City Research Online

City, University of London Institutional Repository

Citation: Dimitrova, M. T. (2002). Instructional Multimedia: Comparison and Enhancement of Expert Evaluation Methods. (Unpublished Doctoral thesis, City, University of London)

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/30849/

Link to published version:

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online: http://openaccess.city.ac.uk/ publications@city.ac.uk

Instructional Multimedia:

Comparison and Enhancement of Expert Evaluation Methods

Maia Tzokova Dimitrova

Submitted for Examination of Doctor of Philosophy

Centre for Human-Computer Interaction Design

City University

London

VOL II

May 2002

Volume 2

Table of Contents

Appendix 4.1.1: Multimedia Cognitive Walkthrough Documentation

Evaluating multimedia presentations

Peter Faraday and Alistair Sutcliffe
Centre for HCI Design, School of Informatics, Northampton Square, London EC1 OHB, UK e-mail: P.M.Faraday@city.ac.uk

From Faraday, P. and Sutcliffe, A. (1997) Evaluating Multimedia Presentations. *The New Review of Hypermedia and Multimedia,* 3, pp 7-37.

The paper reports the basis for a cognitive walkthrough method to support the evaluation of multimedia (MM) expository presentations (e.g. 'how to do it' type tasks). The walkthrough is founded upon an analysis of the cognitive processes and representations formed by the comprehension of an MM presentation. Issues include evaluation of attention, topic focus and information types. The walkthrough provides a series of guidelines for evaluation based on these cognitive models, such as the use of media, scripting and presentation techniques. The value of the guidelines is validated by several empirical studies. An eye tracking study is reported, providing evidence as to how visual attention responds to MM materials. A series of comprehension studies then investigate the effectiveness of a presentation before and after our guidelines were applied. The paper is illustrated with an example evaluation and studies of a commercially produced CD-ROM MM presentation 'The Etiology of Cancer' .

1 Introduction

One of the problems in the design of multimedia (MM) presentation interfaces is the difficulty in using existing iterative prototyping techniques. The costs of authoring may make repeated testing difficult and expensive, whilst simple usability evaluation may fail to uncover problems in usability. A solution may lie in formative evaluation: applying guidelines to evaluate an MM presentation as it is being designed. However current MM evaluation is largely a matter of rules of thumb, or gross measures of usability based on existing metrics. These techniques lead to problems as they fail to accurately pin-point and diagnose presentation problems.

The approach to MM evaluation taken in this paper is based on a number of issues raised by an educational psychologist, Richard Kozma [27], who suggests the need to consider the 'moment by moment' cognitive based evaluation of an MM presentation. He proposes that a 'micro-level' approach will be required to study multimedia design: '...media decisions for integrated multimedia will be micro-level decisions... The moment to moment selection of appropriate media can respond to specific learner needs and task requirements.' Kozma concludes that 'this moment-by-moment collaboration raises a different set of questions for the media researcher: what is the prior knowledge of a particular learner? What symbol systems can best represent various components of the task domain? How do these correspond with the way the learner represents the task? How do they process various symbol systems together? How can the medium process these in a way to support the learner?'

Unfortunately, Kozma does not provide any specific answers to these questions. This paper attempts to fill in some of the blanks, and show how a principled cognitive approach can not only provide a useful way to evaluate an MM presentation, but can also be shaped into a form which could be applied by practitioners, such as MM designers, who have no knowledge of cognitive models.

To do this we propose a 'cognitive walkthrough' method for MM evaluation. A 'cognitive walkthrough' is a means by which complex cognitive models can be delivered for use by designers. The walkthrough focuses on providing key guidelines and principles for evaluation, but masks the detail of the underlying cognitive models upon which they are based. In this paper the walkthrough is founded upon an analysis of the cognitive processes and representations formed by the comprehension of a MM presentation. The method is illustrated with an example evaluation of a commercially produced CD-ROM MM presentation 'The Etiology of Cancer' which is used in a series of empirical studies, comparing actual usability data against the guidelines' predications.

The paper has three main sections. First, existing approaches to MM evaluation are briefly reviewed. The second part presents a cognitive model of attention and comprehension of an MM presentation, and introduces the notion of a cognitive walkthrough for evaluation. Third, the walkthrough and guidelines are described in detail. Each step of the walkthrough is presented, together with paper based representations and guidelines to apply. The walkthrough is illustrated with examples from 'The

Etiology of Cancer' presentation and empirical studies involving eye tracking and tests of comprehension.

2 Existing approaches

The existing literature on MM design can be divided into three categories: guidelines from current HCI literature, from empirical studies of user performance, and guidelines or heuristics specific to MM evaluation.

The guidelines and heuristics which HCI have accrued appear to offer little explicit guidance to producing MM interfaces. Key issues in MM design are the delivery of information to support a task, and how this information will be processed by the user. As an example of the problems in applying conventional HCI guidelines, consider those proposed by Smith & Mosierw. Smith & Mosier offer a typical range of guidelines which describe how to design the visual part of an interface: 'Display data to users in directly usable form', 'Begin every display with a title and header, describing briefly the content or purpose', 'Use short simple sentences', 'Ensure that whatever data a user needs will be available from the display'. Whilst the intention of such HCI guidelines is obviously to provide succinct general design advice, the end result is advice which is high level and says little of specific value as to how to choose, synchronise or design an MM presentation.

Few empirical studies have addressed the effects of varying multimedia design on performance. Rieber [37] investigated the comprehension of a text and animation illustrating the operations of a piston. They suggest that the addition of text enhanced understanding of the presentation as a whole, and they argued that to be effective the presentation must sequence together both pictures and words into a unified whole. However little analysis is given as to how this sequence should be designed within the media themselves, or how such design decisions would have effected their results.

A similar study of MM design was conducted by Large et al [28]. They investigated the use of text, still image with text and animation with text in providing instruction for a descriptive presentation identifying parts of the human heart. The study again only considered all or nothing changes and failed to address issues of how the media were designed. It concluded that 'a difficult text is not necessarily enhanced by an animation even when the animation is well-designed and highly relevant. Complexity is not removed by the addition of more media.' The studies of Large and Rieber offer high level results without considering how media can be combined or designed, or how media should be sequenced.

Several authors have suggested MM evaluation guidelines. A set of principles for MM design are provided by Park & Hannafin [35]. They summarise a range of educational issues in the use of MM. Examples of the type of advice offered include to 'employ screen design and procedural conventions that require minimal cognitive resources, are familiar or can be readily understood, and are consonant with learning requirements' and to 'structure presentations and interactions to complement cognitive processes and reduce the complexity of the processing task'. Whilst the principles of Park & Hannafin clearly state that MM design should take into account the cognitive processes of the viewer, they offer little practical advice for what design requiring 'minimal cognitive resources' would actually entail.

A checklist for MM evaluation is provided by Heller[20]. It focuses on the need to combine media effectively and provides a questionnaire to rate how different media are used in the presentation. An example question would be to rate from 'ineffective' to 'effective' the 'Use of text with sound'. Again, the checklist says little about design of media and gives no details as to how media should be synchronised together. Heller notes that 'what this taxonomy [of design issues] does not address is the design process and the order of use of different media'.

In summary, little existing work addresses the evaluation *of* MM presentations. Studies generally address gross design decisions such as which media to use, and tend to be too high level to be useful.

3 Cognitive model of MM comprehension

The process by which a user comprehends a multimedia presentation is a cognitive one. Thus, in order to produce a set *of* guidelines for the evaluation *of* MM presentations, a model *of* these cognitive processes is required. This section provides a brief summary *of* issues concerning attention, short term memory (STM), working memory (WM), and comprehension *of* different media and combinations *of* media. It is proposed as a framework for cognitive issues in the evaluation *of* MM, rather than an exhaustive discussion *of* human cognition.

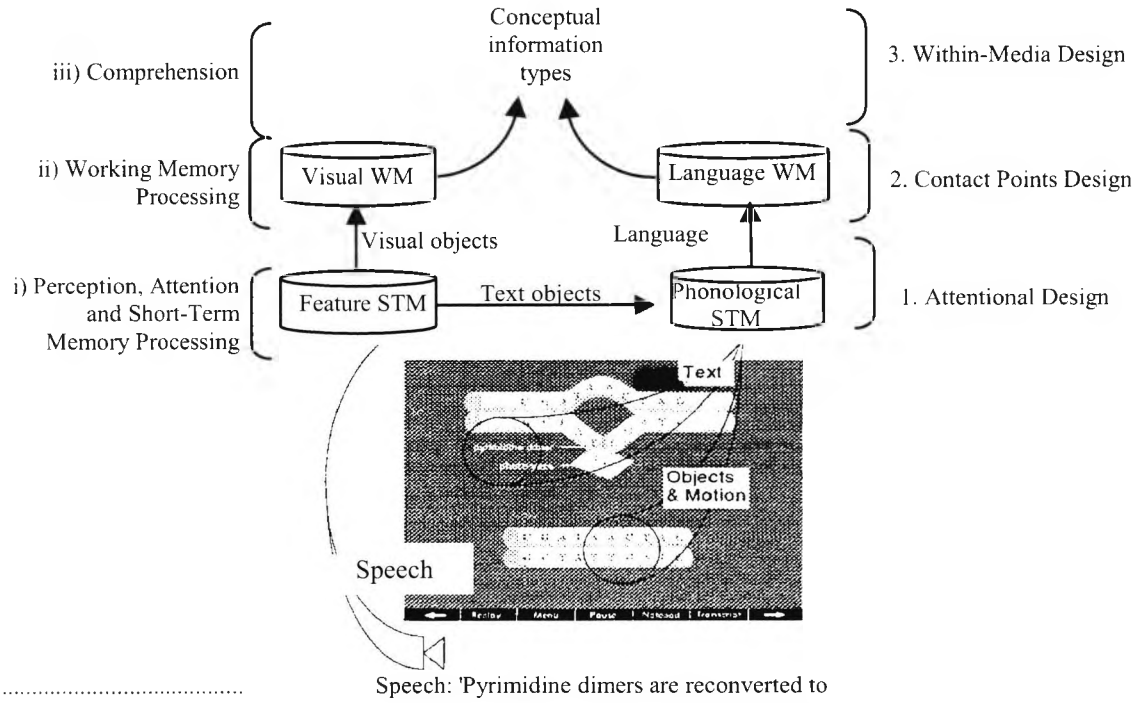


FIG. 1. *Overview of cognitive framework (left) and related evaluation issues (right)*

The structure *of* the framework is based *on* models *of* cognitive architecture and information processing, such as Baddeley4 and Miller & Johnston-Laird32. These models suggest that several processes are required to move from images and language in the world to comprehension in the viewer. The framework is shown in figure 1. It has several components: (i) attention & short term memory, (ii) working memory and (iii) comprehension. The issues raised in the framework are linked to key issues in MM evaluation. The role *of* each component in the framework and its implications for MM evaluation is discussed briefly below, and in more detail in the following sections.

The framework first covers perception and attention to different media, and combinations of media. Attentional processing has *two* separate systems for dealing with auditory and visual input [46]. The auditory part of the presentation is processed by phoneme recognisers, which extract language from other sounds [36]. The phonemes are held in the Short Term Memory (STM) phonological loop, which has a limited duration buffering capacity [3O]. The visual part of the presentation is scanned in serial by visual attention [23]. The visual information is processed by low level visual feature recognisers which extract colour, edges and texture [43]. This is stored in visual STM, which also has a limited duration capacity [43]. The issues *for* MM evaluation concern 'attentional design', how elements of the presentation should be designed to make them more salient, or 'emphasise' them; and what restrictions short term memory places on how auditory and visual media elements can be presented together within a presentation.

The second part of the framework deals with how working memory is used to combine verbal and visual media together. Visual and verbal inputs are translated into propositions, or units of meaning. Working memory acts as a storage buffer for these propositions. This allows coherence within and between these propositions; it is used to resolve co-references such as anaphor«. 'Contact points' describe ways in which visual and verbal media are integrated by co-reference. Issues for MM evaluation include how to bring combine media to form a single whole, and how to effectively design 'contact points'.

Finally, the semantic content of the proposition is constructed within a mental model. A mental model is a representation derived from the contents of the proposition, the user's long term memory (LTM) knowledge and other existing mental models. Mental models are held in a longer duration episodic memory store [24,44]. The construction of mental models is structured around a set of 'information types' based on work by Miller & Johnson-Laird [32]. The information types abstractly specify the type of content to be communicated. The evaluation issues concern 'media selection' (the choice of appropriate types of media to best support comprehension of a particular information type), and 'within media design' (media design to aid comprehension of the information type).

4 Method overview

In order to be of use to a designer, the framework needs to be re-worked as a walkthrough method, with a number of steps guiding the evaluator to ask questions about the multimedia presentation. Each section of the framework is used to provide a step in the method:

i) Evaluating Attentional Design: Section 5 of the paper explores which media will be attended to, and potential problems in sequencing and timing. An attention graph is introduced as a way of analysing which part of a presentation is likely to be attended to and a set of design guidelines are provided to produce and evaluate it.

ii) Evaluating Contact Points: Section 6 asks how well the visual and verbal media are combined together. Contact points are proposed as points at which visual and verbal media co-refer. Guidelines are given which indicate how contact points should be designed, and suggest possible problems.

iii) Evaluating Media Selection & Combination: Section 7 evaluates how the content of the presentation maps to information types, and whether the media used in the presentation are suitable. Information types are introduced which are used to evaluate the content which should be presented against what the presentation media actually provide. Guidelines suggest how different media will support the information types.

4.1 *Case study presentation*

The walkthrough method is illustrated with an evaluation of a commercially produced MM presentation taken from the 'The Etiology of Cancer' CD-ROM. The sequence analysed concerns the repair of damaged DNA. In order to provide the information context for the evaluation, a plain text version of the content was prepared from other source materials:

Repair of DNA by Photoreactivation: The pyrimidine dimers [...] can be repaired by an enzyme called photolyase that causes dimerization to be reversed by photoreactivation in a single step. The presence of a pyrimidine dimer distorts the [DNA] double helix from its normal conformation, and the photoreactivating enzyme [photolyase] binds to this distorted region. Light is not required for binding. The enzyme-DNA complex absorbs a photon of light and the dimerization process reverses. The enzyme [...] does not possess any unique light-absorbing chromophores; it is the DNA-enzyme complex that absorbs light. Once the dimerization reaction has been reversed, the enzyme dissociates from DNA, and the adjacent pyrimidine [monomers] again form [...] in the [DNA] double helix.

Extract from 'Biochemistry', J.D. Rawn, Harper & Row NY. [Brackets indicate added or removed material]

It is assumed that the MM presentation should communicate similar content using of a variety of media such as pictures, animation, speech and text. The presentation given for cancer repair is shown in figure 2. For evaluation purposes it was divided into a sequence of eight frames. These are shown as snap shots; in the presentation they also contain animation. The speech track which accompanies each frame is given below it.

5 Evaluating Attentional Design

The first stage of the evaluation segments the presentation into its constituent units. The evaluator breaks the presentation down into linguistic (speech, text) and visual media (still image, animation) and notes presentation techniques which draw attention to information in the media (such as label, highlight). These are then ordered within a timeline graph, which gives a list of the presentation units and techniques available over time.

Working across figure 3 from left to right, it can be noted that the presentation first shows an object, the DNA, and two labels, 'DNA' and 'Pyrimidine Dimer'. These are marked on the graph from their start time. In the next frame, the photolyase object is revealed, which is in motion. In the following frame the photo lyase attaches to the dimer, and a label is then revealed, 'Photolyase'. The motion of the light energy is then recorded on the graph. Finally the second DNA is revealed and the motion of the photolyase moving away from the DNA is added.

The graph is evaluated using a set of guidelines which provide advice as to which unit will be attended to, and possible contention between units. This aims to evaluate:

- The attentional thread to guide the user's viewing and reading sequence
- Effects and combinations of media to make important facts salient and reinforce the contact points between different media.
- Timing and synchronisation so that the user has sufficient time to assimilate the content.

5.1 *Still images*
IAI: By default objects focused on will be those which are: bright in colour, set apart from other objects, larger in size, shown in more detail, in sharp focus, or nearer the front of the scene. These should be objects which are important in the task.

IA2: Attention is drawn to objects which stand out or are unusual. Use highlighting techniques (colour, shape/size change) to draw attention to important objects43. By default only the scene level will be focused upon unless the user is motivated for closer inspection; or an object in the scene is unusual.

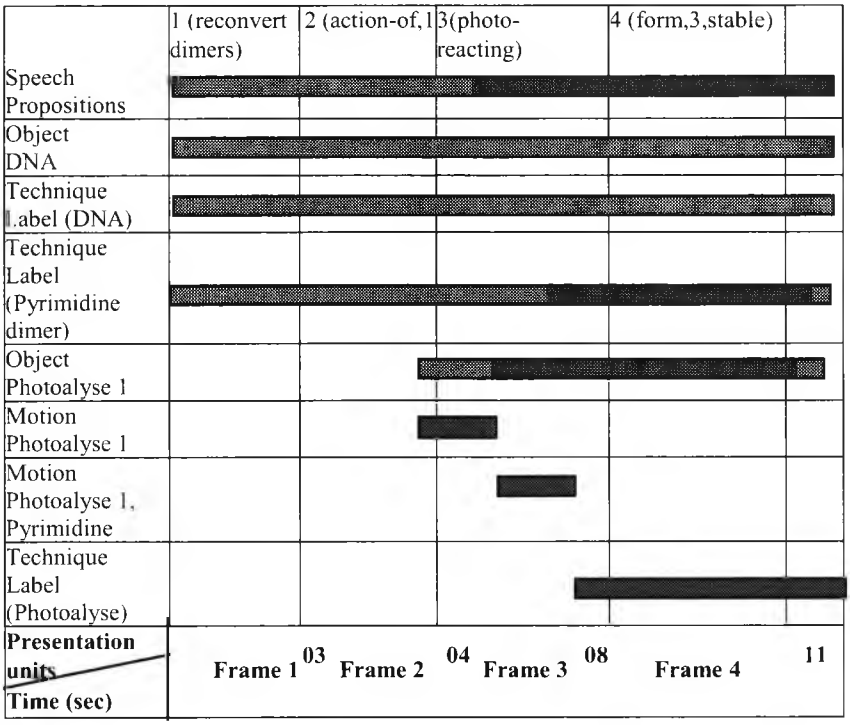


FIG.3: Attention Graph for example presentation

IA3: Gradually reveal objects, labels, symbols to control viewing order. Because attention is attracted to change, revealing each part of the presentation in turn is a useful way of directing order.

IA4: Attention groups visual objects by features, such as colour. To draw attention to a group of spatially distributed objects, set a common visual attribute (e.g. change to the same colour"3). To emphasise a group of co-located objects, highlight the background or draw a box around the objects. The highlighted area will set the granularity of the user's attention.

IA5: Icons or symbols may be placed by or on an object to draw attention [18,26].

IA6: Labels linked to objects in an image are useful for drawing attention and providing supplementary information. Dynamic revealing of labels is particularly effective and can be used to direct the user's reading sequence, drawing attention from the image to the label. .

IA 7: A void showing an object in motion or using a highlighting technique when the user is extracting information from an image; allow at least a second before changing the image [22].

lA8: Beware of using too many highlighting techniques within an image at once; instead sequence highlights to move attention from one object to another.

5.2 *Moving images*
AAl: Motion has a strong effect on attention. Onset of motion will attract attention; but it will also shift focus away from static objects in the scene.

AA2 : If attention is already focused (such as reading a text or label, or attending to another animation) motion onset may not automatically shift attention [21].

AA3: Tracking an object motion will maintain focus but may prevent focus shifting elsewhere. When tracking motion avoid revealing a caption or attentional effect during the animation or display of more than one animation.

AA4: An object which appears or moves across the ground of an image is more attention grabbing than internal movement of an object's components21. Attention will be drawn to the end point of the objects path.

5.3 *Text*

TAl: Because attention is attracted to size, detail and colour, generally an image will be focused on before a text [8]. If focus is required on the text prior to the image, the text should be displayed first, or the text area made larger than the image.

TA2: Reading text is slower than identifying an image. Always allow reading time when a piece of text is displayed [12]. Simple words require 200 msec each. Allow at least 6 seconds for two lines of text, more if the sentence is complex.

TA3: When users are reading text, focus will usually not be able to shift until a break is encountered (e.g. at the end of a clause, sentence or paragraph [10]). A void changing the image or showing animation during this time.

TA4: Use highlighting, bold, large fonts or underlining to make a particular word or clause of text stand out.

TA5: To direct attention to the required part of a text use paragraphs and titles as entry points; bullets and lists to guide attention.

5.4 *Speech & sound*

SA1: Speech is normally dominant over text, requiring no attention to process, and will in general be more alerting [37].

SA2: Speech must be processed as it arrives, it will regularly be overwritten within the phonological loop. If language to be presented is complex or lengthy, then favour text. If the language is simple or short then favour speech [38].

5.5 *Example*

The graph and attentional evaluation heuristics are used together to perform a first pass evaluation of the presentation. Working across the timeline graph in figure 3, the user's attention focus is marked on as thick black line, shown in figure 4. An arrow shows which part of the presentation is given attentional focus. If a part of the presentation is available, but is not being focused upon then it is marked on the graph as a grey line.

An overall problem with the presentation is the lack of text to support the speech material. Since the speech track is highly complex, it should be reinforced with text (SP2). Text is valuable because it allows for review. The following additional problems are also noted:

(i) In frame 1, attention focus is first given to the DNA object and its label, rather than its constituent parts. This is because the DNA is a large object, is shown in detail (IA2), and has a label (IA6). It will then shift to the pyrimidine dimer object and its label (IA6) as the pyrimidine dimer is highlighted and labelled (IAI), attracting attention to it.

A potential problem in frame I is that the dimer object and label should be revealed individually to organise the viewing order of the DNA and photolyase rather than shown all at once (IA6). Focus may be drawn unpredictably to the photolyase object, rather than sequenced from the DNA to the photolyase. Revealing each part of the presentation is a good way of ordering focus.

(ii) In frame 2-4, revealing the photolyase object will first attract attention to it (IA3), and then move focus along a path toward the pyrimidine dimer object (AAI,AA3). The appearance of the photolyase label will then cause a shift of focus (IA3) between the photolyase object and its label (IA5). A problem in frame 4 is that the photolyase and pyrimidine-dimer object should be emphasised as the protein pyrimidine-complex object. A highlight presentation technique should be given to group the two objects as the proteinpyrimidine-complex (IA4, IA2). A label should

also be linked to the protein-pyrimidine-complex object drawing attention to the combination of the pyrimidine dimer and photolyase (IA6). Highlights are useful to draw attention to, and group, important objects.

(iii) In frame 5, attention will be drawn to the light energy being revealed (IAI), it will follow the animation of the light energy object striking the photolyase (AAI). At the end of the animation, attention will be drawn to the end point of the light energy motion, the photolyase object and label (AA4). When the arrow symbol is revealed (IA5) attention will shift and follow the arrow downwards.

(iv) In frame 7-8, attention will then focus on the second DNA object and its label when they are revealed (IA3). Finally, attention will be shifted by the motion of the second photolyase object as it detaches and moves away from the DNA (AAI). A potential difficulty in frame 7 is that more reading time should be given for the repaired DNA label, which is revealed shortly before the animation of photolyase moving away from the DNA is shown (TAl). This may result in the start of the animation being missed if attention is locked on the text (TA3), or that the text may be missed because the animation is being tracked.

5.6 *Empirical evidence*

We performed a study which explored users' attention to a MM presentation using an eye tracking system to match the pattern of subjects' fixations in the visual media. Six novice subjects were drawn from research students within City University. A pupilometer system was fixed upon the subject's eye, whilst they viewed the MM presentation. This produced an x,y location of the pupil every 50 msec. Further details of this analysis are given in14.

Each subject's scan path was analysed into a fixation graph, representing fixation clusters or pauses of the subject's eye at a location. In general, a fixation will indicate attention processing8. The data for the six subjects was synchronised over time to build tip a composite trace. In general, the results followed the predications of the guidelines and provide some validation for their use.

6 Evaluating contact points

When combining verbal media and visual media special care must be taken to ensure that the message thread can be followed between the two media, by 'contact point' references e.g. when the speech track cues a label or object in the visuals. Contact points are added to the attention graph to show linking references and synchronisation between the media streams. Contact points use the same highlighting and attention directing techniques as described above. The key issues for MM evaluation concern how to combine media together to form a contact point:

- How to provide linking references and synchronisation between media streams.
- How to ensure that the message thread can be followed between the two media.

The following set of guidelines are applied to each contact point.

CP1: Use direct contact if the connection between information in an image and language is important e.g. direct the user's attention to the object in the image by highlighting the object which is being spoken about "look at the DNA, [highlight DNA]", or reveal a text caption and an arrow pointing to the object.

CP2: If speech and text are to be combined together, the same wording should be used (e.g. it is vital that the same wording is used in labels as in the speech track).

CP3: Ensure that referents are available when the contact point is made, and are in focus. Reveal text and image elements together if they share a contact point. Use highlighting techniques in visual media when the contact point is cued in speech.

CP4: If verbal media are being presented concurrently with visual media, the message on two media should be integrated. Ensure that the language and image share the same topic.

CP5: Allow time for contact point to be formed e.g. pace the presentation to allow inspection of image or speech.

6.1 Example

The contact point evaluation heuristics are applied to attention graph, searching for links between the speech track and labels shown in the visuals. Contact points are marked onto the graph as lines linking between the speech track and objects or labels, shown in figure 7.

(i) In frame I, speech cues 'pyrimidine dimer' so a contact point is marked from the speech track to the pyrimidine dimer label. The speech directly cues the label (CPl), which is likely to be in focus because of its highlight effect (CP3). A potential problem exists in frame 1 because the DNA object is also labelled,

but does not have a contact point. The attention graph shows that there is a danger the user could be focused on this label rather than the photolyase, confusing the focus (CP3). Since the DNA object is important in the presentation, it would be advisable to update the speech track to provide a contact point with it; or remove its label. In addition, the monomer object is cued in the speech track, but is unshown in the visuals; a problem may exist because the message on two media should be integrated (CP4).

(ii) Next, in frame 3 the speech cues photolyases. Again, a contact point will be formed between the speech and label of the photolyase (CPI). The photolyase object will be in focus in this case because its motion attracts attention.

(iii) In frame 4 a potential problem exists with the speech cue for protein-pyrimidine-complex. The protein-pyrimidine-complex is not labelled or highlighted in the presentation, so a contact point is unlikely to be made (CPl, CP3). Since the protein-pyrimidine-complex is a key object in the presentation, it should be highlighted and labelled when cued in the speech.

(iv) In frame 5, the speech cues the light energy. The attention graph shows that the light energy object is in focus due to its animation (CPl, CP3). This will create a contact point between the two parts of the presentation.

(v) In frame 7, the speech again cues pyrimidine dimer. There is a danger here that focus will be given to the wrong DNA object. This proposition is referring to the repaired DNA object, but the pyrimidine dimer is only shown in the original DNA object. The label should be removed when the repaired DNA object is shown, or a new label shown in the repaired DNA object e.g. 'Dimers reconvert to Monomers' (CP3). A further difficulty may exist with the monomer object which is cued in the speech track, but is unshown in the visuals (CP4).

6.2 Empirical Evidence

Some empirical evidence for the benefits of setting a contact point are shown in the eye track data for frame I. The fixations for the six subjects can be seen to be clustered around the pyrimidine dimer object and its label, rather than the DNA object and label. The density of fixations in figure 8 show that attention was almost all given to the pyrimidine dimer object and label, rather than the DNA label. This is of interest, since the DNA label and photolyase were both shown together at the start of frame 1, but only the photolyase label had a contact point.

Some evidence of the second contact point with the pyrimidine dimer caused by the speech 'converting dimer into monomers' is provided in the results to frame 7-8 (figure 9). Here the subjects return to the original DNA object and refixate the dimer and photolyase region. This indicates that a second contact point may have been formed with the pyrimidine dimer label. However, the results are not clear cut across subjects; and they may simply indicate comparison between the original DNA and the repaired DNA caused by the arrow symbol.

7 Evaluating media selection & combination

Having evaluated the attentional design of the presentation, the second phase of the evaluation deals with how the presentation content is comprehended. This concerns evaluation of the thematic content of the presentation in comparison with the original text source given at the start of the paper. Differences between the content of the presentation and source text will indicate potential problems.

The evaluator works through a set of evaluatory guidelines which suggest the information type produced for a particular media. The key issues here are:

- Which information types represent the content of the presentation How to select or combine media to best present the desired information types
- How to design within the media selected to support the information types

The following information types are proposed to categorise objects and procedures identified in the task model. The main distinctions drawn are between descriptive/operational (how to do it) and concrete/abstract information.

Physical information allows an object to be identified.

Spatial information concerns an object's location, including position, placement and orientation; adjacency to other objects e.g. relative position, above beneath; and path.

Action information concerns action which results in change in location or state by the agent, e.g.

move. This should illustrate how to perform a particular action which 'changes the position' or 'changes the state' of an object.

Role identifies how objects are related with an action. This connects the object affected by an action so a particular object 'is acted upon' as the patient, or 'acts upon another' as the agent.

Procedure is a sequence of actions with control operators, such as an action 'follows', is 'concurrent with', 'next', etc.. This should illustrate iteration for repetition of actions or selection to specify where the flow of control branches in an action sequence.

Causal is the high level, cause-effect relationship explaining how and why some part of the world behaves. This information type combines actions and objects in a model with other knowledge to answer 'why' actions happen 'because' of other actions, and the states which result trom actions.

For each information type, a set of guidelines are provided which suggest how it can be formed, and how the media should be designed to support it.

SR 1 :Physical information is given by visual media showing an identifiable object, or by language which contains a noun or adjective describing an object. If the task requires physical information then prefer a visual media. Language is poor at describing object detail and appearance [6,11].

Caveat: If an object has to be identified which may be difficult for the user to recognise, use language to indicate identity [45].

SR2: Spatial information may be given by a visual image, showing an arrangement of objects; or by language specifying abstract spatial relationships (e.g. 'next-to', 'above'). If the task requires spatial information then prefer visual media [6].

Caveat: Language can be used to identify landmarks and components for spatial information [16].

SR3: Action information may be given in the presentation by an animation, showing an object change in state, or movement along a path; by a still image which shows an object away from its rest position or in an unusual state; or by a verb which describes the action. If the task requires physical action information then prefer visual media:

(a) use animation for complex actions e.g. those with complex paths, or manipulations [42, 34, 25].

(b) use still image for simple actions [42]. For physical action information given as a sequence of still images it is important that breakpoints are shown e.g. start and end points of the action, and any changes in direction or speed are shown within the image sequence for the action to be identified [33].

Caveat: Language should be used to identify and amplify explanation of complex actions, or groups of action; or to qualify actions [7].

SR4: Role information is given implicitly within animations which show an action involving more than one object; and explicitly by nouns which surround a verb identifying the participants in the action such as which object acts on the other. If the task requires role information identifying agents or objects involved in the action then linguistic media should be used to add this information to the image or animation sequence,

SR5: Procedural information is given implicitly within animations which contain several actions in a sequence; and explicitly in cue phrases which order the arrangement of phrases such as 'next', 'then'. If the task requires procedural information then:

(a) prefer linguistic media to provide sequencing relations, particularly when the time frame is not constant. The text may also be structured to indicate procedure order by formatting into lists or frames 3,

(b) animation or image sequences will be useful to support any underlying physical action information.

SR6 : Causal information is given implicitly within animations which contain several actions with a cause-effect relation; and explicitly in cue phrases which order the arrangement of phrases such as 'because', 'in order to':

(a) if the task requires causal infonnation prefer linguistic media to provide key causal relations [31.9.34].

(b) animation or images sequences may provide causal infonnation if the sequence is well known, or not unusual. Symbols such as arrows can be added to provide limited casual information.

7.1 *Example*

The evaluator works across the attention graph prepared earlier (see figure 4) and maps the presentation units in focus at a given point in time onto one or more infonnation types. The guidelines are then used to spot potential gaps in the information provided by the presentation. This analysis gives the following set of models:

(i) In frame I action and physical information are given by speech 'reconvert pyrimidine-dimer into monomers' (SRI, SR3). This integrates physical! spatial infonnation from the pyrimidine-dimer object in the visuals (SRI, SR2) as an initial state in the action with a contact point to speech 'pyrimidine-dimers into monomers' giving physical infonnation as the outcome state (SRI). A problem here might concern the lack of physical and spatial information for the monomers object, which is unshown (SRI, SR2); the viewer must infer that it would in someway replace the dimer to repair the DNA. If spatial infonnation is important then use an image.

(ii) In frame 2 action and physical information are given by speech 'action-of photoreactizating enzyme photolyase' (SRI, SR3). This integrates action, physical and spatial infonnation from the animation of the photolayse and its motion toward the pyrimidine-dimer (SRI, SR2, SR3) with physical and spatial infonnation from pyrimidine-dimer object (SRI, SR2). There is a potential problem here because no specific role infonnation for the photolyase is given within the speech track to indicate what the action of the photolyase is upon (SR4). The role infonnation must be provided in language, or inferred from the animation of the photolyase moving toward the dimer. A better solution would be to cue the role in the speech track e.g. 'action-of photoreactizating enfyme photolyase which binds with the pyrimidine dimer' .

(iii) In frame 3-4 action and physical infonnation are given by speech 'fonn a stable protein pyrimidine complex' (SRI, SR3). This integrates physical and spatial infonnation given by photolyase object (SRI, SR2). A serious difficulty here is that the presentation shows only the initial state of the photolyase and pyrimidine dimer, but not the outcome state of the action with the protein pyrimidine complex being fonned (SR3). If an action is shown in a sequence of images, it is important for each change to be shown. The animation should show the state change in the photolyase and pyrimidine dimer, forming the protein pyrimidine complex, and cue the role in the speech track eg 'action-of photoreactizating enzyme photolyase which binds with the pyrimidine dimer to form a stable protein pyrimidine complex'.

