Learning with an e-learning tutorial: Cognition and the value of errors

M. R. de Villiers*
e-mail: dvillmr@unisa.ac.za

D. Becker*
e-mail: becked@unisa.ac.za
*School of Computing
University of South Africa
Pretoria, South Africa

F. Adebesin
Meraka Institute, CSIR
Pretoria, South Africa
e-mail: fadebesin@csir.co.za

Abstract

E-learning tutorials provide engaging learning activities and feedback as supplementary study material for distance learners. *Relations* is an interactive, CD-based, multi-media tutorial, custom-built to foster mathematical skills for first-level Computer Science students.

Making innovative use of the monitoring and recording technology in a human-computer interaction laboratory (HCI Lab), we investigated how *Relations* supported cognitive learning processes. Participants' error behaviours were studied, distinguishing between errors in using directions, usability errors, and cognitive errors due to misunderstanding the subject matter. Attention was paid to the role of system feedback in recovery from errors.

The results show that several participants scored higher in the post-*Relations* test than in their final examination on the full curriculum, indicating the value of the additional learning resource. Time distribution over learning processes and errors differed considerably between participants, and were unrelated to test scores, showing that *Relations* supports individual learning experiences. The authors advocate for the use of interactive tutorials to provide added value in open distance learning.

INTRODUCTION

In the Open Distance Learning (ODL) environment of the University of South Africa (Unisa), enriching and participative learning activities should supplement the tutorial matter. Interactive electronic tutorials can play valuable roles and

© Unisa Press ISSN 0256-8853 Progressio 34 (3) 2012 pp 81–104

could help to compensate for lack of face-to-face contact. A well-designed tutorial can offer engaging learning experiences, cognitive support, animated concept development, as well as practice. This article describes the innovative use of sophisticated usability testing technology, not to evaluate educational software in, and of itself, but to explore the process of learners interacting with an asynchronous tutorial called *Relations,* that teaches and provides practice in discrete mathematics for Computer Science 1. The study also investigated how students learned through errors and recovery, supported by immediate informative feedback.

The article sets the background and context, explains the research methodology, and refers to literature on relevant current ventures, various dimensions of cognition, and types of errors. Primary data collection involved observing, recording, re-viewing, and analysing interactions and cognitive engagement, as we sampled 17 students who studied through e-learning. The findings relate to three types of errors made by participants -- cognitive errors in particular -- which are part of the learning process. Time spent on these errors and recovery times was also analysed. We also investigated relationships between times spent on tasks and marks (scores) for the tasks.

BACKGROUND AND RELATED WORK

This section considers current related work, and then addresses our view of e-learning before describing the *Relations* tutorial. The evaluation of e-learning systems by usability-testing technology is explained, and the difference between evaluating e-learning applications and conventional software noted.

Related work

The authors take cognisance of current open source material. With the increasing convergence between conventional higher education and distance learning, Web 2.0 technologies are playing important roles. Massive Open Online Courses (MOOCs) are operationalising connectivism as they provide free online courses worldwide in environments where openness, autonomy, interactivity, and diversity are key principles (Mackness, Mark and Williams 2010). Online asynchronous tutorials are available via MITx (MIT 2011) and edX (edX 2012). MITx is a web-based programme launched in 2011, which provides certain Massachusetts Institute of Technology (MIT) courses online via open source software (MITx 2012). It offers self-paced learning platforms worldwide through interactive online laboratories and student-to-student communication, as well as subject-based performance assessment. MITxedX (edX 2012) is a joint venture by Harvard and MIT that offers free interactive Web-based courses in various disciplines, including computer science and engineering. edX also conducts research on students' learning behaviour and the role of technology in transforming learning.

None of these environments present the same functionally and content such as *Relations*, which was custom designed to provide detailed support to distance learners in a complex section of the curriculum. An online search identified no similar public-domain system that offers teaching segments, exercises, and instant feedback in the mathematical domains of reflexivity, symmetry, transitivity, and different types of relations. *Relations* pre-dates all these initiatives, having been released in 2004 and serially evaluated and updated over four cycles. Other South African universities teach similar content, but without associated e-learning tutorials.

E-LEARNING

Definitions of e-learning fall into two categories:

1. Rosenberg (2001, 28) defines e-learning as the 'use of Internet technologies to deliver ... solutions that enhance knowledge and performance'.

2. Masie (2008) similarly defines e-learning as network technology used to design, deliver and administer learning.

The authors of this article hold a broader definition, which includes instruction via all electronic media; including the Internet, intranets, Web-based learning, online courses, multimedia CD-ROMs, and computer-assisted instruction such as interactive tutorials, simulations and educational games (De Villiers 2005; Hung 2010).

Features and foundation of *Relations*

The system used in this study was *Relations*, a multi-media, CD-based, interactive e-learning tutorial. It was custom-built as supplementary material for a first-level Unisa module, COS1501. The subject matter is a section of discrete mathematics relevant to Computer Science. It was rigorously evaluated and refined in four cycles by academics, learners, programmers, and graphic designers in the School of Computing. It is not an online or web-based environment, but is delivered by CD to serve all students, including those in remote areas without continuous Internet connectivity.

Relations conforms to Alessi and Trollip's (2001) definition of a tutorial as software that presents cycles of information presentation, elaboration, questions, learner responses, judgement and instant feedback to learners' answers. Animated step-by-step concept development is demonstrated and instruction is interspersed with graphics and exercises in a theory↔examples↔exercise format (Figure 1). In open exercises, learners provide the fill-in-the-blank responses by entering mathematical characters or terms. Some answers are composite, and more than one answer is correct, provided it meets mathematical conditions. Feedback and explanations are provided to correct and incorrect responses, diagnosing common problems (Figure 2). In line with Boud, Keogh and Walker's (1996) conscious reflection and recapture, learners make second attempts after wrong answers, while theory and definitions are available via hyperlinks.

