



Making sense of audit trail data

Gregor E. Kennedy and Terry S. Judd

The University of Melbourne

In this paper we argue that the use of audit trail data for research and evaluation purposes has attracted scepticism due to real and perceived difficulties associated with the data's interpretation. We suggest that educational technology researchers and evaluators need to better understand how audit trail data can be processed and analysed effectively, and identify three stages of audit trail analysis. We present an investigation of a computer based learning resource as a vehicle for exploring strategies that can assist researchers and evaluators in the analysis and interpretation of audit trail data. The analytical approach we describe is iterative in nature, moving to greater levels of specificity as it proceeds. By combining this approach with primarily descriptive techniques we were able to establish distinct patterns of access to the learning resource. We then performed a series of cluster analyses which, guided by a clear understanding of two critical components of the learning environment, led to the identification of four distinct 'types' or 'categories' of users. Our results demonstrate that it is possible to document meaningful usage patterns at a number of levels of analysis using electronic records from technology based learning environments. The implications of these results for future work are discussed.

Interpreting audit trail data

In educational settings an audit trail describes an electronic record of users' activities within a technology based learning environment. The term 'audit trail' typically refers to the sequence of users' activities within this environment (or their 'history'). But it can also be used to describe records of more discrete user inputs, which make up this history, such as free text entries and responses to fixed questions, instructional activities and tasks.

While some researchers have identified the potential of audit trails as a research and evaluation tool (eg. Evans, Dodds, Weaver & Kemm, 1999; Misanchuk & Schwier, 1992; Judd & Kennedy, 2001a, 2001b; Salter, 1995), scepticism has persisted about their utility. For example, in their recent book on the evaluation of educational technology, Reeves and Hedberg (2003) caution evaluators that:

The analysis of audit trail data within complex multimedia or hypermedia programs is especially challenging. When learners can go wherever they want in any sequence, the possibility of detecting interpretable paths without the input of learners becomes almost impossible (p.182).

Similarly, Salter (1995) suggests that:

Due to the nature of hypermedia, there are some inherent dangers in the interpretation of quantitative analyses... Similar paths may be chosen for quite diverse reasons. Identical movements through a lesson do not necessarily reflect a cognitive correlation (p.460).

While both Reeves & Hedberg (2003) and Salter (1995) imply that audit trails refer only to the historical sequence of users' activities, their point is a valuable one. The foremost challenge of audit trail data is their valid interpretation. While it may be relatively straightforward to record the navigational paths and activities of student users - even in complex learning environments - the interpretation of these pathways and activities can be very difficult.

While at their most basic level audit trails measure the behavioural responses and activities of users, some electronic records of students' activities have a cognitive element to them. Students' textual responses (eg. open text responses to questions or discussion list postings) and their answers to multiple choice questions or other interactive tasks (eg. creating concept maps or completing drag and drop tasks) may be indicative of how the student is thinking (their cognitive processes). For example, students' text based responses or postings on a discussion list can be used to judge their learning processes or their understanding (McKenzie & Murphy, 2000; Swan, 2003).

However, this cognitive component is often absent from electronic records of students' activities within technology based learning environments. Students' movements within programs, their access to its sections, the sequence of their behaviour and the time they spend completing tasks are all devoid of an intrinsically cognitive component. It is this type of audit trail data that researchers and evaluators have particular trouble interpreting, primarily because a single behaviour or action or sequence of behaviours or actions may be indicative of an array of different intentions, processes and outcomes (Misanchuk & Schwier, 1992; Salter, 1995).

While these purely behavioural records can be interesting in and of themselves - for example, access or usage audits can provide very useful evaluation data - often what researchers or evaluators are particularly interested in are the additional, non-behavioural meanings, which are either implied by or associated with the users' behaviours. Often what

educational technology researchers or evaluators are interested in are the intentions of users (why they chose a course of action) or the implications of users' actions (the learning processes and outcomes associated with a chosen course of action). This represents one of the fundamental challenges of audit trail data. Researchers and evaluators are required to determine non-behavioural meanings either directly from the pattern of users' activities and responses or by associating their activities and responses with additional information or measures.