(iv) In frame 5 procedure information is provided by speech 'then uses light energy' (SR5). This will sequence (v) - (viii) as being after (i) - (iii).

(v) In frame 5 action and physical information given from speech 'then uses light energy' (SRI, SR3). This integrates action, physical and spatial information from the animation of the light-energy moving toward the photolyase (SR3). A potential problem here is the lack of an outcome state. The light energy changes the nature of the protien-pyrimidine object; a reaction should be shown to the light energy in the animation. Beware of the danger of showing only before-after in animation if the change is complex.

(vi) In frame 6 causal information is given by the arrow symbol from the photolyase to the repaired DNA (SR6). This will give a causal relationship between (i) - (v) and (vii) - (viii).

(vii) In frame 7-8 action and physical information are provided by speech 'convert dimer into monomers' (SRI, SR3). Again, a problem here might concern the lack of physical and spatial information for the monomers object, which is un shown (SRI, SR2); this may confuse the user's understanding of how the DNA is repaired.

(viii) In frame 7-8 action and physical information are given in speech 'not-disrupt DNA double strands' (SRI, SR3). This integrates physical and spatial information from the repaired DNA and the outcome state of its strands.

Based on this analysis the information types are drawn out as a hierarchy.

In order to evaluate the content of the information types, an ideal hierarchy is also produced based upon the text source given at the start of the paper. This is shown in figure II. The ordering and information types found should be compatible with those given by the presentation if the two are to provide similar content.

This analysis reveals several variations from the thematic content of the ideal script. First, the presentation has DNA repair as it main theme in (vi), given by the causal model interrelating the damaged and repaired DNA. However, in the original source materials, photoreactivation plays a prominent role (Ideal iv), in both the title, and the causal relationship in the first sentence. The theme of photoreactivation is much less obvious in the presentation, being part of the physical information provided about photolyase in the speech track 'a photoreactivating enzyme, photolyase' (frame 2-3). This

is likely to fail because it is poorly signalled within the presentation. A caption should be added which makes clear the higher level causality e.g. 'The effect is called photoreactivation. This is achieved by.'

Secondly, the ideal hierarchy has as the initial cause that DNA contains the pyrimidine dimer distortion (Ideal i). This is not found within the presentation hierarchy which instead describes the action of/' the pyrimidine dimer being converted into a monomer (i). This may confuse the user as to the sequence of DNA repair. The ordering is an important part of the ideal hierarchy. Speech should be added to make the sequence of DNA repair clear e.g. 'Initially the DNA is distorted by a pyrimidine dimer' .

Third, the procedural cue 'then use light energy' causes the presentation structure to be slightly different from the ideal script. This places the light energy as part of a separate procedure to the convert dimer to monomer action. The ideal hierarchy has the action of light energy as the first step in photoreactivation (Ideal v). However, the ordering *of* the actions is the same as in the ideal sequence, so this may not cause a problem.

7.2 *Empirical evidence*

We performed a second study to investigate comprehension *of* the presentation. This used two groups *of* eight subjects. The first group was drawn from research students within City University, who had low domain knowledge; the second from biochemistry post-graduates who had high domain knowledge, but were pre-tested to ensure that they had no specific knowledge *of* the subject matter itself. Both groups were shown the presentation once and were instructed to write down all that they could recall *of* the presentation.

This analysis reveals several variations from the thematic content of the ideal script. First, the presentation has DNA repair as it main theme in (vi), given by the causal model interrelating the damaged and repaired DNA. However, in the original source materials, photoreactivation plays a prominent role (Ideal iv), in both the title, and the causal relationship in the first sentence. The theme of photoreactivation is much less obvious in the presentation, being part of the physical information provided about photolyase in the speech track 'a photoreactivating enzyme, photolyase' (frame 2-3). This is likely to fail because it is poorly signalled within the presentation. A caption should be added which makes clear the higher level causality e.g. 'The effect is called photoreactivation. This is achieved by.'

Secondly, the ideal hierarchy has as the initial cause that DNA contains the pyrimidine dimer distortion (Ideal i). This is not found within the presentation hierarchy which instead describes the action of the pyrimidine dimer being converted into a monomer (i). This may confuse the user as to the sequence of DNA repair. The ordering is an important part of the ideal hierarchy. Speech should be added to make the sequence of DNA repair clear e.g. 'Initially the DNA is distorted by a pyrimidine dimer' .

Third, the procedural cue 'then use light energy' causes the presentation structure to be slightly different from the ideal script. This places the light energy as part of a separate procedure to the convert dimer to monomer action. The ideal hierarchy has the action of light energy as the first step in photoreactivation (Ideal v). However, the ordering *of* the actions is the same as in the ideal sequence, so this may not cause a problem.

7.2 *Empirical evidence*

We performed a second study to investigate comprehension *of* the presentation. This used two groups *of* eight subjects. The first group was drawn from research students within City University, who had low domain knowledge; the second from biochemistry post-graduates who had high domain knowledge, but were pre-tested to ensure that they had no specific knowledge *of* the subject matter itself. Both groups were shown the presentation once and were instructed to write down all that they could recall *of* the presentation.

8 Conclusions

The paper has demonstrated an evaluation method for MM presentations. It was supported by an underlying cognitive model of comprehension which was used to drive a walkthrough and produce a set of guidelines and heuristics. It first analysed attentional processing and the formation of contact points, and then investigated the thematic content and integration of the presentation. An example presentation taken from 'The Etiology of Cancer' was used to illustrate the evaluation and several potential problems were diagnosed.

The papers main contribution is in taking a cognitive framework for MM design, applying it to design problems via a walkthrough method, and validating its usefulness by empirical study. The method steps are strongly rooted in the cognitive processes: the presentation is scanned by attention, co-references formed between verbal and visual media, and comprehended into a hierarchy of mental models. The walkthrough places guidelines around these key steps which enable a designer to conduct a systematic

evaluation of an MM presentation.

Whilst much current work on MM has focused on the automated product of presentations [1,2,17,30], little has addressed either the cognitive processes of MM comprehension, or has provided any method for evaluation. This paper may thus prove useful both to designers and to those who attempt to build tools to provide authoring support for presentation design.

Several problems remain for future work. Firstly, the predictions of the evaluation must be confirmed. Do the errors noted represent real hazards to comprehension? The existing knowledge of the user must be taken into account when diagnosing a problem in the presentation. If the user is already aware of some of the necessary content, then a difficulty found in evaluation may not transpire in comprehension. To validate what are genuine problems, the original DNA repair presentation could be re-authored taking into account the findings of the evaluation. User comprehension could then be tested between the two versions.

The utility of the walkthrough must then be addressed. How will designers of MM presentations make use of the method? It should be noted that the 'DNA repair' sequence amounted to around 20 seconds of play time. The walkthrough needs to be made more accessible. Our current work has investigated producing a computer based authoring system to support the guidelines.5. The tool will help to reduce the burden of applying the heuristics and evaluation techniques, bringing it within the normal design-cycle of an MM product.

References

1. ANDRE, E. & RIST, W. Generating illustrated documents: A plan based approach. *InfoJapan* 90, Vol 2 (1990),163 - 170

2. ARENS, Y., HOVY, E. & VOSSERS, M. On the knowledge underlying multimedia presentations. In: *Intelligent Multimedia Interfaces.* AAAI Press eds. Mark Maybury (1992)

3. ARMBUSTER, B., ANDERSON, T. & MEYER, J. Improving content area reading using instructional graphics. *Reading Research Quarterly,* 393-416, 1991

4. BADDELEY, A. *Human memory.* LEA, 1990.

5. BAGGETT, P. Role of temporal overlap of visual and auditory material in forming dual media associations. *Journal of Educational Psychology,* 76 ,3 (1984),408 - 417 6. BIEGER, G.R. & GLOCK, M.D. 'The information content of picture-text instructions'. *Journal of Experimental Education* 53 (1984), 68-76

7. BOOHER, H.R. Relative comprehensibility of pictorial information and printed word in proceduralized instructions. *Human Factors* 17,3 (1975), 266 - 277

8. BRANDT, H. The psychology of seeing. New York: The Philosophical library (1954)

9. CARPENTER, P.A. & JUST, M.A. Understanding mechanical systems through computer animation. Report No. ONR 92-1, Office of Naval Research, Arlington Virginia, USA, 1992.

10. DUFFY, S.A. Eye movements and complex comprehension processes. In: *Eye Movements and Visual Cognition.* Ed. Rayner, K. Springer-Verlag. 462-,1993

11. DWYER, F.M. The effect of varying the amount of realistic detail in visual illustrations. *Journal of Experimental Education,* 1967, 36, 34-42.

12. D'YDEW ALLE, G. & GIELEN, I. Attention allocation with overlapping sound, image and text. In: *Eye movements and visual cognition* Ed. Rayner, K. Springer Verlag. 415-, 1993

13. FARADAY, P.M. & SUTCLIFFE, A.G. A method for multimedia interface design. In: *People & Computer VIII* (OOs. Alty, J.L. & Diaper, D.) Cambridge University Press, 1993

14. FARADAY, P.M. & SUTCLIFFE, AG. Designing effective multimedia presentations. ACM CHI 97, 272-279,1997

15. Faraday, P.M. & Sutcliffe, A.G. Multimedia: design for the moment. ACM MM '97 (Seattle, USA), 183-193, 1997

16. FERGUSON, E.L. & HEGARTY, M. Properties of cognitive maps constructed from texts. *Memory & Cognition,* 22(4), (1994), 455-473.

17. FIENER, S. & MCKEOWN, KR Generating co-ordinated multimedia explanations. In: *6th IEEE Conference on Artificial Intelligence Applications* (Santa Barbara CA), 1990,290 - 296

18. FLEMING, M. L. Designing pictorial and verbal instruction: some speculative extensions from research to practice. The Psychology of lllustration : Basic Research, Vol 2, Springer New York (eds. Willow, DM & Houghton, HA), 1987, 136 - 157

19. GRIMES, T. Mild auditory-visual dissonance in TV news may exceed viewer attention. *Human Communication,* Research 18,2 (1991), 268 - 298

20. HELLER, R.S. & MARTIN, C. A media taxonomy. *IEEE Multimedia,* Winter 1995, 36-45.

21. HILLSTROM, AP. & YANTIS, S. Visual motion and attentional capture. *Perception & Psychophysics,* 55(4), (1994), 399-411.

22 HOCHBERG, J. & BROOKS, V. The perception of motion pictures. In: *Handbook of Perception and Human Performance* X, 1986.
23. JOHNSON, W.A & DARK, V.I. Selective attention. *Annual review of psychology,* 37, 43-45. 24. JOHNSON-LAIRD, P. *Mental Models.* Harvard Press, 1983. 25. KAISER, M., PROFFITT, D. WHELAN, S. & HECHT, H. Influence of animation on dynamic judgements. *Journal of Experimental Psychology. Human Perception & Performance,* 1992, 18(3),669-690,1992. 26. KENNEDY, J. M. Metaphor in pictures. *Perception,* 11 (1982),589 - 605 27. KOZMA, RB Learning with media. *Review of Educational Research.* 61(2), 179-211, 1991
28. LARGE, A. BEHESHTI, J. & BRELEUX, A. Effect of animation in enhancing descriptive and procedural texts in a multimedia learning enviroment. *Journal of American Society for Information Science.* 437-448, 47(6), 1996.
29. LEVIE, W.H. & LENTZ, R Effects of text Illustrations: a review of research. *ECTJ,* vol 30, N4, pages 195-232, 1990.
30. MAYBURY, M. T. Planning multimedia explanations using communicative acts In: *Proceedings of 9th National Conference on Artificial Intelligence,* 1991.61 - 66
31. MAYER, R.E. & ANDERSON, R.B. The instructive animation: Helping students build connections between words and pictures in multimedia. *Journal of Education Psychology,* 84(4), 444-452,1992.
32. MILLER, G. A. & JOHNSON-LAIRD, P. Language and perception. Cambridge: Cambridge University Press, 1976
33. NEWTSON, D. & ENGQUIST, G. The perceptual organisation of ongoing behaviour. *Journal of Experimental Social Psychology,* 12, 436-450, 1976.
34. PARK, O. & HOPKINS, R. Instructional conditions for using dynamic visual displays: a review. *Instructional Science* 21 (1993),427 - 449
35. PARK, I & HANNAFIN, J. Empirically-based guidelines for the design of interactive multimedia. *Education Technology Research & Development,* 41(3),63-85,1994.
36. REMEZ, E.R., RUBIN, P.E., BERNS, S.M, PARDO, J.S. & LANG, J.M. On the perceptual organisation of speech. *Psychological Review,* 101(1), (1994),129-156. 37. RIEBER, L. P. Animation in computer based instruction. *Educational Technology Research & Development* 38 (1990),77 - 86
38. RUBIN, A. A theoretical taxonomy of the differences between oral and written language. In: *Theoretical Issues in Reading Comprhension.* Eds Spiro, R, Bruce, B. & Brewer, W. LEA 1980
39. SALAME, P. & BADDELEY, A. Disruption of STM by unattended speech: Implications for the structure of Working Memory. *Journal of Verbal Learning & Verbal Behoviour,* 21, (1982), 150 - 164
40. SMITH, S.L., & MOSIER, J.N. , Design guidelines for the user interface software, technical report ESD-TR-84-190 (NTIS No. AD AI54 907), U.S. Air Force Electronic Systems Division, Hanscom Air Force Base, Massachusetts, September 1984.
41. SUTCLIFFE, A.G. & FARADAY, P.M. Designing presentation in multimedia interfaces. In: *Proceedings CHI'94 Human Factors & Computer systems* (Boston), (1994),92-98
42. SWEZEY, R. W. Effects of instructional strategy and motion presentation conditions on the acquisition and transfer of electromechanical troubleshooting skill. *Human Factors,* 33, 3 (1991), 309 - 323
43. TREISMANN, A. Features and objects. *Quarterly Journal of Experimental Psychology,* 40 (2), (1988), 201 - 237.
44. V AN DIJK, T.A. & KINTSCH, W. (1983) *Strategies of discourse comprehension.* New York: Academic Press.
45. WADDILL, P. & MCDANIEL, M. Pictorial enhancement of text memory: limitations imposed by picture type and comprehension skill. *Memory & Cognition,* (1992), 20(5), 472-482.
46. WICKENS, C. D. Processing resources in attention. In: *Varieties of Attention* (eds. Pararsuraman, R. & Davies, D.) New York: Academic Press, 1983

Appendix 4.1.2: Interactive Multimedia Checklist Documentation

Evaluating Interactive Multimedia Courseware-A Methodology

Philip Barker and Terry King

'Interactive Systems Research Group, Human-Computer Interaction Laboratory, School of Computing and Mathematics"" University of Teeside, Cleveland TSI 3BA and 'Computers in Education Research Group, Department of Information Science, University of Portsmouth, Southsea, Hants, England

From *Computers and Education.* Vol. 21, No.4, pp. 301-319, 1993

Abstract: Learning design is an important aspect of multimedia courseware production since it significantly influences the quality of the courseware that is produced. In order to help formulate a model for learning design a wide range of courseware products have been examined using a standard method of assessment. The purpose of the evaluations was to try to identify any 'hallmarks of quality' that seemed to characterize good learning products. This paper describes the steps involved in deriving the evaluation methodology and its use within the context of assessing the quality of learning design embedded within a range of multimedia courseware products that have been published on interactive compact disc.

BACKGROUND

Through its DELTA programme the European Commission is currently funding 22 major research and development projects covering a range of technical and educational issues. The project in which we were involved was intended to examine the Integration of Learning Design in Interactive Compact Disc (ILDIC). One of the main aims of the ILDIC project was to develop practical models of learning for interactive multimedia which would enable the building of well-designed and innovative prototype educational products-leading eventually to the creation of standards in the design and development of learning materials. These models of learning were to be based on both learning theory and current practice: Early ILDIC activities provided both insights in learning theory, from educational and psychological perspectives, and a theoretical methodology in learning design. Our research dealt with current practice in design by seeking to assess the quality of existing interactive multimedia learning products. This paper describes the objectives of the assessment and the evaluation methodology that was developed. Details of the results obtained by applying the evaluation methodology to actual products are presented elsewhere [1].

OBJECTIVES OF ASSESSMENT

The broad aims of the evaluation were to investigate learning design in a range of multimedia learning materials mainly published on the medium of compact disc (CD)-but not exclusively so. It was appreciated that other media such as interactive video (IV) or simply multimedia disc-based systems could offer insights into learning design which could affect future software design in emerging media such as CD extended architecture (CD-ROM XA), or CD Interactive (CD-I). Our aim would therefore be to try to find pointers to excellence in the design of multimedia materials on whatever publishing medium or hardware platform that happened to be available.

Within these aims, we identified three clear objectives for the evaluation: (I) that it would focus specifically on issues relating to learning design; (2) it would enable desirable design features to be identified and assessed; and (3) it would be sufficiently flexible as to enable innovatory features to reveal themselves. However, it was felt that if specific design features were truly innovatory then they could not be built into the evaluation methodology from the beginning.

When assessing the different CD products, the evaluation methodology would also promote consideration of the impact on learning design of the additional features that have become available on CD. These are detailed in Copeland [2]. They can be summarized in terms of the capabilities of four basic technologies: (I) CD-ROM which uses established design features from conventional disc-based systems (such as text, still and animated graphics) enhanced with good quality but non-synchronized audio, more

use of photographic quality, still photographs, and very limited use of small moving video images; (2) CD-ROM XA which uses enhanced CD-ROM features with greater use of motion video and synchronized audio; (3) CD-I which is, in many ways, similar to CD-ROM XA in terms of possible features but with greater possibilities for part-screen, moving video images and with the mode of interaction limited to a number of different types of pointing device; and (4) Commodore's Dynamic Total Vision (CDTV) system which is very similar to CD-I in terms of the features available but with the possibility of being expanded into a full computer system by the addition of a keyboard and other peripherals. Digital video interactive (DVI) was also considered to be important but learning and training products embedding this technique are currently not widely available in Europe.

DEVELOPMENT OF THE EVALUATION METHODOLOGY

When we initially started work on the development of our evaluation methodology we immediately identified two broad problem areas.

First, the difficulty of drawing up an evaluation scheme for learning design which was capable of evaluating all interactive educational products across the widest range, and the lack of credibility which such evaluations seemed to hold amongst educators. Tucker [3] seemed to speak for many [4-6] in his aptly titled paper 'Software evaluation: who cares?' when he complained that the majority of courseware evaluations were not only very flawed and largely inappropriate, but were far from being truly objective (as many claimed) and were, in fact, wholly subjective.

Second, there seemed to be no really standard approaches to the type of evaluation which we wished to undertake. For example, a recently published paper from the ECOSET project [7] which collected data from 19 organizations in 10 member states of the EC on attempts to evaluate courseware, concluded that the evaluation activities, both summative and formative, differed widely. Some groups did not use formal instruments at all; other groups used formal instruments but with varying criteria and without carrying out user trials; with a final group carrying out user trials with groups of participating educators. The main function of the results across the evaluations varied widely.

In the initial stages of our project, the purpose of the evaluation activities from our research were identified as summative - in that they would lead to a report which would comment on the quality of learning design issues in a number of completed products. However, from the summary recommendations in this report, decisions might be made which would affect the learning design in the ILDIC prototypes. This extended the role of the evaluation to include some formative aspects [8].

Initial considerations of the evaluation scheme tended to focus on trying to make results as scientifically valid as possible. Various researchers have explored different ways of developing 'objective' methods for the evaluation of computer-assisted learning (CAL) products. A common choice is to use a set of criterion-referenced categories with a simple scoring system to obtain a scored result for each product[6,9-15]. All these schemes are self-anchoring [I 6], as respondents are asked to anchor their points of view graphically or numerically. Other methods depend on having an accurate knowledge of the learning objectives of a piece of software or its typology, so that learning outcomes can be measured [17,18]. All such schemes are, however, open to criticism on a number of issues [4,5,10,1 6]-particularly, that the results of such evaluations are not applicable to the design or development processes [19,20] nor useful for decision making [21,22]. Furthermore, they are unlikely to lead to the recognition of educational innovation [23,24], they are only suitable for large scale evaluations with pre- and post-testing [25], and they are only applicable to software with explicit learning objectives or where the desired type of learning is explicit[5].

Such criticisms described the antithesis of what we were trying to achieve in our own evaluation. Objective evaluations were also concerned with reliability, in the sense that the evaluation instrument should be designed in such a way that it would produce similar ratings for a product from all evaluators [6]. As well as being open to the criticism that it is impossible for evaluators to have such a disinterested or unbiased approach [25] or to divorce evaluation from professional judgment leaving inevitably to ambiguities and unreliability [5], this approach leaves no room for a model which includes aspects of criticism and connoisseurship [26] where a synthesis of differing views can result in a holistic view and creative conclusions.

As a result of this analysis, the legitimacy of alternative evaluation methodologies became apparent, especially where these can include both summative and formative aims [21,27]. We were interested in methods which were suitable for informal, small scale evaluations[25], could focus on innovation [21], locate problems of usability [28] and, although not as rigorous as the experimental methods, could provide results applicable to development and design. By using a method combining different strategies, e.g. interviews, observations, questionnaires, results aggregated over many users can have sufficient validity.

A method of competitive analysis was chosen as being particularly apt for our task. Nielsen [29] describes competitive analysis as a two stage process which is part of usability engineering. First,

analysing existing products heuristically, e.g. asking opinions, using common sense, according to established guidelines, then performing empirical user tests with these products. The aim is to provide ideas for the design of new products and *ad hoc* guidelines for approaches that seem to work and others that should be avoided. By analysing the strengths and weaknesses of existing products, new products will have improved design. Nielsen provides data which shows that 77% of usability engineers use competitive analysis to look at existing, competing products and find that it has real impact on improving usability design. The guidelines to be used in the analysis were derived from well-known principles of user interface design.

The evaluation of learning design in interactive courseware can involve a range of different dimensions of interest. Several researchers have developed a basic set of categories by which instructional software could be evaluated [11, 12, 30-32]. For example, Morrison [33] placed maximum emphasis on the facet of interactivity which he defined as 'the learner in conversation with himself over the material to be learned'. He introduced the idea of 'adaptivity' where, in one sense, artificial intelligence might be used in the software to accommodate a learner's preferred learning style or, in another way, an adaptable program is provided which 'gives control back to the learner'. Morrison stated that well-designed products would offer users the opportunity to find individual pathways through material and encourage students to develop the higher cognitive skills of self-assessment and evaluation of their own learning style and patterns. Such facets of good learning design were developed more extensively by Plowman [34] in her consideration of learning design for interactive video. Important categories for evaluation which emerge here are considerations of navigation through material, mobility through the software, physical aspects of human-computer interaction, e.g. mouse or keyboard, different media and how they interact, issues of screen presentation, such as legibility, use of colour, access structure, screen layout and typography, quality of user interactions through consistency of appearance and functionality, visual clarity, and legibility, access to support in the form of 'help' and 'next' options, availability of a bookmark facility and user support tools like a glossary or other reference items, and the style of user interface which might embody windows or other devices to maximize the amount of easily available information.

In our approach to evaluation it was decided to seek a category-based method, with just a few basic categories of evaluation. Using such dimensions of interest (as have been explored above) a list of 12 basic categories was drawn up. These embodied the essential principles of good learning design listed in Table I. The context in which these facets were used is documented in Appendix I.

In order to maintain consistency, a descriptive passage was added to each of the categories which described (as fully as possible) which essential design features were included.

Table I. Facets of the evaluation

1. Engagement
2. Interactivity
3. Tailorabilily
4. Appropriateness of multimedia mix
5. Mode and style of interaction
6. Quality of interaction
7. Quality of end-user interaction
8. Learning styles
9. Monitoring and assessment techniques
10. Built-in intelligence
11. Adequacy of ancillary learning support tools
12. Suitability for single user/group/distributed use

When considering the structure of the evaluation tool, reference was made to the approach used in the development of the MEDA Tool[34] for the evaluation of training software. From the starting point of the actual evaluation practices of 40 European organizations, a typology of evaluation practices was constructed, and these used to construct an evaluation tool which first defined the general evaluation concerns of the user and then facilitated their focus into the specific, in the form of a list of detailed questions that the evaluator might ask. However, it was felt that the evaluation tool for ILDIC could not be so rigidly structured and should try to be 'self-focusing'. If each category was kept broad, then there would be room within the evaluation scheme for important issues to develop during the evaluation and for innovation to be made apparent, along the lines described by Stake[21] and Prosser[27].

As a step towards establishing our evaluation methodology we developed an enhanced version of the questionnaire approach previously used by Giller for the evaluation of electronic books [15,36]. In Giller's research, a check-list of desirable design features was drawn up and these were used to establish three evaluation categories, containing a number of more elementary evaluation facets. Each of the facets was expressed as a phrase (ease of use, readability of text, etc.), from which a question could be formed.

The methodology for the ILDIC evaluation replaced the facets with a list of prompt questions for each category which the evaluator administers alone, to individuals or to groups of users who have been invited to assess particular courseware products. The prompt questions, which are listed by category in Appendix 2, were designed to explore all aspects of the particular categorization under consideration. The results obtained from individual product assessments were to be collated and analysed with a view to identifying the outstanding features of particular products and/or their major limitations.

The initial draft of the evaluation questionnaire was critically examined to assess if the likely results gained from using this evaluation tool would be valid and meet the evaluation aims of the research activity. T-hree techniques were used to reduce the subjectivity of the questionnaire.

First, a form of item analysis was used to improve the objectivity of the set of prompt questions. Given that the basic set of categories was properly researched and the category descriptions would induce conformity of approach, a panel of experts then considered each category in turn and offered their own prompt question(s) to be considered by the others. These questions were then discussed and placed in order of agreed relevance to the category. In this way a small set of questions was agreed for each category which best expressed the quality of the category and would allow effective evaluation of a product without a scoring system being used as such. This approach followed that used by Bitter and Wighton [32) where experts agreed a list of criteria and ranked them in order of importance so that a valid questionnaire could be constructed for the evaluation of CAL products.

Second, wherever possible all products would undergo evaluation by a minimum of two researchers.

Third, a program of user trials would be directed at products chosen by specific criteria; for example, products regarded by the research group to be at the extremes of effectiveness or poor quality.

TESTING THE LOGISTICS OF THE METHODOLOGY

A member of the research group used the evaluation questionnaire to assess the product 'Suilleq-the Greenland Project'. This is a multimedia product based upon a CD-ROM for sound material and an IV disc for visuals that had been developed by the Danish partner involved in the ILDIC project. The evaluation took place over 2 days and numerous discussions were also held about the design, development, and user trialling of the product. As a result of this first prototype evaluation, minor modifications were made to the wording and order of some of the prompt questions in order to improve the usability of the evaluation tool.

Finally, a report on the methodology of learning design from the ILDIC Project raised four important concepts for learning design: engagement, interaction, ownership, and transferable skills [37]. While it was felt that the first two concepts would be assessed fully by the agreed questionnaire, one prompt question was modified and a further added to ensure that the final concepts of ownership and transferable skills were adequately addressed.

The final modification to the evaluation tool arose after consideration of the methodology to be used for the user trials. There was concern that the prompt questions prepared for use by experts, when more freely or less jargonistically worded to prompt non-expert users for their comments, could result in significant variation between evaluators. To avoid this, a special novice user version of the questionnaire was produced which contained the same prompt questions but expressed much more simply. All evaluators when working with non-expert users would use these questions (rather than those in Appendix 2) and, in so doing, would ensure that a standard approach was applied. The prompt questions to accompany the novice user evaluation questionnaire are presented in Appendix 3.

METHODS OF ASSESSMENT

Two methods of assessment were agreed upon: (a) self-evaluation of all products by our research group using the check-list; or (b) we would conduct user trials of as many products as possible.

Method (a) could involve two possible approaches: either, individual researchers would assess tile same or different products; or, researchers would act as an evaluation team and evaluate a product together and produce a combined assessment.

Method (b) could also involve two possibilities depending upon whether expert users or novice users were employed in an evaluation.

In the first type of user trialling, experts with some experience of computer-aided learning would be asked to complete the check-list themselves. They would be invited to undertake several realistic tasks using the

system which would enable the evaluators to view the system in its entirety [38]. The emphasis on using expert opinion in this way has shortcomings-such as subjectivity and possible inconclusiveness [39]-but it becomes a viable tool for product evaluation if supplemented by accurate information obtained systematically[4]. It can also provide a clear framework for the evaluation which includes not only the points mentioned, such as object identification, presenting questions to experts and getting opinions, and synthesizing expert opinions, but also cross-sampling of experts' views[40]. Thiagarajan [4] suggests that a range of six different types of expert should be used for the evaluation process: language, media, format, content and target experts. In our evaluations we tried to cover each of these areas as best we could.

In the second approach to user trialling, comments and verbal observations from novice users would be noted and supplemented by the simpler prompt questions[16]; the check-list would then be completed by researchers. Dick and Carey[42] favoured this one-to-one evaluation between user and evaluator, where data is collected through such means as observation, interview and questionnaire. However, they recommend the selection of at least three users: one of average ability, one who is above average and one who is below average. In our investigation we attempted to use evaluators with various levels of ability and who had different amounts of CAL experience.

When completing the check-list, evaluators would consider each category question in turn but would be free to comment as they felt appropriate. So replies to prompt questions might vary from very short to lengthy comments which may in some circumstances draw together several prompt questions in one complex response. The intention was to free evaluators to consider all aspects of learning design which seemed re relevant to them. In this way, facets which were not formally considered during the drawing up of the questionnaire would be revealed.

APPLYING THE EVALUATION METHODOLOGY

During the ILDIC investigation a total of 43 products were evaluated by members of the research team. Out of the total, the research group obtained 18 products for evaluation at either the University of Teesside or the University of Bradford. User trials were then carried out for 11 of these and a recent report of user testing was available for another. That is, of the accessible products, 67% were user trialled. Of the remaining 25 products, where visits had to be made to external sites for evaluation, reports of field trials were obtained for 4 products. In total, results from user trials were available for 16 out of the 43 products (37%). Of the 27 products for which no information on user trials was available, 16 were evaluated by at least two members of the research group, leaving only II evaluated by one member of the team (25%). This percentage, while larger than hoped, included 'Suilleq-The Greenland Project' which could only be evaluated in Denmark and 8 other products where team members had to travel at least 200 miles to carry out the evaluation.

The user trials were varied; products were assessed by CAL experts and/or users with either no experience or a limited amount of experience in computing. Full details of the expert and novice users employed in our evaluations can be found in [1].

In keeping with the aims of the study and the competitive analysis approach, we chose to report only those results which indicated particular strengths or weaknesses of a product. Varied comments by experts and novice users, although not substantiated by others, were nevertheless regarded' as valid and included in the summary of findings. Obviously, it was not the object of a competitive analysis to report the detailed results of all findings, nor did we think that this would be necessarily useful for reasons given by Dudley-Marling and Owston[11] that subjective evaluations tended to give a relatively normal distribution of software quality with few rated exemplary or unacceptable.

After each product had been evaluated, a short report was prepared which detailed findings about each product and which highlighted any interesting or exceptional points. A summary of the points of interest which emerged follows, while a copy of the detailed findings can be obtained from the authors[1].

SUMMARY OF FINDINGS

A wide range of different types of learning and training product were evaluated during the ILDIC evaluation activity using the methodology described in this paper. The functionality of these products varied considerably - from simple CAL resources published on CD through various types of electronic encyclopedia to complex simulations based upon the use of high fidelity sound and motion pictures.

As a result of evaluating the 43 products involved in this assessment exercise we were able to identify a number of important factors which we believed acted as pointers to quality as far as the current ILDIC research project was concerned. Obviously, the evaluations revealed any common or generic issues such as interface design, monitoring and feedback, adaptability, structure and content, view, which applied across the board to all products, On the other hand, quite a number of issues turned out to be very highly

product-specific and related to individual classes of product, such as language learning materials or electronic encyclopedia. Unfortunately, we cannot give a detailed analysis and discussion of our overall findings here; the interested reader is therefore referred to Ref. [1]. However, it is worth pointing out, in summary form, one or two of our major findings.

Generally, the results obtained from the investigation seemed to suggest that the quality of end-user interface design was of paramount importance in producing a quality product.

Indeed, it seemed that in the products that were rated highly by the evaluators, the designers had followed well-established rules and guidelines for developing their end-user interfaces. End-user interface design is important because it affects how users perceive a product, what they can do with it and how completely it engages them.

As it happened, engagement (facet 1 in Table 1) seemed to be one of the most popular hallmarks of a quality product-over 50% of the exceptionally good products were rated highly in this category. Interestingly, in the context of engagement, most evaluators thought that the use of audio and motion video made a product noticeably more engaging.

Interactivity was another important hallmark of quality. Thus; products that actively involved users in participatory tasks seemed to rate much more highly than those that exhibited low interactivity and-which just presented information to users.

Another important attribute which good products seemed to possess was tailorability-the ability to allow end-users to configure them and change the_ to meet particular individual needs.

Overall, our results seem to suggest that of the items listed in Table 1, there were about five or six (numbered I, 2, 3, 6, 7 and 8) that seemed to playa key role in defining the characteristic hallmarks of quality courseware products. Other facets seemed to be only of marginal importance and, in some cases, some factors seemed not important at all.