A brief overview of learning theory is provided to articulate the paradigm(s) underlying *Relations:*

1. *Behaviourism*, based on behavioural 'conditioning', is somewhat rigid. Learning occurs via observable environmental events and learning content is presented using a stimulus-response-reinforcement paradigm.

2. *Cognitivism*, based on an 'information processing' paradigm, emphasises the mental processes of learning, positing that information is initially stored in short-term memory and must be applied and 'encoded' in order to be transferred via working memory to long-term memory. New knowledge must be incorporated into the network of prior learning.

3. *Constructivism*, an open-ended, multi-perspective view, which posits that learners personally construct knowledge. Current learning theory is geared towards *constructivism*. For constructivists, the most important reality is a learner's individual interpretation of what he/she perceives (Alessi and Trollip 2001; De Villiers 2007).

The learning paradigm underlying *Relations* is mainly behaviouristic in nature, due to the fixed mathematical subject matter. However, the manner in which concepts are approached from multiple perspectives (graphics, evolving animations, and examples) and the user control available for running the tutorial bring a constructivist flavour that fosters personal styles of interpretation, leading to the comprehension of realities. Due to its logical mathematical domain, *Relations* requires critical thinking on the analysis, synthesis, and evaluation levels of Bloom's (1998) Taxonomy, hence cognitive learning also plays a role.

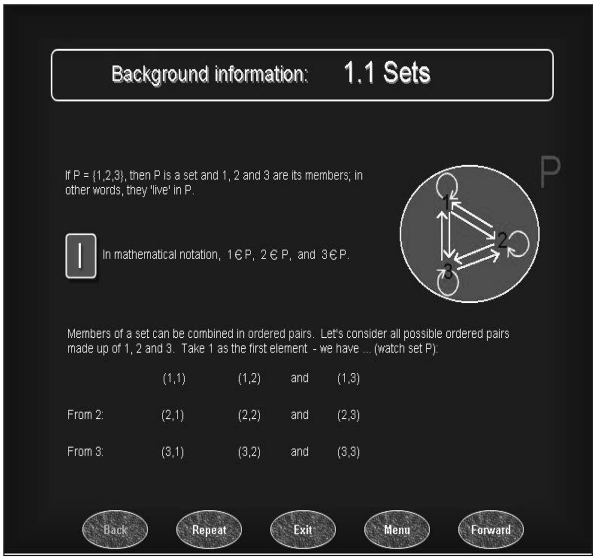

Figure 1: Graphics and animation

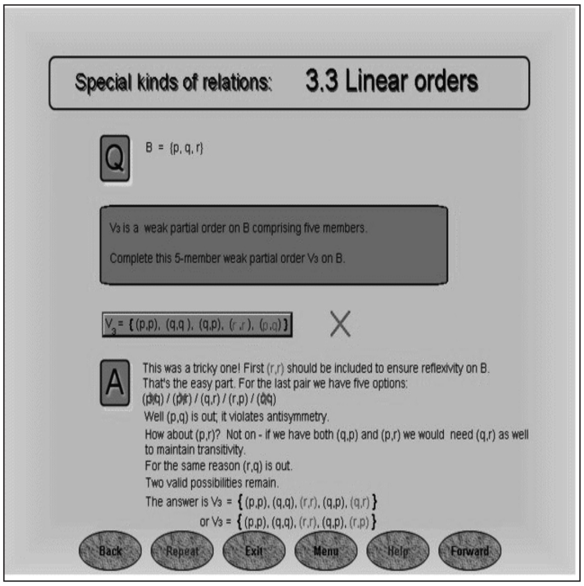

Figure 2: Feedback to an incorrect answer

Evaluation of e-learning

Evaluation is part of the iterative development of e-learning systems. Various usability evaluation methods (UEMs) exist to assess the effectiveness and efficiency of systems and users' satisfaction. In a longitudinal action-research process *Relations* was serially evaluated by questionnaires, interviews, heuristic evaluation by experts, and post-tests (Masemola and De Villiers 2006; De Villiers 2007). Each stage was followed by reflective refinements, resulting in a system that was almost problem-free, thus providing an appropriate environment for our study of learning through e-learning.

Usability testing

The UEM of usability testing (UT) involves observation and experimentation in controlled HCI laboratories with monitoring and recording facilities (Dix, Finlay, Abowd and Beale 2004). UT is a formal technique that traditionally measures the performance of users doing defined tasks on a target system, with the goal of identifying system problems (Barnum 2002; Rubin and Chisnell 2008). Researchers observe through glass and on split-screen monitors, which simultaneously show the participant's hands on the keyboard, facial expressions, and the screen being viewed. Actions are monitored and video-recorded for re-viewing and detailed analysis. Verbalisation, called 'think-aloud' contributes to the study of emotions. Classic usability metrics include time to complete a task, degree of completion, number of errors, and the time taken to recover from errors (Dix et al. 2004; Sharp, Rogers and Preece 2007).

In 2005, the School of Computing acquired an HCI Lab for the controlled evaluation of systems by quantitative performance measurement, eye-tracking and live observation, including research on e-learning. Using the new equipment, interactions with *Relations* were studied again, not to evaluate or redesign *Relations*, but to apply UT technology in an innovative way to study participants' interactive learning processes and distribution of time -- navigating; gaining knowledge during experiences with an asynchronous tutorial; and an initial sub study of errors (De Villiers 2009). This research is a more extensive and in-depth follow-up, focussing, in particular, on the nature of errors made.

