (mark) for the question on relations	78% (7% more than in control group)	71%	61% (4% more than in control group)	57%
What percentage of students obtained a pass mark for that question?	87% (12% more than in control group)	75%	73% (6% more than in control group)	67%

If effectiveness is measured by grades and performance, then *Relations* resulted in effective learning. Table 1 shows that 12% more of the survey participant group passed the relevant question in 2004 and 6% more in 2005. The average mark obtained by participants was 7% higher for that question in 2004 and 4% higher in 2005, while their averages for the full examination were 4% and 2.5% higher respectively.

For the 2004 data, scores were statistically investigated to determine whether there was a linear association between achievement in the exam as a whole and the mark for the question on relations. Following the exclusion of four outlying cases, the p-value was greater than 0.05, indicating that the correlation was not significant and there was no direct proportional relationship between the two marks. This indicates that several participants in the study attained higher marks for that question than would have been expected from their overall performance, i.e. those who undertook this optional learning experience enhanced their skills in that section of the curriculum. A similar analysis was not done on the 2005 data, since the exam question on relations in 2005 was not as close to the content of the *Relations* lesson as was the case in 2004.

NB: Participation in the questionnaire survey was voluntary, so the respondents do not represent all the students who used *Relations.* Thus the 'other students' group includes some who were users of *Relations*, but not participants in the survey. If they had completed the questionnaire, and thus been identified as subjects within the study group, it is probable that the 'other' group, then comprising only genuine non-users of *Relations*, would have achieved yet lower performance.

Conclusion and Recommendations

In an action research approach, *Relations* has undergone a continuous process of evaluation, reflection, correction, and refinement. This occurred with its initial development in the mid-1990s, when it was formatively evaluated by peers of the designer. However, the evaluation research was particularly rigorous throughout its redesign and releases in 2004 and 2005. Formative and summative evaluations were conducted; functionality, usability and educational effectiveness were investigated. Each new cycle revealed usability problems or subtle content flaws, which were corrected. The problems identified became increasingly minor over time.

This article presents findings about the functionality and usability of *Relations,* which lead finally to discussions, generic reflections and lessons. Approaches to the design of CAI applications are briefly addressed, several of which are relevant to e-learning in general. A further contribution is the consideration of the complementary roles of different evaluation techniques in the multi-method evaluation of *Relations*.

Design and Evaluation of CAI/L: Reflection and Lessons

The role of traditional computer-aided instruction and learning in the WWW era

The development team of *Relations* believes that conventional computer-aided instruction and learning (CAI/L) has a role to play in the milieu of e-learning. Classic CAI can present efficient instruction, motivate and engage learners, challenge them with meaningful exercises, and can support effective learning. The students' requests for more such tutorial and practice environments should be noted. *Relations'* greatest strength is its excellent diagnostic feedback, attested to by learners and expert evaluators alike. Courseware authoring systems such as *Quest 7.0*, which was used to develop *Relations*, have powerful facilities that can be used by developers to assess the forms of learner-input and provide appropriate, detailed, tailor-made feedback. This can be done in web-based learning (WBL), using specialized web-programming languages, but it is more complex. A further obstacle to the use of WBL at UNISA is that many UNISA students still lack broadband Internet access.

An expert evaluator and some students requested hi-tech effects – *Flash* animations, 3-D modeling, etc. Such suggestions must be considered, but not all special effects can be implemented with conventional authoring tools. Moreover, designers should be careful not to use special effects, ' bells and whistles', for their own sake, but rather to use them to illustrate concepts in ways that enhance cognition. Technology should be the medium and not the message. The blue-water recreational theme was well received by most learners and evaluators, who acknowledged its role in providing brief interludes of diversion and relaxation in a demanding cognitive domain.

Lessons about usability

The concepts of usability and interaction design from the discipline of human-computer interaction (HCI) are receiving increasing attention in the development of commercial and corporate software. It is equally, or even more, important to produce usable applications in educational contexts where the users are not professionals in the workplace, but learners who must first be able to use and interact with a system before they can even commence learning.

Some learners approach e-learning experiences after exposure to commercial software. As far as possible, learning applications should use operations and keypresses that support the HCI principles of predictability and consistency with familiar systems. The *Quest* authoring tool does not naturally support the generation of software that is identical in appearance and functionality to commercial systems. However, *Relations* was found to be easy to learn and use. It adheres to the fundamental principle of internal consistency, where a system's own internal operations are characterized by predictability and visibility.