CONCLUSIONS

Certain findings of the evaluation exercise suggest that it was successful in what it attempted to do in line with the expected outcomes from a critical evaluation and the specified objectives. Innovations and design features were revealed which were not expected during the drawing up of the questionnaire. Other topics became more clearly focused as the assessment proceeded. The controlled subjectivity of the assessment methodology provided necessary validity to the exercise, however, in the nature of such naturalistic research, the range of design issues thought to be contentious was large and only subjective guidance was available as to their relative importance to design considerations. Also, unfortunately very little support was offered from software developers in the form of documentation on design or user trialling. Only two producers were prepared to discuss design issues. This left the evaluators with only the limited range of user trialling carried out in this programme. This necessarily reduced the validity of the results.

REFERENCES

1. Barker P. G. and King T. R., Report on evaluation and user trials, Deliverable 3b, Workpackage 1.4, DELTA Project 2012-ILDIC (August 1992).
2. Copeland P., The multimedia mix. *Educ/ Training Technol. Int.* 28, 154-163 (1991).
3. Tucker R., Software evaluation: who cares? In *Aspects of Educational Training and Technology-Volume XXII: Promoting Learning* (Edited by *Bell* C *et al.),* pp. 8-16. Kogan Page, London (1989).
4. Bates T., Towards a better research framework for evaluating the effectiveness of educational media. *Br. J. Educ/ Technol.* 12, 215-233 (1981).
5. Borich G. and Jemelka R., *Evaluation in Computer-based Instruction 0 State of the Art Assessment* (Edited by O'Neil H. F.), pp. 161-209. Academic Press, New York (1981).
6. Miceeri T.. Pritchard W. H. and Barrett A. J., Must computer courseware evaluation be totally subjective? *Br. J. Educ/Techno!.* 20, 12(H28 (1989).
7. Bonekamp L. W. F., Summative and formative evaluation activities regarding courseware in Europe: results of the ECOSET ProjecI, Educational Computing Consortium BV, P.O. Box 217, 7500 AE Enschede, Holland (1991).
8. Stufflebeam D. M., Evaluation as a basis for decision making. Paper presented to the American Educational Research Association, Chicago, Ill. (1967).
9. Legree P. J. and Gillis P. D., Product effectiveness evaluation criteria for intelligent tutoring systems. *J. Computer-based Instruct.* 18, 57-{j2 (1991).
10. Rengger R. E., A preliminary design for a methodology for experimentally measuring usability. National Physical Laboratory, Teddington. UK, NPL Report, DITC 16/3/90 (1990).
11. Dudley-Marling C. and Owston R. D., The state of educational software: a criterion-based evaluation. *Educ/ Technol.* 27, 25-29 (1987).

12. Rothwell J., Buying and using off-the-shelf courseware. *Interactive Learning Int.* 4, 5-8 (1987).
13. Akahori K., Evaluation of educational computer software in Japan: methods and results. *Programmed Learning Educl Technol.* 25, 46-56 (1988).
14. Azizan R. and Meerah T. S. M., Evaluating educational software: a proposed model. *J. Sri. Math. Educ. Southeast Asia* 10, 52-{j7 (1987).
15. Giller S., Design guidelines for electronic book production. M.Phil. thesis, Univ. of Teesside, Cleveland, U.K. (1992).
16. Knussen C, Tanner G. R. and Kibby M. R., An approach to the evaluation of hypermedia. *Comput'" Educ.* 17, 13-24, (1991) 17. Sankar Y., Evalualion of computer-assisted instruction. *Programmed Learning Educl Technol.* 25, 314-321 (1988).
18. Kulik J. A., Kulick C. C and Cohen P. A., Effecliveness of computer-based college teaching: a meta-analysis of findings.*Rev. Educ/ Res.* SO, 525-544 (1980).
19. Bosco J., An analysis of evaluations of interactive video. *Educl Technol.* 26, 7-17 (1986).
20. Bayard-White C, Program assessment and evaluation checklist. *Videodisc Monitor* 7, 24-25 (1989).
21. Stake R. E., Program evaluation, particularly responsive education. Occasional Paper No.5, Evaluation Center, Western Michigan Univ. (1975).
22. Reeves T. C, Research and evaluation models for the study of interactive video. *J. Computer-based Instruction* 13, 102-106 (1986).
23. Apple M. W., The process and ideology of valuing in educational settings. In *Educational Evaluation: Analysis and Responsibility* (Edited by Apple M. *et al.),* pp. 3-34. McCulchan, Berkeley, Calif. (1974).
24. Scriven M., The concept of evaluation. In *Educational Evaluation: Analysis and Responsibility* (Edited by Apple M. *et al.),* pp. 55-66. McCulchan, Berkeley, Calif. (1974).
25. Harris N. D. C. and Bailey J. G., Conceptual problems associated with the evaluation of educational technology courses. *Br. J. Educ/ Technol.* 13, 4-14 (1982).
26. Eisner E., *The Educational Imagination.* Macmillan, New York (1979).
27. Prosser M. T., Towards more effeclive evaluation studies of educational media. *Br. J. Educ/ Technol.* 15,33-42 (1984).
28. Nielsen J. and Molich R.. Heuristic evaluation of user interfaces. In *Empowering People: Human Factors in Computing Systems, CHI'90 ConJerence Proceedings* (Edited by Chew J. C. and Whiteside J.), Seattle, Washington, pp. 249-256 (1990).
29. Nielsen J., The usability engineering life cycle. *IEEE Computer* 25, 12-22 (1992).
30. Whiting J., The nature of computer-assisted learning. In *Aspects of Educational Technology-Volume XVIII: New Directions* in *Education and Training Technology* (Edited by Alloway B. S. and Mills G. M.), pp. 33-37. Kogan Page, London (1985).
31. Dejoy J. K. and Mills H. H., Criteria for evaluating interactive instructional materials for adult self-directed learners. *Educ/ Technol.* 29, 39-41 (1989).
32. Bitter G. G. and Wighton G., The most important criteria used by the educational software evaluation consortium. *Computing Teacher* 14, 7-9 (1987).
33. Mnrrison D. J., Interactive learning systems and the learner. In *Aspects of Educational Technology-Volume XX: Flexible Learning Systems* (Edited by Percival F. *et 01.),* pp. 134-138. Kogan Page, London (1987).
34. Plowman L., Learning from learning theories: an overview for designers of interactive video. *Interactive Learning lot.* 5, 165-174 (1989).
35. Machell J. and Saunders M., MEDA: an evaluation tool for training software. Centre for the Study of Education & Training, Univ. of Lancaster (1991).
36. Barker P. G. and Giller 5., Design guidelines for electronic book production. Final Report submitted to Learning Technologies Unit, Training, Enterprise and Education Directorate, Department of Employment, Sheffield, pp. 28-33 (1992).
37. Good M., Shaw S. and Tabberer 5., EC DELTA Project D20121LDlC, Deliverable I, report on learning design, Part 3: a methodology for learning design. Cambridge Training and Development Ltd, Cambridge CBI 4FB, pp. 11-12 (1992).
38. Johnson G.I., CleggC. W. and Ravden S. J., Towards a practical method of user interface evaluation. *Appl. Ergonomics* 20,255-260 (1989).
39. Chinien C. and Hlynka D., Formative evaluation of prototypical products: from expert to connoisseur. *Educ/ Training Technol. Int.* 30, 60-66 (1993).
40. Nevo D., Expert's opinion: a powerful evaluation tool. Paper presented at the *Annual Meeting oj the American Educational Research Ass.,* Chicago, 111. (1985).
41. Thiagarajan 5., Instructional product verification and revision: 20 questions and 200 speculations. *Educ/ Commun. Technol. J.* 26, 133-141 (1978).
42. Dick W. and Carey L.' *The Systemalic Design oj Instruction,* 3rd edn. Scott, Foresman and Co., Glenview, 111. (1990).

APPENDIX I

EC DELTA Projeci ILDIC Integrating Learning Design in Interactive Compact Disc The Evaluation Check-list

The check-list presented below can be used in order to gauge the quality of a computer-based interactive learning facility with respect to its basic aesthetics, the nature of the learning environment which it provides and the types of pedagogy involved. Individual products should be assessed with respect to each of the broad evaluation categories presented in the check-list.

Basic Evaluation Check-list

(1) Engagement
(2) Interactivity
(3) Tailorability
(4) Appropriateness of multimedia mix
(5) Mode and style of interaction
(6) Quality of interaction
(7) Quality of end-user interfaces
(8) Learning styles
(9) Monitoring and assessment techniques
(10) Built-in intelligence
(11) Adequacy of ancillary learning support tools
(12) Suitability for single user/group/distributed use
(13) Availability in terms of cost and delivery platforms
(14) Outstanding strengths and attractive features
(15) Outstanding limitations and weaknesses

Notes are provided in the questionnaire which outline what is involved in applying each of the evaluation categories listed above. These notes should be used as an aide memoire; they should be read prior to conducting either a single or a batch of assessment experiments. Included with the notes is a list of key questions which are best felt to encompass the learning design features of each evaluation category. Such questions should be applied to each software product and used to form the overall conclusion about that product.

It is important to realize that no attempt is being made to assess the learning effectiveness of products since this would require running more extensive controlled experiments involving both pre-tests and post-tests.

Notes and Prompt Questions for Basic Evaluation Check-list

When using the broad evaluation categories to assess software products, the following notes and questions can be used as guide-lines.

(1) *Engagement*

This category seeks to assess whether the product engages the user's interest, or involves the user because of factors which are particularly motivating, enjoyable or challenging.

1.1 Is the content appropriate for the user group for which it was written?
1.2 Does the content offer sufficient diversity?
1.3 Is the content and experience authentic for the learner?
1.4 Are the goals, objectives and stages always clear to the user?
1.5 Are the tasks relevant to the needs of the user?
1.6 Can users identify with the goals and objectives and build in their own personal plan of achievement?
1.7Can the objectives of learning activities be met?
1.8 Is the presentation effective in maintaining interest?
1.9 Is the presentation enjoyable?
1.10 Does the product hold the user's attention?
1.11 Can the user's concentration be maintained consistently?
1.12 Is the product challenging? .

(2) *Interaclivily*

In this category the product is assessed on whether it offers both passive and active interactions with the user and whether it provides the means by which a high degree of user involvement can be established.

2.1 What level(s) of interaction does the product support?
-passive (e.g. looking, watching, listening)
-trivial (that is, no cognitive load)
-active (e.g. choosing, entering, learners making their own choices)
-reactive (e.g. user takes on novice role in relation to a .tutor-like' piece of software)
-excessively active (i.e. very high cognitive load)
2.2 Are the level(s) of interaction well balanced?
2.3 Is the complexity of user response minimized? .
2.4 Is feedback offered hi a consistent manner?
2.5 Are learners able to navigate through the system making real and personally relevant decisions about order, topic, level, pace of material presented?

(3) *Tailorability* .
This facet looks at features which either enables the product to be personalized or to be adapted to different hardware environments (e.g. networks, varied CD-ROM drives) 'I'd therefore be made more accessible to a group of users.
3.1 Does the user have control over product parameters such as colour settings, design layouts or sound volume settings?
3.2 Can the new product configuration be saved for the user?
3.3 Can the content be altered or adapted to the needs of the user (e.g. by means of editing, provision of a notebook facility)?
3.4 Can the user select the level of information presentation (e.g. novice, experienced user)?
3.5 For guided learning, is the level of fixed pathing appropriate?
3.6 Can the user establish new paths through the product?
3.7 Can the user set up 'short cut' facilities such as macros?
3.8 Can the product be adapted to different hardware environments?
3.9 Can the product make use of information on a user's background provided by the user?

(4) *Appropriateness of the multimedia mix*

Questions in this category seek to assess whether various multimedia features work well in relation to the educational aims of the product and in relation to one another.
4.1 Do the multimedia techniques complement one another or compete for the user's attention?
4.2 Is the scale of the graphics or video display appropriate for the information content?
4.3 Is the level of audio quality appropriate to the information content transferred?
4.4 It the multimedia mix on screens used responsibly or does over-crowding of multimedia objects reduce the necessary effectiveness of some media?

4.5 Does the multimedia mix support the learning style involved?
4.6 Do the multimedia techniques enhance or inhibit engagement?

(5) *Made and style of interaction*
This category seeks to comment on the nature of the interface in terms of the mode of interaction (e.g. keyboard, mouse, trackerball, etc.) and the style of interaction, such as, the way choices are made or options selected (e.g. using menus, icons, etc.). ..
5.1 Is the mode of interaction appropriate to the end-user's tasks?
5.2 Is the style of interaction appropriate to the needs of the user?
5.3 Is the multimodality of interaction offered to users, parallel, discrete or both?
5.4 Are facilities such as accelerator keys, hot-keys or short-cut key combinations made available to users?
5.5 Does the mode of interaction encourage engagement?
5,6 Does the style of interaction encourage engagement?
5.7 Is the hardware base chosen appropriate to the style of the interaction?

(6) *Quality of interaction*
Regardless of the quality of the interactivity displayed by the product, this facet seeks to assess the quality of the end-user's interaction with the system. Here we consider the nature of the control that is given to the user, and the ease of use of the product in terms of the help and support systems and their general level of accessibility. It also includes possible comment on the capacity of the product to offer the user the facility to make real decisions about his/her route through the product, thereby enhancing a sense of ownership in relation to the learning process.
6.1 Is the user interaction always appropriate to the circumstances as the software is running?
6.2 Does the product always display a high level of consistency in user interactions?

6.3 Are user interactions always clearly defined?
6.4 Are user interactions always effective?
6.5 Is the user always presented with good visual and spatial access to information?
6.6 Does the product offer support or help to the user as needed?
6.7 Does the product develop a sense of ownership in a user, in relation to his/her own learning process, through the development of their own route to a chosen goal?
6.8 Are user interactions secure?
6.9 Is the courseware accessible to users generally with reference to cultural background, gender, disabilities, etc.?

(7) *Quality of end-user interface.,*

This facet addresses the issue of the type of interface that is presented to the user and with which the user will have to interact. Features might include the use of colour and graphics, windowing, the design of icons and the positioning of information.

7.1 Does the interface offer consistent features to the user (e.g. screen style, icons, menu bars, colours, language positioning of information)?
7.2 Does the interface have generic features?
7.3 Does the interface display good visual and spatial attributes?
7.4 Is the interface integrative with regard to multimedia techniques?
7.5 Does the interface allow links to be made between related areas of information?
7.6 Is the way in which options are presented to the user consistent and of high quality (e.g. icons, menus?)
7.7 Does the end-user interface enhance the provision of support to the user?
7.8 Is the interface intuitive?

(8) *Learning styles*

In this section, comments may be made on the learning style chosen for the subject content under consideration, and the views held by the learner and designer about the conceptual learning model and the design aims of the product. In this category it is also possible to comment on whether the learning material recognizes the existing skills of users and encourages the transfer of these skills into the new learning situation.

8.1 Does the product offer a multiplicity of learning styles? Which ones?
-focused
-browsing
-guided
-structured
-unstructured
-shallow
-deep
-other ...
8.2 Is the style of learning under user control?
8.3 Does the learning material tend to facilitate the transference of existing skills?
8.4 If the learning styles are restricted, is this appropriate for the course content?
8.5 What conceptual model does the user have of the product after using it?
8.6 What conceptual model did the software developers have in mind when developing the system?

(9) *Monitoring and assessment techniques*

Does the product monitor the progress of its users, offer formal assessment, or provide the means for self-assessment? Does this product have any low-level 'intelligent' features like making use of monitored information on a user in order to provide support or enhance engagement?

9.1 Does the product monitor performance?
9,2 Can the product assess performance?
9.3 Are students advised of the effectiveness of their performance?
9.4 Can the monitoring or assessment processes be made transparent for the benefit of the tutor?
9,5 Is there a facility for long term monitoring of student progress?
9,6 Does the product facilitate review and study of user's own learning?
9,7 Does the product empower the user to create new assignments?
9,8 Does the product facilitate the drawing up of self-assessment protocols?
9,9 Can the product make use of monitored information a user to provide support or enhance engagement?
9.10 Does the product support the creation of personalized maps or navigational aids by the use of

monitoring?

(10) Built-in intelligence

This section may be used to comment on any 'intelligence' in the product; for example, whether it includes features of an intelligent tutoring system (such as learner modelling), uses artificial intelligence techniques with knowledge base(s), uses an expert system, or uses monitored information to provide user support such as advisement strategies or prompting.

10.1 Does the product give personalized help to the user when a route is to be chosen through the software (advisement strategies)?
10.2 Does the product prompt the user according to personal circumstances when 'option choices are to be made (prompting)?
10.3 Does the product offer the facilities of an expert system?
10.4 Does the product offer any features of an intelligent tutoring system such as learner modelling?

(11) Adequacy of ancillary learning support tools

Comments here assess the suitable provision and quality of on-line support tools (such as a dictionary, calculator, atlas or specialist ancillary software) or other user support services, such as a notebook facility. Comment can be made here on how essential or trivial such tools are felt to be.
11.1 Does the product provide links to needed tools (e.g. dictionary, atlas, calculator)?
11.2 Does the product provide links to specialist software (e.g. spelling and style checkers, design tools)?
11.3 Are the tools provided essential for the learning tasks undertaken by the user?
11.4 Is there free access to such tools by the user?
11.5 Does the interface support the linking in of such tools in an easy fashion?

(12) *Suitability for single user/group/distributed use*

Comment can be made here on the suitability of the product for use in a range of educational settings.
12.1 Is the product adaptable to a number of educational settings (e.g. single-user or group use)?
12.2 If the product is applicable for use by a group as groupware or in a distributed network,
-Does it facilitate the construction of group goals for learning?'
-Does it facilitate continued cooperation between members of the group?
-Does it enable. the group to assess what needs to be done?
-Does it enable the self-assessment of the learning process by members of the group?
12.3 For a distributed network, where necessary, is anonymity assured?

(13) Availability in terms of cost and delivery platforms

This section attempts to comment on the publishing medium both in terms of cost, and the specific hardware platform (e.g. CD-ROM or CD-ROM XA) and software delivery platform(s) (e.g. Apple Macintosh, IBM PC, UNIX, etc.).
13.1 For the multimedia features chosen, is the publishing medium selected the most widely accessible for users in terms of cost and hardware availability?
13.2 Does the end-user interface chosen, maximize the possible number of software delivery platforms?

(14) *Outstanding strengths and attractive features*

Comments in this section may be drawn from sections above but include features which are noted as being exceptionally good, or may make general comments on the software in its entirety.
14.1 Does this product have any outstanding features which would tend to enhance the learning process?
14.2 Does this product have any unusual or innovative features which would tend to enhance the learning process?

(15) Outstanding limitations and weaknesses

Comments in this section may be drawn from sections above but include features which are noted as being exceptionally
poor, or may make general comments on the software in its entirety.
15.1 Does this product have any outstandingly bad features which would tend to inhibit the learning process?
15.2 Does this product have any unusual or regressive features which would tend to inhibit the learning process?

APPENDIX 3

Prompt Questions for Non-expert Users

When using the basic evaluation check-list with non-expert user groups (during user trials of software), the following questions can be used as prompts. These questions are directly related to the key expert questions (Appendix 2) but expressed less jargonistically for the benefit of non-specialists.

(1) Engagement

Were the tasks you were asked to do about the right level for you (or the group aimed at)?
Was there enough variety?
Did the product seem relevant to real life?
Was it always clear where you were in this product?
Was it always clear where you were going?
Was it enjoyable?
Was it interesting?
Was it challenging?
Did you feel that you had achieved (or could achieve) something with this product?
Would you want to use the product again?

(2) *Interactivity*
Did !he product give you enough opportunity to choose and enter things?
When you weren't making choices, was it still interesting?
Did you ever feel that you were making choices just for the sake of it?
Did you feel that you could decide what you wanted to do?
Could you choose which route you wanted to take that was most relevant to you? When you were using the product, did you feel that you were basically in control? Did you get good feedback when needed (for example, when something went wrong)? Was it easy to make choices with the equipment provided?

(3) *Tailorability*
Did you feel that you could alter things sometimes to suit you personally? For example,
-turn down the volume?
-change the colours?
-set up new ways of doing things?
-save your own work?
-work at the right level?
-get just the right level of help (say, for a new user)?

(4) *Appropriateness of the multimedia mix*
Did you feel that the various features (e.g. text, sound, graphics, video, etc.) worked well together, or did one (or two) crowd the others out?
If there was video or graphics, did the size seem OK?
If there was audio, was the sound quality good enough?
In this case, do you think that having extra features (like sound, video, etc.) would help you to learn?
Do you think that having extra features (like sound, video, etc.) made the product more enjoyable and interesting?

(5) *Mode and style of interaction*
Did the use of the mouse, keyboard or trackerball (etc.) seem natural when using the product or did it restrict you in any way?
Did the use of the mouse, keyboard or trackerball (etc.) make using the product more enjoyable or interesting? Did you like the way that you made choices (e.g, from the menu, etc.)?
Did the way choices were made alter how you felt about the software (that is, did they make it more or less interesting or enjoyable)?

(6) *Quality of interaction*
Were you always certain how to proceed?
How do you feel about the screen displays?
Was everything always clearly laid out?
Did you find enough help built into the product?
Could you always find out what you wanted from the help system?

Did you get the sense of where this product might fit into your own personal educational scheme?
Did !he software ever bother you or confuse you or just mystify you, perhaps because of the language or examples used? Was it racist or sexist?

(7) *Quality of end-user interfaces*
How did you feel about the way in which information was presented to you? Did you always know how to make important choices, like,
-moving from one part to another?
-calling up the help system?
Did you feel that there was a regular way of moving about the package?
If icons were used,
How did you feel about using the icons?
Was it fairly obvious how to use each one?

(8) *Learning styles*
What do you think this product was trying to be?
-a hook?
-a tutor?
-a reference book?
-a tool?
Etc.
Were you able to bring to this product other skills you had already acquired in the outside world or from other computer
packages?

(9) *Monitoring and assessment techniques*
Did this product help you assess how you were getting on?
Do you think this product would help you to set goals for yourself?

(10) Built-in intelligence
Did you get any extra help from the product because it seemed to remember what you did in the past?

(11) *Adequacy of ancillary learning support tools*
If there were learning support tools in the product,
Did you notice what they were tools (a dictionary, notepad, bookmark, calculator, etc.)?
Did you use any of them?
Were they easy to get to?
Were they easy to use?
Were they helpful?

(12) *Suitability for single user/group/distributed use*
If used with a group of people,
Did this software help you work together as a group?
Did it help you co-operate with one another?
Did it help you to assess how well you were doing?

(13) Availability in terms of cost and delivery platforms
No questions for this section.

(14) *Outstanding strengths and attractive features*
Is there one thing that stands out as being really good about this product?

(15) *Outstanding limitations and weaknesses*
Is there one thing that stands out as being really bad about this product?

Appendix 4.1.3: Multimedia Taxonomy Documentation

Using a Theoretical Multimedia Taxonomy Framework

Rachelle S. Heller and C. Dianne Martin
The George Washington University
Washington DC, USA
Sheller@seas.gwu.edu; Diannem@seas.gwu.edu

In *Proceedings of ED-MEDIA 1999*, June 1999, USA, AACE Press, pp 415-420.

Abstract: An expanded multimedia taxonomy is proposed to provide a framework for understanding and conceptualizing interactive multimedia for the purposes of teaching, research and project design. The taxonomy is an especially helpful tool for the generation of evaluation materials. Specific examples of how to generate evaluation categories and concerns is included.

WHAT IS MULTIMEDIA?

Multimedia is a polysemous term - a term with many definitions and in this case many roots. In this paper, multimedia is defined as the seamless integration of two or more medium. A multimedia taxonomy is proposed to help organize the discipline based on a previous media taxonomy (Heller & Martin, 1995). The taxonomy helps to classify the space called multimedia and helps to draw attention to difficult issues. While multimedia is an emerging field, the fact that we can provide a taxonomy indicates that it has reached some level of stability and maturity to be reckoned with.

THE TAXONOMY

The multimedia taxonomy can be visualized as a three dimensional matrix (Heller and Martin, 1998). Media type, the rows of the matrix, pertains to the various media involved: text, sound, images, motion and combinations. Media expression, the columns, refers to the level of abstraction portrayed using these media: elaboration, representative and abstraction. Context, the third dimension, portrays the various roots in multimedia: disciplines, interactivity, audience, aesthetics, quality and usefulness. As originally expressed the taxonomy had 90 cells in three dimensions to be defined (Heller and Martin, 1995). In the beginning work, it was observed that evaluation or design questions often repeated themselves, regardless of the medium form or expression format. For example, questions about text size are appropriate no matter the format of textual presentation. Based on this observation, the taxonomy was expanded to provide a category in the format called general - this category is used for questions that are constant over the format space. Hence there are 120 cells provided by the taxonomy and each can be expressed as a series of questions that can be used as a development or evaluation guide. The newly expanded taxonomy becomes a floor, not a ceiling, for a series of guidelines that can be used to generate a series of questions about an application.

Media type is arranged as a series of individual medium of increasing complexities (i.e. storage). The various media should be clear. Text is the presentation of information using an alphabetical symbol system. This includes prose is various languages as well as presentations in such forms as mathematics or other symbol systems. Sound is the inclusion of spoken word as well as generated tones forming music or other audible information. Images include photos and hand drawn items while motion can be motion pictures or animations. Multimedia, as defined above, is the inclusion of any two or more of these.

Media Form	General	Media Expression Elaboration	Representation	Abstraction
Text	Math functions were correct Is it clear which input device to use Word choice was unclear/clear Grammar was correct Reading characters on screen (very hard to very easy) Spelling was correct (No to Yes)	Font for bold and italics is readable		
Sound	Sound quality (unclear to clear) Sound speed (too fast/slow.. just right) Sound volume (too loud/soft.. just right)	Speech is Captured without stutters	Sound effects are recognizable	
Graphics	Quality of image (low to high) Size of image (too large/small to just right)			Icons are clear Iconic Metaphor is sharp
Motion	Quality of motion (jerky to smooth)			
Multi media	Media Integration (Uncoordinated to coordinated) Media technical synchronization			

FIGURE 1: Sample Questions from the Quality Dimension

Contextual categories that are not applicable to the specific product being evaluated can be eliminated from the protocol. The questions (Figure 1) are framed either as a Likert scale question with five gradations or an open ended questions requiring a free form user response. Those that are intended to be considered using a Likert scale are presented with the categorical extremes listed in parenthesis. For example, to evaluate the item listed as use of motion one can generate a question for the form:

	Jerky			Smooth
Use of motion	1.....2.....	3.....	4.....	5

The Multimedia Taxonomy is best understood in its application. The new media space is 6x4x5, where each cell can be used to address specific aspects of media design or evaluation. For example, the aesthetics category might encompass the following concerns: In the general expression of text, we might ask, is the font size appealing, while in the elaboration expression of text, it is appropriate to ask the aesthetic question concerning the richness of the literary expression on the continuum from dry to poetic. In representational expression of text, we might ask whether the format of the outline is logically well-structured and in considering the abstract expression of text within the aesthetic categories, we could ask if the textual image as an icon is pleasing.

Similarly, when considering questions of audience in the general category for text we might ask if the instructions are clear, while in elaborative or representational expressions of text we might ask evaluative questions about whether the feedback to the audience is appropriate or if there is a compatibility between the reading level of the audience and the level of expression of the textual description. The usefulness category prompts us to assess the value added by a medium type or expression. The extended Multimedia Taxonomy elicits questions about sound quality as well as sound impact.

Under the category of Usefulness, the evaluation protocols contained questions about whether a medium like sound was functional, by asking users to rate the sound on a Likert scale that ranged from annoying to helpful. The protocols also contained open-ended questions about what the user remembered about the sound in the application being evaluated. Thus, the newly expanded Multimedia Taxonomy becomes a floor, not a ceiling, for a series of guidelines that can be used to generate evaluation questions about a multimedia application. Parts of the taxonomy not relevant to the specific product being evaluated can be ignored when developing the evaluation protocol and data-gathering instruments.

AESTHETIC

The guidelines in this section are intended to examine the appearance, artistic look or impression of the IMM. Questions related to design, rather than function, are the focus in this section Figure 2 is an example of questions for aesthetics.

AUDIENCE

Audience guidelines are intended to direct the evaluator to address issues of how the IMM relates to the audience and how the audience members might be able to process the medium form in a specific medium format. For example in the area of text as elaboration, it is necessary to ask whether the text is at the reading level for the intended audience. On the other hand, while text as representation does require the user to read, there tend to be fewer words and grammatical constructions in this category but the representation (e.g. outline) might not be familiar to the particular audience. This section requires the evaluator to know to whom the IMM is intended. Before preparing an evaluation protocol for a particular IMM, the evaluator should know who the intended audience is.

DISCIPLINE

Guidelines in this section are intended to identify the content specific material in each of the media forms within a specific medium format. For example in reviewing text that is presented in representational form (outlines or lists) is necessary to ask here whether the lists are complete and accurate within the content area. This section is often best evaluated by experts in the content area rather than by IMM designers or members of the intended audience.

Media Type	General	Elaboration	Representation	Abstraction
Text	Font size is appealing Free from stereotypes Media overuse? Overall impression Screen design Screen spacing Use of color	Literary Expression (Dry to Poetic)	Format (i.e. outline) is well placed Position of messages (appropriate)	Metaphors are consistent within the application Text image or icon
Sound	Sound (annoying to helpful) Sound (noisy to tuneful) Volume (too soft/loud to just right) What did you think of the sound?		Background sound (pleasing)	Sound effects (not pleasing to pleasing) Sound effects (unclear to clear)
Graphics	Still images attractive? Use of color pleasing?		Background images (too faint)	Icons pleasing?
Motion		Video or animation (not pleasing to pleasing)		

Figure 2: Sample questions based on aesthetics issues

INTERACTIVITY

In this category, evaluators should be addressing aspects of control, navigation and linking. Aleem (1997) is examining the relationship between the attribute of interactivity to the media type and expression. He has further subdivided the attribute of interactivity into four categories: Passive, Reactive, Proactive and Directive. With the Passive aspect of interactivity, the user has no control, but instead, all control is

embodied in the application (i.e., a PowerPoint presentation) Reactive interactivity provides limited response for the user within a scripted sequence. Proactive interactivity allows the user to play a major role in the design and construction of situations, typically by manipulating values for valuables. Multimedia that has interactivity at the Directive level allows the user both to respond to and to initiate actions within the application as well as to tailor aspects of the environment, such as selection of color choice, feedback choice, and so on.

QUALITY

This section refers to the technical, reproductive aspects of the IMM. Quality issues relate to clear images, correct synchronization and timely delivery.

USEFULNESS

This category refers to the value of the material presented in the IMM as well as the ease of use of this material. Since Interactivity covers the navigational issues and Usefulness covers the ease with which a user can operate the application there seems to be a muddy ground between the two categories. But there is a distinction. Usefulness also includes whether the user can operate the equipment necessary to make the application run. For example, questions such as can the user operate the head tracker or the roller ball as well as is there a need for external devices to run the IMM are appropriate in the ease of use category.

As can be judged from the foregoing paragraphs, not all of the 120 cells contained within the three-dimensional Multimedia Taxonomy are completed. Some are being examined in detail, while others are still to be developed. Students in courses in the Seminar on Multimedia Evaluation in the Department of Electrical Engineering and Computer Science at The George Washington University (Heller, 1994) have been successful in designing evaluation protocols using the Multimedia Taxonomy as a guideline. Figure 3 represents a sample of the survey comments developed from the formative evaluation of Ada Mentor, an online web site for the study of the Ada programming language. The Multimedia Taxonomy helped pinpoint such problems as too many forms within the web site and the inability for a user to see the entire page. They were able to suggest on the bases of the evaluation, that the Ada Mentor developers reduce the navigational capabilities and clarify the functionality of navigational icons.

CONCLUSIONS

We believe the multimedia taxonomy is a reasonable organizing framework. Admittedly there is room for extension of the multimedia taxonomy if new ways of thinking about multimedia are presented. The extensibility of the taxonomy is demonstrated by the ability to take a cell or series of cells and, using them, form new detailed sub-taxonomies. This is demonstrated by the work presented by Aleem (1997).

The multimedia taxonomy is a solid framework - the details of each of the cells needs to be completed. Currently we are beginning usability studies using the taxonomy. Specifically, we are interested in whether evaluators can use the taxonomy as a guideline to improve the coverage of their existing evaluation protocols or does the fragmentation of the taxonomy interfere with, or distract from, the evaluation. In continuing efforts to establish quality evaluation products we are also beginning to review the use of object oriented design techniques and metaphors as a possible format for expressing multimedia evaluation. As a beginning we are considering each medium (text, sound, stills and motion) as a specific object and each object has its specific attributes (size, color, form declaration, representation and form) and behaviors (interactions). Objects can form clusters and it is these clusters that can have different relations and interdependency. Investigations into the use of cluster analysis as a methodology for multimedia evaluations is just beginning.

As an organizing principle the Multimedia Taxonomy presented here can be used to understand both design and content messages. The next step is to review the taxonomy in light of various studies in the psychological and cognitive aspects of understanding of multimedia applications to determine whether this taxonomy can shed light on these areas. Questions such as how we come to understand an image and how that understanding is different from our understanding of text can be answered in part by using the taxonomy. We have presented a few examples of the impact of the Multimedia Taxonomy on the design and implementation of evaluation protocols for multimedia products. More work in this area remains to be done. Finally, however, the Multimedia Taxonomy represents an attempt at the formalism that is needed to provide the both a theoretical and practical framework for the new and rapidly growing field of interactive multimedia.