Figure 3 shows the users' and observers' stations in the laboratory. Researchers observe and listen through a glass partition and via a triple-screen monitoring view (Figure 4) to see users' keystrokes, face- and-body language, and the screen with which they are interacting.

Figure 3: The participant's work station

Figure 4: The researcher's view>

Evaluation studies of e-learning systems

Studies on e-learning have certain notable differences from evaluation studies of conventional systems. Conventional usability evaluation assesses the usability, defined by ISO 9241-11 (1998) as '... the extent to which a product can be used by specified users to achieve specified goals with *effectiveness*, *efficiency* and *satisfaction* in a specified context of use'. Task-based systems have products as output, e.g. a payroll or report. Evaluation assesses whether their functionality and interfaces are easy to use and operate as expected. Evaluation of e-learning, however, has notable differences (Masemola and De Villiers 2006):

- E-learning applications are focused more on a process than a product, the process being learning.
- Evaluation of e-learning should consider both usability and pedagogic issues. As well as being computing artifacts, many e-learning applications include course material.
- With conventional systems, the faster a task is done, the better. With e-learning, efficiency cannot be determined by rapid completion. Users should learn using personal styles and at their own pace.
- Users of conventional systems should make as few errors as possible, but when educational systems are used, it is not always desirable to minimise errors. Researchers should distinguish between peripheral *usability-related errors* and *cognitive errors*, which are part of the learning experience (see literature review).

RESEARCH DESIGN AND METHODOLOGY

Research questions

This study applied usability testing technology in an innovative way -- not to evaluate the e-learning tutorial, but to investigate participants' learning processes as they did specified tasks on *Relations*. The research questions are:

- How can electronic tutorials support cognitive learning processes?
- What types of errors did participants make, and what was their role in the learning experience?

Pilot study

Boshier (2006, 52) emphasises the importance of a pilot study with a sample similar to that of the main study. A pilot study was conducted with five participants to confirm the efficacy of the research approach, and this resulted in marginal refinements to the data collection instruments and analysis process.

Main empirical study and participants

In the ODL situation, participants were recruited by SMSs, tutorial letters, and an invitation during a tutorial class. The study was conducted in the HCI Lab of Unisa's Muckleneuk campus. The 17 participants in the main study were a heterogeneous group, all living within the 60 km radius of Unisa. Seventeen may appear small as a sample size, but it is large in the context of a UT study, with its time-intensive analysis and considerably higher than Hwang and Salvendy's (2010) suggestion of 8 to 12 participants. Most participants were under 30, which corresponded well with the cohort, where 75 per cent were under 30. Sixty-five per cent were males, whilst the cohort constituted 40 per cent males. Ninety four per cent were African and 6 per cent White, while the cohort was 64 per cent African, 21 per cent White, and 15 per cent Coloured or Asian. All were computer literate and conversant in English, the medium of tuition. A printed tasklist directed the participants to do certain units of *Relations,* involving theoretical concepts, graphical demonstrations, worked examples, and exercises where they entered mathematical symbols or terms. Sometimes more than one answer was correct, provided it met specific mathematical conditions. The tutorial provided diagnostic feedback to answers and second attempts were required after wrong answers. The researchers determined how the participants distributed their learning and practicing time by observing the cursor pointing on the screen; recording response times and listening to think-aloud. Details of these aspects are reported in a parallel study (De Villiers and Becker 2012) that also discusses the qualitative findings from observations and interviews with the same participants.

Data collection and analysis

In *quantitative measurement*, we recorded times spent on fine-grained sub-activities, including times on different types of errors and time studying feedback. We tabulated the numbers of correct and incorrect responses, times to recover from errors, as well as scores from overall performance in *Relations* sessions. Times spent on major sub-activities varied considerably amongst participants. In *qualitative research*, we studied the various types of errors, by real-time observation during sessions and by re-viewings the video recordings. Analysis was facilitated when participants were comfortable in thinking out loud.

Ethical aspects

Ethical clearance was obtained from the College (Faculty) Ethical Clearance Committee. The procedures and lab environment were explained to participants, who signed informed consent forms. No ethical problems occurred during the study and, overall, the participants were enthusiastic regarding their experiences.

Triangulation, validity, reliability

Cohen, Manion and Morrison (2005) define *data triangulation* as using two or more data collection methods in studying some aspect of human behavior. This article discusses the quantitative findings, while a parallel study discusses the qualitative findings from observations and interviews with the same 17 participants (De Villiers and Becker 2012). The two studies triangulated the data and the two sets of findings tallied with each other. This study, in itself, adopted a mixed-methods approach, comprising the collection of quantitative measures supported by qualitative observation.

Validity and reliability relate respectively to the appropriateness of the research methods and the quality of the data obtained. According to Cohen, Manion and Morrison (2005), *validity* entails showing that the research instruments measure what they are supposed to measure. *Reliability* in quantitative studies involves consistency, accuracy and replicability over time, instruments, and groups of participants; while in qualitative studies, reliability means a good fit between the data recorded by the researcher and what occurs in the natural situation being investigated (Cohen, Manion and Morrison 2005).

The impact of the implementation of triangulation, as well as the ways in which validity and reliability were ensured, are addressed in the conclusion.

LITERATURE REVIEW AND APPLICATION

Earlier sections cited literature on asynchronous tutorials and evaluation. This section overviews various dimensions of cognition and errors and relates them to *Relations*.