This paper argues that researchers and evaluators can document meaningful usage patterns using electronic records from technology based learning environments and presents empirical evidence to support this. In addition, we explore some of the strategies researchers and evaluators may employ to assist their interpretation of audit trail data. While ultimately the interpretation of fundamentally behavioural audit trail data will, most likely, rely on external measures, we maintain that researchers and evaluators need to better understand how to construct systems which allow audit trail data to be accessed easily and how the returned data can be processed and analysed effectively. In our work with audit trail systems and the data they return we have recognised a number of stages in the process of data analysis and interpretation.

The first stage in the analysis and interpretation process is to prepare the data returned from the audit trail system for analysis. This is not always easy and the degree to which the raw audit trail data is suitable for analysis and the amount of parsing and pre-processing required will depend on how the data is originally collected by the audit trail system. A major concern about audit trail data is that it requires prohibitively extensive data processing and management (Misanchuk & Schwier, 1992; Reeves & Hedberg, 2003). Targeting data collection to specific aspects of a technology based learning environment mitigates this difficulty. In addition, carefully designing investigations and having a clear idea about the use audit trails will be put to before their implementation will help ensure that the appropriate type of data is collected.

Once the data has been prepared for analysis, the second stage involves determining whether any distinct patterns of user interactions emerge. We have used two strategies to help us identify patterns of usage. First, we use our detailed knowledge of the specific learning tasks users engage with to guide our analysis. In general we assess the degree to which users' behaviour corresponds to the structure of specific learning activities. Second, if distinct patterns of behaviour emerge for discrete learning activities we look for associations between these patterns of use across a number of learning activities. The aim of this exercise is to build profiles of different 'types' of users for a particular learning environment. A simple

example would be a computer facilitated learning module that has 'theoretical' and 'practical' sections. One set of users might review the theory before applying this in practice, while a second set of users may prefer to attempt the practical tasks before verifying what they found in the theoretical section.

The final stage in the process of audit trail data analysis and interpretation is to seek external verification for any internally established patterns of usage. If distinct patterns of usage do emerge from the second stage of analysis, external measures (observation, think aloud protocols, interviews or questionnaires) are typically needed to verify any implied meaning (Reeves & Hedberg, 2003). This may involve seeking evidence of whether different patterns of usage are based on different sets of user intentions, or whether different learning processes or outcomes are associated with different patterns of usage.

As mentioned above, this paper explores some of these issues associated with the analysis and interpretation audit trail data (particularly in Stage Two). Our discussion centres on an audit trail system we have developed (Judd & Kennedy, 2001a) and a computer based learning resource used by medical students at the University of Melbourne.

The learning environment: *Communicating with the Tired Patient*

Communicating with the Tired Patient is a stand alone computer based learning resource in the medical curriculum at the University of Melbourne (Harris, Keppell & Elliott, 2002). The program aims to assist medical students with their clinical communication skills and to help them develop an integrated biopsychosocial approach to identifying patients' problems. The learning design of *Communicating with the Tired Patient* is broadly consistent with constructivist principles of teaching and learning. More specifically, the program follows a situated model of learning by requiring students to play the role of a doctor in a simulated clinical interview (Brown, Collins & Duguid, 1989). As the doctor, students decide on questions to ask a patient and observe their response (see AUTC, 2003). Structurally the program can be thought of as a decision tree comprising a series of decision points or 'nodes' where students determine the direction in which their interview will go. The branches on the tree represent single, discrete doctor-patient interactions (a doctor's question and a patient's response).

There are two critical learning activities in the program: interviewing and note taking. The interviewing process centres on each node. At each node students can listen to or preview between two and four audio options of

questions the doctor could potentially ask the patient. Students select the question they believe to be most appropriate given the current state of their interview. After selecting a question, the patient responds via a second audio-visual display. After viewing the patient's response students have three options: (i) they can reselect a different doctor question on the basis of the patient's response; (ii) they can replay the patient's response for a more thorough examination of it; or (iii) they can continue to a new screen which contains 'Expert Comments and Questions' about the last doctor-patient interaction. A summary of this series of interactions is presented in Figure 1.

In the second critical learning activity, note taking, students are challenged to think about their interview. Students are presented with 'Expert Comments and Questions' after each doctor-patient interaction, an example of which is, "You reflected on Mrs Nacarella's [the Patient] previous response and then asked her a closed and focused question. Can you think of difficulties associated with asking this type of question?". Terms that may be unfamiliar to students (e.g. 'closed question') are hyperlinked to a glossary in an independent window. A text entry field is provided for students to make notes in. When students have finished their notes for a particular doctor-patient interaction they continue their interview. Students can view a transcript of their interview at any time which enables them to see the questions they have asked and the patient's response, the expert comments made and their own notes.