REFERENCES

Aleem, T. A. (1997). World Wide Web site: http://www.erols.com/aleem/mediatax.html or http://www.erols.com/aleem/interact.html, work in progress.
Heller, R. S. and Martin, C. D. (1995) A media taxonomy, *IEEE Multimedia* , 2, 4, 36-45.
Heller, R.S. and Martin, C.D. (1998) Multimedia Taxonomy for Design and Evaluation. In Fruht, B (Ed) *Handbook on*
Multimedia Computing. CRC Publications.
Heller, R. S. (1994) Creating an advanced degree program in multimedia, *Proceedings of the ASE*E, New Orleans, January
1994, 353-356.

AdaMentor Formative Evaluation Comments for Open Questions

Keyword	Actual Comments of Concern	Actual Comments of Praise
Text	Hard to read; could be bigger Too much technical language	Not too wordy
Navigation	Hard to find exercises; Buttons didn't load in some cases Need to move inherent icons Need better button description Trouble getting back to where I was	Link design is efficient
Speed	Took forever to load Slow on T1 line Slowness more trouble than its worth	
Screen Design	Can't see entire page Screen size for exercises are too small	Clean setup Easy to look at Nice graphic layout
Interactivity	Not much better than a book Book could have been read anywhere Exercises are not interactive	
Content	Need better explanations Too much technical language Only include security information when necessary Make it more fun	Information on each topic was sufficient
Audience	Get it off the net Show Ada's practical application	Good tutorial for learning Ada First of its kind I've seen Easy to access Potential to be good tutorial

FIGURE 3: Selected Responses to Formative Evaluation Protocol Using the Taxonomy

Appendix 4.2.1: Multimedia Designer Qualification Questionnaire

This questionnaire is designed to collect background information about you, the evaluator, to help us interpret the evaluation findings. For each question, tick in the box next to the appropriate answer or write your response in the space provided.

1. **What is your job title?** _____

2. **Please briefly describe your main job activities.**

   _____

   _____

   _____

3. **Are you familiar with any of the following types of evaluation method? Please write down the name of the method/s.**

   - ❑ Heuristics Evaluation _____
   - ❑ User Testing _____
   - ❑ Checklist Evaluation _____
   - ❑ Guidelines Evaluation _____
   - ❑ GOMS _____
   - ❑ Cognitive Walkthrough _____
   - ❑ Other, please specify _____
   - ❑ None

4. **Have you used any of the following types of evaluation method prior to this experiment? Please write down the name of the method/s.**

   - ❑ Heuristics Evaluation _____
   - ❑ User Testing _____
   - ❑ Checklist Evaluation _____
   - ❑ Guidelines Evaluation _____
   - ❑ GOMS _____
   - ❑ Cognitive Walkthrough _____
   - ❑ Other, please specify _____
   - ❑ None

5. **Had you been involved in evaluating software before this experiment?**

❑ yes ❑ no *(If 'no', proceed to question 6)*

5.1 If 'yes', please specify the number of projects on which you had been involved in evaluation.

Number of projects: _____

6. **If you have any formal Human-Computer Interface Design training or qualification, please write down what it is.**

_____

_____

7. **If you have any formal Software Engineering training or qualification, please write down what it is.**

_____

_____

8. **Have you been involved in designing multimedia applications?** By multimedia applications we mean ones that integrate two or more media resources, such as text, images, animation, video, speech or sound in representing system information.

❑yes ❑ no *(If 'no', proceed to question 9)*

8.1 If 'yes' for how long?

Number of months _____ OR Number of years _____

9. **Have you been involved in developing multimedia applications?**

❑yes ❑ no *(If 'no', proceed to question 10)*

9.1 If 'yes' for how long?

Number of months _____ OR Number of years _____

10. **Had you been involved in evaluating multimedia applications before this experiment?**

❑yes ❑ no *(If 'no', proceed to question 11)*

10.1 If 'yes', please specify the number of projects on which you had been involved in evaluation.

Number of projects: _____

10.2 If 'yes', please specify what evaluation method/s you used.

_____

_____

_____

11. Are you familiar with any of the following evaluation methods?

- ❑ Multimedia Taxonomy, developed by Rachelle Heller & Dianne Martin (1999)
- ❑ Cognitive Walkthrough, developed by Pete Faraday & Alistair Sutcliffe (1997)
- ❑ Evaluation Checklist, developed by Philip Barker & Terry King (1993)

12. Have you used any of the following methods to evaluate multimedia applications before?

- ❑ Multimedia Taxonomy, developed by Rachelle Heller & Dianne Martin (1999)
- ❑ Cognitive Walkthrough, developed by Pete Faraday & Alistair Sutcliffe (1997)
- ❑ Evaluation Checklist, developed by Philip Barker & Terry King (1993)

13. Were you familiar with Maths CD-ROM prior to this experiment?

❑ yes ❑ no

14. How would you rate your knowledge of Exponential Graphs prior to this experiment?

❑ none ❑ basic ❑ intermediate ❑ advance

Appendix 4.2.2: Subject Matter Expert Qualification Questionnaire

This questionnaire is designed to collect background information about you, the evaluator, to help us interpret the evaluation findings. For each question, thick in the box next to the appropriate answer or write your response in the space provided.

15. What is your job title? _____

16. Please briefly describe your main job activities.

_____

_____

_____

17. If you have a formal qualification in mathematics, please write down what it is.

_____

_____

18. How would you rate your knowledge of Exponential, Logarithmic and Hyperbolic Graphs?

❑ basic ❑ intermediate ❑ advance

19. Have you got any teaching experience?

❑ yes ❑ no *(If 'no', proceed to question 6)*

20. If 'yes', for how long?

Number of months _____ OR Number of years _____

21. Have you got any experience in teaching mathematics?

❑ yes ❑ no *(If 'no', proceed to question 7)*

6.1 If 'yes,' for how long?

Number of months _____ OR Number of years _____

6.2 If 'yes,' at what level?

❑ GCSE ❑ University undergraduate

❑ A-Level ❑ University postgraduate

❑ Other, please specify _____

22. If you have a formal pedagogical or teaching qualification, please write down what it is.

_____

_____

23. Have you been involved in designing or developing software applications?

❑yes ❑ no *(If 'no', proceed to question 9)*

8.1 If 'yes' for how long?

Number of months _____ OR Number of years _____

24. If you have any formal Software Development training or qualification, please write down what it is.

_____

_____

25. Have you been involved in evaluating educational software before this experiment?

❑yes ❑ no *(If 'no', proceed to question 11)*

10.1 If 'yes', please specify the number of projects on which you had been involved in evaluation.

Number of projects: _____

10.2 If 'yes', please specify what evaluation method/s you used.

_____

_____

_____

26. Were you familiar with any of the following evaluation methods prior to this experiment?

❑ Multimedia Taxonomy, developed by Rachelle Heller & Dianne Martin (1999)

❑ Cognitive Walkthrough, developed by Pete Faraday & Alistair Sutcliffe (1997)

❑ Evaluation Checklist, developed by Philip Barker & Terry King (1993)

27. Had you used any of the following methods to evaluate multimedia applications prior to this experiment?

❑ Multimedia Taxonomy, developed by Rachelle Heller & Dianne Martin (1999)

❑ Cognitive Walkthrough, developed by Pete Faraday & Alistair Sutcliffe (1997)

❑ Evaluation Checklist, developed by Philip Barker & Terry King (1993)

28. Were you familiar with the Maths CD-ROM prior to this experiment?

❑yes ❑ no

11.1 If 'yes', have you used the Maths CD-ROM for teaching?

❑yes ❑ no

Appendix 4.3: Context Description of IMM Software

MathWise is a computer-based learning environment for the study and practice of mathematics. It comprises mathematical modules, reference material, assessments and resource tools. MathWise modules are based on the European Engineering syllabus, comprising mathematical topics taught in pre-university and first year university, together with a number of key topics in second-year university Science and Engineering courses.

Main Aim of "Exponential Graphs" Section

The aim of the section is to describe the main graphical features of the following elementary graphical functions:

- The exponential and hyperbolic functions
- The logarithmic functions.

The section also aims to explain the effect of translation, scaling and reflection on those functions.

Target User Group:

Mathwise is designed to be used by foundation, first year and second year university students, both male and female. The students can be of various ages, all older than 18 years of age, who also vary in their cultural backgrounds. They are required to have prior maths knowledge at A-levels or GCSE level, however no prior knowledge of the module material is assumed. The students can differ in their prior knowledge in the material studied, ranging from none to advance knowledge. The users' familiarity with computers and computer literacy can also vary – from novices to expert users.

Target Learning Environment:

Assume that the application will be used primarily for self-study in a computer lab (available from a computer network) or at the student's home, where immediate lecturer support will not be available. The material in the software is designed to complement the lecture material taught in foundation and first year mathematics courses.

Appendix 4.4: Evaluator Task Instruction Sheet - Study 1

The task you need to perform is to use the evaluation method as described in the documents to evaluate the effectiveness of the *Exponential Graphs* section of the multimedia Maths CD-ROM. The aim of the section is to enable students to learn more about the different types of exponential graphs.

While evaluating, could you write down an evaluation report, specifying:

(1) Any comments you have and any problems that you find with the application. Please include even the smallest errors, like misspellings.

(2) For each comment or error, could you please write a short description of it.

(3) For each comment or error, please specify the number or code of the criterion that led you to the problem.

Appendix 4.5: Expert Post-Evaluation Questionnaire

Thank you for taking part in this study. We hope you found today's session interesting. To help up establish how effective and easy to use the evaluation method is, please take a few moments to answer the following questions. For each question please indicate the answer, which most closely matches your opinion, or write down your answer.

1. Your general experience with the evaluation method

1.1 I found it easy to evaluate the application using the evaluation method.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

1.2 The evaluation method was sufficiently documented.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

1.3 I found it easy to grasp the concepts of the evaluation method.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

1.4 The documentation describing the method and how to use it was adequate for the evaluation task.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

1.5 I thought the concepts of the evaluation method were complex.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

2. The process of evaluation

2.1 Please list the steps you followed in evaluating the software.

_____

_____

_____

_____

_____

2.2 Were instructions provided by the method as to how to review the application and how to apply the evaluation criteria?

❑yes ❑no *(proceed to question 2.2.3)* ❑no opinion (*proceed to question 2.3)*

If 'yes':

2.2.1 I found the instructions easy to follow.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

2.2.2 The steps prescribed by the method were adequate for the evaluation of the IMM application.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

If 'no':

2.2.3 I believe it was necessary to have instructions provided for more effective evaluation.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

2.2.4 I was happy to adopt an approach to evaluation which I considered appropriate.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

2.3 How many of the evaluation criteria did you find relevant to the application?

❑ 0 - 20% ❑ 21 - 40% ❑ 41 - 60% ❑ 61 - 80% ❑ 81 – 100%

2.4 While evaluating the application did you use all questions or guidelines, or did you select the ones which are relevant to the IMM application?

❑selected the relevant criteria ❑ used all criteria *(proceed to question 2.5)*

2.4.1 If you selected relevant criteria, how did you do that?

❑following guidance provided by method ❑using your own judgement ❑both ❑neither

If you used method guidance:

2.4.1.1 I found the method's guidance helpful.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

If you used your own judgement:

2.4.1.2 I found it easy to decide which criteria were relevant.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

2.4.1.3 I would prefer to have selection guidance provided by the method.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

2.5 Did the method encourage you to generate questions and issues specifically tuned to the perceived use of the software?

❑yes ❑no *(proceed to question 2.6)* ❑no opinion *(proceed to question 2.6)*

If 'Yes':

2.5.1 I felt that the method provided sufficient guidance as to how to generate such questions.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

2.5.2 I found the process of generating such questions somehow tedious using method's guidance.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

2.6 Did the method encourage you to modify the existing evaluation criteria to match the particular context of use, subject matter content, particular user characteristics, or the target learning environment?

❑yes ❑no *(proceed to question 2.7)* ❑no opinion *(proceed to question 2.7)*

If 'yes':

2.6.1 I found it easy to adapt the evaluation criteria to the particular context aspects of the Maths IMM application.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

2.6.2 I felt that the method provided sufficient guidance as to how to modify the evaluation criteria.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

2.6.3 I found the process of modifying evaluation criteria somehow tedious using method's guidance.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

2.7 In evaluating the application did you consider any of the following contextual aspects? Please tick all that are appropriate.

❑ users' prior knowledge in mathematics;
❑ other user characteristics, such as age or computer literacy;
❑ what the learner needs could be, e.g. to extend their knowledge in a subject matter, or to sustain their interest in learning with the application (motivation);
❑ users' learning styles;
❑ aspects of the target learning environment, like classroom or self-study use, or teacher involvement;
❑ aspects of the particular subject matter content;
❑ how students can learn with the multimedia application;
❑ the nature of the learning task/s;
❑ cognitive task requirements, i.e. the cognitive processes required to be performed by the user in order to achieve the learning task.
❑ other, please specify

2.8 Were you prompted by the method to consider any of the following contextual aspects? Please tick all that are appropriate.

❑ users' prior knowledge in mathematics;
❑ other user characteristics, such as age or computer literacy;
❑ what the learner needs could be, e.g. to extend their knowledge in a subject matter, or to sustain their interest in learning with the application (motivation);
❑ users' learning styles;
❑ aspects of the target learning environment, like classroom or self-study use, or teacher involvement;
❑ aspects of the particular subject matter content;
❑ how students can learn with the multimedia application;
❑ the nature of the learning task/s;
❑ cognitive task requirements, i.e. the cognitive processes required to be performed by the user in order to achieve the learning task.
❑ other, please specify

3. The evaluation criteria, such as checklist questions or guidelines

3.1 I found every evaluation criterion well defined and adequately described.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

3.2 I found every evaluation criterion clearly and unambiguously worded.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

3.3 Did you have to assign values of certain criteria while evaluating the application?

❑yes ❑no *(end)*

If 'yes', for the criteria where a range of values were provided and you had to choose one:

3.3.1 I thought the range of values was clearly defined.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

3.3.2 I found it easy to choose a value for the criterion.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

3.3.3 I found the range of values appropriate to the criterion.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

If 'yes', for the criteria where a range of values were not provided, and you had to use your own judgement in order to assign a rating:

3.3.4 I found it easy to assign a value to the criterion.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

3.3.5 I would prefer to have a range of values to choose from provided by the method.

❑strongly disagree ❑disagree ❑neither ❑agree ❑strongly agree

Appendix 4.6.1: Multimedia Cognitive Walkthrough Interview Questions

Cognitive Framework:

1. How easy to understand did you find the cognitive framework?

2. Did you find the cognitive processes clearly and sufficiently defined? Were there any processes you did not quite understand?

3. Was it clear from the cognitive model how people process cognitively information presented in different single media, e.g. images or text? Was is clear from the cognitive model how people process cognitively information presented in integrated multiple media resources?

4. The cognitive framework is aimed to influence the design of the different media resources. Did you find the implications of each component in the framework on the relevant aspects of the design of various media clear and well defined?

5. Did you find the cognitive framework aided you in evaluating the various aspects of the media resources? Were evaluation issues clear in each step, and did the framework support the evaluation of these issues?

Evaluation Instructions:

6. Did you find the three steps of evaluation sufficiently defined in the documents provided to you?

7. While evaluating the application, did you follow the three steps suggested in the paper? If not, why, and what did you do differently?

Evaluation of Attentional Design:

8. Did you find it easy to draw an attentional graph?

9. Did the graph aid you in identifying focal points, i.e. what objects will be attended to at any point in time and for how long?

10. Was the relationship between the attentional guidelines and the juxtaposition of the presentation units on the attentional graph clear?

11. Did you find that the guidelines supported you in identifying the influence of media characteristics on users' attention? If yes, did you specify any adverse effects on users' attention in the evaluation report?

Evaluation of Contact Points:

12. Was the concept of contact points clearly defined?

13. Did you experience any difficulties identifying contact points between verbal and visual media resources?

14. Did you find that the guidelines and the attentional graph aided you in verifying when contact points are required and how to establish contact points?

15. Did the guidelines aid you in identifying the effect the contact points will have on users' ability to integrate information presented in different verbal and visual media? If yes, did you specify any adverse effects the contact points may have on users in the evaluation report?

Evaluation of Media Selection and Combination:

16. In this step you had to allocate an information type to each piece of information on the attentional graph, e.g. descriptive, spatial, procedural information type. Did you find it easy to decide which information types represent the mathematics content?

17. Was the description of each information type clear and understandable? Do you think that the suggested information types represent adequately the content included in the IMM application?

18. Did you find the guidelines helped you to validate whether the media used in the application are appropriate to represent the maths content?

19. Did you find it easy to map the media types and attributes provided in the guidelines onto the media representations included in the IMM application?

20. Did the media selection and combination guidelines aid you in identifying how the attributes of different media could influence users' comprehension of information? Did you specify any adverse effects the media may have on users' comprehension in the evaluation report?

Selecting Relevant Guidelines:

21. Did you consider all guidelines, or did you select only those, which were relevant to the IMM application? If you selected the relevant guidelines, how did you decide which ones were relevant and which were not?

22. Do you think that a procedure for selecting relevant guidelines should have been included in the evaluation method?

Tailoring Guidelines to Application Context:

23. While evaluating the application, did you consider any contextual aspects, such as who the target users will be, how they are likely to use the software for learning mathematics, or in what environment?

24. Did the evaluation method encourage you to consider such contextual issues? In what way?

25. Did you modify some or all guidelines according to certain contextual aspects? How did you approach the process?

26. Did the evaluation method encourage you to modify some or all guidelines? Did you find the method helpful in tailoring the guidelines to the contextual issues?

Usability of Evaluation Guidelines:

27. Do you think that adequate description was provided for each guideline?

28. Did you find it difficult to understand the meaning of any of the guidelines because of the way they are worded? If yes, please specify what was the nature of the difficulty, for instance double meaning, unknown terminology, ambiguous wording.

29. Would you like the guidelines to be improved, and in what respect?

Appendix 4.6.2: Interactive Multimedia Checklist Interview Questions

Evaluation Procedure:

1. Did you find the procedure for evaluation clear and sufficiently defined in the paper provided?

2. Would you prefer more guidelines regarding how to go about evaluating the software to have been included in the method's documentation?

Evaluation Categories:

3. Did you find the category names clear?

4. Do you think that adequate description was provided for each category?

Evaluation Questions:

5. Did you find the evaluation questions clear and understandable?

6. Did you find it difficult to understand the meaning of any of the questions because of the way they are worded? If yes, please give examples and specify what was the nature of the difficulty, e.g. double meaning of a term, unknown terminology, ambiguous wording?

7. Were all questions relevant to the category they were in?

8. A number of questions use specific software design or instructional terminology, e.g. *user needs (1.5), mode of interaction (5.1), visual and spatial access to information (6.5), learning styles (8.1).* Did you find it difficult sometimes to establish the precise meaning of these terms?

9. Did you find the sub-categories within some of the questions understandable and appropriate to the question? Example sub-categories are *focused*, *browsing*, *guided* and *structured* learning styles, which appear in question 8.1.

10. Did you experience any difficulties choosing one or more sub-categories where required? If 'yes', what was the nature of the difficulty?

11. Was it always clear what aspect or function of the software the question was referring to?

12. Were there any questions which you found difficult to answer? If 'yes', what was the nature of the difficulty?

Consideration of Application Context:

13. While evaluating the application, did you consider any contextual aspects, such as who the target users will be, how they are likely to use the software for learning mathematics, or in what environment?

14. Did the evaluation questions encourage you to consider such contextual issues? If 'yes' in what way?

15. In the questions where you had to imagine the users' reaction to the software, how easy did you find it? For instance, imagining whether the product develops a sense of ownership in a user (6.7), or if the user can identify with the topic's goals and objectives (1.6).

16. In your evaluation report did you specify any effects that the various aspects of the software may have on users' cognitive abilities, such as attention to or comprehension of information?

17. Did the evaluation questions aid you in identifying such effects, and how?

Selecting Relevant Questions:

18. While evaluating the application, did you follow the sequence of questions suggested in the paper? If not, why, and in what order did you answer the questions?

19. Did you consider all questions, or did you select only those, which are relevant to the IMM application? If you selected the relevant ones, how did you decide which ones were relevant and which were not?

20. Do you think that a procedure for selecting relevant questions should have been included in the evaluation method?

Appendix 4.6.3: Multimedia Taxonomy Interview Questions

Media Taxonomy:

1. How easy to understand did you find the media taxonomy?

2. Did you find the five *Media Types* clearly and sufficiently defined?

3. Were the five *Media Type* categories sufficient to represent all media resources included in the IMM application? Do you think that more categories are needed?

4. Did you find the different *Media Expression* categories clearly and sufficiently defined? Were the category names meaningful? Did the category names match their respectful descriptions?

5. Did media resources included in the application map easily in one of the *Media Expression* categories? Did you experience any difficulties allocating media resources included in the application to different media expressions?

6. Was it clear from the documentation what expressions of each medium will be able to represent information in a way that will increase users' understanding and retention of information?

7. How clearly and sufficiently defined did you find every category within the *Context* dimension? Did the category names represent their respective descriptions clearly and unambiguously?

8. How clearly defined did you find the four levels of *Interactivity*? How well do you think they are integrated with the rest of the Multimedia Taxonomy?

Generation of the Evaluation Protocol:

9. Was there sufficient guidance regarding how to create *Evaluation Protocols* provided in the method's documentation? Would you like more information to be provided, and regarding what aspects?

10. Did you find it easy to generate questions regarding specific media design aspects in each cell? Did you experience any difficulties while generating such questions, and what was the nature of the difficulties?

11. Could you rate the clarity of the relationship between the following dimensions:

- Media Type and Media Expression ❑very unclear ❑unclear ❑neither ❑clear ❑very clear
- Media Type and Context ❑very unclear ❑unclear ❑neither ❑clear ❑very clear
- Media Expression and Context ❑very unclear ❑unclear ❑neither ❑clear ❑very clear
- All three dimensions ❑very unclear ❑unclear ❑neither ❑clear ❑very clear

12. How easy did you find generating questions regarding the effectiveness of particular media combinations, e.g. text and images, within the *Multimedia* category of the *Media Type* dimension?

13. Were there any evaluation questions which you could not decide in which cell to place?

14. Were there any design aspects, which you wanted to consider not included in the taxonomy?

Selecting Relevant Cells:

15. Did you generate questions in every cell, or did you select only those, which are relevant to the application? If you selected only the relevant ones, how did you decide which ones were relevant and which were not?

16. Do you think that procedure for selecting the relevant cells should have been included in the evaluation method's documentation?

Considering the Application Context:

17. While evaluating the application, did you consider any contextual aspects, such as who the target users will be, how they are likely to use the software for learning mathematics, in what environment, or the subject matter content?

18. Did the evaluation method encourage you to consider such contextual issues? In what way?

19. Did you find the method helpful while generating questions regarding contextual issues?

20. Did you consider the contextual aspects individually, or did you consider them in relation to each other? If the latter, please give examples.

21. Did the Multimedia Taxonomy aid you in identifying relationships between the various contextual aspects? If yes, please give examples.

22. Did you consider relationships between any of the following aspects? If yes, please tick in the relevant box.

    - ❑ Audience and subject matter content;
    - ❑ Audience and interactivity;
    - ❑ Purpose and interactivity;
    - ❑ Audience and usefulness;
    - ❑ Audience and aesthetics.

23. Did the method explain the above relationships and how clearly and sufficiently?

24. Do you have any additional comments on the evaluation method?

Appendix 4.7: User Task Description

In order to help us evaluate the Maths CD-ROM we would like you to perform a set of four tasks with the software. The application is a computer-based learning application for the study and practice of mathematics.

Imagine that as part of the Induction in Mathematics course you need to learn more about the main concepts of Exponential functions. In particular, your aim is to learn more about the main graphical features of *Exponential functions*, including the three types of transformations which can be performed - reflection, scaling and translation - using the Maths CD-ROM.

Thus, the four main tasks you need to do are:

> *Task 1:* Find out about what *exponential functions* are, where they are used and how their functions and graphs represented.
>
> *Task 2:* Explore the principles of the two types of *Reflection* – horizontal and vertical - and their effect on exponential functions and their graphs.
>
> *Task 3:* Explore the principles of the two types of *Scaling* – horizontal and vertical - and their effect on exponential functions and their graphs.
>
> *Task 4:* Explore the principles of the two types of *Translation* – horizontal and vertical - and their effect on exponential functions and their graphs.

You will find information about the above topics in the *Graphs of Functions* module in the Maths CD-ROM.

REMEMBER: The aim of the experiment is to evaluate the software and NOT to test you!!!

Appendix 4.8: User Post-Evaluation Interview Questionnaire

This questionnaire contains questions regarding the extent to which you, the user, find the Maths CD-ROM easy and enjoyable to use. The information you provide will help us identify the good and the bad features of the software, and inform us how to improve the software and make it more suited to the way you would like it.

What two things about the Maths CD-ROM did you like most?

What two things about the Maths CD-ROM did you like least?

1. Engagement

1.1 Did you find the tasks you had to do using the Maths CD-ROM:
❑ too difficult for you ❑ difficult ❑ easy ❑ too easy ❑ just right ❑ don't know?

1.2 Were the goals and stages of each section of the Maths CD-ROM:
❑ very unclear ❑ not very clear ❑ sometimes clear ❑ always clear ❑ don't know?

1.3 How relevant did you find the Maths CD-ROM to the Foundation Maths course you are doing?
❑ not relevant at all ❑ not very relevant ❑ fairly relevant ❑ very relevant
❑ don't know

1.4 How enjoyable did you find using the Maths CD-ROM?
❑ not enjoyable at all ❑ not very enjoyable ❑ enjoyable ❑ very enjoyable
❑ don't know

1.5 How interesting was the material in the Maths CD-ROM?
❑ not interesting at all ❑ not very interesting ❑ fairly interesting
❑ very interesting ❑ don't know

1.6 Did you find the tasks you did with the Maths CD-ROM:
❑ very boring ❑ boring ❑ challenging ❑ very challenging ❑ don't know?

1.7 Did you feel that you learned something with this product?

1.8 Would you want to use the Maths CD-ROM again?

User Interface

2.1. How consistent was the information presented on the screens?
❑ very inconsistent ❑ fairly inconsistent ❑ consistent ❑ very consistent
❑ don't know

2.2. When you had to make important choices (e.g. moving from one part to another), did you find making such choices:
❑ always very difficult to make ❑ fairly difficult ❑ fairly easy ❑ very easy
❑ don't know?

2.3. Were the icons on the buttons:
❑ very unclear ❑ fairly unclear ❑ fairly clear ❑ very clear ❑ don't know?

2.4. Was the way information was laid out on the screen:
❑ very unclear ❑ fairly unclear ❑ fairly clear ❑ very clear ❑ don't know?

2.5. What do you think about the messages, which appeared on the screen? Were they:
❑ very confusing ❑ confusing ❑ clear ❑ very clear ❑ don't know?

2.6. Were the colours used in the application:
❑ very awful ❑ fairly awful ❑ fairly nice ❑ very nice ❑ don't know?

2.7. Did the background colours aid you in browsing through information?

2.8. Did you find any of the information presented culturally inappropriate or in contradiction with your culture?

Media Design

3.1. Was the text presented:
❑ very hard to read ❑ hard to read ❑ readable ❑ very readable
❑ don't know?

3.2. Was the quality of the graphics:
❑ very bad ❑ bad ❑ OK ❑ good ❑ very good ❑ don't know?

3.3. Was the size of the graphics:
❑ too large ❑ large ❑ small ❑ too small ❑ just right ❑ don't know?

3.4. Do you think that the graphs in the demos and tests help you understand the different types of transformation?

Media Integration

4.1. Do you feel that the text on the left-hand side usually supported or explained what was displayed in the graphs on the right-hand side of the screen?

4.2. How did you find the speed at which animated text messages appeared on the screen? Were they:
❑ too fast to read ❑ fast ❑ slow ❑ too slow ❑ just right ❑ don't know?

4.3. Were the graphics and text put together:
❑ in an unattractive way ❑ in an attractive way ❑ neither ❑ don't know?

4.4. When text was displayed dynamically together with graphs, did you find this:
❑ very distracting ❑ distracting ❑ it didn't bother you ❑ neither ❑ don't know?
❑ very irritating ❑ irritating ❑ pleasing ❑ neither ❑ don't know?

4.5. Was it always clear to you which graph line related to which equation on the demos and tests?
❑ always unclear ❑ sometimes unclear ❑ usually clear ❑ always clear
❑ don't know?

User Interaction

5.1. Did you find inputting information (e.g. typing in a number or pressing a button):
❑ always difficult ❑ sometimes difficult ❑ usually easy ❑ always easy
❑ don't know?

5.2. When clicking on a button, the result was:
❑ always unexpected ❑ sometimes unexpected ❑ usually expected
❑ always expected ❑ don't know?

5.3. When something when wrong, did you find the error messages:
❑ very unhelpful ❑ unhelpful ❑ helpful ❑ very helpful ❑ don't know?

5.4. Was the way the system responded:
❑ always unexpected ❑ sometimes unexpected ❑ usually expected
❑ always expected ❑ don't know?

5.5 Was changing the scale of the graphs on the demos and tests:
❑ very difficult ❑ difficult ❑ easy ❑ very easy ❑ neither ❑ don't know?

5.6 Did you find changing of scale of the graphs helped you view and understand them better?

5.7 When you were using the software, did you feel that you were in control? Did you feel that you could do what you wanted to do with the application?

5.8 Did the software ever bother you or confuse you in any way? When was it? What exactly confused you?

Mode of Interaction

6.1. Did the use of the mouse and the keyboard seem:
❑ very restricting ❑ restricting ❑ natural ❑ very natural ❑ neither ❑ don't know?

6.2. How clear was it to you when to use the mouse, when to use the keyboard and when you can use both:
❑ always unclear ❑ sometimes unclear ❑ usually clear ❑ always clear
❑ don't know?

Navigation

7.1. Did you ever feel lost when you were working with the Maths CD-ROM?
❑ always felt lost ❑ sometimes ❑ rarely ❑ never felt lost ❑ don't know

7.2. How certain were you about where to go next?
❑ always uncertain ❑ sometimes uncertain ❑ usually certain ❑ always certain
❑ don't know

7.3. When you wanted to proceed to the next screen, was the next screen:
❑ always unpredictable ❑ sometimes unpredictable ❑ usually predictable
❑ always predictable ❑ don't know?

7.4. When you wanted to go back to be previous screen, was it:
❑ always impossible ❑ sometimes impossible ❑ usually possible ❑ always possible
❑ don't know
❑ always unpredictable ❑ sometimes unpredictable ❑ usually predictable
❑ always predictable ❑ don't know?

7.5. Was the information organised in a:
❑ very illogical way ❑ illogical way ❑ logical way ❑ very logical way ❑neither
❑ don't know?

7.6. Did you feel at every point in the application that you had a clear idea of how much of the task you have completed?

7.7. Do you think you could choose which route to take that was most relevant to you?

7.8. Did you feel that there was a regular way of moving about the application?

Maths Content

8.1. Did you find the maths material:

❑ very difficult to understand ❑ difficult to understand ❑ easy to understand

❑ very easy to understand ❑ sometimes difficult – sometimes easy (50-50)

❑ neither ❑ don't know?

8.2. Was there help easily available when you needed explanations of unfamiliar words? If there was, was it:

❑ very unhelpful ❑ unhelpful ❑ helpful ❑ very helpful ❑ neither ❑ don't know?

8.3. Did the maths material match any maths conventions you were already familiar with? If not, which it didn't?

Monitoring and Assessment Techniques

9.1. How helpful did you find the software in assessing how you were getting on with the task?

❑ very unhelpful ❑ unhelpful ❑ helpful ❑ very helpful ❑ neither ❑ don't know

Learning Support

10.1. When you needed help (e.g. when you did not understand some of the material, or had difficulties doing the tests) was the help provided by the application:

❑ very insufficient ❑ insufficient ❑ sufficient ❑ absolutely sufficient ❑ neither

❑ don't know?

10.2. Do you think that the feedback given to you while doing the tests was:

❑ very inadequate ❑ inadequate ❑ adequate ❑ very adequate ❑neither

❑ don't know?

10.3. Did the feedback help you find the correct solution for the test?

❑never ❑ rarely ❑ sometimes ❑ always ❑don't know

10.4. Were the instructions about how to do the demos and tests:

❑ very unclear ❑ unclear ❑ clear ❑ very clear ❑ neither ❑ don't know?

10.5. Did you notice whether there are additional tools, such as glossary, calculator or graph builder?

10.6. Did you use any of these tools?

10.7. How helpful did you find them?

❑ very unhelpful ❑ unhelpful ❑ helpful ❑ very helpful ❑neither ❑ don't know

Tailorability

11.1 Was it possible to alter things to suit you personally, for example:

- change the colours,
- set up new ways of doing things?