Memory, remembering and learning

There are various forms of memory (Kidd 1973). *Recognition* means selecting something that has been learned, while *recall* is remembering. When a product lends itself to recognition rather than recall, it contributes to its usability (Nielsen 1994). Kidd proposes *repetition with meaning* as an approach to practicing skills. He addresses the difference between *rote learning* and *meaningful* *learning*, emphasising that *retention* is strengthened when material is presented meaningfully. In discussing *cognitive learning*, Kidd cites Jean Piaget and Jerome Bruner, explaining that cognition relates to locating information in memory and retrieving it.

Cotton (1995) mentions the learning styles of memorisation, understanding, and doing. *Memorisation* is a daunting task for adult learners, which can be supported by visual memory as in *Relations*, where theory is either on the same screen as exercises or available via hyperlinks. Learning by *understanding* is an individual process, whereby learners need different time spans to grasp concepts. Cotton recommends using different forms of explanations, as done in the *Relations* environment where concepts are approached from multiple perspectives: text, graphics, and evolving animations -- bringing constructivist understanding. Learning by *doing* is implemented in *Relations* when mathematical skills are learned, firstly, by on-screen demonstrations of annotated, colour-coded examples, and secondly, by doing hands-on exercises.

Conceptualising and categorising

Kidd (1973) stresses the role of concept formation in adult education. Order should emerge in a rational process as learners grasp a field and its phenomena, objects, definitions, and problem-solving strategies. This is an objective process that articulates the concepts that give order and meaning to objects and the relationships between concepts. This occurs in *Relations*, which operates in a closed mathematical domain (De Villiers 2005), where answers are fixed and not open to options, except in cases where more than one answer is correct if it meets fixed mathematical conditions. Some sections in *Relations* define and illustrate the basic constructs of mathematical relations, while more advanced sections introduce composite relationships between basic elements.

Learning through cognition and errors

Cognition occurs as adults learn in various ways-- remembering, understanding and doing, supported by experience, reflection, theoretical knowledge and preparation (Cotton 1995). Errors play a role in this process. In seminal work on the usability of 'educational software', Squires and Preece (1999) distinguish between usability errors (system-related) and cognitive errors, which are part of the learning process, particularly in complex domains. *Usability errors* pinpoint systemic problems that should be avoided and rectified in cases where they do occur. In contrast, *cognitive errors* relate to students' misunderstanding of major learning issues. They are useful and should be permitted, along with mechanisms to promote recognition and recovery. Students learn from these mistakes.

According to Mayes and Fowler (1999, 485), 'seamless fluency of use is not necessarily conducive to deep learning ... the software must make learners think'.

In classifying errors made when learning to use software, Kay (2007) cites the general errors proposed by Reason (1990)-- *slips* (incorrect execution of a correct plan), *lapses* (memory-related), *rule-based* (when users apply inappropriate rules), and *knowledge-based* (resulting from incorrect representations of problems). Kay asserts that *HCI errors* (e.g. usability errors) are low-risk occurrences. They do not lead to serious financial or organisational time losses, but are frustrating and consume personal time. Kay's study identified six types of errors, occurring when users were:

a. conducting an *action*; *orienting themselves* in a search;

b. *processing* knowledge;

c. *seeking* information; in an error-prone *state*; or

d. working in a distinctive *style*.

An advantage of computer-based instruction is that it can individualise the learning process by identifying students' error patterns and diagnostically providing more activities on that instructional content (Grant and Courtoreille 2007). This feature is not present in *Relations*, where fixed questions are administered to each user, although users have an option of bypassing theory.

FINDINGS AND DISCUSSION

After a brief orientation in the laboratory, participants signed informed consent forms. In line with UT practice, they did not tackle the entire system, but were given a printed tasklist directing them to specified tasks on the *Relations* CD, each of which included theory, examples, and exercises. The first three related to Reflexivity, Symmetry and Transitivity, respectively, while the fourth was a capstone on Equivalence Relations, which integrated the first three.

RELATIONS AND ITS FEEDBACK

Boshier (2006) found that adult learners perceive criticism and lack of encouragement in a negative way. In contrast, *Relations* provides immediate positive feedback to correct and incorrect responses. It is supportive, and aims to support recovery from cognitive errors by mechanisms to promote a recognition-diagnosis-recovery cycle. *Relations* is not an intelligent tutor and does not implement artificial-intellligence techniques, as in the system of Gonzalez, Guerra, Sanabria, and Moreno (2010), which differentiates between errors and detects causes. However, there are limited diagnostic capabilities providing customised feedback to particular types of responses via pattern matching or by calculations. In other cases, generic feedback and hints are offered. Second attempts are permitted when answers are incorrect. Scaffolding and feedback support cognition, in line with Kay's (2007) emphasis on handling errors by detection, diagnosis, and correction.

Categories of errors

In this study on learning by e-learning in the *Relations* environment, the categories of errors are based on the errors defined by Squires and Preece (1999); namely, usability-related errors and cognitive errors with an additional category defined for errors that occurred when participants used the tasklist:

1. *Errors in using tasklist*: mistakes and delays from consulting ancillary material, i.e. when participants misread directions on the printed list;

2. *Usability errors*: caused by system weaknesses; and

3. *Cognitive errors*: they are part of the learning process.

***Errors in using tasklist*:** CD-based applications are self-contained with built-in instructions. However, for experiments and evaluation, participants undertake specified tasks only. Some participants found it difficult to focus both on the printed tasklist and the screen -- see table 1.