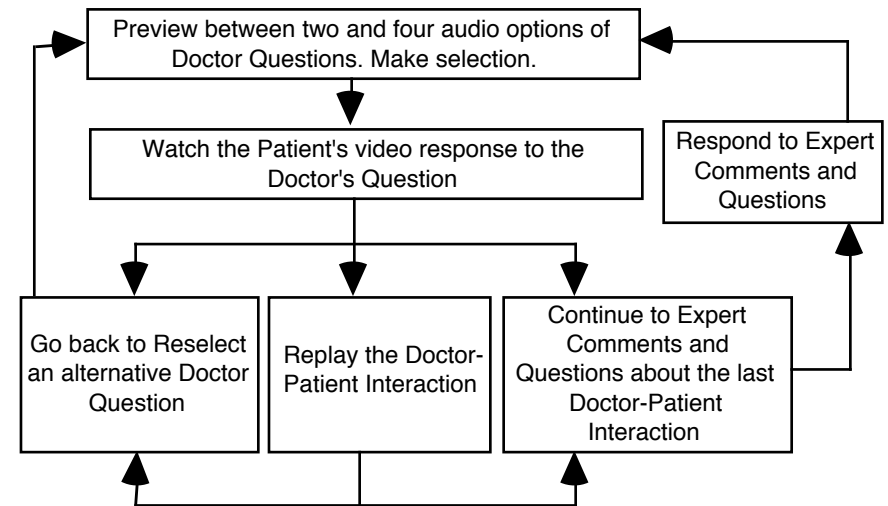


Figure 1: A summary of potential interactions when carrying out an interview

These two aspects of the program - interviewing and note taking - are central to the learning design of the courseware. The interviewing processes of previewing, replaying and reselecting allow students to compare, contrast and review clinical interviewing techniques and microskills. Note taking in response to targeted questions aims to promote reflection. By employing audit trails, we aimed to not only investigate general patterns of program usage, but also whether distinct patterns of interviewing and note taking behaviour emerged.

An audit trail system was embedded into *Communicating with the Tired Patient* and configured to use both object and event driven data collection methods. Data objects were used to capture navigational paths and notes entered by users during each stage of each attempted interview. An event driven component of the audit trail system was employed to capture a chronological record of specific user activities comprising selection, previewing, replaying and reselection of video and audio clips and selection of the glossary and transcript components during each interview. Each event captured by the system consisted of a timestamp (in seconds since the program started), an event descriptor (e.g. PreviewDoctorAudio1, SelectNode, OpenTranscript), and, depending on the event, an additional event related parameter. For example, a hypothetical 'SelectNode' event might generate the following raw data (descriptor in italics added for clarity):

213(*time stamp*) SelectNode (*event*) = 112(*node identifier*)

Data analysis and interpretation

Communicating with the Tired Patient was released on March 20, 2001. First year medical students were told the program was available and it was included in the weekly list of learning resources. While the program was designated as a resource for one week of the curriculum, audit trail data was collected for three weeks as we acknowledged that students might not access the resource in the exact week it was highlighted in the curriculum. Over this three week period the program was accessed 187 times.

We adopted an iterative approach to the analysis and interpretation of the audit trail data in an attempt to establish meaningful patterns of usage in *Communicating with the Tired Patient*. This iterative approach involved conducting a series of 'passes' at the data set. Each pass involved looking at usage patterns at a greater level of specificity and involved further refinement and reduction of the data set. As such, each pass reflects a different level of analysis (see Figure 2).