11.2 Would you like to be able to do such things?

Appendix 4.9: Pre-Exposure Knowledge Test

1. A-level or GCSE level of maths?

2. Are you familiar with co-ordinate systems?
   ❑ yes ❑ no

3. Do you know what a mathematical function is?
   ❑ yes ❑ no

4. If 'yes', what types of function are you familiar with?
   - linear,
   - quadratic,
   - trigonometric,
   - exponential,
   - logarithmic,
   - other:

5. Do you know how to draw a graph for a function?
   ❑ yes ❑ no

6. Have you ever drawn such a graph?
   ❑ yes ❑ no

7. If 'yes', what kind of function have you drawn graphs of?
   - linear,
   - quadratic,
   - trigonometric,
   - exponential,
   - logarithmic,
   - other:

8. Have you heard about *exponential* functions?
   ❑ yes ❑ no

9. If 'yes', what do you remember about them?

10. Have you drawn graphs of *exponential* functions?
    ❑ yes ❑ no

11. Are you familiar with the 3 types of transformation of graphs:
    - reflection,
    - scaling,
    - translation.

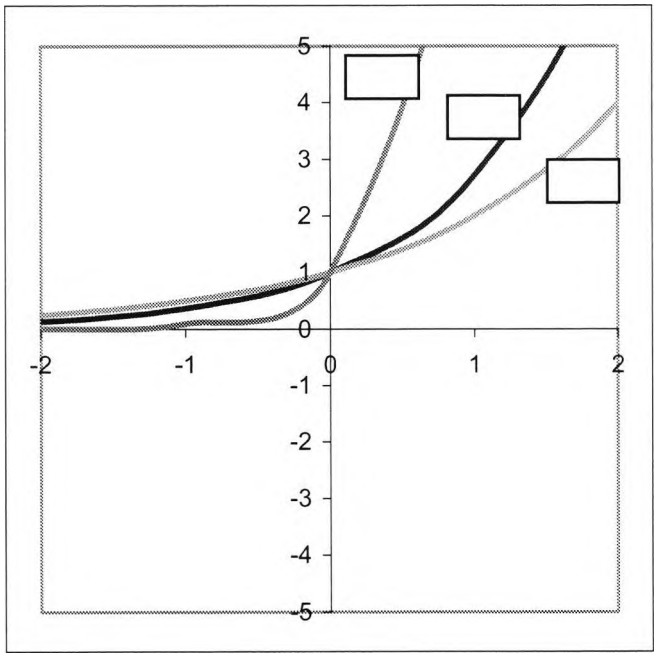
12. If 'yes' of which type of functions?
    - linear,
    - quadratic,
    - trigonometric,
    - exponential,
    - logarithmic,
    - other:

Appendix 4.10: Post-Exposure Knowledge Test

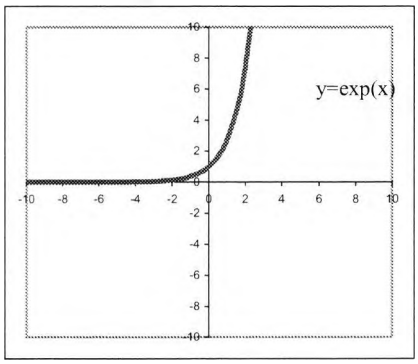
1. In the list of functions given below place a tick (✓) only next to the ones which are **exponential**. You can tick more than one if necessary.

- ❑ $y = x + 1$
- ❑ $y = x^2 + 4x + 5$
- ❑ $y = \sin 2x$
- ❑ $y = \exp(x) + 3$
- ❑ $y = e^{4x}$
- ❑ $y = \log_2 x$

2. **Label the graphs** in the figure below by writing which of the following three functions: e^x, 2^x or 10^x it corresponds to in the boxes provided.



3. The graph shown below is the graph of $y = exp(x)$.



3.1 In the boxes below, describe the **type** and **direction** of the transformation involved in transforming
$y = exp(x)$ into one of the following exponential functions.

a) y=exp(-x)

b) y=exp(x+2)

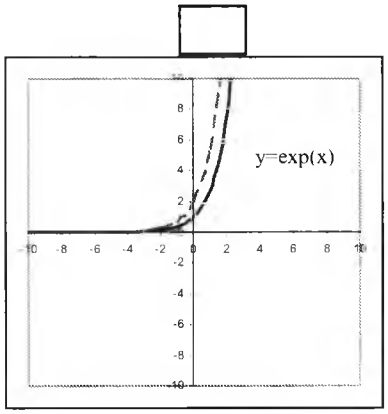
c) y=2 exp(x)

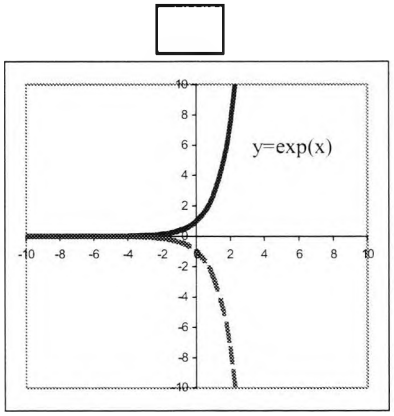
d) y=exp(2x)

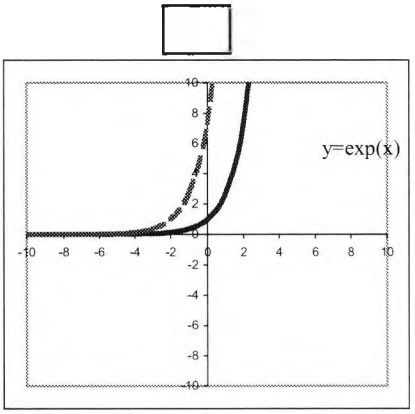
e) y=exp(x) + 2

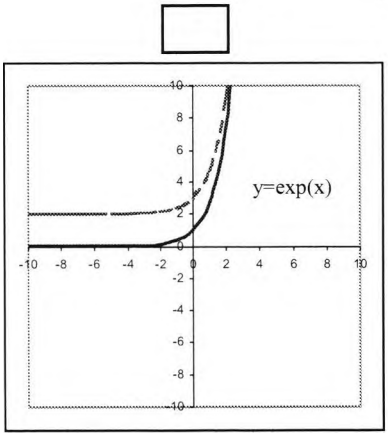
f) y=-exp(x)

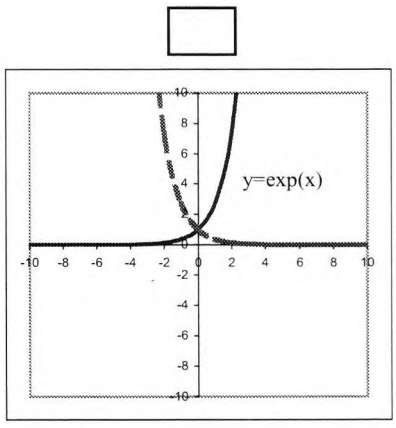
3.2 Now, using the principles of how the different types of transformation move the graph of $y = exp(x)$, **match** each of the exponential functions listed above with its corresponding graph, which is shown below. For your convenience the graph of $y = exp(x)$ is shown as well with the solid line, whereas the new graph is shown with an intermitted line. To match the graphs with its corresponding function, just write the letter the graph corresponds to in the box above the graph, as shown in the first graph.

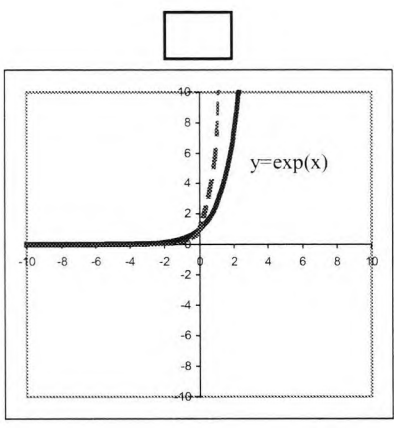






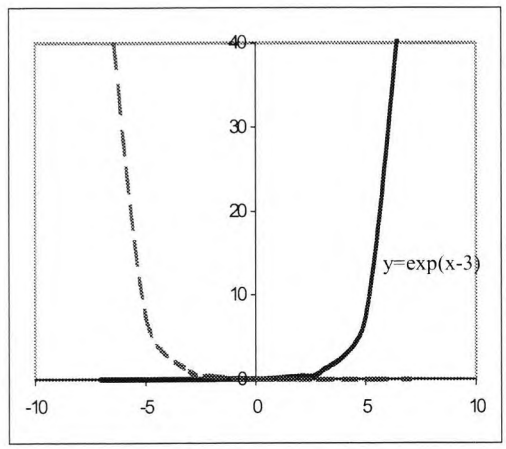




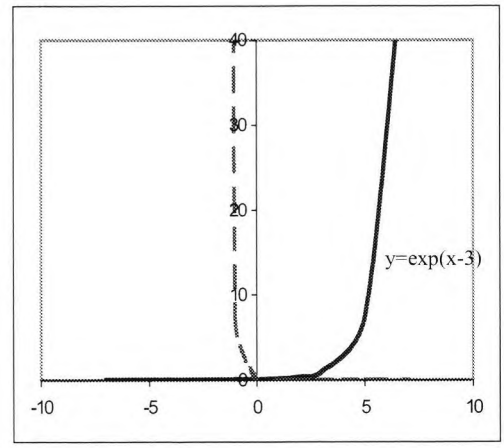


4. The graph of the function $y = exp\ (x-3)$ is sown below with the solid line, together with four transformed graphs of the same function, which are shown with intermitted lines.

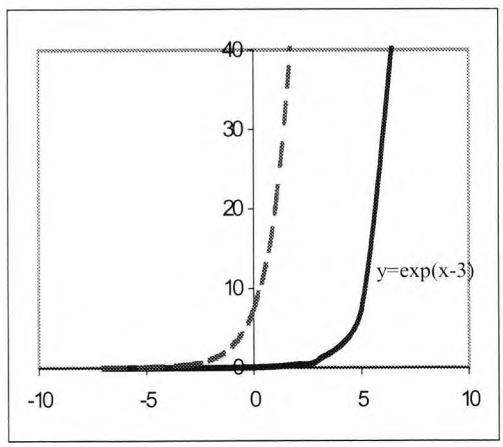
(a)



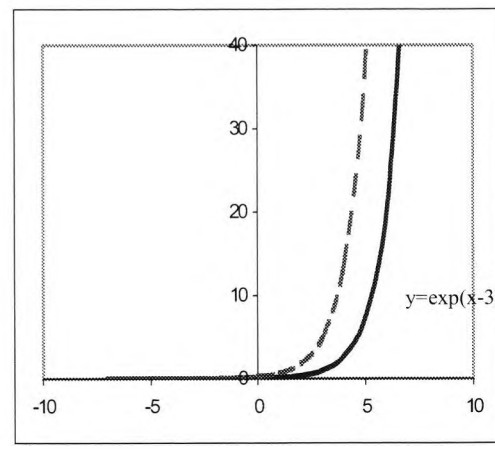
(b)



(c)



(d)



4.1 Describe the **type** and **direction** of the transformation which moved the graph of $y = exp\ (x-3)$ in each one of the transformed graphs.

(a) _____ (b) _____

(c) _____ (d) _____

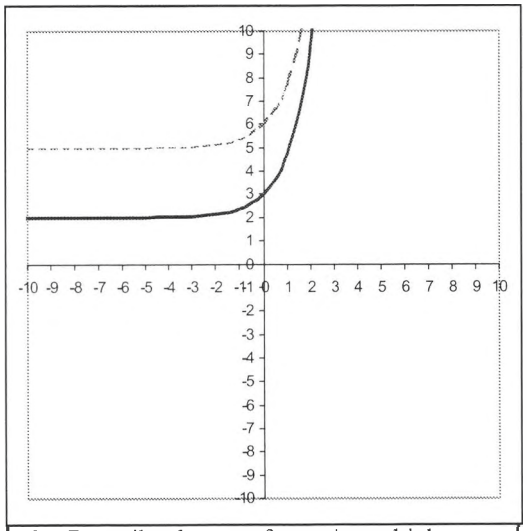
4.2 Now, write the **equation** for each of the transformed graphs.

(a) _____ (b) _____

(c) _____ (d) _____

5. The graph of the function $y = exp(x) + 2$ is shown in each of the four figures below. Underneath each figure, a transformed function is also given. Follow the instructions underneath the graphs.

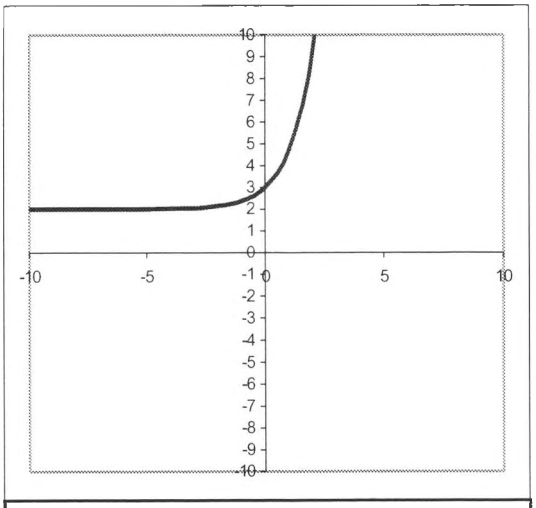
(a)



5.1 Describe the transformation which transforms the graph of $y=exp(x)+2$ into the graph of $y=exp(x)+5$.

5.2 Roughly sketch the graph of $y=exp(x)+5$ on the co-ordinate system provided above.

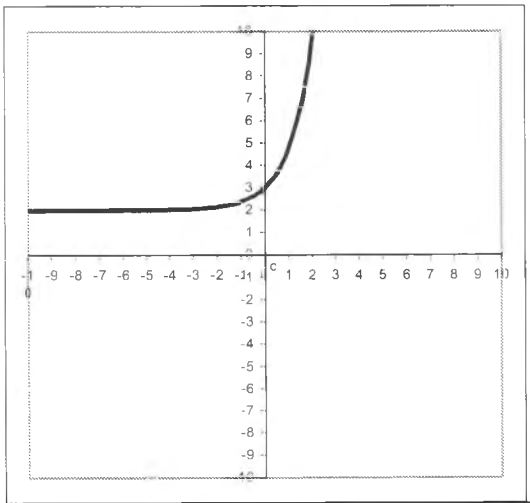
(b)



5.1 Describe the transformation which transforms the graph of $y=exp(x)+2$ into the graph of $y=exp(3x)+2$.

5.2 Roughly sketch the graph of $y=exp(3x)+2$ on the co-ordinate system provided above.

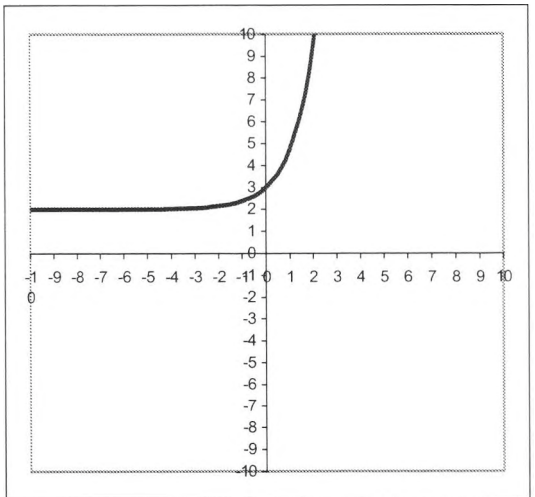
(c)



5.1 Describe the transformation which transforms the graph of $y=exp(x)+2$ into the graph of $y= - exp(x)+2$.

5.2 Roughly sketch the graph of $y= - exp(x)+2$ on the co-ordinate system provided above.

(d)



5.1 Describe the transformation which transforms the graph of $y=exp(x)+2$ into the graph of $y= - 3\ exp(x)+6$.

5.2 Roughly sketch the graph of $y= - 3\ exp(x)+2$ on the co-ordinate system provided above.

6. Given the exponential equation $y = exp(-10x) - 8.2$, and its vertically translated equation $y = exp(-10x) + 70$

6.1 What would be the **distance** between the two graphs if drawn on a co-ordinate system? Please tick the box, which you think contains the right answer.

❑ 8.2 ❑ 78.2

❑ 70 ❑ 10

6.2 Which **direction** would the translation take? Please tick the box, which you think contains the right answer.

❑ Up ❑ Down

7. Given the exponential equation $y = exp(-13x) - 1.2$, and its horizontally translated equation $y = exp(-13\ x\text{-}26) - 1.2$

7.1 What would be the **distance** between the two graphs if drawn on a co-ordinate system? Please tick the box, which you think contains the right answer.

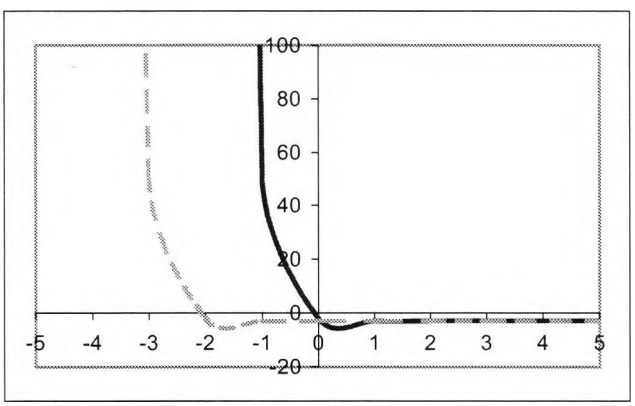
❑ 1.2 ❑ -13

❑ 2 ❑ -26

7.2 Which **direction** would the translation take? Please tick the box, which you think contains the right answer.

❑ Right ❑ Left

8. Below in is an exponential graph, which has been horizontally translated. Calculate the **distance** between the two graphs.



❑ 1 ❑ 3

❑ 2 ❑ 4

Appendix 4.11: Sample Transcripts from Semi-Structured Interviews

SMM3 – Multimedia Cognitive Walkthrough

A1: I looked at it before, but for using the method I didn't refer to it. I think I found it fairly easy to understand. The concepts aren't very complicated, so perhaps they could have been presented in a simpler language.

A3: I didn't feel the need to look at that. I think the method went on to focus on the attention and contact points, and information types, so perhaps it didn't need you to require these cognitive processes. I didn't feel that it's essential.

A6: I thought it just gave you a set of guidelines, quite a lot of them, certainly for the first bit. So perhaps it would be helpful to have more structure to that, to give you more guidance about what to do in this attentional design thing. For each method stage it seems to have just a list of guidelines.

Q: so you didn't draw an attentional graph? Do you think it may have helped?

A: I would feel it would be an overkill, because in general I though the method wasn't appropriate for this type of presentation. I thought there was too much text in here. And with it being graphs rather than images, and the movement being only the graphs moving up or down, I didn't think this method is the right approach. I though the package included more direct manipulation stuff, so guidelines for direct manipulation or hypertext would be more useful, it seemed quite hypertexty, pages appearing here and there. And maybe something like this would be more appropriate, where there's more movements more images and less text, so the text is more simple sentences and maples rather than full paragraphs or descriptions like it was here. I thought that more general sense and general HCI that you can use.

A11: Not really very much. I think I only picked up one or two of the guidelines and the rest didn't really seem appropriate, because this is a different type of presentation. So here one will just read through the text and maybe click on this, but this won't be happening at once, and you choose when to click on it. So it was more hypertext than multimedia.

A12: I generally understood what was meant by contact points. But because I wasn't really using this guidance very much I didn't develop a deep understanding of it. But that's maybe more because I didn't look to develop a deeper understanding of it.

A 13: Yes, I think I mentioned that, the first one a bit. Yes, I said not good contact between graph and functions.

A 14: Well, in general I thought that the pages are quite similar so they make the same problems/mistakes. So in general, when they were referring to a graph function in one of the demos they should make a better link to it. I don't think I've talked about solutions, I only talked about problems, so I just said not good contact, I haven't said what they should do about it.

A15: I don't think I decided to look that deeply, for me it seemed enough that the problem have been highlighted, and I didn't look at proposing solutions in a more detailed way.

Q: did you look at the media selection guidelines and the information types?

A: I didn't find the information types helpful, because it was all a maths thing, it all seems abstract to me. I mean it was non physical and it was numbers and the guidelines here said that perhaps you should use text rather than images. So I think that seemed a bit contradictory, because the graphs seemed useful for representing that. But I couldn't say that was because it was spatial information, because it wasn't really spatial. So I don't feel it could be applied to this application very easily. And the media selection guidelines just carried on from that. So I found it a bit unhelpful that it seemed to be suggesting that you should be using text for these numbers and not images, when my common sense says that the image was better.

A18: the guidelines said linguistics. Because there's no sound at all in the application (this section) I didn't consider sound very much, it seemed that they've excluded it altogether. I think it would take a bigger decision to put sound in there, it's at a higher level decision to say if you haven't got it anywhere else. If you had sound at some places, then I could say you should use sound here as well.

A20: Not, because this leads from the fact that this domain didn't really fit in the information types very well or maybe the other way round, there wasn't a suitable information type in which to fit the info to be presented. It didn't really help when to use the graphs and when not. All seemed to fit in this qualifying information type.

A21: I skimmed through them, and generally when I was going through them the first one or two would be relevant, and they would get increasingly more specific and less relevant. So the still image one talks about brightness, but I didn't see an issue of that here. It talks about attention and gradually revealing objects, and the rest seem to get more specific, like too many highlighting techniques or labels. I think I used the label one. Whichever seemed relevant, but most of them didn't seem immediately relevant. There's a long list of them and it wasn't an easy way to work your way around them for each possible issue.

A23: I think that was at the beginning, that was in the back of my mind that they wouldn't have a trainer and that they would be interested in this maths area, which I don't seem to remember any of it, and also have some background as far as A-Levels. When I was reading this I didn't try to understand it, because I didn't really remember. I was just assuming that the user would be at a level to understand the content.

A24: No, it seemed to focus more at detailed level of presenting information in a coherent way. No.

A27: I thought the description was sometimes too abstract and difficult to understand. But there was so many of them, so I think I don't want more text on each guideline because that would be difficult to digest, maybe a bullet point that was very simple, that if you needed more information you could get it.

A28: I can't think of anything specific, but in general I found it a bit abstract, terms like direct contact is a bit abstract. I think at one point context is mentioned, and it's not clear what is meant by context.

A1: Yes, I think so. It was a very dry paper, not very easy to read.

A2: Not with something like a checklist.

A3: Yes, they were fine.

A4: Yes, that was fine.

A5: the majority of them yes, but the terminology of some of them was tricky.

A7: Yes.

A9: I had to guess a little bit, it wasn't quite spelt out, but it was fine.

A10: I did have to hesitate a little bit. On both questions, but the one on interactivity in particular, because I didn't know whether I was allowed to thick all of them.

A11: Yes.

A12: 8.5 and 8.6 were impossible, I couldn't understand them at all. I put N/A to lots of them which I think were not applicable to that software. And on some of them I had to put 'don't know' partly because I don't have the expertise to answer them or because I didn't know from the part that we've seen whether the question is applicable or not. Otherwise fine.

A13: The only thing I did consider was how different the experience would be if you actually knew what they were talking about. As I didn't understand any of it I wondered whether I was able to thoroughly evaluate it. I know it was more about usability, but maybe to have been able to understand more of what was being said might have given me more indication of the usability. That mean that the target users I just had to take one's word for it, that's it's 1st year undergraduate students.

A14: No, I think with regard to the end user, there were only a couple of user at the beginning about engagement, is it appropriate for the user group. Well, I presume yes, but I don't know. And some of these questions were difficult to use because of those issues that have been raised.

A15: Not very easy.

Q: Do you think that if you had more information given to who the target users will be, that would have helped?

A: I think that just the context of the product is so much outside my knowledge, I that was the bit what made it difficult for me. I can put myself in the shoes of an undergraduate learning something, but not maths.

A16: I mentioned something under the learning styles, that although a path is suggested and you have the buttons and you're meant to click on them, I think it would be very easy to do what I did to just browse through it without taking anything in at all. Because, yes, there is a certain level of interactivity even with the tasks and the tests, but if you get the wrong answer it doesn't actually prompt you to go back and really seriously figure out what the right answer is. It would be very easy to do it in a mindless fashion.

A18: Yes.

A19: Basically intuitively, whatever I could answer I guest was applicable, and whatever I couldn't answer I just guessed wasn't applicable. For a few of them I said 'I didn't know' as I wasn't entirely sure whether it was applicable or whether it was me not knowing. Basic intuition I'm afraid.

A20: Yes, I might have been a nice idea.

Appendix 4.12.1: Problems Predicted Using the Expert Evaluation Methods

No	Problem Description	UEM - # of experts	Severity	Type	Diagnosed	User Problem Match - Strength
1.	Page 1 and in leaflet – the contact between text and graphs of functions is not good.	MMCW-1	Minor	MEDINT	US-NE	U38 -2
2.	Page 1/demo – Contact point cannot be formed between graph and text 'comparing f(x) with 2^x and 10^x', as graph is moved on too fast, and the text is changed too fast.	MMCW-2	Medium	MEDINT	US-NE	U4 -3
3.	Page 1 - Back button on graph maybe too indirect and not obvious. Should use maybe direct type for moving back.	MMCW-1	Cosmetic	NAVIG	US-NE	
4.	Page 1 – Contact point between prompting text to re-draw graph for 2^x and button is very indirect.	MMCW-1	Minor	MEDINT	US-NE	U1-3 C1-1
5.	The language used is mathematical and quickly jumps to core maths. The scene has not been set sufficiently to have an understanding level high enough for the following panes.	MMCW-1	Cosmetic	CONT	FS	False alarm
6.	Pane 2 - Although navigation through the leaflet is fairly straightforward, there is never a direct indication of where to advance to next.	MMCW-1	Minor	NAVIG	US-NE	False alarms
7.	Pane 2 – If I do the demo/test in the second pane I remain confused as to what I should be doing. I click on start which is fine, click the buttons if I got it wrong it is not clear to me, it simply claims I should read the notes again – which notes?	MMCW-1	Cosmetic	USESUPP	FS	U36-3
8.	The application jumps around a lot, notably when leaflets are employed. User focus is disjoined, and a concept of where you are and how much you have completed of the work eludes you.	MMCW-1	Minor	NAVIG	FS	False alarm
9.	There is a natural progression, but buttons do not take precedence over each other.	MMCW-1	Invalid	NAVIG	US-NE	
10.	Most of the application used indirect connections, allowing user freedom to move around the work, but rarely focusing the user's attention on the area.	MMCW-1	Minor	ATTENT	US-NC	U6-1
11.	You lose the concept of where you are in the bigger picture. And the ability to move into a leaflet while doing a demo-test may cause problems – to indirect.	MMCW-1	Minor	NAVIG	US-NC	
12.	No objects are highlighted anywhere, other than by sequence as being more important than others. Very few highlighting techniques are used.	MMCW-1	Cosmetic	ATTENT	US-NE	
13.	All text, apart from headings, has the same typeface, none is highlighted above the rest as being more important. Bullet points may help here.	MMCW-1	Cosmetic	ATTENT	US-NE	
14.	All objects are revealed at once in screens, viewing is controlled by the assumption that the user will read all text then move to demos. Buttons provide additional information, but this is not controlled in terms of the sequence the user accesses it.	MMCW-1	Cosmetic	ATTENT	FS	U9 - 1 U10-1
15.	Dialogue boxes appear in awkward places obscuring the view of the thing they are about.	MMCW-1	Medium	SCRLAY	US-NE	U40-4

No	Problem Description	UEM - # of experts	Severity	Type	Diagnosed	User Problem Match - Strength
16.	Graph drawing animation is too fast, the animation is pretty much irrelevant given their speed – maybe slowing them down will help. Graph lines are not revealed gradually. Graphs are shown too quickly – should be shown one at a time to allow users to read and look at each one before displaying the next.	MMCW-2	Medium	ATTENT	FS	
17.	Many of the demos and tests have text on the graphs, which appears with an animation. The text is ignored by the user until the animation is complete, some of this text disappears after the end of the animation meaning it may not get read by the user.	MMCW-2	Medium	MEDINT	FS	U4 -4 U38-3
18.	Leaflet pages - Should give instructions with each change and not all at once.	MMCW-1	Medium	TASKINS instructional	US-NE	
19.	Page 2 leaflet – The two graphs are confusing - should distinguish between them better. Perhaps graph line for *k* should be in language form.	MMCW-1	Minor	ATTENT MEDSEL	US-NE	U20-2
20.	The user has no control over product parameters such as colour settings, design layouts or sound volume settings.	IMMC-2	Cosmetic	TAILOR	US-NE	
21.	In some of the demos, you are told you can click the positive bounds of the graphs to change the scale. On some of the graphs, e.g. in page 1, it seemed to do nothing when tried to click in response to that same prompt.	IMMC-1	Invalid	MISLFUNC	US-NE	False alarm
22.	The goals, objectives and stages are not always clear. No, it's not obvious what the module is trying to achieve.	IMMC-3	Cosmetic	USEGOAL	US-NE	U39-4
23.	The user cannot identify with the goals and objectives and build in their own personal plan of achievement.	IMMC-2	Medium	USEGOAL instructional	US-NC	
24.	The presentation is not effective in maintaining interest. It's too static and formal.	IMMC-2	Minor	ENGAG	US-NC	U44-3
25.	I believe the format may be frustrating if the learner experiences difficulties.	IMMC-1	Minor	PRESNT	US-NC	
26.	The presentation is not enjoyable, because I'm not interested in the subject matter.	IMMC-1	Minor	ENGAG	US-NC	U43-4
27.	The product can't hold user's attention. I think frustration will set in for users with difficulties.	IMMC-2	Minor	ENGAG	US-NC	
28.	User's concentration cannot be maintained consistently.	IMMC-2	Invalid	ENGAG	US-NC	False alarm
29.	The product is not challenging.	IMMC-2	Invalid	ENGAG	US-NC	U45-4
30.	A greater range of levels of interaction is required.	IMMC-1	Invalid	USEINT	R.S.	
31.	Dialogue boxes appear too quickly and conceal the graph being drawn.	IMMC-1	Medium	SCRLAY	US-NE	U40-4
32.	The feedback is not offered in a consistent manner. There wasn't much feedback.	IMMC-1	Minor	FEEDBK	US-NE	U36-3

No	Problem Description	UEM - # of experts	Severity	Type	Diagnosed	User Problem Match - Strength
33.	The learners are not able to navigate through the system making real and personally relevant decisions about order, topic, level, pace of material presented.	IMMC-1	Invalid	NAVIG	US-NC	
34.	New configuration cannot be saved for the user.	IMMC-1	Cosmetic	MISSFUNC	US-NE	
35.	The content cannot be altered or adapted to the needs of the user.	IMMC-2	Invalid	TAILOR	US-NE	
36.	There's no much opportunity for the user to change the pace/level of the material.	IMMC-1	Invalid	TAILOR	US-NC	
37.	The style of interaction is not necessarily appropriate to the needs of the user.	IMMC-1	Invalid	USEINT	US-NE	
38.	The user cannot select the level of information presented.	IMMC-3	Invalid	MISSFUNC	US-NE	
39.	The user cannot establish new paths through the product.	IMMC-1	Invalid	MISSFUNC	US-NE	
40.	The user cannot set up short-cut facilities such as macros.	IMMC-2	Invalid	MISSFUNC	US-NE	
41.	The product cannot make use of information on a user's background.	IMMC-2	Invalid	MISSFUNC	US-NE	
42.	The multimedia mix is not used responsibly or over-crowding of multimedia objects reduces the necessary effectiveness of some media.	IMMC-1	Invalid	MEDINT	US-NE	U40-2
43.	The size of display used is small; the graph scale may be adjusted to appropriate values.	IMMC-1	Minor	PRESNT	US-NE	
44.	More use of audio would help to support the learning style involved.	IMMC-1	Minor	PRESNT	R.S.	U43-1
45.	Multimedia techniques slightly inhibit engagement.	IMMC-1	Invalid	ENGAG	US-NC	False alarm
46.	You can't use the keyboard with the calculator, you can only use the mouse. So, it would be nice to allow you to have the keyboard as well.	IMMC-1	Minor	MODEINT	US-NE	
47.	Facilities such as accelerator keys are not made available to users.	IMMC-3	Cosmetic	MISSFUNC	US-NE	
48.	The mode and style of interaction does not encourage engagement.	IMMC-2	Invalid	ENGAG	US-NC	
49.	There was some inconsistency in the interaction, in terms of the backgrounds. The use of a background colour wasn't always consistent, sometimes the graphs were drawn on a green background, sometimes they were white, and sometimes the proof were on a green background the same as the graphs. So some kind of colour coding would have been nice. The graphs were always on the same colour, the proofs were always on the same colour would have helped.	IMMC-2	Minor	PRESNT	US-NE	
50.	There was some inconsistency in the interaction, in terms of the moving forward.	IMMC-1	Invalid	USEINT	US-NE	
51.	An inconsistency of user interaction is that they've decided not to have any pop-up windows for sections other than what's found in the resource browser. So if you're going in to a scaling leaflet it comes up over and in the same window as the previous section, which is in terms of the use of the back button.	IMMC-1	Cosmetic	PRESNT	US-NE	
52.	An inconsistency of user interactions is that there wasn't always labelling of the icons.	IMMC-1	Cosmetic	PRESNT	US-NE	U50-2
53.	Interface not consistent – when working through exponential graphs, everything is consistently mapped on	IMMC-1	Cosmetic	SCRLAY	US-NE	