Table 1: Errors in using tasklist (times in seconds)

Actions (n = 17)		Task 1: Reflexivity	Task 2: Symmetry	Task 3: Transitivity	Task 4: Equivalence relations
Number who took ≥30 seconds		5	0	0	7
Time reading printed direction	Range:	5--63	0--29	5--30	10-60
	Average:	29	15	14	34
	Outlier:	111			
Time on errors	Range:	0--64	0--45	-	0--15
	Average:	19	11	0	2

Number who made errors	10	8	0	3
Number who needed help	4	3	0	2

Table 1 shows that the joint use of a list and the CD initially caused hesitation and errors, but the situation improved. The number of participants requiring more than 30 seconds to read instructions were 5 for Task 1, followed by 0 for Tasks 2 and 3. The number increased to 7 for Task 4 in a different section of *Relations*. The average time wasted by errors in reading the list decreased through 19, 11, and 0 seconds, and increased slightly to 2 for Task 4. The numbers of participants who made errors (10, 8, 0, 3), and who asked the researchers for help (4, 3, 0, 2) showed a similar tendency. These findings suggest that participants adapted easily to the situation.

Usability errors: These result from incorrect use of the system. They cause frustration and also waste time.

Table 2: Usability errors (times in seconds)

	Actions (n = 17)		**Task 1: Reflexivity**	**Task 2: Symmetry**	**Task 3: Transitivity**	**Task 4: Equivalence relations**
	Time on errors	Range:	0-57	0-15	0-15	0-24
		Average:	**17**	**1**	**1**	**2**
		Outlier:	84			
	Number who made errors		11	2	2	3

Due to rigorous evaluation by various UEMs, *Relations* had very few usability errors. Table 2 shows that participants were able to overcome the few barriers to usability. The number of errors and the time spent on errors declined drastically as participants grew accustomed to *Relations.* The total average time spent on usability errors for all four tasks was only 17+1+1+2 = 21 seconds, and the most time wasted by an individual was 84 seconds. A similar decreasing pattern was observed for the number of participants who made errors (11, 2, 2, 3).

Cognitive learning and cognitive errors: Table 3 shows the times that participants spent respectively on cognitive learning and on cognitive errors. How do we know how they distributed their time? We looked at the screens on which they were working and the activities they were performing, which indicated whether they were engaging cognitively with theory, controlling an animated model, entering answer to exercise, or studying the valuable feedback which become immediately after they completed each exercise. The analysis was also supported by listening to their think-aloud (done better by some than by others), noting their facial expressions, and tracking mouse movements. It was thus possible to determine the durations of activities and learning processes as they engaged interactively with reflective reading, evolving animations, or step-wise worked examples. The <Repeat> function was available for re-doing theoretical sections or activities. All but one participant were engaged, even engrossed, in the learning experiences. The respective times spent on learning activities and on cognitive errors resulting from misconceptions or forgetting, were recorded. The 'Time spent on cognitive errors' in Table 3 includes reading the feedback to their responses and reflecting on it.

Table 3: Cognitive errors (times in seconds)

Actions (n = 17)		Task 1 Reflexivity	Task 2 Symmetry	Task 3 Transitivity	Task 4 Equivalence relations
Time spent on cognitive learning -- active engagement	Range:	111-279	188-496	201 - 717	157-662
	Average:	220	376	565	516
	Outliers:	375, 515	652, 831	1025,1263	810, 918
Time spent on cognitive errors in exercises	Range:	0-83	0- 78	0-384	0 -160
	Average:	46	85	182	101
	Outliers:	127, 165	310, 323	431, 437	277, 322
Number who needed help		3	5	6	1

The extent of cognitive errors did not decrease as was the case with the other types of errors. Although participants settled into practical aspects of using *Relations*, they did not reduce the occurrence of cognitive errors. Each unit of academic material presented new challenges, and the times spent on cognitive learning and errors increased with increasing complexity and extent of work in successive tasks. Times increased over Tasks 1 (Reflexivity) and 2 (Symmetry), peaking at Task 3 (Transitivity, the most difficult). For these four tasks, times taken manifested similar patterns for learning: (220, 376, 565, 516) and errors: (46, 85, 182, 101), respectively. The complexity of transitivity is reflected by the longer times on learning and errors and by the most participants (6) asking for help. Task 4 (Equivalence relations) was a capstone, integrating the three previous concepts.

Due to the complexity of Transitivity and the consolidation role of Equivalence Relations, Tasks 3 and 4 included two major exercises, while Tasks 1 and 2 had one each, i.e. six exercises in total in the defined tasks. To investigate post-error conditions in these exercises, reflection on the feedback and recovery from errors was analysed.

Table 4: Time on exercises and error recovery

Opportunity	Exercises											
	1		2		3		4		5		6	
	Number	Average time (secs)	Number	Average time (secs)	Number	Average time (secs)	Number	Average time (secs)	Number	Average time (secs)	Number	Average time (secs)
First_ attempt correct	**10**	40	**8**	18	**11**	61	**6**	66	**9**	64	**16**	87
Second attempt correct	**3**	21	**2**	40	**3**	58	**1**	96	**5**	63	**1**	211
Second attempt incorrect	**4**	65	**7**	53	**3**	53	**10**	40	**3**	29	**0**	--

Table 4 shows six exercises (one each in Tasks 1 and 2; two in Tasks 3 and 4). There were 6x17 = 102 first attempts across the six exercises. Of these, 10+8+11+6+9+16 = 60 (59%) of first attempts were correct. Of the second attempts 36 per cent (15/42) were correct following the feedback, attesting to its value. No pattern is evident in the average times spent, nor in the relationship between average times and correct/incorrect responses. Furthermore, we investigated the relationships between times spent on *Relations* and total marks. Figure 5 shows that there was no relationship between total times spent on the four tasks and total marks for the tasks. Scores were distributed in a non-systematic pattern, in which high scores were attained by participants who took less time, who took average times, as well as some who took spent the most times doing the tasks. Similarly, low scores were obtained by participants who worked at varying paces. This lack of a relationship between times and scores emphasises the role of tutorials in self-paced learning and in fostering personal learning styles. Three participants each took about an hour, and their scores ranged from the lowest (25%) to high (83%) -- as shown in both Table 4 and Table 5.