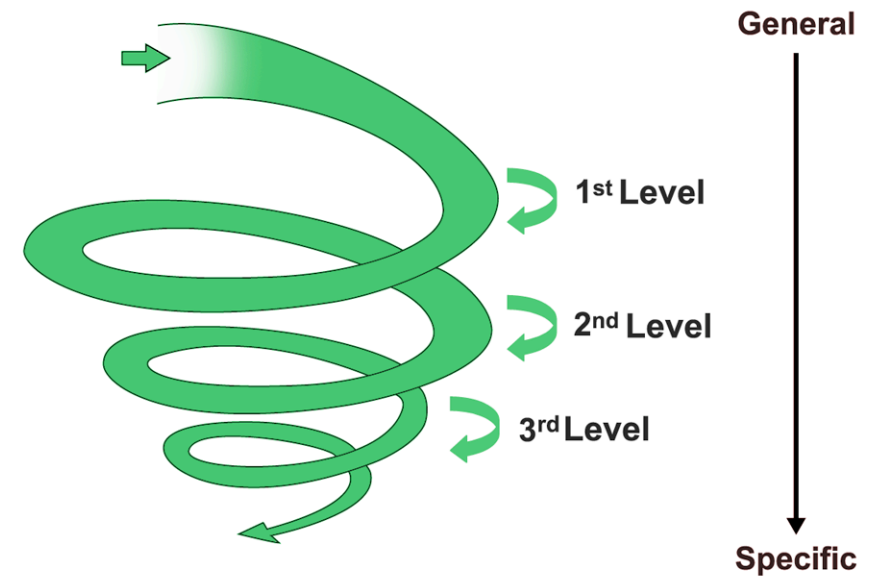


Figure 2: The iterative process of data analysis and interpretation showing different levels and specificity of analysis

First level of analysis: General access patterns (program)

In an attempt to gain an understanding of how users interacted with the program, our first level of analysis aimed to ascertain general access patterns. In order to do this we generated criteria that were used to judge each user session. The indicators were (i) the number of interviews the user started; (ii) the number of nodes visited in each interview; and (iii) the time users spent in the program. Using these criteria we determined a number of categories of users' interactions, the defining features of which are presented in Table 1. Of the 187 sessions, 32 users (17.1%) launched the program but quit from it before starting an interview, 26 (13.9%) started an interview but did not persist to an extent that was meaningful, 36 (19.3%) made a meaningful attempt at an interview but did not complete it, and 93 (49.7%) completed at least one interview.

Second level of analysis: General usage description

For the next stage of the analysis, the focus moved from documenting the general access to the program to a general description of its usage. Given our primary interest was in the two critical learning activities of users

within an interview, we restricted the second level of analysis to data from users who had completed at least one interview (n=93).

Table 1: Frequency and percentages of interviews attempted and completed by users

Category	Definition	N	%*
No Interview	The user launches the program but does not begin an interview.	32	17.1
One Interview	The user completes one interview.	61	32.6
One Interview [None]	The user begins one interview but reviews two nodes or less and /or spends less than five minutes interacting with the program.	23	12.3
One Interview [Attempt]	The user begins one interview, but does not complete it. The user visits more than two nodes and spends more than five minutes interacting with the program.	35	18.7
Two Interviews	The user completes two interviews	3	1.6
Two Interviews [None]	The user begins two interviews but reviews two nodes or less each time and /or spends less than five minutes interacting with the program	3	1.6
Two Interviews [One]	The user begins two interviews but only completes one. The second interview involves visiting two nodes or less.	22	11.8
Two Interviews [Attempt]	The user begins two interviews, but does not complete either. The user visits more than two nodes in each interview and spends more than five minutes interacting with the program.	1	0.5
Three Interviews	The user completes three interviews.	1	0.5
Three or more Interviews [One]	The user begins three or more interviews but only completes one. Subsequent interviews involve visiting less than two nodes.	2	1.0
Three Interviews [Two]	The user begins three interviews but only completes two. The third interview involves visiting less than two nodes.	4	2.1
Total		187	100%

* Refers to the percentage of the sample.

Variables of interest were obtained by parsing the raw audit trail data. Frequency counts for the variables 'nodes visited' 'reselection' 'replays' 'transcript visits' and 'glossary visits' were relatively easy to parse out as

these activities are based on users' button selections. The preview variables required more elaborate parsing as the difference between the length of an audio 'clip' and the time a user spent listening to it had to be calculated. 'Incomplete previews' were characterised by users spending more than one second previewing a particular audio clip but less than one second less than the clip's duration. 'Complete previews' were characterised by users activating the audio clip for a period of time greater than or equal to one second less than the clip's duration. Similarly, the variables 'time writing' and 'time on interview' were calculated as the time differences between key or sequential events. Word counts were derived from the data objects dedicated to recording any user entered notes.

After the critical variables were parsed out from the raw data for the first time we went through a short iteration of data cleaning. Anomalies in the parsed data were checked with the original audit trail histories of the session and the parsing routine was refined where necessary. Descriptive analyses were completed on the data output from the parser and a small number of outliers were removed from the analysis. The final descriptive statistics for the critical variables are presented in Table 2.