No	Problem Description	UEM - # of experts	Severity	Type	Diagnosed	User Problem Match - Strength
	page until reaching logarithmic graphs. Then left-hand side of window is different from previous pages – only access to leaflets, no information on the page itself.					
54.	User interactions are not always clearly defined.	IMMC-2	Cosmetic	USEINT	US-NC	U39-3
55.	User interactions are not always effective.	IMMC-2	Invalid	USEINT	US-NC	U13,14-1
56.	The user is not always presented with good visual and spatial access to information.	IMMC-2	Minor	PRESNT	US-NE	
57.	The product does not offer support or help to the user as needed.	IMMC-2	Minor	USESUPP	US-NE	U36-1
58.	The product doesn't develop a sense of ownership in a user, as it doesn't ask for their name at the beginning.	IMMC-3	Cosmetic	ENGAG	FS	
59.	User interactions are not secure.	IMMC-1	Invalid	MISSFUNC	US-NE	
60.	The interface is not integrative with regard to multimedia techniques.	IMMC-1	Minor	MEDINTG	US-NE	U38-1
61.	The interface does not allow links to be made between related areas of information.	IMMC-3	Medium	PRESNT	US-NE	
62.	The way in which options are presented to the user is not consistent and of high quality.	IMMC-1	Minor	PRESNT	US-NE	
63.	Help is accessible, but is not obvious, and once there it's not useful.	IMMC-1	Minor	USESUPP	US-NE	
64.	The interface is below averagely intuitive.	IMMC-1	Minor	PRESNT	US-NE	False alarm
65.	The style of learning is not under user control.	IMMC-3	Invalid	TAILOR	US-NE	
66.	The learning material does not tend to facilitate the transference of existing skills.	IMMC-1	Invalid	CONT	US-NE	
67.	When incorrect answers are input, there are no useful hints provided to assist user. They can only use the 'show me' button which does not actually assist in correcting understanding.	IMMC-1	Medium	USESUPP	FS	U36-4
68.	The product does not monitor performance.	IMMC-3	Minor	MISSFUNC	US-NE	
69.	The products cannot assess performance.	IMMC-3	Minor	MISSFUNC	US-NE	
70.	The students are advised of the effectiveness of their performance only locally, immediately after you've answered the question. So there's no recording of scores.	IMMC-2	Minor	ASSESS	US-NE	
71.	The students are not advised of the effectiveness of their performance.	IMMC-1	Minor	ASSESS	US-NE	
72.	The monitoring and assessment processes cannot be made transparent for the benefit of the tutor.	IMMC-3	Minor	MISSFUNC	US-NE	
73.	There isn't a facility for long term monitoring of student progress.	IMMC-3	Minor	MISSFUNC	US-NE	
74.	The product does not facilitate review and study of user's own learning.	IMMC-4	Minor	MISSFUNC	US-NE	
75.	The product does not empower the users to create new assignments.	IMMC-4	Invalid	MISSFUNC	US-NE	
76.	The product does not facilitate the drawing up of self-assessment protocols.	IMMC-4	Minor	MISSFUNC	US-NE	
77.	The product cannot make use of monitored information on a user to provide support to enhance engagement.	IMMC-3	Minor	MISSFUNC	US-NE	
78.	The product does not support the creation of personalised maps or navigational aids by the use of monitoring.	IMMC-3	Minor	MISSFUNC	US-NE	
79.	The product does not give personalised help to the user when a route is to be chosen through the software.	IMMC-3	Invalid	MISSFUNC	US-NE	

No	Problem Description	UEM - # of experts	Severity	Type	Diagnosed	User Problem Match - Strength
80.	The product does not prompt the user according to personal circumstances when opinion choices are to be made.	IMMC-3	Minor	MISSFUNC	US-NE	U13-2
81.	The product does not offer the facilities of an expert system.	IMMC-4	Invalid	MISSFUNC	US-NE	
82.	The product does not offer any features of an intelligent tutoring system.	IMMC-4	Invalid	MISSFUNC	US-NE	
83.	The product does not provide links to specialist software.	IMMC-2	Cosmetic	MISSFUNC	US-NE	
84.	The product is not applicable for use by a group as groupware or in a distributed network.	IMMC-1	Invalid	MISSFUNC	US-NE	
85.	The product does not facilitate the construction of group goals for learning.	IMMC-2	Invalid	MISSFUNC	US-NE	
86.	The product does not facilitate continued co-operation between members of the group	IMMC-1	Invalid	MISSFUNC	US-NE	
87.	The product does not enable the group to assess what needs to be done.	IMMC-2	Invalid	MISSFUNC	US-NE	
88.	The product does not enable the self-assessment of the learning process by members of the group.	IMMC-2	Invalid	MISSFUNC	US-NE	
89.	The product is not adaptable for a number of educational settings.	IMMC-1	Invalid	MISSFUNC	US-NE	
90.	Text changes too fast to read. If the text was slower it would be easier to see what was going on. Page 5, leaflet, page 1 – the text which changes beside the demo button is too fast to read.	MMT-1	Medium	ATTENT	FS	U4-4
91.	Page 1, demo- It says 'finally click again to see the graph', and I clicked on the graph pane, not on the button, which is confusing. I guess in the past you could only use buttons, but we have whole layers, which can be active.	MMT-1	Minor	TASKINS	FS	U1-3
92.	Page1, demo - 'now click again for the graph of 2x', well I clicked again and there's something funny about that. I'm not sure that it's good to assume that my reading speed and focus on that top left corner is sufficient. I might not have been reading it at all, it seems very unlikely actually that I'll be reading the text at all, I would have been struck immediately by the graph. They are all things that cause 30 sec of confusion, they are not things that are major faults, maybe they affect people's confidence.	MMT-1	Minor	ATTENT	FS	U4-3
93.	Page1, demo - 'now click again for the graph of 2x', well I clicked again and there's something funny about that. So it should say 'click the button on the left of this window'. My suggestion would be, unless there is some help facility somewhere else, the button should be up front and the label should be used at the beginning rather than the end. They are all things that cause 30 sec of confusion, they are not things that are major faults, maybe they affect people's confidence.	MMT-1	Medium	MEDINT	US-NE	U1-1
94.	No information about what is next.	MMT-1	Minor	NAVIG	US-NE	False alarm
95.	This box keep jumping on top of my graph. It's just a nuisance.	MMT-1	Medium	SCRLAY	US-NE	U40-3
96.	It's difficult to feel in control of the application.	MMT-1	Cosmetic	USECONT	US-NE	False alarm
97.	The feedback to audience is not always appropriate. Test, feedback–when wrong answer is typed in, the	MMT-1	Medium	FEEDBK	US-NE	U36-4

No	Problem Description	UEM - # of experts	Severity	Type	Diagnosed	User Problem Match - Strength
	right answer could be explained to correct understanding.					
98.	Navigation misguides. There are no explanation of the structure and sequence of pages. At the beginning you would expect an introduction to the structure, and there is no introduction to the structure of the section. First we do that, and then do that. Don't do that, because you may think that you know it, but you have to do it first. It's supposed to be a tutor, and there is nothing personal about it.	MMT-1	Medium	NAVIG	US-NE	False alarm
99.	The narrative of problem solution explanation is out of sequence.	MMT-1	Minor	USESUPP instructional	US-NE	
100.	Some errors in the narration lead to confusion, i.e. text on both sides of screen, different background colours lead to confusion about its significance. In leaflets, when they use the area on the right-hand side to display text for what, it is not clear whether the text is a continuation of the information in the left-hand pane, or an explanation of it. And if it's a continuation of this, the same background colour should be used.	MMT-1	Minor	PRESNT	US-NE	False alarm
101.	The size of graphs, the font, the colours, the size and location of buttons can be made customisable.	MMT-1	Minor	TAILOR	US-NE	U46-2
102.	Need to make more space proactive to mouse clicks. In introduction of the topic they could have used the whole space to be active, so clicking anywhere on the screen can take you forward, instead of having to go to the bottom and click on a button.	MMT-1	Medium	USEINTR	US-NE	U24-2
103.	The product did not encourage the use of the tools that exist.	MMT-1	Medium	MISSFUNC	US-NE	U41-1
104.	In front screen the word 'Mathwise' is not anti-aliased. Aliasing only occurs in cases – use antialise tools for the graphics.	MMT-1	Cosmetic	MEDDESN	US-NE	
105.	Location of graphs is predictable, but inappropriate. In the leaflets the demo graphs would have been better to be in the bottom half on the left-hand side and by pressing on the demo button to have the graph activated. The way it is now, the user can skip the demo. The k line doesn't mean anything to me. I didn't read the text, the problem is how does this help me to come here.	MMT-1	Cosmetic	SCRLAY	FS	U5-1 U8-2
106.	Linear delivery is increasingly outdated and uninteresting.	MMT-1	Cosmetic	NAVIG	US-NE	U44-3
107.	The location of the text was found to be inappropriate in a number of areas. The text should have a more prominent role in the interface than the buttons on the right-hand side. The buttons could be smaller, because the aim is not to hit the buttons but to read the text.	MMT-1	Minor	SCRLAY	US-NE	
108.	Use of 'back page' and 'back section' buttons is inappropriate. Same applies for forward button. I am confused about 'back page' and 'back section' buttons, because I would expect them to have a similar function. But sometimes the beginning of the topic is the level behind.	MMT-1	Minor	NAVIG	US-NE	U6-3
109.	Use of buttons needs serious reconsideration based on current WWW practices. The application takes no account of web browser conventions. Now we are used to having the navigation buttons at the top left-hand corner, and the menus to be on the left and the text to be on the right. And this is not taken into account.	MMT-1	Minor	SCRLAY	US-NE	

No	Problem Description	UEM - # of experts	Severity	Type	Diagnosed	User Problem Match - Strength
	Now I'm browsing the Internet and then coming into something like this is a bit of a problem.					
110.	Main menu page – bold text doesn't look like hypertext. Maybe if it was blue, because there is a convention on the Internet to use different colours.	MMT-1	Cosmetic	PRESNT	US-NE	
111.	The icons are outdated.	MMT-1	Cosmetic	MEDDESN	US-NE	
112.	Page 1 – graph icon is not too clear as to its function. Does it display an example?	MMT-1	Minor	MEDDESN	FS	U50-4
113.	Final instruction 'you can alter the scale of the graph …' is unclear,	MMT-1	Cosmetic	TASKINS	FS	U29-1
114.	Page 2, subsection headings – the characters in the titles seem to overlap slightly, giving the impression of a slightly unprofessional product.	MMT-1	Minor	MEDDES	US-NE	
115.	Page 2, leaflet - clicking on the number line is quite difficult – the clickable area is too small, however you can get used to it.	MMT-1	Cosmetic	MODEINT	FS	False alarm
116.	Navigation buttons are a bit confusing having no tool tips.	MMT-1	Minor	NAVIG	US-NE	U41-3 U6-2
117.	Page 4 & 5 – in the first line 'e ^{mx}' is not explained at all. The 'e ^{mx}-n' is not mentioned in the demo/test either.	MMT-1	Major	CONT	US-NE	U34-2
118.	Page 5, leaflet, page 2 – in the information window 'proof of relation' the formulas are confusing to look at as it is difficult to see which formula is on which line, or any line at all.	MMT-1	Major	MEDDES instructional	FS	
119.	Line spacing seems to be random, reducing flow.	MMT-1	Invalid	MEDDES	US-NE	
120.	Some of the equations could be better presented, possibly by using a sort of hypertext pop-up explanation of what is being displayed.	MMT-1	Minor	MEDDES	R.S.	
121.	Page1, demo - The graphs should be different colours, if one is red and the other one blue you know that you saw the red one first and then the blue. If I hadn't been concentrating hard, I wouldn't know which was the first of the 3. They could also be labelled, or there could be a key attached. If this is supposed to be for people who haven't seen these functions at all before, all those things could be important.	MMT-1	Minor	ATTENT	FS	C1-1
122.	Scaling leaflet - They can leave the reference graph bold, and the transformed one lighter, and they haven't done that here. Just for consistency it seems like a good idea.	MMT-1	Minor	ATTENT	US-NE	
123.	Page1, demo, final comment - I would have a paragraph break there, because there are two separate points they are making.	MMT-1	Cosmetic	MEDDESN	R.S.	
124.	Page 1, demo, changing of scale – Error message 'that is too large' is not understandable enough, a better message would be to say what's the upper bound, so I've got no idea how far I'm allowed to go. The next error message 'try saving this book..' is even more incomprehensible. The other point is that I asked for 2000 and it gave me 20, and the graph's disappeared. It would be better if it said what the range of allowed	MMT-1	Cosmetic	FEEDBK	FS	

No	Problem Description	UEM - # of experts	Severity	Type	Diagnosed	User Problem Match - Strength
	bounds were, but again it's just a time saving thing.					
125.	Reflection leaflet, p.1 - There is a minor ambiguity whether that little thing (-) is mathematics symbol, or a piece of punctuation in the text. It is obvious to me that it is a minus sign, but it's quite small and the gap between the minus sign and the x is no smaller than the gap between the minus and the text. It would solve the whole thing by putting every equation in a different font or colour, or in braces. This is exactly the thing which if I didn't know enough about these functions I would be irritated. Why do I have to worry about these ambiguities when I'm trying to understand mathematics?	MMT-1	Minor	CONT instructional	FS	
126.	Reflection leaflet, p.1 - There's two concepts jumbled together there. They are talking about reflection here, and it is mentioned with the minus, but a lot of attention is taken up with this translation, which isn't mentioned at all. They are both trivial. If you apply to a graph a minus, you reflect it in the horizontal axis, and if you add a constant to the argument, you shift it backwards and forward. But no one of those things should have precedence over the other. Those two ideas are both fairly basic, and somehow the heading is about reflection and half of the text is about translation. And the icon is about translation. It seems to me that there is a bit of distraction from the main point by talking about another point, which is equally trivial, or equally complex, or equally interesting. You can just take away that k and put it somewhere else.	MMT-1	Medium	CONT	US-NE	
127.	Vertical translation - They've just changed the notation from exp(x) to e^x without any comment. If I was a novice that would throw me completely. They haven't said that you can write exp() as e^x. There is a number e, they haven't said anywhere that e is a number, they can say e=2.7… and give the first 6 figures. This has to be said in the beginning, unless there is a hyperlink to another part. If this is for novices, it ought to say that there are two different notations, and what they are. Otherwise it is just asking for trouble, someone will look at that and ask what on earth is going on. The problem with this mismatch between notations is exasperated by having different notation in these two places (tests and text on left-hand side).	MMT-1	Major	CONT	FS	U34-3
128.	Horizontal translation - A minor point, a convention is to put a little bar |--> if you're saying that a functions maps the value of x on to the value of x+k.	MMT-1	Cosmetic	CONT	US-NE	C19-1
129.	Horizontal translation - exp(mx+k) should be exp (mx+km) That's a minor mistake. It should either say m(x+k) or mx+mk. They've got the brackets right in the tests. For consistency they should use the brackets.	MMT-1	Medium	CONT instructional	US-NE	
130.	Horizontal translation, demo/test - It so happens that the scale is such that the graphs are not very clear. Now I know that this can be cured by using bigger bounds or shifting the graph along, but that's only because of my knowledge of the function. It's OK to generate the equations randomly, but they should make sure that the first few have recognisable movement.	MMT-1	Minor	MEDINT	US-NE	U17-4 U21-4
131.	Horizontal translation, demo/test - There should have been a different icon to say we are now playing a quiz game.	MMT-1	Minor	MISLFUNC	R.S.	False alarm

No	Problem Description	UEM - # of experts	Severity	Type	Diagnosed	User Problem Match - Strength
132.	Page2 & Horizontal translation, demo/test - There's no cancel button! It's quite hard to cancel the test once started – there are no 'cancel' buttons. If you click a button accidentally, you cannot cancel the action.	MMT-2	Cosmetic	USECONT	US-NE	U28-4
133.	Vertical translation - I would say the shortest distance also becomes more and more horizontal. They could expand on that a bit. For someone who hasn't done that before, they can get lost with the use of pronouns. I think you can improve the wording there. And I would take out more and more horizontal. As an enhancement you could have a picture with two graphs and two vertical lines at different points. And a line here in red and another line here and that would be visible. And you can even have dotted lines to show the shortest distance. Basically a picture would probably work better that the text here.	MMT-1	Minor	CONT	US-NE	C17-1
134.	Vertical translation, demo - The resolution's not very good. The graph doesn't look like a curve here, it looks like three straight lines. And I guess if you are completely new to exponential graphs, you may think there's a bug in the software.	MMT-1	Minor	MEDDESN	FS	U49-2
135.	Scaling, demo/test - There is a general problem here. Graph are superimposed, and I guess that's why I can't see the other graph. And there should be some explanation somewhere that the exp function is a curve. I suppose I'm able to see what's happening by zooming in. But would a person who is unfamiliar with these functions know that? Perhaps there should be a hint to say – if the graph doesn't look like a curve, you can zoom in by changing the scale to understand the shape better. But it would be even better, if the scaling in this test is dynamically chosen for every set so you can get a sense of the shape. That can be improved again. In summary, the default bounds on the axis should be selected in such a way that you can get a sense of a curve. Otherwise people can be puzzled by the two lines. That's just a bit of programming, 30-40 lines of code that someone 3needs to do.	MMT-1	Minor	PRESNT	US-NE	U17-4
136.	Page 3, Vertical Scaling - There's a mistake in the first sentence. It says '… replaces x by ky', that x should be y. It's just a typo.	MMT-1	Medium	CONT	US-NE	C9-1
137.	Page 3, Scaling, Test - That comment was a bit confusing – 'next comes the vertical scaling represented by the second graph.	MMT-1	Invalid	USEINTR	US-NE	U19-4
138.	Page 5 - That's completely out of place. I've just been trough a number of pages of stuff where I needed to know that. And also it can be a bit more complete, there's no mention of e again. This needs to be put in the introduction, it could be hyperlinked here, but there's no point of telling me this here.	MMT-1	Major	CONT Instructional	US-NE	
139.	Page 5- That's a mistake. There shouldn't be a minus sign there. It should say log 1/x. That's the kind of a mistake that can make a beginner give up. It's quite serious	MMT-1	Major	CONT instructional	FS	
140.	Page 5, main leaflet, demo - The graphs have come and the equations have gone, and the two should be occurring together. Lack of synchronisation of text and graphs.	MMT-1	Minor	MEDINT	US-NE	
141.	Page 5, main leaflet –I should be looking at a picture as I'm reading this text, rather than read the text first.	MMT-1	Cosmetic	PRESNT	US-NE	

No	Problem Description	UEM - # of experts	Severity	Type	Diagnosed	User Problem Match - Strength
142.	Page 5, main leaflet, demo – The labels on graph lines are pretty illegible.	MMT-1	Minor	PRESNT	US-NE	False alarm
143.	Page 1, main leaflet - The exponential graphs 2, 10 and e are not explained. Something could be said about why they are important. You can take the exponential of any positive number, it doesn't have to be 2, 10 or e. It seems that a big part of the story is missing. 10 is important because it's the basis of our counting system, 2 is important because of computers. You can take the log to any number and it's the inverse of the power function.	MMT-1	Minor	CONT instructional	US-NE	
144.	Equivalence between vertical scaling and horizontal translation for x -> exp(x) is never mentioned or illustrated visually in the tool. This means a missed opportunity in the tool design. A geometric property of the exponential curve that has interest is never explicitly pointed out and which would require very little extra work. It cross-links two of the sections of the tool in an interesting way. To be more specific: y=exp(x+k) is a horizontal translation leftward by k units; y=c exp(x) is a vertical scaling by a factor c. if c=exp(k), e.g. k=3 and c=20.09, then these two apparently very different transformations produce exactly the same result. It would not be difficult to add text, a static graph or even another interactive demo to draw attention to this geometrical property, which is a basic defining feature of an exponential curve. The value of these is obvious. It provides an illustration, more generally, that mathematics is really all about studying these kinds of connections and relationships.	MMT-1	Medium	CONT instructional	US-NE	
145.	Pages 2 & 3 – After clicking on the 'show me' button no graphs are shown, which is confusing.	MMT-1	Minor	MISLFUNC	FS	U13-4

Appendix 4.12.2: Problems Predicted Using Experts' Own Judgement

No	Problem description	Type
	Problems Predicted by the Experts Who Applied the MMCW	
1.	Navigation buttons at bottom right hand corner have no tool tips. If I was to take the perspective of someone who doesn't know anything about computers, I would say they would get lost, especially since none of the buttons have any tool tips.	NAVIG
2.	No bound check on scales, it kind of crashes if you set a very high number.	MISSFUNC
3.	Page 1, demo - Expected to click on image for other graphs not on icon.	USEINPT
4.	Page 2 and Leaflet - It's not clear which graph line is for which function. Lines could be labelled.	MEDINTG
5.	Page 2, demo/test - Doesn't say which the second equation is.	MEDDES
6.	Page 2 – Buttons are confusing – sometimes they worked and sometimes not. Got lost in pages available.	NAVIG
7.	Page 2, leaflet – Expected graph line of *k* to be in the same window as the graph, rather than separate.	SCRLAY
8.	Page 2, demo/test – took while to start.	
9.	Page 3 – Test is slow when asked for factor value, there was no response.	
10.	Page 3 – not clear why there are two different leaflets on horizontal and vertical scaling.	CONT
11.	Page 3 – nesting of sets of pages is confusing. Perhaps there is too much nesting in these sets of pages.	NAVIG
12.	Page 3 - Confusion between demo in leaflet and the main demo.	CONT
13.	I would expect menus to be on the left-hand side.	SCRLAY
	Problems Predicted by the Experts Who Applied the IMMC	
14.	Page 1, demo pane – system asks user to 'click again for the graph of 2^x', but it's not clear where to click.	TASKINS
15.	Couldn't move main window around as it is centred. So it means that if I had the calculator or a glossary windows open I can't see them both in their entirety. And also it may have been nice to have both the glossary and the calculator up at the same time, but you can only have one of them. Can't have more than one tool or help window up at the same time.	SCRLAY
16.	Page 1 –function of graph icon not clearly labelled 'demo test'. On the second page the demo/test icon was labelled, but before that I didn't know whether that was an icon that led on to something else or whether that was a graphic.	MEDDESN
17.	Page 2 – The letter 'e' of reflection not properly drawn.	MEDDESN
18.	There was some sound used at some places, but only when you reached the bottom of a scroll box. So the system has the capabilities to use sound, but sound was not used at all anywhere else. So for example in the text section, it would have been nice to have some feedback in sound for a wrong or right answer, or just throughout.	MEDSEL
19.	Page 2 of Leaflet Reflection – you're able to click on the line number *k* to change the value of k without first having clicked on the demo icon to see the graph, which resulted from changing the value of *k*. So, you shouldn't be able to click on the line number *k* without having clicked on the demo icon beforehand.	USEINTR
20.	In the help section they used hypertext to enable you to go from one page to another, and they used the word continue for that hyperlink, but that word wasn't distinguished from any other word on the screen by use of colour or underlining.	PRESNT
21.	I started by thinking underlined text was hyperlinked. Too much web use?!	PRESNT
22.	From the Resource Browser couldn't select 'module map'.	MISSFUNC
23.	The bibliography didn't work.	MISSFUNC
24.	Minor point – glossary scrolling was too fast.	USECONT
25.	Page 2, Reflection - How does the leaflet differ from main application screen? It's not clear why there is extra information in a leaflet and not just in a follow on page.	CONT
26.	It took me a couple of clicks to work out how to get back from leaflet to main page.	NAVIG

No	Problem description	Type
27.	Horizontal scaling leaflet, p. 1 – when you change value of *k*, it redraws the sine curve, *x* gives written information at the top of the box, too – I didn't notice the information changing until about my 5th click to change *k*.	MEDINT
28.	The information pop-up boxes have headings that look like buttons.	PRESNT
29.	Page 5, Main leaflet, Page 2 – There are some bugs in the application. For instance, that says *–ln(1/x)* is equal to *-ln(x)*. And on this green information screen it says that *–ln(x)* is the same as *ln(1/x)*, so there shouldn't be a minus in the main screen. It's confusing.	CONT
	Problems Predicted by the Experts Who Applied the MMCW	
30.	Page 3, test - It's not clear which graph it is talking about.	
31.	It's quite easy to get lost, and lose track of where you are and where you have been.	NAVIG
32.	You can loose track, because you have no indication of a tree. The only indication is in the top bar and I missed it. They could have used tree information for navigation. Some of use browse information, and when we do that we want to have a point of reference to know where we are in a tree.	NAVIG
33.	System is quite slow.	
34.	In test pop-up dialogue box appeared almost immediately after a graph was drawn, which hides the graph before user has had the chance to read the graph. You could move the dialogue box away, but some of the users may not have known that you can do that.	SCRLAY
35.	Page 3 – it's a pity that you can't move the window around to see the whole screen. Main window is unmoveable, but as there is no print, copy and paste functions, this does not matter too much. For typing notes it would be a nightmare.	SCRLAY
36.	Layout of information screens is slightly different.	SCRLAY
37.	As the buttons look like windows standard buttons, I'd expect them to behave like the windows buttons, and its irritating when they don't (although you get used to it).	PRESNT
38.	The colour scheme is bad.	PRESNT
39.	At the begging of the module if you know where you want to go, you have to do 5 clicks to get to the screen you want. Sometimes you may not want to read the introduction of the module page.	NAVIG
40.	This is a bit confusing. I don't know what they like everybody to read sequentially, or to go from one screen to another. It depends on how the course is designed, but I would expect that the navigation buttons at the bottom right corner have no validity. For me on the Intro page if I click on the forward button would mean that I would take me to a deeper on that thread not in parallel introduction rather than reflection.	NAVIG
41.	Page 1 - Back button is blanked out. Back always have to be one as long as you start the first page, you have to have a back button. On all leaflets, page 1- Here I would expect to have the back button available	NAVIG
42.	On module menu screen - the three words are bold, now because of the Internet we expected it to be underlined. For me because the bold word was first in the paragraph I couldn't think it is a link. Possibly if they had it underlined it would have been better.	PRESNT
43.	Page 2, Leaflet page 1- I would expect the leaflet information to take over some the right-hand space not to loose my reference to the previous point. The bar on top could have said – exponential graph:reflection:examples. So maybe we should have the top-bar information in a more prominent space on the top left-hand corner. We take the top part for granted, we don't look at it a lot.	SCRLAY
44.	Scaling Leaflet – the demo button caption is slightly at different places on page 1 and page 2.	PRESNT
45.	The buttons on the right-hand side are usually highlighted by a line drawn around them, but possibly something else could have been used, e.g. a different colour.	PRESNT
46.	Page 7, Hyperbolic Graphs Leaflet – The information on the right-hand side is obviously a comment for the demo, but how can you give me the comment before you give me the graph. I would expect to have nothing on the right-hand side, to have the graph and then close the graph to see the information about the relation between *cos* and *sin*. Or to have the graph on the left and have the explanation beside. Here it looses the sequence.	SCRLAY

No	Problem description	Type
47.	Scaling Leaflet - There is something which hasn't been brought up here, which is very important about the sine waive, which is the vertical scaling. If you invert the vertical scaling it's equivalent to inverting the phase. So a negative number on the vertical scaling is equivalent to changing the scaling on the horizontal scaling. That's a point, which is very important about the nature of how sin waive works.	CONT
48.	Scaling pages - What scaling does to a function is more complicated than that. This concentrates on what it does in terms of expanding the scale or changing the height or looking at the offset how it goes through the zeros.	CONT
49.	Page 5, Main leaflet - All this can be explained graphically, but it's not. This is by looking at this slide drawn. This is about how log's develop, it's all to do with that, having a method for doing graphical multiplication of numbers. And some of the mathematics of the history of how log's develop off that. So this is just like another equation. What you don't see is the fact that log(a) is equal to log(b) equals that. And without something else coming and saying this is very important, these are very important relationships between the numbers. Crucially important aspects about what you want to teach engineering students, the whole of decibels is based on this.	MEDSEL
50.	We have all these relationships between the functions, and the block here shows what they are, but what's the next step? What is that for? What does it lead on to? I think a lot of effort is spent at the beginning looking into the geometry of the functions – translation, scaling, etc., which are all pertinent to the screen ? of any quadratics functions. And what happens is we end up having very little on exponential graphs, apart from this basic (page).	CONT
51.	We have identities coming up to cos, sin and tang in terms of exponential, but there's no much talk about that in terms of sine functions, they don't even mention about cos function. There's no identity sin square cos + sin square = one, there's no mention of that aspect. Basically this thing logs utility, logs in terms of being able to multiply, and you can show ratios of treat numbers in a graphical way, which brings the relevance of these functions very clearly. So the whole thing is talking about identities of log numbers in terms of log of 1 over x = -x, that's identity, which is much better shown using a slide draw, where you can see how it is. For a student who is poor in mathematics or of mixed ability, this is not going to be significant. But you actually making up the whole schema, you have to be making sure you are actually covering all interconnectedness. You are actually covering not just this material, you are actually interdicting the exponential functions.	CONT
52.	To bring out the graphical relationships between exponential functions, you really want to be talking about changing the skills from linear skills to exponential skills. That way you can linearise the equations, and show what you can actually do with these exponential and logs, how that can be useful in terms of working with arithmetic for start.	CONT
53.	And also these graphs can easily be made dynamic. It will be nice to put a couple of icons and put decrease y scale or increase the offset, and you can actually see that happen dynamically. Mathematic relationships are dynamic in a way, relations between variables is being able to change the value, so there's no much significance in terms of fixed values. So we are going to have a scaling factor, which is actually a variable. Here they are constants, but you want to know how people can change these things, we are talking about variables in terms of the use. So it will be nice to be able to play with them. So then students can learn intuitively what something looks like. If I can actually see a square wave pattern, intuitively at the back of my head comes up with fully transformed of that. My brain can do that because of lots of practice.	MEDSEL

Appendix 4.13: User Interaction Problems Experienced by the Users

No	Problem Description	No. of Students	Type	Severity	Predicted by
1.	Page 1, Demo - Student has difficulties identifying where they should click for the graph of 2^x.	2	AFFORD	Minor	MMCW – SMM1 MMT-SEDU3
2.	Page 1 - The difference between the three graphs on the demo and why 10^x is steeper that the other two graphs, is not explained.	1	CONTENT	Medium	
3.	Page 1 – Student reads the whole text, but skips the demo button.	1	MISSINTR	Major	MMT-SMM4
4.	Page 1, Demo - Text inside the demo pane was too fast to read.	1	ATTENTION	Medium	MMCW-SMM1,3 MMT-SMM4
5.	Reflection leaflet – Skipped demo, although read the part about changing the value of *k*. Student explained that they didn't notice the demo button.	1	MISSINTR	Major	MMT-SMM7
6.	Reflection leaflet, Page 2 – Difficulty going back to main reflection page. Clicked on 'back page' button, which took the student back to p.1 of leaflet not to main page. Tried again and went back to main page.	1	NAVIG	Minor	MMCW-SMM1 MMT-SMM7
7.	Reflection leaflet, Page 1 – Clicks on 'horizontal reflection' button in the right-hand pane to find out more about the horizontal reflection of exponential functions, but nothing happened. Student expected a graph.	1	MISLFUNCT	Minor	
8.	Reflection leaflet, Demo after being prompted – Student was unsure whether and how they can change the value of *k*, although a couple of minutes ago they said it they can. It is just a suggestion when the mouse goes down in this part maybe it'll be put 'change the value of k', because in a first time. Yes, it is written, but it's not obvious.	1	AFFORD	Cosmetic	MMT-SMM7
9.	Page 2& 3 – Skipped demo because student didn't look where the button is.	1	MISSINTR	Minor	MMCW-SMM1
10.	Page 2 & 4– Students skipped leaflets initially.	2	MISSINTR	Minor	MMCW-SMM1
11.	Scaling leaflet, Page 2, Demo – Student can see that the graph changes when changing the value of *k*, but is not sure why is that.	1	COMPR	Minor	
12.	Scaling leaflet, Page 2 – student had difficulties understanding the text description about k> or <1. It wasn't clear to them how the graph would change in both cases. Student didn't notice the text on the left-hand side prompting to do the demo. They suggested that the text can be made a different colour or a small hand can be displayed over the line of *k* to show that the values can be changed.	1	COMPR	Major	
13.	Page 2, Test – Clicked on the 'show me' button a few times to display where the graphs of the two equations are, but no graphs were shown.	2	MISLFUNCT	Minor	IMMC-SEDU4 MMT-SEDU2

No	Problem Description	No. of Students	Type	Severity	Predicted by
14.	Page 2, Test – If graphs for a pair of equations have already been displayed and the student starts the same test again, same messages are displayed, but the graphs display is not reset, i.e. after 'now comes the graph of the second equation' no new graphs are drawn and student cannot distinguish between the graphs on the screen.	1	MISLFUNC	Medium	IMMC-SEDU1,4
15.	Page 2, Test – Has difficulties understanding the last question asked – 'which is the reflected graph, the one below or one above'.	2	USEROUTP	Minor	
16.	Page 2, Test – Student initially has difficulties understanding the concepts of the two types of reflection. Next tests more are successful.	1	CONTCOMP	Medium	
17.	Page 3, Test – Two graphs are superimposed – original and horizontally scaled one. After displaying the scaling factor, there is a message that the system will display the third graph (vertically scaled one). But as the first graphs are one on top of each other, student wondered how to find the horizontally scaled graph.	1	MEDDESN	Medium	MMT-SEDU3
18.	Scaling leaflet – student is a bit confused that the demos use trigonometric functions rather than exponential functions, although they can still see how the graphs change.	1	CONTENT	Minor	
19.	Scaling, Test – Difficulties understanding what final-but-one message means – 'finally is the graph of the vertical scaling'.	1	USEROUTP	Medium	MMT-SEDU3
20.	Scaling, Test – At the end of a test, student tries to zoom in to see graph better, but it is not clear to them which graph is which as there is one graph in bold and two graphs presented with same line.	1	MEDDESN	Medium	MMT-SEDU3
21.	Scaling, Test – student notices that scale changes every time, but it would be nice if such a scale is chosen to let you see the graphs properly, as student may not be able to see which graph is which.	1	MEDDESN	Major	MMT-SEDU3
22.	Page 4, Test – Student initially did not understand the first question – ' please type the vertical distance between the two equations'.	1	USEROUTP	Minor	
23.	Translation leaflet, Page 1, Demo – Student wasn't sure what translation is after reading the text in leaflet demo.	1	CONTCOMP	Minor	
24.	Translation leaflet, Page 2 – After reading the information box on the right-hand pane, student clicked on the grey pane underneath it to bring it forward, but nothing happened. In students' college programs behave in a way so when you click on a window it comes forward. And student didn't see the 'close' button on the pane.	1	AFFORD	Cosmetic	MMT-SMM7
25.	Page 5 – Student wants to open the leaflet about horizontal translation, but the first page of the leaflet is about vertical translation and the student is confused why that happened.	1	CONTENT	Minor	
26.	Page 5 – Student was confused when viewing the leaflet about Vertical Translation. They had covered horizontal one as well, and now there is another page about horizontal translation.	1	CONTENT	Minor	
27.	Leaflet demos – Clicked on the demo button a few times expecting to see more demos.	2	MISLFUNCT	Minor	
28.	Tests – student wants to cancel the test in the middle, but cannot.	3	MISSFUNCT	Minor	MMT-SEDU3

No	Problem Description	No. of Students	Type	Severity	Predicted by
29.	Students found changing the scale confusing as they weren't sure what 'positive bounds' means. Maybe there can be a dialog to say what kind of scale would you like, or a hand on the number to show that you can perform an action. Or may be there should be a button on the side saying zoom in/out, because a student may not be sure what scale is better.	4	AFFORD	Minor	MMT-SMM4/7 MMCW-SMM3
30.	Page 5, Test – after changing the scale student wasn't sure what exactly can be achieved by changing the scale.	1	UNCLFUNC	Cosmetic	
31.	Page 5, Test – Student would like to see only the positive quadrant of the co-ordinate system, but that is not possible.	1	MISSFUNC	Minor	
32.	Page 5, Test – Student closed the test by accident and wanted to go back to what they had just done, but it was not possible, they had to start another test. In Internet Explorer the student is used to having the back button always there.	1	USERCONTR	Minor	
33.	Translation leaflet – Student skipped the information buttons on both pages.	1	MISSINTER	Minor	
34.	The number *e* is not explained anywhere.	1	CONTENT	Major	MMT-SEDU3 MMCW-SMM1 IMMC- SMM2
35.	Student found only half of the content understandable.	1	CONTCOMP	Medium	MMT-SEDU3
36.	Tests – When having difficulties getting a correct answer, students clicked on 'show me' button, but could not understand the right answer shown. The feedback from the tests did not help student find the correct answer, as it just said 'wrong'. Students prefer if the system can explain the answer. Answer usually becomes clearer from more examples than the text or feedback.	3	SCAFF	Medium	MMT-SMM7
37.	Student got confused about which axis is the horizontal and which is the vertical - small problem with the English language.	2	LANGUAGE	Cosmetic	
38.	Student found text and graphs coming up dynamically at the same time distracting.	1	ATTENTION	Medium	MMCW-SMM1,3 MMT-SMM4 IMMC- SMM2
39.	The goals were sometimes clear, but not always.	1	LEARNGL	Minor	IMMC-SEDU1,4, SMM2
40.	Pop-up messages regarding the test obscure the view of how the graphs are constructed. It would be better if they are displayed on the side rather than covering the graph.	2	SCRLAY	Medium	MMCW-SMM1 MMT-SEDU1 IMMC-SMM2

No	Problem Description	No. of Students	Type	Severity	Predicted by
41.	Student hadn't tried clicking on the tools button during the session, and they weren't sure what the button does.	1	MEDDESN	Cosmetic	MMT-SMM4/7
42.	On tests, when the graphs are going up or coming down maybe an arrow can be used to show which direction it is going.	1	MEDDESN	Cosmetic	
43.	Mathwise wasn't very very enjoyable, but if there were more pictures, it would have made it more enjoyable for younger people especially. Also include some more cheerful statements, like 'well done'.	1	ENGAG	Minor	IMMC-SMM2
44.	It wasn't interesting or fun, but it made the student understand the material.	1	ENGAG	Minor	IMMC-SEDU1,4
45.	It wasn't very challenging as it didn't set really difficult questions.	1	ENGAG	Medium	IMMC-SEDU1,4
46.	The colour of the pop-up messages could be changed to a brighter colour rather than grey, as student thought that computer had stopped. Other programs use same grey colour to show an error message.	1	PRESNT	Cosmetic	MMT-SMM4 IMMC- SMM2
47.	Leaflet demos – Students chose to close the demo pane, then change the value of k and then open demo to see the effect.	2	USERINTER	Cosmetic	
48.	Maybe some animations can be included. For example the application of exponential graphs, like the voice waves in electronics.	1	MEDDESN	Medium	
49.	The quality of the graphics is not very good. Graphs quality could be improved, in some graphs the lines were too thick.	2	MEDDESN	Minor	MMT-SEDU3
50.	The demo icon is more like a graph, not like a test, but it usually tells you the meaning in the label.	1	MEDDESN	Minor	MMT-SMM4 IMMC-SMM2
51.	The colours are a bit too light, and one cannot work with the application for a long time.	1	PRESNT	Minor	

Appendix 5.1: Cause-Breakdown-Outcome Descriptions of User Interaction Problems

Problem # 1:

Cause: task instruction is not clear as to where to click for the graph of 2^x.