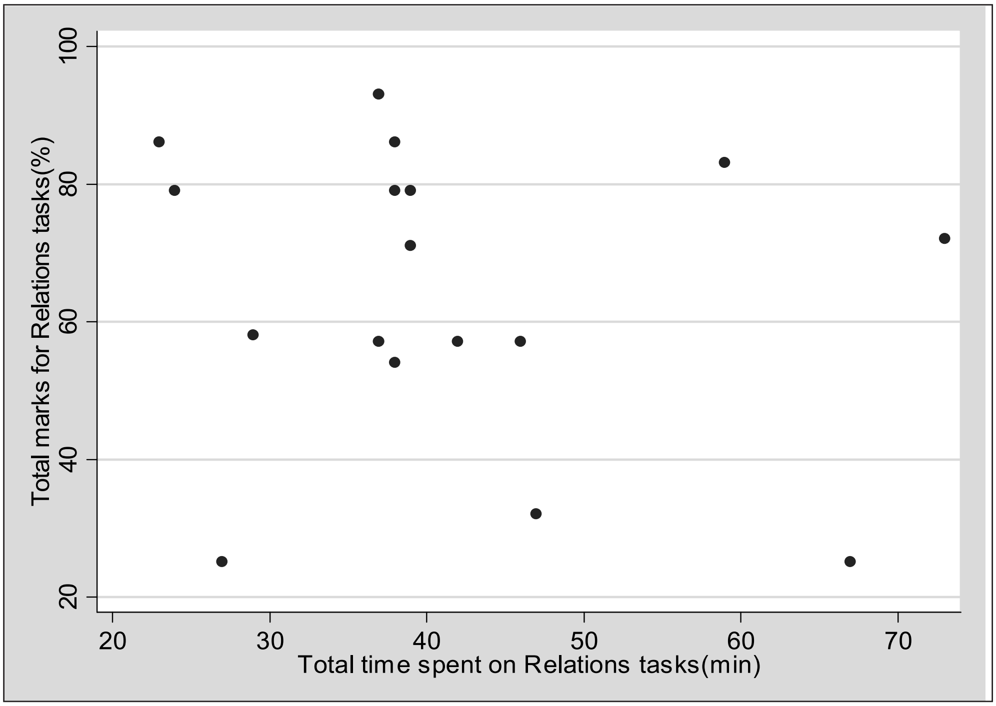

Figure 5: Times on Relations versus scores

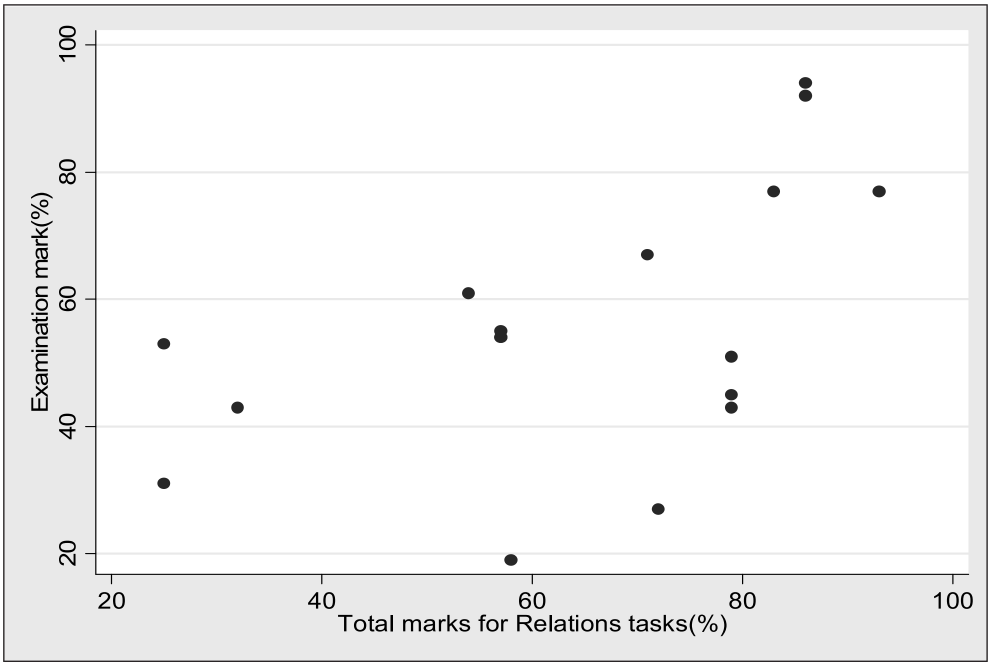

Figure 6: Total marks on Relations tasks versus examination mark

Figure 6 plots the end-of-year examination mark against the accumulated mark for the four tasks. For n=17: r=0.5226 with p=0.0314, indicating a statistically significant correlation. Although there is a strong tendency among 11 of the participants towards a linear relationship between examination marks and performance on *Relations* tasks, six lie outside the linear relationship, achieving higher scores in *Relations* tasks than their natural level in the examination. We contend that the additional support in the *Relations* tutorial contributed to this success.

Table 5 shows this data in tabular form, presenting total scores on the four tasks (including all the exercises), and examination marks.