Table 2: Descriptive statistics for critical variables associated with users' sessions (one interview)

	Mean	SD	Low	High	Mode	Median
Nodes visited	14.46	2.22	8	18	14	15
Incomplete previews	7.01	4.05	0	20	5	7
Complete previews	32.56	14.14	0	63	30	35
Reselections	9.02	8.81	0	36	0	6
Replays	2.88	4.73	0	31	0/1	1
Transcript visits	1.11	1.32	0	6	0/1	1
Glossary visits	1.67	2.54	0	14	1	1
Words written	206.17	279.14	0	1411	0	76
Time writing*	800	836	65	3730	270	476
Time on interview*	1582	963	328	5158	501	1331
* Times given in seconds						

The relatively simple statistics presented in Table 2 provide a great deal of information about the degree of variation in the way the program was used. The minimum number of nodes a user could visit to complete an interview is eight. While five users completed the interview using the shortest possible route, on average users visited around 14 nodes. The number of complete previews was higher than incomplete ones and while one user completed 63 previews, five users completed none. The average number of reselections was nine, however the mode and the median show

that a number of users made few or no reselections. In comparison to reselections, users were not employing the replay function to a great extent. Similarly, while being used, overall the glossary and the transcript were not being accessed extensively by users. There was a great deal of variation in the number of words written and the time users spent writing their notes. The mode for 'words written' was zero and indicates that 29 users made no notes at all. Finally, there was wide variation in the amount of time users spent completing their interview; five and a half minutes was the shortest time while one user spent close to an hour and a half on a single interview.

Third level of analysis: Specific usage patterns (critical learning tasks)

The third level of the analysis centred on the critical learning activities in the program: interviewing and note taking behaviour. We were interested in specific usage patterns, and used our knowledge of the learning task to restrict our analysis to users' previewing, replaying and reselecting behaviour (see Figure 1). To investigate the relationship between these components of the interview more closely a cluster analysis was performed.

The frequency with which each user was previewing, replaying and reselecting was divided by the total number nodes he or she visited. This allowed us to get an indication of each user's activity while not biasing the sample to those who had visited more nodes in the course of their interview. We then constructed a dissimilarity matrix (using the squared Euclidean distance similarity measure) from these three new variables and performed a cluster analysis (using Ward's method) to determine whether distinct patterns of behaviour would emerge. The results of this analysis showed the cluster solutions were not being discriminated by 'Replay'. That is, all clusters were characterised by approximately the same amount of replaying behaviour regardless of how many clusters were specified. As a result replaying was removed from further cluster analysis (representing a further refinement of our data analysis).

A second cluster analysis was completed using only 'Previewing' and 'Reselecting' scores. The same similarity measure and clustering method were applied and a four cluster solution seemed most appropriate from the resulting dendrogram. Differences between the clusters were verified using a MANOVA with previewing and reselecting scores as dependant variables. A significant multivariate effect was recorded ($F(6,178) = 79.55$; $p < .001$) and univariate tests indicated significant differences between many of the clusters. The mean scores for each cluster on previewing and reselecting, and where differences between clusters occur are summarised

in Table 3. Previewing and reselecting mean scores indicate the number of previews or reselections users made per node (for example, a score of .33 would indicate a reselection every three nodes). The average number of audio previews available at each node is 2.6. Thus, if a previewing score is greater than 2.6 it is indicative of a user fully previewing all 'Doctor Questions' audio options at each node.

Table 3: Mean scores and standard deviations on Previewing and Reselecting for the four clusters

	N	Previewing		Reselecting	
		M	(SD)	M	(SD)
Cluster 1	9	.08 ^a	(.08)	1.15 ^a	(.75)
Cluster 2	38	2.03 ^b	(.36)	.30 ^b	(.32)
Cluster 3	22	2.87 ^c	(.32)	.22 ^b	(.18)
Cluster 4	24	2.84 ^c	(.52)	1.25 ^c	(.33)

abc: Mean scores with different superscripts within columns indicate differences between clusters (p<.001)

The clusters identify four distinct patterns of usage in *Communicating with the Tired Patient*. Users in Cluster One did almost no previewing but were reselecting their choice of doctor question at each node. The majority of users did engage in previewing behaviour, which can be seen in Clusters Two, Three and Four. Users in Clusters Two and Three showed similar patterns of behaviour: solid previewing and limited reselection. However, users in Cluster Three were previewing significantly more frequently than those in Cluster Two. Moreover, as the mean previewing score for users in Cluster Three was above 2.6, these users, unlike those in Cluster Two, were fully previewing all 'Doctor Questions' audio options at each node. Finally, users in Cluster Four fully previewed all options at each node and were reselecting more than once per node.