Cognitive breakdown: user is confused as to where to click for the graph of 2^x.

Behavioural outcome: user tries to click on the graph pane and the graph lines a few times, without achieving the goal.

Performance outcome: most users needed external help to identify that they need to click on the icon to the left of the graph in order to complete the task. However, this confusion may have also distracted the students from achieving the main task of finding out the differences between the three functions and their graphs.

Problem # 2:

Cause: the difference between the three graphs on the demo and why 10^x is steeper that the other two graphs are not explained.

Cognitive breakdown: student has difficulties understanding why the graphs were different.

Behavioural outcome: the student redoes the demo as a result of that and reads the text again.

Performance outcome: Users spend more time on task. Two of the students had problems recognising which graph is which on the comprehension test.

Problem # 3:

Cause: possibly icon is not clear as to its function.

Cognitive breakdown: user is not sure what action icon affords and ignores it.

Behavioural outcome: Student reads the whole text, but skips the demo button.

Performance outcome: Student misses out important demo initially. They needed external help to accomplish task.

Problem # 4:

Cause: Text inside the demo pane was too fast to read.

Cognitive breakdown: user's attention is not drawn to text, and user may not associate that the graph is for the function of e^x.

Performance outcome: the student cannot label the graph of e^x correctly. Student got the 10^x wrong as well, as they swapped the two.

Problem # 5:

Cause: The demo button is placed in the bottom right hand corner and is not labelled sufficiently.

Cognitive breakdown: Student didn't notice the demo button.

Behavioural outcome: Student skipped the demo, although read the part about changing the value of *k*.

Performance outcome: Student needed external help to view demo.

Problem # 6:

Cause: The two buttons 'back page' and 'back to main section' in the Reflection leaflet are confusing as to their function.

Cognitive breakdown: The user was not sure which button to click on to go back to main Reflection page.

Behavioural outcome: Clicked on 'back page' button, which took the student back to p.1 of leaflet not to main page. Tried again and went back to main page.

Performance outcome: User recovered after a second attempt and managed to achieve their goal.

Problem # 7:

Cause: The 'horizontal reflection' button n the right-hand side is not clear to its function.

Cognitive breakdown: User wrongly assumes that by clicking on the 'horizontal reflection' button graphs or more information about Reflection will be displayed.

Behavioural outcome: Student clicks on 'horizontal reflection' button in the right-hand pane a couple of times to find out more about the horizontal reflection of exponential functions, but nothing happened.

Problem # 8:

Cause: The text explaining how to change the value of k is not clear and is not placed close enough to the value line. Furthermore, outside of the context of the demo, changing the value to k does not make much sense, but is presented before the users view the demo.

Cognitive breakdown: Student was unsure how to change the value of *k*.

Behavioural outcome: Student skipped the demo initially. He had to be prompted to view the demo and he needed external help to change the value of k for a first time, as he had not paid attention to the text when he read it.

Performance outcome: Student needed external help to complete the task.

Suggested solution: When the mouse goes down in this part maybe it'll be put 'change the value of k'.

Problem # 9:

Cause: The demo buttons are not highlighted enough, placed in an obscure place and are not labelled.

Cognitive breakdown: The student did not notice the demo buttons on pages 2 & 3.

Behavioural outcome: Student skipped the demos on pages 2& 3.

Performance outcome: Student needed external help to complete the task, otherwise they would have missed important information.

Problem # 10:

Cause: The leaflet buttons are not highlighted enough or clear as to their function.

Cognitive breakdown: The students did not notice the leaflet buttons on pages 2 & 4, or they were not sure what the buttons are for.

Behavioural outcome: Students skipped leaflets on pages 2 & 4 initially.

Performance outcome: Students needed external help to complete the task, otherwise they would have missed important information.

Problem # 11:

Cause: student read the explanation of moving k on the right hand side, but did not find it understandable enough – connected to prob # 12. Text and graphs are not linked together to aid student in linking the explanation about changing the graphs and the actual graphs. In fact the graphs cover the text after they are displayed.

Cognitive breakdown: Student can see that the graph changes when changing the value of *k*, but cannot understand why is that.

Behavioural outcome: student continued changing the value of k and looking at the corresponding graphs.

Performance outcome: User spends more time on task. The student had problem with scaling on the comprehension tests.

Problem # 12:

Cause: possible causes are that the text is not understandable and the text is not clearly linked to demo graphs.

Cognitive breakdown: student had difficulties understanding the text description about k> or <1. It wasn't clear to them how the graph would change in both cases. Also caused frustration to the user, they feel they are unable to understand the material.
Behavioural outcome: student reads the text a couple of times and then tries to view demos, but can't understand only from the text. However the graphs in the demos obscure the explanation text.

Problem # 13:

Cause: The 'show me' function does not work.

Cognitive breakdown: The user assumed that by clicking on the 'show me' button they the test would be done for them.

Behavioural outcome: In Page 2 on the test the student clicked on the 'show me' button a few times to display where the graphs of the two equations are, but no graphs were shown.

Performance outcome: User fails to achieve their goal.

Problem # 14:

Cause: The functions of the two buttons on the tests 'new test' and 'start' are not clear.

Cognitive breakdown: Student does not know that they had to press 'new test' button first before they can start a new test. They probably assumed that 'start' means reset and start a new test not just start.

Behavioural outcome: Page 2, Test – If graphs for a pair of equations have already been displayed and the student starts the same test again, the same messages are displayed, but the graphs display is not reset.

Problem # 15:

Cause: cause can either be that the actual question is worded ambiguously, or that the student did not understand the concepts of reflection which is needed to understand and answer the question. But student needs such understanding to answer the previous question, but the knowledge is different – one is equation transformation, which is represented in the same way in text explanation, and the other is visual/graphical transformation, which student has not seen before, and which could be the cause of this problem.

Cognitive breakdown: Student has difficulties understanding the last question asked – 'which is the reflected graph, the one below or one above'.

Behavioural outcome: Student re-reads question 2-3 times, moves the pop-up box next to graph pane and looks at graphs, but student gets initial answers wrong.

Performance outcome: Both students who experienced this problem seem to have overcome it and did well on the tests on reflection, and the problem did not have a bad effect on their comprehension of the material in the long term.

Problem # 16:

Cause: Probably the student did not understand the textual description on Reflection.

Cognitive breakdown: Page 2, Test – Student initially has difficulties understanding the concepts of the two types of reflection.

Behavioural outcome: Student gets answers to questions on the test wrong initially. He then tries more tests.

Performance outcome: Student recovers, and starts getting more answers right.

Problem # 17:

Cause: The two graphs are superimposed – original and horizontally scaled one. After displaying the scaling factor, there is a message that the system will display the third graph.

Cognitive breakdown: As the first graphs are one on top of each other and are not visible as separate graphs, student wondered how to find the horizontally scaled graph, which is hidden behind the original graph.

Behavioural outcome: Student looked a bit puzzled and then tried a new test, which was the same. On neither occasions did the student try to change the scale to see the graphs properly.

Performance outcome: Student missed seeing one of the graphs. Although the task appears to be completed successfully, student may not have understood the difference between the original and horizontally scaled graph. Consecutively, the student did badly on scaling on the comprehension test.

Problem # 18:

Cause: The material in the leaflet is about a different type of function than that in the main section.

Cognitive breakdown: In the Scaling leaflet the student has difficulties understanding why the demos use trigonometric functions rather than exponential functions, although they can still see how the graphs change.

Problem # 19:

Cognitive breakdown: In test on Scaling student has difficulties understanding what the final-but-one message means – 'finally is the graph of the vertical scaling'.

Behavioural outcome: Student re-reads question 2-3 times, moves the pop-up box next to graph pane and looks at graphs, but student gets initial answers wrong.

Problem # 20:

Cause: At the end of a test, student tries to zoom in to see graph better, but one graph in bold and there are two other graphs presented with the same line none of which is labelled.

Cognitive breakdown: It is not clear to the user which graph corresponds to which equation. User is frustrated.

Performance outcome: As these demo/tests are supposed to demonstrate to the students how graphs are transformed horizontally and then vertically, such ambiguous presentations fail to do so, thus the principles of transforming graphs may not be clear to the students, which is supported by the comprehension test results.

Problem # 21:

Cause: The scale chosen is inappropriate for the example.

Cognitive breakdown: On the test on Scaling the scale of the graphs is such that the student is unable to see which graph is which.

Problem # 22:

Cause: Perhaps the student did not understand the concepts of vertical intercept explained in the textual description.

Cognitive breakdown: In the test on Page 4 the student initially did not understand the first question – ' please type the vertical distance between the two equations'.

Problem # 23:

Cause: Translation leaflet, Page 1, Demo –student reads text, but text is not understandable.

Cognitive breakdown: Student could not understand what translation is after reading the text. Student developed only partial understanding of the material.

Behavioural outcome: Student re-reads text a couple of times and tries demos.

Performance outcome: Student did better on recognising the equations of vertical translation than on transforming the graphs.

Problem # 24:

Cause: User prior knowledge of using computer systems is different from the design.

Cognitive breakdown: Student assumed that by clicking on the grey area around the pop-up box it will be closed, as in student's college programs behave in a way so when you click on a window it comes forward.

Behavioural outcome: Translation leaflet, Page 2 – After reading the information box on the right-hand pane, student clicked on the grey pane underneath it to bring it forward, but nothing happened. The student then clicked on the 'close' button.

Performance outcome: Student recovered and accomplished task successfully after second attempt.

Problem # 25:

Cause: The first page of the leaflet about horizontal translation is actually about vertical translation.

Cognitive breakdown: Student wants to open the leaflet about horizontal translation, but is confused about why the information presented in about vertical translation.

Behavioural outcome: Student clicks on 'horizontal translation' button to display leaflet about it.

Performance outcome: Student recovers and achieves goal.

Problem # 26:

Cause: The leaflets about horizontal and vertical translation are put together in one although the main pages are separate.

Cognitive breakdown: Page 5 – Student was confused when viewing the leaflet about Vertical Translation. They had covered horizontal one as well, and now there is another page about horizontal translation.

Behavioural outcome: Student closed leaflets and went back to main page.

Problem # 27:

Cause: Leaflet demos – Student wanted to see more demos and clicked on the demo button a few times to achieve that rather than on the line of k. Maybe they expected that because it was similar to selecting a new demo on the main screens or is just more intuitive.

Cognitive breakdown: User assumed that by clicking on the demo button new demos will appear, but none were appearing on the screen.

Behavioural outcome: User kept clicking on the demo button. And was perhaps a bit confused.

Performance outcome: to continue with the task the user needed external help to show them how to get more demos.

Problem # 28:

Cause: Tests – student wants to cancel the test in the middle, but there is no cancel button available.

Cognitive breakdown: User cannot identify the operator which will enable them to cancel the test.

Behavioural outcome: User is confused and continues with test.

Performance outcome: User does not achieve the goal of cancelling the test.

Problem # 29:

Cause: The terms used in the message prompting to change the scale of the graphs were ambiguous to all students.

Cognitive breakdown: Students found the message about changing the scale confusing as they weren't sure what 'positive bounds' means.

Behavioural outcome: One student re-read the instructions, then started clicking on the positive side of x axis, but nothing happened. Another student clicked on the axis, but again nothing happened. Students needed external help to change the scale.

Performance outcome: Not being able to see graphs properly as default scale is not always appropriate can form wrong understanding or incomplete view of graphs.

Suggested solution: Maybe there can be a dialog to say what kind of scale would you like, or a hand on the number to show that you can perform an action. Or may be there should be a button on the side saying zoom in/out, because a student may not be sure what scale is better.

Problem # 30:

Cause: The option for changing the scale is not supported adequately as the students were not sure how to use it and what scale is better. Students probably did not have experience with plotting graphs.

Breakdown: Page 5, Test – after changing the scale, student wasn't sure what exactly can be achieved by changing the scale. Student does not know what scale is better and therefore the result of the action was not clear to them.

Behavioural outcome: Student experimented with different scales, often making it worse to view. Student suggests that zoom in and out would have been easier to understand.

Performance outcome: students not being able to set adequate bounds, as a result of which they won't be able to see graphs properly and how they relates to each other, ultimately result in lack of understanding of transformations.

Problem # 31:

Cause: A function to zoom only in one quadrant is not available.

Cognitive breakdown: Page 5, Test – Student would like to see only the positive quadrant of the co-ordinate system, but that is not possible.

Performance outcome: The intended goal was not achieved.

Problem # 32:

Cause: A function to go back to where test was left is not available.

Cognitive breakdown: Page 5, Test – Student closed the test by accident and wanted to go back to what they had just done, but it was not possible. In Internet Explorer the student is used to having the back button always there.

Behavioural outcome: Student started another test.

Problem # 33:

Cause: The information buttons are not highlighted enough or clear as to their function.

Cognitive breakdown: The students did not notice the information buttons, or they were not sure what the buttons are for.

Behavioural outcome: In horizontal translation leaflet the student skipped the information buttons.

Performance outcome: Student needed external help to complete the task, otherwise they would have missed important information.

Problem # 34:

Cause: The number e is not explained anywhere.

Performance outcome: Lack of understanding about fundamental principle of exponential functions.

Problem # 35:

Cognitive breakdown: Student found only half of the content understandable.

Performance outcome: Low learning performance on comprehension test.

Problem # 36:

Cause: When having difficulties getting a correct answer, students clicked on 'show me' button to get the right answer, but the feedback from the tests did not help student find the correct answer, as it just said 'wrong'.

Cognitive breakdown: The students could not understand the right answer shown.

Behavioural outcome: Students reflected on the results a bit and did more tests.

Performance outcome: The students may have none or formed only partial understanding of the material which is not supported by the feedback provided, thus the task of understanding the graph transformations is not aided sufficiently and is not successfully completed. However, students usually reported that further tests are clearer so maybe they are made to reflect and find explanations of the answers themselves.

Problem # 37:

Cause: Problem with the English language.

Cognitive breakdown: Student got confused about which axis is the horizontal and which is the vertical.

Performance outcome: Possible effect on learning as all types of transformation use the terms.

Problem # 38:

Cause: in the leaflet demos text and graphs sometimes change dynamically at the same time, which competes for users' attention.

Cognitive breakdown: Student's attention was distracted by text and graphs coming up dynamically at the same time, and text usually gets ignored.

Performance outcome: student misses seeing relevant formulas and explanations.

Problem # 39:

Cognitive breakdown: The goals were sometimes clear, but not always.

Problem # 40:

Cause: Pop-up messages regarding the test are displayed over the graphs they and obscure their view.

Cognitive breakdown: Students is unable to see part of the graphs as they are constructed as they are obscured by the pop-up messages.

Behavioural outcome: Student moves pop-up box away.

Suggested solution: It would be better if they are displayed on the side rather than covering the graph.

Problem # 41:

Cause: The function of the tools button is unclear.

Cognitive breakdown: Student is unsure as to what the tools button does.

Behavioural outcome: Student does not click on the button and does not go to see the tools available.

Performance outcome: Student missed interacting with support tools, particularly an opportunity to plot their own graphs.

Problems #42 and 48 are redesign suggestions only, problems #43, 44 ,45 are affective cognitive breakdowns.

Problem # 46:

Cause: Pop-up messages look like Windows error messages, as they are in grey.

Cognitive breakdown: The student thought that the computer had stopped.

Suggested solution: The colour of the pop-up messages could be changed to a brighter colour rather than grey.

Problem # 47:

Cause: The instructions as to how to view graphs and change the value of k are not clear.

Cognitive breakdown: Probably because the line of k is outside the demo pane and the students sees it first he assumed that he needs to have the graph window closed before he can change the value of k.

Behavioural outcome: Leaflet demos – Students chose to close the demo pane, then change the value of *k* and then open demo to see the effect. Student did that repeatedly.

Performance outcome: Task is achieved successfully but inefficiently.

Problem # 49:

Cause: The quality of the graphics is not very good. In some graphs the lines were too thick.

Problem # 50:

Cause: The demo icon is more like a graph, not like a test, but it usually tells you the meaning in the label.

Cognitive breakdown: Student was initially unsure what action the icon affords.

Behavioural outcome: Student clicked on icon.

Problem # 51:

Cause: The colours are a bit too light.

Cognitive breakdown: User cannot work with the application for a long time.

Appendix 5.2: Cause-Misconception-Outcome Descriptions of Comprehension Problems

Problem # 1:

Context: The two students had no prior knowledge of exponential graphs before using the application.

Cause: Graph lines are not labelled on the demo on p.1. Some of the text explaining which graph is being displayed is changed too quickly. (Also possible is that the principles of translating exponential equations into graphical representations are not well explained.)

Misconception: (Students either did not pay enough attention to which graph refers to which equation, or they lacked sufficient understanding of the principles of how exponential functions are translated into graphs, and the connection between algebraic and graphical representations.)

Comprehension Outcome: Two of the students had difficulty recognising the difference between the three basic exponential graphs, e^x, 2^x, 10^x.

Problem # 2:

Context: Students did not have sufficient knowledge of the principles of reflection.

Cause: Graph lines of the demo/tests in Refection section are not labelled adequately. The connection between exponential equations and their graphs are not explained adequately. Answers to test questions are not highlighted sufficiently.

Misconception: (Students possibly failed to understand how equations get plotted into their respective graphs. They also may not have noticed the answers to the questions given to them.)

Comprehension Outcome: Two students had difficulties translating the equation of an exponential function into its vertically reflected version while given the original equation and the resulting graphs of both the original and reflected equations.

Problem # 3

Context: (During the tests the students are asked to recognise the type of scaling from the equations, however it depends on how many tests the students did for scaling.)

Cause: Grpah lines are not labelled adequately. Connection between textual explanation of both types of scaling and the graphical representations during tests is not direct or obvious. The differences between the two types of scaling are not explicitly explained.

Misconception: (The students could not grasp how exponential equations are translated into scaled equations. The students also may not have paid enough attention to the differences between the two types of scaling.)

Comprehension Outcome: All 4 students had difficulties recognising the type of horizontal scaling from its equation. One of them recognised that it was scaling, but swapped the two types - horizontal and vertical.

Problem # 4

Context: Students prior knowledge of graph extrapolation could have been insufficient.

Cause: Graphs are plotted automatically for the students and they are not asked to practice that skill.

Misconception: (Task requires a good grasp of exponential graphs, scaling, and plotting graphs is general.)

Comprehension Outcome: Two students had problems extrapolating horizontally scaled graphs from their equations.

Problem # 5 (as #3)

Comprehension Outcome: Two students had problems recognising the type of vertical scaling from its equation.

Problem # 6

Context: Students did not have considerable prior knowledge of how graphs are scaled vertically.

Cause: The difference in the position of the vertical intercept is only explained in the text and not highlighted on the graphs.

Misconception: (Students were confused between the two types of scaling - vertical and horizontal, as they appear similar, and also did not pay attention to the differences in the vertical intercept in both types. Students also confused the principles of vertical translation and scaling.)

Comprehension Outcome: Three students had difficulties recognising the types of vertical scaling from looking at the graph of a vertically scaled exponential graph. One student recognised the type - scaling but not vertical, and another recognised that it was a vertical transformation, but didn't recognise the type - scaling.

Problem # 7

Context: The principles of this type of reflection are difficult as it actually combines two types of transformation, and assume good understanding of the principles of these types of transformation, which students may not have had.

Cause: Principles of horizontal negative scaling are only briefly explained in a text format in a small pop-up box.

Misconception: (Students could not grasp the concepts of translating equations of horizontal scaling with a negative factor.)

Comprehension Outcome: No student could recognise the equation of horizontal scaling with a negative factor.

Problem # 8

Context: The principles of this type of reflection are difficult as it actually combines two types of transformation, and assume good understanding of the principles of these types of transformation, which students may not have had.

Cause: The principles of horizontal negative scaling are only briefly explained in a text format in a small pop-up box. No graphical representation of it is presented.

Misconception: (Students could not grasp the concepts of translating exponential graphs into horizontal scaled ones with a negative factor.)

Comprehension Outcome: Three students could not recognise the graph of horizontal scaling with a negative factor.

Problem # 9 (as #7)

Comprehension Outcome: Two students could not recognise the equation of vertical scaling with a negative factor.

Problem # 10 (as # 8)

Comprehension Outcome: Three students could not recognise the graph of vertical scaling with a negative factor.

Problem # 11

Context: Students' did not have any prior knowledge of translation of equations.

Cause: The notation of representing exponential equations is different to the one used in the previous pages. Textual explanation of the transformation of the questions is not clear or sufficient.

Misconception: (Students could not grasp the concepts of translating exponential equations into vertically translated ones.)

Comprehension Outcome: Two students could not recognise the type of vertical translation from its equation.

Problem # 12

Context: Students' did not have any prior knowledge of translation of graphs.

Cause: Graph lines are not labelled adequately. The explanation of how graphs are moved when horizontally scaled may not be clear or sufficient. Not enough links between textual description and graphical representation. Answers to test questions are not highlighted sufficiently.

Misconception: (Students could not grasp the concepts of translating exponential graphs into vertically translated ones with a negative factor.)

Comprehension Outcome: Three students could not recognise the graph of vertical translation.

Problem # 13

Context: Students did not have sufficient prior knowledge of horizontal translation.

Cause: The calculation of the distance in not explained in this section, it is only given in text format in the previous page, which is regarding vertical translation, although the principles applies to both types of translation. Graphs are sometimes plotted is such a way that they overlap, i.e. making it hard to perceive the actual distance. The concept of the distance between graphs is not highlighted.

Misconception: (Students could not grasp the principle of calculating distance between original and horizontally translated graphs.)

Comprehension Outcome: Three students could not calculate the vertical distance between the graphs.

Appendix 6.1.1: Educational Multimedia Cognitive Walkthrough Method
Usability Specialist Version

The evaluation method described in this document aims to aid you, the evaluator, in analysing how effective multimedia presentations are in supporting students to learn while using educational multimedia software. The method comprises three steps of evaluation: evaluation of attentional design, evaluation of media integration and evaluation of individual media design. A set of questions and supporting design principles are provided in each step to help you identify usability problems. The method also specifies a procedure for identifying and describing usability problems.

General Guidelines for Learning with Multimedia Software

Using a combination of two or more media is generally more effective in supporting students to learn concepts than using only one medium. Particularly effective is combining visual (still images or animations) and verbal (text or speech) media, where the verbal medium is used to reinforce or explain the message presented in the visual medium. For example, in Figure 1 the water cycle is described by showing an image with all nature elements involved, and giving a more detailed description of how the cycle works in the text. You can also play a short animation with voice narration. The text alone cannot explain the process sufficiently well.

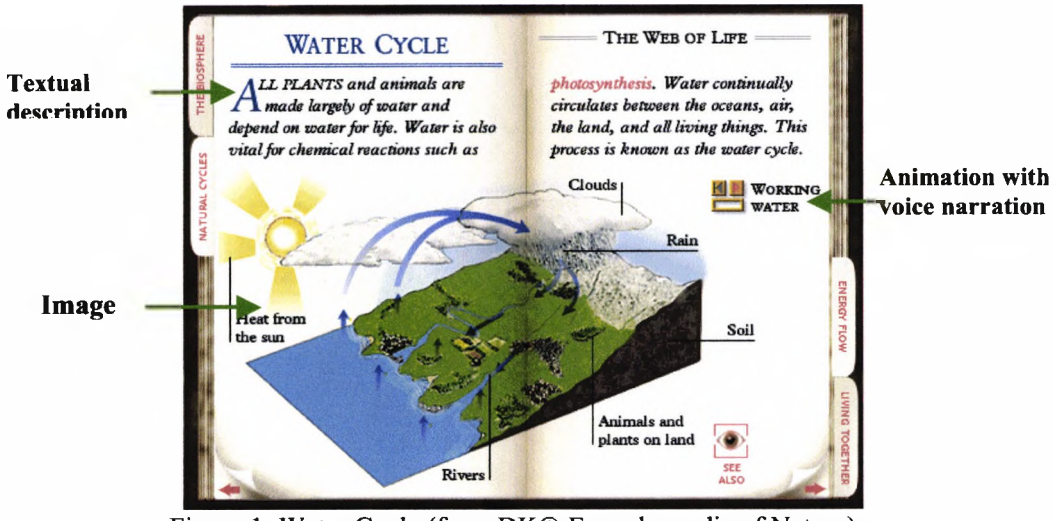


Figure 1: Water Cycle (from DK© Encyclopaedia of Nature)

Effective multimedia presentations need to ensure that:

1) The user will first **focus on** important concepts and details presented either in the text or image. This is important to ensure that the user will read and view all important information and will not miss out essential details in the presentation. To ensure that, different highlighting techniques, such as presenting text in a different colour or in bold, can be used to draw user's attention to important words or concepts.

2) Secondly, when concepts are presented in two different media, the user will need to **integrate** the information presented in both visual and verbal media. If a number of

different media are used to describe the subject matter, the user can easily get confused or loose the thread in the message if the media is not integrated well. Imagine if different terminology is used in the text and on the image, or in the voice narration. To avoid confusing the users, they will need to clearly understand the relationship between corresponding information in the image and the text. 'Contact points' or links can be designed to point the user to corresponding information in the text and the image. Contact points may include using the same label designating an object both in an image and in text, or allowing the user to click on a word in the text, which will highlight the related section in an image.

3) Finally, only if the first two processes have been successful the user will be able to **understand** the concepts being presented in the media. How much the student will understand will also depend on how well each individual medium has been design, and on much the user already knows about the concept being taught. For example, a child who has just started learning about nature and natural cycles may find it difficult to understand the principle of the water cycle if the nature elements were not labelled on the image and the blue arrows were not added to represent a cycle.

Therefore, successful learning using educational multimedia software depends on how well the media design supports student's in attending to, integrating and understanding important information presented in each medium. If the user is hindered in any of the process they may fail to understand the material.

Evaluation Method Overview

To help you assess how effective a multimedia presentation is in supporting these three processes of learning and in identifying problems in the design, the following three steps were formulated.

Step 1: **Evaluation of Attentional Design:**
In this section you need to explore which media will be **attended to** first and whether the user's attention is drawn to important information so they do not miss any important details of the presentation.

Step 2: **Evaluation of Media Integration:**
In this section you need to assess how well the visual and verbal media are combined together to allow the user to **integrate** the information presented in them.

Step 3: **Evaluation of Individual Media Design:**
In this section you need to evaluate whether the design of individual media makes it easier or harder for the user to **comprehend** the information presented.

Each step contains a set of questions, which aim to aid you in analysing the effectiveness of different aspects of the multimedia presentation and to predict likely problems users may experience while learning with the presentation. The questions have two parts. The first part points you to aspects of the design, which can potentially cause problems to the user, and the second part enables you to identifying what effect the problem may have on the user's cognitive processes. For some questions the second part is presented in a separate question indented underneath the original question. The questions are complemented with brief design guidelines to provide you with more explanation on basic principles of media design and integration.

Using the questions, you need to identify potential usability problems and determine how they may hinder user's learning processes and their performance in achieving their task.

To achieve best results, it is suggested that you analyse each individual section of the presentation following the three steps, before going to the next section. For each section assess the effectiveness of the attentional design, media integration and the media design before assessing the next section. This is aimed to ensure complete analysis of the multimedia user interface, tracking user's cognitive processes from attending to information through to comprehending the information presented in the media resources.

Specifying Usability Problems

In order to help the developers of the application to improve the design of the presentation, you need to describe the problems as fully as you can. This includes describing what the problem is, specifying its cause and where in the presentation it occurs, specifying the effect of the problem on the user's cognitive processes of attention, information integration or comprehension. The questions in each step should help you in determining these effects. After that, you need to describe the likely *actions* the users can take as a result of the problem and the likely *outcome on the user performance* and the achievement of their task.

The *cause* of the usability problem is usually the design feature that causes it. The design feature has an *effect* on the user, e.g. user's attention may not be drawn to an important concept, or s/he may have difficulties linking information in the text and the image, or understanding part of the text.

As a result of the problem the users can take different *actions*. For example, the user can press the wrong button, or s/he may repeat an action in an attempt to overcome the problem, or the user may give up. When deciding what the likely user actions might be, consider how critical the effect is and what the user's response may be, i.e. what s/he would do after experiencing such a problem. What the user may do will also depend on the options available in the interface and how familiar s/he is with the system.

Performance outcomes are the effects on the user's work and the achievement of their tasks resulting from the problem. Some example performance outcomes are that the user may spend a long time on the task, or s/he may fail to achieve their task at all (e.g. will not be able to understand the concept, or will only gain partial understanding of the concept), or if the problem is minor the user will be able to continue with the task and successfully achieve it.

Summary of the Procedure for Evaluation:

For each section of the presentation:

1. **Review the media design following the three steps and answer the relevant questions in each step.**
2. **Identify likely problems users may experience.**

For each problem you identify:

3. **Specify its *cause* and *location*.**
4. **Specify the likely *effect* it may have on the user's attention, information integration or comprehension.**
5. **Describe the likely *actions* the user may perform as a result of the problem.**
6. **Describe the likely *effect on the user's performance* the problem may have.**

STEP 1 : Evaluation of Attentional Design

The aim of this section is to identify potential problems in directing user's attention to important facts and concepts in the presentation. The main aspects to consider concern whether the user's attention is attracted to information which is important for their learning, and whether highlighting techniques are used effectively to guide the user to view important facts and concepts. This is vital to ensure that the user focuses on important concepts and does not miss essential details in the presentation.