Table 5: Total time spent on Relations tasks; total marks for tasks; examination mark

Participant	**Total time on 4 tasks (minutes)**	**Total score: four tasks (%)**	**Examination marks (%)**
1	11+19+26+17=**73**	72	27
2	3+5+8+7=**23**	86	94
3	4+9+13+11=**37**	93	77
4	6+4+10+18=**38**	86	92
5	9+7+12+10=**38**	54	61
6	13+18+17+19=**67**	25	31
7	6+9+18+13=**46**	57	54
8	7+6+14+12=**39**	79	45
9	6+5+18+9=**38**	79	43
10	4+6+9+10=29	58	19
11	10+8+24+17=**59**	83	77
12	6+5+6+7=**24**	79	51
13	11+11+15+10=**47**	32	43
14	7+10+11+9=**37**	57	54
15	6+9+17+10=**42**	57	55
16	7+5+6+9=**27**	25	53
17	9+9+12+9=**39**	71	67
Average	**8**+**9**+**12**+**12**=**41**	65	56

Performance scores in Table 5 show that 10 participants obtained similar scores in their *Relations* tasks and in the examination. Of the seven participants whose two scores differed notably, six demonstrated their potential by scoring higher in the tasks than in the examination. Four of those six failed the examination, and two obtained distinctions for *Relations*, and this indicated the level of their performance in a topic where they had additional support and self-pacing. The average score for the *Relations* tasks (65%) is higher than the average examination mark (56%). The sample is small, but ability was demonstrated, giving a qualitative indication that six of the 17 achieved above their norm. If similar tutorials had been available in other topics, they might have obtained similar scores across the board.

There was clustering around the median -- seven participants took 37--39 minutes and did well in the tasks.

RESEARCH QUESTIONS REVISITED

This study used more participants than is usually the case for usability-testing studies, yet it is a small scale to confidently generalise the findings. Nevertheless, valuable information has been obtained on cognitive processes and errors made while participants were learning with an e-learning tutorial. We re-visit the research questions and provide concise findings, before concluding with discussion and recommendations.

1. *How can electronic tutorials support cognitive learning processes*?
   Investigations in the HCI Lab confirmed the value of *Relations* as a supplementary mode of presentation over and above printed material. Participants' gazes and think-aloud confirmed that the graphics and animations supported cognition and the interactive activities engaged them. In post-UT interviews, all 17 were in favour of using technology to supplement traditional teaching, and praised the enriched perspectives gained from multiple perspectives, authentic examples, and graphical representations. In particular, the self-paced nature of the e-learning tutorial supported individual styles of learning. The absence of any relationship between times spent and scores obtained, attests to this individualisation. Participants studied the feedback intently, suggesting that it added value to their learning experiences.

2. *What types of errors did participants make, and what was their role in the learning experience*?
   Some participants initially misread the tasklist or the fine print on the screen. More than half made initial errors by applying printed instructions in tandem with the electronic system, but adjusted rapidly.

Regarding usability and ease of learning, participants learned quickly how to use *Relations*. After Task 1, an average of only 7 seconds was wasted on usability errors, indicating that in the well-designed and rigorously-evaluated tutorial there were few such usability problems. Users easily overcame the minor hindrances.

By contrast, they continuously made cognitive content-related errors as they encountered new challenges and complexities in each task. However, in the second attempts made on major exercises after using the feedback incorporated in *Relations*, 15 of the 42 second attempts (36%) were correct. These participants learned from their mistakes and the feedback, and were able to self-correct.

CONCLUSION AND RECOMMENDATIONS

The use of usability-testing technology proved its worth in this investigation of learning and the different kinds of errors made while using an e-learning tutorial. Repeated viewing of video recordings allowed the researchers to study usage and errors perceptively, providing information and insights beyond what could be obtained with field studies. Over and above iterative design and development, a series of comprehensive evaluations had been conducted on *Relations* to remove usability errors and improve learning content, thus making it an appropriate system for studying students' cognitive errors, unimpeded by system errors. It demonstrated its value in providing an effective interactive learning experience.

We briefly re-visit triangulation, validity and reliability (Cohen, Manion and Morrison 2005), mentioned in the section Research Design. Data triangulation was implemented by a mixed-methods approach; namely, the collection of quantitative metrics regarding times spent on activities, scores in the post-task tests during lab sessions, as well as results of final examinations. The findings were supported by qualitative observation and in-depth analysis of session activities.

Validity involves proving that the research instruments measure what they are intended to measure, while quantitative reliability involves consistency and accuracy. These are supported by the equipment in the HCI Lab, which has been accurately calibrated and is in regular use. A competent lab facilitator assists researchers with the technological techniques. As mentioned, we conducted a pilot study in which we ascertained the appropriateness of the methods, tasks and instruments, thus ensuring validity of data processing in the main study, and this resulted in minor revisions to the tasks and interview questionnaire (the latter for the separate qualitative study). The reliability of the findings was confirmed by re-viewings of the video recordings, at least twice by each of the three researchers. Approximately two man-days were spent analysing each of the 17 video clips. In qualitative studies, reliability means a good fit between the recorded data and what occurs in the natural situation. In the interviews participants were asked whether they were put off by the laboratory environment and being observed. Only three felt daunted, while the others were as relaxed as they were at home, owing to the warm and informal tone of the researchers.

Finally, the findings from the parallel qualitative study (De Villiers and Becker 2012) confirmed the results of this study, as well as provide complementary information beyond the present scope.

Contribution of this research

First, in answering its research questions it trod new paths in highlighting personal approaches to learning and the varying times used by different learners in interactions and cognitive engagement with the same e-leaning tutorial. These styles were independent of the marks they got, indicating that personal learning styles differ across the range of intellectual capabilities.