Table 4: A cross-tabulation of user category and extent of note taking showing observed and standardised residual scores for each cell (cell percentages are in parentheses)

User Cluster	Note taking (words per node)							
	< 1		about 12		about 29		about 54	
	O (%)	Res.	O (%)	Res.	O (%)	Res.	O (%)	Res.
Cluster 1	4 (4.3)	- 0.3	3 (3.2)	0.7	0 (0)	- 1.3	2 (2.2)	1.0
Cluster 2	25 (26.9)	2.6*	7 (7.5)	- 1.0	5 (5.4)	- 0.4	1 (1.1)	- 2.3*
Cluster 3	10 (10.8)	- 0.4	4 (4.3)	- 0.7	6 (6.5)	1.8	2 (2.2)	- 0.5
Cluster 4	7 (7.5)	- 2.3*	8 (8.6)	1.3	3 (3.2)	- 0.4	6 (6.5)	2.3*

* p < .05

The final stage of our analysis assessed whether users' patterns of interviewing were related to the other critical learning activity: note taking. As the availability of note taking was dependent on the number of nodes users visited, word counts for each interview were expressed as a proportion of the number of nodes visited in the interview. A cluster analysis was again performed and four clusters emerged that simply reflected increasing numbers of mean words written at each node (see Table 4).

A chi-square test was used to determine whether there was an association between the four patterns of interviewing behaviour and the degree to which users took notes on their interview. The chi-square test was significant ($\chi^2(9) = 17.12$; $p < .05$) and Table 4 presents the observed scores and the standardised residual scores for each cell. Residual scores greater than two are indicative of cells where the observed score is significantly different from the expected score. While not overwhelming given the low frequencies in some cells, the residual scores indicated that users in Cluster Two were over-represented in the low note taking and under-represented in the high note taking category. Conversely, users in Cluster Four were over-represented in high note taking and under-represented in the low note taking category.

Discussion and future directions

The results show, in some detail, how *Communicating with the Tired Patient* was subject to different types of use. At the first level of analysis we identified four types of general access to program: (i) those who started the program but did not start an interview; (ii) those who started the program but did not make a meaningful attempt at an interview; (iii) those who started the program, made a meaningful attempt at an interview but did not persist to its conclusion; and (iv) those who started the program and completed a least one interview. These patterns of access are of interest particularly given *Communicating with the Tired Patient* is an optional, self directed learning resource in a problem based curriculum. This 'access' analysis shows that audit trails can be used by program and curriculum developers to determine how well technology based resources are integrated into the curriculum. That is, such an analysis can be used to effectively show the degree to which resources are used and by how many students. This type of analysis would not only be useful in the evaluation of discovery based curricula, but also in the evaluation any curricula where students are provided with supplementary course material via the web (eg. pre-practical exercises, reading lists, class notes or quizzes).

Second and third levels of analysis were restricted to users who had completed at least one interview and the third level of analysis

concentrated on two critical learning activities. We established four distinct types of interviewing patterns based on users' previewing and reselecting behaviour. In addition, we found an association between patterns of interviewing and note taking behaviour, which allowed us to compile a more detailed picture of the four user 'types'. The first type of user completes an interview primarily using the strategy of reselection; unexpectedly this type of user does little or no previewing. The second type of user carries out an interview primarily using the strategy of previewing but does not compare all the audio options at each node. These users are also more likely to make fewer notes about their interview. This pattern of behaviour may be indicative of users who are less engaged with the program than others. The third type of user completes an interview by previewing all the audio options at each node. These users are exhibiting the 'expected' interviewing behaviour: full previewing and moderate reselection. Users in the final category preview all the audio options at each node and review their selection at every node. These users were also likely to make more notes on their interview. This pattern of behaviour may be indicative of a more conscientious type of user who is more engaged with the program than others.