Lessons for new CAI environments

The teaching approaches in *Relations* present multiple perspectives on the theoretical concepts and multiple representations of content. These methods were successful and could be supplemented by multi-modal presentations in contexts where they would add value. There is a degree of learner-control, permitting choices about sections, sequence, repetition, and whether to omit theoretical segments and head straight for practice. The hyperlinks incorporated into *Relations* V2.2 were a success. This feature of double-clicking on content hotspots for definitions, elaborations or cross references should be regularly employed.

Concerns were raised regarding some aspects of control and navigation. Where pedagogically appropriate, linear sequences should be supplemented with hyperlinked access. Exit and re-entry facilities, which are not present in *Relations*, are desirable. User-control should be provided over the level of difficulty, so that users can choose particular exercises. This would facilitate different ways of usage at different times. Some users did not appreciate *Relations*' system control that deactivates certain buttons in particular situations. This was carefully designed for appropriate contexts, but some users saw it as inconsistent. Ways of implementing such tactics must be carefully chosen and designers must exercise discernment in deciding when to provide the facilities that users request and when to overrule for pedagogical reasons.

Lessons from the use of different evaluation techniques

The multi-method evaluation of *Relations* deserves mention. The complementary use of four evaluation techniques and triangulated data provided a synergic framework, generating information superior to that obtained by any single method. The *heuristic evaluations by experts* were valuable sources of critique and suggestions, particularly in the mathematical domain. As leading subject matter experts, two tested the feedback to its limits, revealing flaws and compromises of mathematical rigour. HEs are cost-effective and can be done in a short time frame. The experts moved beyond the given criteria, identifying omissions and suggesting further features.

The discernment, insight, and long-standing expertise of the heuristic evaluators accurately pinpointed the strengths and weaknesses of *Relations,* as well as subtle mathematical inaccuracies. For overall evaluation, the team of heuristic evaluators should include experts with subject matter skills and those with usability expertise.

The *students' questionnaire* responses were very useful with regard to educational aspects. In particular, the questionnaire confirmed the value of the detailed feedback, showing that the time spent painstakingly developing it was worthwhile. The 2004 questionnaire survey elicited particularly useful information and identified some errors that were corrected even before V2.1 underwent heuristic evaluation. The 2005 survey did not provide as much value, mainly because there was not much to add to the comprehensive information obtained in 2004.

Questionnaire design is vital for the elicitation of useful information. An inadequate questionnaire can trivialize an evaluation. Likert-scale questionnaires with specified options may exclude some important aspects, but the inclusion of open-ended questions for qualitative responses can be used to compensate, to elaborate responses, to enrich the information obtained, and to motivate the Likert options selected. In general, students used these to express praise and appreciation of *Relations*. Unanticipated aspects that emerge can be probed further in interviews or subsequent questionnaires.

The *interviews* among end-users (students) explored novel avenues, following up on usability problems. The interviews were used to probe for explicit identification of features that fostered learning for each individual and the features they enjoyed most. Since interviews can be done only with a limited number of subjects, the sample should ideally be heterogeneous and scientifically selected. UNISA's distance-teaching context makes personal contact difficult and a sample of volunteers was used in this study, representing a fairly homogenous group of young fulltime students, typical of the shift in learner profile. However, contact with a stereotypical group is informative in its own right and, in this case, was very rewarding, due to the responsible and serious attitude evidenced by these five young people towards their studies and the subsequent exam.

To optimize on interviews, the structure should be flexible. A good semi-structured interview should be based on a set of common core questions, presented by a researcher who has the discernment to know how to present good follow-up questions and when just to listen. This approach can elicit rich information.

The *post-tests* gave quantitative measures that could be statistically analysed. However, scores in tests remain debatable instruments for measuring the effectiveness of an artefact or other intervention in improving learning. Tests frequently show no significant difference between methods of instruction, even when other evaluation methods generate favourable reports of the intervention. When a test, therefore, does indicate improved academic performance, it would appear that the method under investigation indeed enhances learning. This was the case in the present evaluation.

The participative action research process of designing, developing, evaluating and refining the e-learning tutorial, *Relations* has taught the designer and development team a great deal – not just about the application being studied, but also lessons and generic principles for the design of e-learning applications as well as lessons about the rich complementary roles of different evaluation methods and techniques.

Acknowledgements

Great appreciation is expressed to the enthusiastic, capable and diligent development team, who went beyond the limits in their efforts! The researcher acknowledges the valuable contributions of colleagues who served as formative evaluators and as expert evaluators. Finally, thank you to the students who completed the questionnaires and to those who volunteered for interviews.

References

Albion, P. R. (1999). Heuristic evaluation of educational multimedia: From theory to practice. Retrieved December 2005, from http://www.usq.edu.au/users/albion/papers/ascilite99.html.