Second, we recommend the use of well-designed e-learning tutorials as supplementary study material to foster liberating and, most importantly, individual self-paced experiences in the ODL context. This is in line with the call by Boud et al. (1996), who advocate learning approaches that liberate adults from habitual ways of thinking and acting. Design and development of e-learning materials by multi-disciplinary teams is time-consuming and expensive, calling for financial resources and multi-disciplinary development teams. Owing to the demand on resources, the research should be undertaken only in modules with stable content. Moreover, existing tutorials could be extended with further questions. In line with the formats suggested by Grant and Courtoreille (2007), the fixed-item question format could be supplemented with response-sensitive questions that give students further problems of the sorts they had answered incorrectly.

Relations can be viewed as a pioneering venture that could be used by other higher-education institutions. Its value as a supplementary interactive learning tool in an ODL environment has been confirmed by positive feedback from participants. Boud et al. (1996) assert that learners who enter a learning experience from a background of success in similar situations are better prepared to transfer to the next context. For the isolated distance learner, a positive encounter with one tutorial would raise the expectation of enhanced learning with others. Let us rise to the challenge and develop interactive, engaging, and motivational study material that enhances the learning experience.

REFERENCES

Alessi, S. M. and S.R. Trollip. 2001. *Multimedia for learning: Methods and developments*. Third edition. Allyn & Bacon.

Barnum, C. M. 2002. *Usability testing and research*. Allyn & Bacon.

Bloom, B. S. 1998. *Bloom's Taxonomy of the cognitive domain*. Available at: http://classweb.gmu.edu/ndabbagh/Resources/Resources2/bloomstax.htm (accessed 4 October 2011).

Boshier, P. 2006. *Perspectives of quality in adult learning*. London: Continuum International.

Boud, D., R. Keogh and D. Walker. 1996. Promoting reflection in learning. In *Boundaries of adult learning*, eds. R. Edwards, A. Hanson and P. Raggatt, 32--56. New York: Routledge.

Cohen, L., L. Manion and K. Morrison. 2005. *Research methods in education.* Fifth edition. London: Routledge.

Cotton, J. 1995. *The theory of learning: An introduction.* London: Kogan Page.

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

------. 2007. An action research approach to the design, development and evaluation of an interactive e-learning tutorial in a cognitive domain. *Journal of Information Technology Education* 6: 455--478.

------. 2009. Applying controlled usability-testing technology to investigate learning behaviours of users interacting with e-learning tutorials. In *Proceedings of E-Learn 2009, World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education.* Canada: Association for the Advancement of Computing in Education.

De Villiers, M. R. and D. Becker. 2012. Qualitative findings from usability testing of an e-learning tutorial. In *Proceedings of E-Learn 2012, World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education.* Canada: Association for the Advancement of Computing in Education.

Dix, A., J. Finlay, G. D. Abowd and R. Beale. 2004. *Human-computer interaction*. Third edition. Pearson Education.

edX. 2012. *About edX*. Available at: http://goo.gl/fUxWD (accessed 21 September 2012).

Grant, L. K. and M. Courtoreille. 2007. Comparison of fixed-item and response-sensitive versions of an online tutorial. *The Psychological Record* 57(2): 265--272.

Gonzalez, C. S., D. Guerra, H. Sanabria and L. Moreno. 2010. Automatic system for the detection and analysis of errors to support personalized feedback. *Expert Systems with Applications* 37(1): 140--148.

Hung, J. L. 2010. Trends of e-learning research from 2000 to 2008: Use of text mining and bibliometrics. *British Journal of Educational Technology.*

Hwang, W. and G. Salvendy. 2010. Number of people required for usability evaluation: The 10±2 rule. *Communications of the ACM* 53(5): 130--133.

ISO 9241-11, International Organization for Standardization. 1998. *Ergonomic requirements for office work with Visual Display Terminals (VDTs) -- Part 11: Guidance on usability.* Geneva, Switzerland.

Kay, R. H. 2007. The role of errors in learning computer software. *Computers & Education* 49(2): 441--459.

Kidd, J. R. 1973. *How adults learn*. Association Press.

Mackness, J., S. Mark and R. Williams. 2010. The ideals and reality of participating in a MOOC. *Proceedings of the 7th International Conference on Networked Learning*. Available at: http://goo.gl/8OBgQ (accessed 21 September 2012).

Masemola, S. S. and M.R. de Villiers. 2006. Towards a framework for usability testing of interactive e-learning applications in cognitive domains, illustrated by a case study. In *Service-oriented software and systems. Proceedings of SAICSIT2006*, eds. J. Bishop and D. Kourie, 187--197. ACM International Conference Proceedings Series.

Masie, E. 2008. What is the meaning of the *e* in e-Learning? *American Society for Training and Development: Handbook for Workplace Learning Professionals*, ed. E. V. Biech, 377--380.

Mayes, J. T. and C. J. Fowler. 1999. Learning technology and usability: A framework for understanding courseware. *Interacting with Computers* 11(5): 485--497.

MIT. 2011. *MIT launches online learning initiative*. Available at http://goo.gl/kWsDN (accessed 21 September 2012).

------. 2012. *MITx, MIT's new online learning initiative*. Available at http://mitx.mit-edu/ (accessed 29 August 2012).

Nielsen, J. 1994. Heuristic evaluations. In *Usability inspection methods*, ed. J. Nielsen and R. L. Mack, 25--62. New York: John Wiley & Sons.

Rosenberg, M. J. 2001. *E-learning: Strategies for delivering knowledge in the digital age.* McGraw-Hill Professional.

Rubin, J. and D. Chisnell. 2008. *Handbook of usability testing: How to plan, design, and conduct effective tests.* Second edition. Indiana: Wiley Publishing Inc.

Sharp, H., Y. Rogers and J. Preece. 2007. *Interaction design: Beyond human–computer interaction*. Second edition. New York: John Wiley & Sons.

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