We have shown that it is possible to document meaningful usage patterns at a number of level of analysis using electronic records from technology based learning environments. To add further meaning to these categories we need to complete further analyses. The next stage of our work in this area will involve seeing whether the categories of users we have established can be replicated with other cohorts of student users. This will allow us to make a strong argument about the stability of these usage patterns. Following this we will use external measures to gather information about the intentions of users in each of the categories we have established. We expect this will lead us to an understanding of why users interact with *Communicating with the Tired Patient* in the way they do; both at the level of general access and at the specific level of interviewing and note taking behaviour. Ultimately, we will look for concomitants to the user categories we have established. We are interested in whether users in different categories show differences in learning processes and outcomes in the areas of clinical interviewing and microskills, or in general learning strategies such as reflection.

Finally we are interested in extending the types of analyses completed on the audit trail data described in this paper. While our analyses so far have concentrated primarily on two discrete learning activities, an alternative analysis could consider the sequence of nodes users select over the course of their interview. This type of analysis would allow us to determine the pathways users commonly follow and if any nodes are not being accessed. Finally, in this paper our analysis of users' note taking was entirely

quantitative. A qualitative analysis of students' comments at each node could also be completed which may allow us to see where they were having difficulties or showing misunderstandings.

References

- AUTC (2003). ICT-based learning designs. Exemplar: *Communicating with the Tired Patient*. [verified 27 Jan 2004]
<http://www.learningdesigns.uow.edu.au/exemplars/info/LD46/index.html>
- Brown, J.S., Collins, A. & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32-42.
- Evans, G., Dodds, A., Weaver, D. & Kemm, R. (1999). Individual differences in strategies, activities and outcomes in computer-aided learning: A case study. In R. Jeffery & S. Rich (Eds), *Proceedings of AARE-NZARE Conference*. Melbourne, Australia: Australian Association for Research in Education. [verified 27 Jan 2004] <http://www.aare.edu.au/99pap/dod99220.htm>
- Judd, T. & Kennedy, G. (2001a). Extending the role of audit trails: A modular approach. *Journal of Educational Multimedia and Hypermedia*, 10(4), 377-395.
- Judd, T. & Kennedy, G. (2001b). Flexible audit trailing in interactive courseware. In C. Montgomerie & J. Viteli (Eds), *Proceedings of ED-MEDIA 2001 World Conference on Educational Multimedia, Hypermedia & Telecommunications* (pp.943-948). AACE. Charlottesville, USA.
- Harris, P.J., Keppell, M. & Elliott, K.A. (2002). Integration of multimedia in the problem-based learning curriculum. *Journal of Medical Education*, 6(4), 469-475.
- McKenzie, W. & Murphy, D. (2000). "I hope this goes somewhere": Evaluation of an online discussion group. *Australian Journal of Educational Technology*, 16(3), 239-257. <http://www.ascilite.org.au/ajet/ajet16/mckenzie.html>
- Misanchuk, E.R. & Schwier, R. (1992). Representing interactive multimedia and hypermedia audit trails. *Journal of Educational Multimedia and Hypermedia*, 1(3), 355-372.
- Reeves, T.C. & Hedberg, J.G. (2003). *Evaluating interactive learning systems*. Athens, GA: University of Georgia, College of Education.
- Salter, G. (1995). Quantitative analysis of multimedia audit trails. In J.M. Pearce, A. Ellis, G. Hart & C. McNaught (Eds), *Learning with Technology*. Proceedings ASCILITE 1995. (pp.456-461). Melbourne: Science Multimedia Teaching Unit, The University of Melbourne.
<http://www.ascilite.org.au/conferences/melbourne95/smtu/papers/salter.pdf>
- Swan, K. (2003). Developing social presence in online course discussions. In S. Nadia (Ed), *Learning & teaching with technology: Principles and practices* (pp.147-164). London: Kogan Page.

This article received an Outstanding Paper Award at ASCILITE 2003, gaining the additional recognition of publication in AJET (with minor revisions). The reference for the Conference version is:

Kennedy, G. and Judd, T. (2003). Iterative analysis and interpretation of audit trail data. In G. Crisp, D. Thiele, I. Scholten, S. Barker and J. Baron (Eds), *Interact, Integrate, Impact*. Proceedings 20th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education, pp273-282. Adelaide, 7-10 December: ASCILITE.
<http://www.adelaide.edu.au/ascilite2003/docs/pdf/273.pdf>

Gregor E. Kennedy and Terry S. Judd
Biomedical Multimedia Unit
Faculty of Medicine, Dentistry and Health Sciences
The University of Melbourne, Parkville Vic 3010, Australia
gek@unimelb.edu.au, tsj@unimelb.edu.au