Alessi, S. M., & Trollip, S. R. (2001). *Multimedia for learning: Methods and development* (3rd ed.). Massuchusetts: Pearson Education.

Baskerville, R. L., & Wood-Harper, A. T. (1996). A critical perspective on action research as a method for information systems research. *Journal of Information Technology, 11,* 235-246.

CEDEFOP (European Centre for the Development of Vocational Training). (2002). *E-learning and training in Europe*. Luxembourg: Office for Official Publications of the European Communities.

Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education* (5th ed.). London: RoutledgeFalmer.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper & Row.

Costabile, M. F, De Marsico, M., Lanzilotti, R., Plantamura, V. L., & Roselli, T. (2005). On the usability evaluation of e-learning applications. In *Proceedings of the 38th Hawaii International Conference on System Science,* 1-10. Washington: IEEE Computer Society.

De Villiers, M. R. (2004). Usability evaluation of an e-learning tutorial: Criteria, questions and case study. In G. Marsden, P. Kotze, & A. Adesina-Ojo, *Fulfilling the promise of ICT* (pp. 284-291). Proceedings of SAICSIT 2004. ACM International Conference Proceedings Series.

De Villiers, M. R. (2005a). Three approaches as pillars for interpretive information systems research: Development research, action research and grounded theory. In J. Bishop & D. Kourie, *Research for a Changing World* (pp. 142-151). Proceedings of SAICSIT 2005. ACM International Conference Proceedings Series.

De Villiers, M. R. (2005b). e-Learning artefacts: Are they based on learning theory? *Alternation, 12*(1b), 345-371.

De Villiers, M. R. (2006). Multi-method evaluations: Case studies of an interactive tutorial and practice system. *Proceedings of the 9th Annual Informing Science + IT Education (InSITE) Conference,* June 2006. Manchester, UK. University of Salford. Available at http://proceedings.informingscience.org/InSITE2006/ProcDeVi225.pdf

Dick, B. (1999). What is action research? Retrieved December 2005, from http://www.scu.edu.au/schools/gcm/ar/whatisar.html

Dix, A., Finlay, J., Abowd, G. D., & Beale, R. (2004). *Human-computer interaction* (3rd ed.). Harlow: Pearson Education.

Du Poy, E., & Gitlin, L.N. (1998). *Introduction to research: Understanding and applying multiple strategies* (2nd ed.). St Louis: Mosby.

Hodges, C. (2004). Designing to motivate: Motivational techniques to incorporate in e-learning experience. *The Journal of Interactive Online Learning, 2,* 3.

Jonassen, D. (1999). Designing constructivist learning environments. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: A new paradigm of instructional theory Volume II*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Keller, J. M. & Suzuki, K. (1988). Use of the ARCS motivational model in courseware design. In D. H. Jonassen, (Ed.), *Instructional designs for microcomputer courseware*. N.J.: Lawrence Erlbaum Associates.

Nielsen, J. (1994). Heuristic evaluation. In J. Nielsen & R. L. Mack (Eds.), *Usability inspection methods.* New York: John Wiley & Sons.

Preece, J., Rogers, Y., & Sharp, H. (2002). *Interaction design*. New York: John Wiley & Sons.

Reeves, T. C., & Hedberg, J. G. (2003). *Interactive learning, systems evaluation*. Englewood Cliffs, N.J.: Educational Technology Publications.

Reeves, T. C., & Reeves, P. M. (1997). Effective dimensions of interactive learning on the World Wide Web. In B. H. Khan (Ed.), *Web-based Instruction*. Englewood Cliffs, N.J.: Educational Technology Publications.

Squires, D. (1997). An heuristic approach to the evaluation of multimedia software. Retrieved December 2005, from http://www.media.uwe.ac.uk/masoud/cal-97/papers/squires.htm

Squires, D., & Preece, J. (1999). Predicting quality in educational software: Evaluating for learning, usability and the synergy between them. *Interacting with Computers, 11*(5), 467–483.

Wesson, J. L., & Cowley, N. L. (2003). The challenge of measuring e-learning quality: Some ideas from HCI. *IFIP TC3/WG3.6 Working Conference on Quality Education @ a Distance,* 231-238, Geelong, Australia, February 2003.

Zuber-Skerrit, O. (1992). *Action research in higher education*. London: Kogan Page.

Biography

Ruth de Villiers is a professor in the School of Computing of the University of South Africa. She has an MSc, an MEd and a PhD (Computer-Integrated Education). Ruth teaches courses in human-computer interaction, theoretical computing and e-learning. Her cross-discipline research areas – in which she also supervisors postgraduate students – are design, usability and evaluation of educational applications; e-learning in general; HCI education; and meta-research.