Attention to Text

In this section you need to evaluate whether the design of text will direct the user to read and concentrate on parts of the text which are essential for them to learn.

Evaluation Questions:

1.1 Are important words, phrases or concepts highlighted to ensure the user will pay particular attention to them?

1.2 Are there important words and concepts which are not highlighted at all?

1.3 If so, do you think that the user may fail to read them?

1.4 Are there any parts of the text which are highlighted but are not important for the user's task?

1.5 Is that likely to divert user's attention from more important parts in the text?

> *Design principle: In order to ensure that the user will extract all important information their attention needs to be drawn explicitly to key words and phrases. Effects which can be used to make particular words or phrases stand out include bold, large fonts or underlining.*

1.6 If there are other media resources, such as still images, presented at the same time as important text, is the image likely to take user's attention away from the text before s/he has finished reading it?

> *Design principle: Because attention is attracted to size, detail and colour, generally an image will be focused on before a text. If focus is required on the text prior to the image, the text should be displayed before the image or make the text area larger than the image.*

1.7 Is there any event in the user interface, which is likely to interrupt the user while s/he is reading a piece of text? Such an event may include displaying an image or the onset of an animation.

1.8 Is that event likely to distract the user from finishing reading the text?

> *Design principle: When users are reading text their focus can easily be shifted to another medium, thus showing an animation or changing the image should be avoided at that time.*

1.9 If text is displayed dynamically, does the rate of changing it allow the user sufficient time to read the whole text?

> *Design principle: Reading text is slower than identifying an image. As a rule simple words require 200msec each to read, and at least 6 seconds should be allowed for two lines of text, and more if the scene is complex for the users.*

Attention to Still Images

In this section you need to evaluate whether the design of still images will allow the users to focus on important objects within the image. You also need to identify potential problems in the design likely to distract user's attention from inspecting important parts of the image.

Evaluation Questions:

2.1 Are sufficient highlighting techniques (e.g. use of bright colour, arrows pointing to an object) used to emphasise objects which are important for the user's to learn about?

2.2 Are the highlighting techniques sufficient to attract user's attention to such important objects?

2.3 Are there any objects which are particularly important for the user's tasks that are not highlighted or are left outside the main scene?

2.4 Is it likely that the user may fail to view these objects?

2.5 Are highlighting techniques used to emphasise any unimportant objects or minor details in the image?

2.6 If so, is this likely to distract the user from the more important objects or parts of the presentation?

> *Design principle: It is important to highlight information important for the user's task in an image to ensure they do not miss important concepts or details. Important objects can be highlighted by presenting them in bright colour, by setting them apart from other objects, making them larger in size, show more detail or moving nearer to the front of the scene.*

2.7 If a number of objects are to be displayed in an image, is each object revealed gradually, one at a time, or are they revealed all at once?

2.8 If they are revealed at the same time, are the objects going to compete for the user's attention?

2.9 If objects are displayed gradually in turn, is the user given sufficient time to inspect each object, or are the objects displayed too quickly?

2.10 If they are displayed too quickly, is this likely to confuse the user as to which object to view first?

Design principle: To control the order in which the user will view information it is important to reveal objects, labels and symbols gradually, one at a time, rather than all at once. Because attention is attracted to change, revealing each part of the presentation in turn is a useful way of directing order.

STEP 2: Evaluation of Media Integration

The aim of this section is to identify any problems users may experience in integrating the concepts presented in different media, such as text and image. Important aspects to consider include whether sufficient *contact points* are established between related information in both media to help the user integrate the information into a single concept. Remember that successful learning **will only occur** if the students are able to build connections between corresponding aspects of the visual and the verbal representations. If they are unable to integrate these aspects, this will only cause them confusion and will result in partial or miss-understanding of the material.

Evaluation Questions:

4.1 Considering how much the user already knows about the material, is the scene set sufficiently in the text before the user views the image?
4.2 Given the level of detail presented in the text, will the user be able to understand the concept presented in the image after having read the text?

4.3 Are important objects, details or concepts depicted in the image described sufficiently in the text?
4.4 Considering the user's prior knowledge of the concepts, are they able to understand the particular objects or concepts depicted in the image?

> *Design principle: If text is being used to set the scene, the text should be presented before visual media. It is also important that the correct level of detail is provided, i.e. the whole scene needs to be described sufficiently. Detailed textual descriptions of images are essential to help students with little prior knowledge of the material to understand what is depicted in the image.*

4.5 Does the information in the text direct the user's attention to important parts of the image?
4.6 Is that sufficient to ensure the user will not miss seeing them?

> *Design principle: When language is combined with visual media, language should be used to direct the user's attention to information within the image. For example, direct attention to the image by placing a message "look at the blue arrow in figure 1" in the text.*

4.7 If there is information about an important concept presented in an image and described in text at the same time, are there contact points established between relevant pieces of information in each medium?
4.8 Will the user be able to perceive the connection between the relevant aspects in each medium and is s/he able to integrate the related information?

4.9 If there is information about a concept which is unfamiliar to the user, are there contact points established between related concepts in the presentation?

4.10 Will the user be able to build associations between related concepts?

Design principle: Use explicit contact points if the connection between information in an image and language is important. For example, direct user's attention to an object in an image by highlighting the object which is being spoken about, or reveal a text caption and an arrow pointing to the object.

STEP 3 : Evaluation of Individual Media Design

After you have assessed whether the user will be able to focus on important information and whether s/he will be able to integrate the information presented in different media, in the final stage you need to evaluate whether the design of individual media makes it easier or harder for the user to comprehend the information presented.

The main consideration at this stage is to ensure that the subject matter content is presented in the most efficient way to the user to aid their comprehension.

Comprehension of Text

In this section you need to evaluate whether the textual messages and descriptions are easy to comprehend.

Evaluation Questions:

5.1 Are the messages clear enough for the user to understand, considering hers/his familiarity with the subject matter and their language abilities?

5.2 Is there any terminology used that is not familiar to the user, again considering hers/his familiarity with the subject matter and their language abilities?

5.3 Are the concept descriptions at the right level for the users to understand, given their prior knowledge of the material?

> Design principle: The level of the content should be suitable for the intended users – it should be higher but not too high above their current level of knowledge. Less knowledgeable students will need more explanation and greater details of the concepts being taught and of related concepts.

5.4 Does the text structure reflect the structure of concepts in the domain?

5.5 Is the concept structure clearly indicated so that even unfamiliar users/students can understand it?

5.6 Is the order of concepts indicated clearly?

5.7 Is the order of concepts clearly indicated so that even unfamiliar users/students can understand it?

> Design principle: Clear content structure is vital to aid students, especially those who are not familiar with the material, to understand the correct structure of events and concepts. Good design of content structure should specify the different parts of the content and their relationships. This can be aided by dividing the content into sections and sub-sections, and by using bullet points and numbered lists.

Comprehension of Still Images

In this section you need to assess how clear and easy to understand the still images are.

Evaluation Questions:

6.1 Are the objects in the image labelled appropriately to aid the user in identifying the objects?

6.2 Are labels placed next to the objects they identify to help the user in identifying them?

> *Design principle: Labels of objects improve objects' identification, thus labels should be placed next to the objects they identify.*

6.3 Is the design of the icons clear for the user to understand what they represent, considering the user's prior experience with the application and other computer systems?

6.4 Can the user distinguish the icons from other symbols presented on the screen?

> *Design principle: Well designed icons should be: easily associated with the message they represent, distinguishable from other symbols, not overly complex, suitable for different cultures, and in accordance with international or accepted standards.*

6.5 If motion information is shown in a still image, are the start and end points of the motion designated clearly?

6.6 Are any changes of direction clearly represented on the images to aid the user in imagining what the motion is or what the changes resulting from the action are?

> *Design principle: For motion information presented in still images it is important that breakpoints are shown, e.g. start and end point of the motion, and any changes in direction or speed are shown within the image sequence for the motion to be identified.*

Appendix 6.1.2: Educational Multimedia Cognitive Walkthrough Method
Instructional Specialist Version

The evaluation method described in this document aims to aid you, the evaluator, in analysing how effective multimedia presentations are in supporting students to learn while using educational multimedia software. The method comprises three steps of evaluation: evaluation of attentional design, evaluation of media integration and evaluation of individual media design. A set of questions and supporting design principles are provided in each step to help you identify design problems. The method also specifies a procedure for identifying and describing design problems.

General Guidelines for Learning with Multimedia Software

Using a combination of two or more media is generally more effective in supporting students to learn concepts than using only one medium. Particularly effective is combining visual (still images or animations) and verbal (text or speech) media, where the verbal medium is used to reinforce or explain the message presented in the visual medium. For example, in Figure 1 the water cycle is described by showing an image with all nature elements involved, and giving a more detailed description of how the cycle works in the text. You can also play a short animation with voice narration. The text alone cannot explain the process sufficiently well.

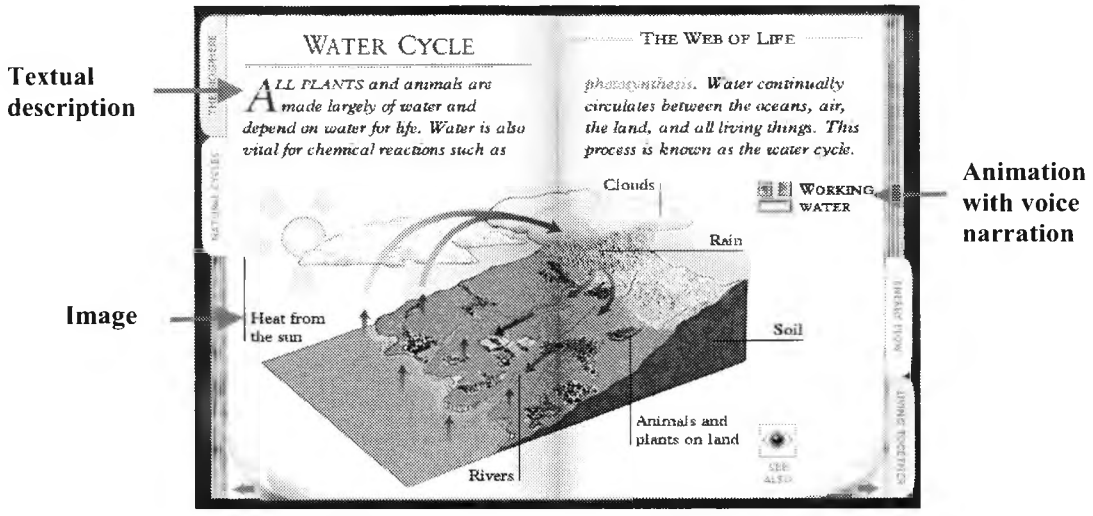


Figure 1: Water Cycle (from DK© Encyclopaedia of Nature)

Effective multimedia presentations need to ensure that:

4) The student will first **focus on** important concepts and details presented either in the text or image. This is important to ensure that the student will read and view all important information and will not miss out essential details in the presentation. To ensure that, different highlighting techniques, such as presenting text in a different colour or in bold, can be used to draw student's attention to important words or concepts.

5) Secondly, when concepts are presented in two different media, the student will need to **integrate** the information presented in both visual and verbal media. If a number of different media are used to describe the subject matter, the student can easily get

confused or loose the thread in the message if the media is not integrated well. Imagine if different terminology is used in the text and on the image, or in the voice narration. To avoid confusing the students, they will need to clearly understand the relationship between corresponding information in the image and the text. 'Contact points' or links can be designed to point the student to corresponding information in the text and the image. Contact points may include using the same label designating an object both in an image and in text, or allowing the student to click on a word in the text, which will highlight the related section in an image.

6) Finally, only if the first two processes have been successful the student will be able to **understand** the concepts being presented in the media. How much the student will understand will also depend on how well each individual medium has been design, and on much the student already knows about the concept being taught. For example, a child who has just started learning about nature and natural cycles may find it difficult to understand the principle of the water cycle if the nature elements were not labelled on the image and the blue arrows were not added to represent a cycle.

Therefore, successful learning using educational multimedia software depends on how well the media design supports student's in attending to, integrating and understanding important information presented in each medium. If the student is hindered in any of the process they may fail to understand the material.

Evaluation Method Overview

To help you assess how effective a multimedia presentation is in supporting these three processes of learning and in identifying problems in the design, the following three steps were formulated.

Step 1: **Evaluation of Attentional Design:**
In this section you need to explore which media will be **attended to** first and whether the student's attention is drawn to important information so they do not miss any important details of the presentation.

Step 2: **Evaluation of Media Integration:**
In this section you need to assess how well the visual and verbal media are combined together to allow the student to **integrate** the information presented in them.

Step 3: **Evaluation of Individual Media Design:**
In this section you need to evaluate whether the design of individual media makes it easier or harder for the student to **comprehend** the information presented.

Each step contains a set of questions, which aim to aid you in analysing the effectiveness of different aspects of the multimedia presentation and to predict likely problems students may experience while learning with the presentation. The questions have two parts. The first part points you to aspects of the design, which can potentially cause problems to the student, and the second part enables you to identifying what effect the problem may have on the student's cognitive processes. For some questions the second part is presented in a separate question indented underneath the original question. The questions are complemented with brief design guidelines to provide you with more explanation on basic principles of media design and integration.

Using the questions, you need to identify potential design problems and determine how they may hinder student's learning processes and their performance in achieving their task.

To achieve best results, it is suggested that you analyse each individual section of the presentation following the three steps, before going to the next section. For each section assess the effectiveness of the attentional design, media integration and the media design before assessing the next section. This is aimed to ensure complete analysis of the multimedia student interface, tracking student's cognitive processes from attending to information through to comprehending the information presented in the media resources.

Specifying Design Problems

In order to help the developers of the application to improve the design of the presentation, you need to describe the problems as fully as you can. This includes describing what the problem is, specifying its cause and where in the presentation it occurs, specifying the effect of the problem on the student's cognitive processes of attention, information integration or comprehension. The questions in each step should help you in determining these effects. After that, you need to describe the likely *actions* the students can take as a result of the problem and the likely *outcome on the student performance* and the achievement of their task.

The *cause* of the design problem is usually the design feature that causes it. The design feature has an *effect* on the student, e.g. student's attention may not be drawn to an important concept, or s/he may have difficulties linking information in the text and the image, or understanding part of the text.

As a result of the problem the students can take different *actions*. For example, the student can press the wrong button, or s/he may repeat an action in an attempt to overcome the problem, or the student may give up. When deciding what the likely student actions might be, consider how critical the effect is and what the student's response may be, i.e. what s/he would do after experiencing such a problem. What the student may do will also depend on the options available in the interface and how familiar s/he is with the system.

Performance outcomes are the effects on the student's work and the achievement of their tasks resulting from the problem. Some example performance outcomes are that the student may spend a long time on the task, or s/he may fail to achieve their task at all (e.g. will not be able to understand the concept, or will only gain partial understanding of the concept), or if the problem is minor the student will be able to continue with the task and successfully achieve it.

Summary of the Procedure for Evaluation:

For each section of the presentation:

7. **Review the media design following the three steps and answer the relevant questions in each step.**
8. **Identify likely problems students may experience.**

For each problem you identify:

9. **Specify its *cause* and *location*.**
10. **Specify the likely *effect* it may have on the student's attention, information integration or comprehension.**
11. **Describe the likely *actions* the student may perform as a result of the problem.**
12. **Describe the likely *effect on the student's performance* the problem may**

STEP 1 : Evaluation of Attentional Design

The aim of this section is to identify potential problems in directing student's attention to important facts and concepts in the presentation. The main aspects to consider concern whether the student's attention is attracted to information which is important for their learning, and whether highlighting techniques are used effectively to guide the student to view important facts and concepts. This is vital to ensure that the student focuses on important concepts and does not miss essential details in the presentation.

Attention to Text

In this section you need to evaluate whether the design of text will direct the student to read and concentrate on parts of the text which are essential for them to learn.

Evaluation Questions:

1.10 Are important words, phrases or concepts highlighted to ensure the student will pay particular attention to them?

1.11 Are there important words and concepts which are not highlighted at all?
1.12 If so, do you think that the student may fail to read them?

1.13 Are there any parts of the text which are highlighted but are not important for the student's task?
1.14 Is that likely to divert student's attention from more important parts in the text?

> *Design principle: In order to ensure that the student will extract all important information their attention needs to be drawn explicitly to key words and phrases. Effects which can be used to make particular words or phrases stand out include bold, large fonts or underlining.*

1.15 If there are other media resources, such as still images, presented at the same time as important text, is the image likely to take student's attention away from the text before s/he has finished reading it?

> *Design principle: Because attention is attracted to size, detail and colour, generally an image will be focused on before a text. If focus is required on the text prior to the image, the text should be displayed before the image or make the text area larger than the image.*

1.16 Is there any event in the student interface, which is likely to interrupt the student while s/he is reading a piece of text? Such an event may include displaying an image or the onset of an animation.
1.17 Is that event likely to distract the student from finishing reading the text?

> *Design principle: When students are reading text their focus can easily be shifted to another medium, thus showing an animation or changing the image should be avoided at that time.*

1.18 If text is displayed dynamically, does the rate of changing it allow the student sufficient time to read the whole text?

> *Design principle: Reading text is slower than identifying an image. As a rule simple words require 200msec each to read, and at least 6 seconds should be allowed for two lines of text, and more if the scene is complex for the students.*

Attention to Still Images

In this section you need to evaluate whether the design of still images will allow the students to focus on important objects within the image. You also need to identify potential problems in the design likely to distract student's attention from inspecting important parts of the image.

Evaluation Questions:

2.11 Are sufficient highlighting techniques (e.g. use of bright colour, arrows pointing to an object) used to emphasise objects which are important for the student's to learn about?

2.12 Are the highlighting techniques sufficient to attract student's attention to such important objects?

2.13 Are there any objects which are particularly important for the student's tasks that are not highlighted or are left outside the main scene?

2.14 Is it likely that the student may fail to view these objects?

2.15 Are highlighting techniques used to emphasise any unimportant objects or minor details in the image?

2.16 If so, is this likely to distract the student from the more important objects or parts of the presentation?

> *Design principle: It is important to highlight information important for the student's task in an image to ensure they do not miss important concepts or details. Important objects can be highlighted by presenting them in bright colour, by setting them apart from other objects, making them larger in size, show more detail or moving nearer to the front of the scene.*

2.17 If a number of objects are to be displayed in an image, is each object revealed gradually, one at a time, or are they revealed all at once?

2.18 If they are revealed at the same time, are the objects going to compete for the student's attention?

2.19 If objects are displayed gradually in turn, is the student given sufficient time to inspect each object, or are the objects displayed too quickly?

2.20 If they are displayed too quickly, is this likely to confuse the student as to which object to view first?

> *Design principle: To control the order in which the student will view information it is important to reveal objects, labels and symbols gradually, one at a time, rather than all at once. Because attention is attracted to change, revealing each part of the presentation in turn is a useful way of directing order.*

STEP 2: Evaluation of Media Integration

The aim of this section is to identify any problems students may experience in integrating the concepts presented in different media, such as text and image. Important aspects to consider include whether sufficient *contact points* are established between related information in both media to help the student integrate the information into a single concept. Remember that successful learning **will only occur** if the students are able to build connections between corresponding aspects of the visual and the verbal representations. If they are unable to integrate these aspects, this will only cause them confusion and will result in partial or miss-understanding of the material.

Evaluation Questions:

4.11 Considering how much the student already knows about the material, is the scene set sufficiently in the text before the student views the image?

4.12 Given the level of detail presented in the text, will the student be able to understand the concept presented in the image after having read the text?

4.13 Are important objects, details or concepts depicted in the image described sufficiently in the text?

4.14 Considering the student's prior knowledge of the concepts, are they able to understand the particular objects or concepts depicted in the image?

> *Design principle: If text is being used to set the scene, the text should be presented before visual media. It is also important that the correct level of detail is provided, i.e. the whole scene needs to be described sufficiently. Detailed textual descriptions of images are essential to help students with low prior knowledge of the material to understand what is depicted in the image.*

4.15 Does the information in the text direct the student's attention to important parts of the image?

4.16 Is that sufficient to ensure the student will not miss seeing them?

> *Design principle: When language is combined with visual media, language should be used to direct the student's attention to information within the image. For example, direct attention to the image by placing a message "look at the blue arrow in figure 1" in the text.*

4.17 If there is information about an important concept presented in an image and described in text at the same time, are there contact points established between relevant pieces of information in each medium?

4.18 Will the student be able to perceive the connection between the relevant aspects in each medium and is s/he able to integrate the related information?

4.19 If there is information about a concept which is unfamiliar to the student, are there contact points established between related concepts in the presentation?

4.20 Will the student be able to build associations between related concepts?

Design principle: Use explicit contact points if the connection between information in an image and language is important. For example, direct student's attention to an object in an image by highlighting the object which is being spoken about, or reveal a text caption and an arrow pointing to the object.

STEP 3 : Evaluation of Individual Media Design

After you have assessed whether the student will be able to focus on important information and whether s/he will be able to integrate the information presented in different media, in the final stage you need to evaluate whether the design of individual media makes it easier or harder for the student to comprehend the information presented.

The main consideration at this stage is to ensure that the subject matter content is presented in the most efficient way to the student to aid their comprehension.

Comprehension of Text

In this section you need to evaluate whether the textual messages and descriptions are easy to comprehend.

Evaluation Questions:

5.8 Are the messages clear enough for the student to understand, considering hers/his familiarity with the subject matter and their language abilities?

5.9 Is there any terminology used that is not familiar to the student, again considering hers/his familiarity with the subject matter and their language abilities?

5.10 Are the concept descriptions at the right level for the students to understand, given their prior knowledge of the material?

Design principle: The level of the content should be suitable for the intended students – it should be higher but not too high above their current level of knowledge. Less knowledgeable students will need more explanation and greater details of the concepts being taught and of related concepts.

5.11 Does the text structure reflect the structure of concepts in the domain?

5.12 Is the concept structure clearly indicated so that even unfamiliar students can understand it?

5.13 Is the order of concepts indicated clearly?

5.14 Is the order of concepts clearly indicated so that even unfamiliar students can understand it?

Design principle: Clear content structure is vital to aid students, especially those who are not familiar with the material, to understand the correct structure of events and concepts. Good design of content structure should specify the different parts of the content and their relationships. This can be aided by dividing the content into sections and sub-sections, and by using bullet points and numbered lists.

Comprehension of Still Images

In this section you need to assess how clear and easy to understand the still images are.

Evaluation Questions:

6.7 Are the objects in the image labelled appropriately to aid the student in identifying the objects?

6.8 Are labels placed next to the objects they identify to help the student in identifying them?

> *Design principle: Labels of objects improve objects' identification, thus labels should be placed next to the objects they identify.*

6.9 Is the design of the icons clear for the student to understand what they represent, considering the student's prior experience with the application and other computer systems?

6.10 Can the student distinguish the icons from other symbols presented on the screen?

> *Design principle: Well designed icons should be: easily associated with the message they represent, distinguishable from other symbols, not overly complex, suitable for different cultures, and in accordance with international or accepted standards.*

6.11 If motion information is shown in a still image, are the start and end points of the motion designated clearly?

6.12 Are any changes of direction clearly represented on the images to aid the student in imagining what the motion is or what the changes resulting from the action are?

> *Design principle: For motion information presented in still images it is important that breakpoints are shown, e.g. start and end point of the motion, and any changes in direction or speed are shown within the image sequence for the motion to be identified.*

Appendix 6.2: Context Description of IMM Software - Study 2

MathWise is a computer-based learning environment for the study and practice of mathematics. It comprises mathematical modules, reference material, assessments and resource tools. MathWise modules are based on the European Engineering syllabus, comprising mathematical topics taught in pre-university and first year university, together with a number of key topics in second-year university Science and Engineering courses.

Main Aim of "Exponential Graphs" Section

The aim of the section is to describe the main graphical features of the following elementary graphical functions:

- The exponential and hyperbolic functions
- The logarithmic functions.

The section also aims to explain the effect of translation, scaling and reflection on those functions.

Target Learning Environment:

Assume that the application will be used primarily for self-study in a computer lab (available from a computer network) or at the student's home, where immediate lecturer support will not be available. The material in the software is designed to complement the lecture material taught in foundation and first year mathematics courses.

Appendix 6.3: Detailed Users Profile

The Maths application is designed to be used by students attending foundation, first or second years Mathematics modules at British Universities typically enrolled on Engineering courses. The students who use the software are both male and female, and are of various ages, but all older than 18 years of age.

The students may also be of diverse ethnic origin and can come from different parts of the world. Therefore, their English language abilities and, in particular, knowledge of Maths terminology in English can vary. Some students also may not have studied Maths in English before. However, their command of the English language in general will be good, as this is a requirement by most universities within the U.K.

As a prerequisite, the students are required to have Maths knowledge equivalent to GCSE or A-Level Maths, however no prior knowledge of Exponential Functions should be assumed, although some students may have acquired some knowledge of the topic at their schools. Familiarity with related topics, such as co-ordinate systems and ability to plot graphs, is expected, however no familiarly with the concepts of graph transformations, e.g. scaling, should be assumed.

The students' familiarity with and frequency of use of computers will also vary. The majority of the students use a computer once a day or once a week, but there are some who do not use a computer as often. Thus, some students may not be very competent at or comfortable with using a PC. A vast majority of the students will be familiar with packages such as MS Word and Excel, Internet Explorer and Netscape, and some email packages.

The students generally would not have used the Maths application before starting their Maths courses.

Students typically are asked to use the Maths application during their tutorials. They are also encouraged to use the software as part of assignment preparation or for exam revision alongside other textbooks and lecture notes. However, as students' interest in learning Maths varies, their motivation in using the software outside tutorial times will also vary.

Appendix 6.4.1 Usability Problem Report Format

Question No.:

Problem Description:

Cause:

Location:

Effect:

Actions:

Performance outcome:

Question No.:

Problem Description:

Cause:

Location:

Effect:

Actions:

Performance outcome:

Question No.:

Problem Description:

Cause:

Location:

Effect:

Actions:

Performance outcome:

Appendix 6.4.2 Comprehension Problem Report Format

Question No.:

Problem Description:

Location:

Learning Difficulty:

Question No.:

Problem Description:

Location:

Learning Difficulty:

Question No.:

Problem Description:

Location:

Learning Difficulty:

Appendix 6.5: Evaluation Task Instruction Sheet - Study 2

The aim of this study is to evaluate the usability of a small section of an application for teaching and learning Maths. The section you need to evaluate is Exponential Graphs – only the first 5 pages and any other pages branching from them, e.g. the demo/tests and the leaflets. But do not evaluate pages 6 and 7 on Logarithmic or Hyperbolic graphs.

1) Firstly, familiarise yourself with the Exponential Graphs Section. While browsing through the pages please identify:

- important maths concepts described in the text or depicted on the graphs;
- the different types of media that are used to represent the Maths material, e.g. still and animated text, graphs, icons;
- any highlighting techniques used, e.g. text in bold or italics;
- similar or the same information represented in the text and the graphs and any contact points between such information.

2) After you have familiarised yourself with the Exponential section, please follow steps 1-4 of the suggested procedure for evaluation, in order to identify any usability problems. This includes identifying the **cause** of each problem you have identified and specifying **the cognitive effect** it may have to users while they are interacting with the particular part of the application (steps 3 and 4).

3) After you have done this for each page, you will be given additional information which will help you predict the likely user behaviour and performance outcomes. Please revisit each problem you have identified and specify any **likely user actions** or **performance outcomes** you may think of. Please note that not every problem may, i.e. in some cases the users may not perform any action or they may choose to give up with what they are doing as a result of the problem.

Appendix 6.6: EMMCW Post-Evaluation Interview Questions

1. How clear was it from the description how people learn with multimedia?

   - Not clear at all - unclear - reasonably clear - very clear - don't know

2. Did the explanation at the front help you in the evaluation of the multimedia presentation? Was it useful?

3. Was the concept of contact points clearly defined?

4. Did you experience difficulties identifying contact points between verbal and visual media?

5. How easy to follow did you find the procedure of evaluation? Why? Which part was particularly difficult?

   - Very difficult to follow - difficult follow - easy to follow - very easy to follow
   - Don't know

6. How clear was the problem format?

   - Not clear at all - unclear - reasonably clear - very clear - don't know

7. How easy to specify were the different parts of the format?

   - Cause – very difficult – difficult – easy – very easy – don't know
   - Breakdown/Misconception– very difficult – difficult – easy – very easy – don't know
   - Actions – very difficult – difficult – easy – very easy – don't know
   - Performance/Learning outcome– very difficult – difficult – easy – very easy – don't know

8. How understandable were the questions?

   - Very difficult to understand - difficult to understand - easy to understand
   - Very easy to understand - don't know

9. Did you have difficulties understanding some of the questions? Which ones?

10. Were there any unfamiliar terms used in the method? Which ones?

11. Did the questions help you to identify where misconceptions can occur?

12. Did you find the user description helpful to predict the likely learning outcomes?

13. Which user characteristics did you consider? Please list them.

14. On average how many of the questions did you use in the evaluation?

    - 10% - 30% -50% -70% - 100%

15. Did you find the design principles helpful?

16. Do you have any other comments and would you recommend any improvements to the method?

Appendix 6.7.1: Problems Predicted by Multimedia Designers using the EMMCW Method

No	Problem Description	User Problem Match
1	**Question #:** 1.4 **Description:** the emphasis is on particular functions, e.g. $y= x^{10}$, than on the overall idea that $x<30$ demonstrates slow exponential growth. **Cause:** separating equation into new paragraph **Location:** page 1 **Effect:** user may ponder on what equation means rather than taking in general idea of example. **Actions:** **Performance outcome:** caused me to think about equation much more than I ought, perhaps idea was unclear in general.	C1
2	**Question #:** 4.1, 4.2 **Description:** Graph displays general idea of exponential equations, but may have been more helpful in the next pane where description is about now steep an exponential graph is. **Cause:** Information revealed too soon. **Location:** Page 1 graphs. **Effect:** not helpful at demonstrating the concept. **Actions:** opening graph again to see it demonstrating steepness of growth. **Performance outcome:**	U2
3	**Question #:** 4.3, 4.4 **Description:** text notes that exponential graphs grow slowly, citing a value of 30, but graph only goes up to units of 2 and 5. When tried to increase values to 50 was told did not have sufficient memory. **Cause:** **Location:** Page 1, graphs. **Effect:** text description not demonstrated by graph. **Actions:** closed graph. **Performance outcome:**	
4	**Question #:** 4.7 **Description:** no visual display for steepness delay of exponential equations. **Cause:** graph does not address. **Location:** Page 1, graphs. **Effect:** concept not fully understood. **Actions:** **Performance outcome:**	U2
5	**Question #:** 5.2 **Description:** Phrase 'for $x<30$ it is completely outstripped' is ambiguous. **Cause:** unclear what the words 'it' and 'outstripped' are referring to. **Location:** Page 1 **Effect:** confusion as to point of concept. **Actions:** **Performance outcome:** time spent trying to figure out sentence.	
6	**Question #:** 6.1 **Description:** after lines are drawn on graph no labelling stating what they represent. **Cause:** no labels on lines. **Location:** Page 1, graphs. **Effect:** unclear information, slight confusion. **Actions:** **Performance outcome:**	U2 C1

No	Problem Description	User Problem Match
7	**Question #:** 6.3 **Description:** the user is told to click to see another graph drawn, but not directed where to click. **Cause:** action is not directed. **Location:** Page 1, graphs. **Effect:** some frustration of not being able to see demo properly. **Actions:** clicked several places before desired result was achieved. **Performance outcome:**	U1,4
8	**Question #:** 6.3 **Description:** user given option to change values of graph but unclear hot to do so. **Cause:** no change in font, text colour to show where to click. **Location:** Page 1, graphs. **Effect:** **Actions:** several clicks on different areas of screen before result achieved. **Performance outcome:**	U29
9	**Question #:** 5.1 **Description:** the curve y=exp(-x) is described as 'decay curve' but does not explain how it shows decay. **Cause:** term 'decay curve' is not explained. **Location:** page 2. **Effect:** wondering what term means **Actions:** going back to top of page to see if inverse of horizontal reflection is any special termed graph. **Performance outcome:**	
10	**Question #:** 6.5 **Description:** graphs are explained but when values of variables are changed the graphs may overlap making it unclear which line represents positive and which negative values. **Cause:** lines are not marked. **Location:** page 2 **Effect:** not sure what is being demonstrated. **Actions:** redrawing graph several times to figure out the demonstration. **Performance outcome:** time spent wondering about graph.	U12, 17
11	**Question #:** 1.7 **Description:** pop-up window prompting for actions obscures graph it is referring to. **Cause:** poorly placed pop-up window. **Location:** Page 2 **Effect:** graph obscured, but annoying. **Actions:** moved box out of way. **Performance outcome:** time spent moving graph, attention not paid to concept but to obstruction.	U15, 40
12	**Question #:** 5.2 **Description:** phrase 'vertical intercept' appears for first time without description, hence key part of concept left unclear. **Cause:** no definition given. **Location:** page 3. **Effect:** concept unclear. **Actions:** time spent pondering concept. **Performance outcome:**	C6
13	**Question #:** 1.5 **Description:** text appears in grey navigation section for first time. Unsure if I should read it first and how I should apply it to the graph, which then obscures text. **Cause:** text in new area o screen. **Location:** page 3 leaflet. **Effect:** not sure which text on page should receive priority. **Actions:** read all text a few times, drew graph, closed graph to see if text was relevant, redrew graph. **Performance outcome:** time spent wondering about how text fits into rather than absorbing its meaning.	U23

No	Problem Description	User Problem Match
14	**Question #:** 1.5 **Description:** same as above, less so now as it occurs for a second time. **Cause:** **Location:** page 4 and page 5 leaflet. **Effect:** **Actions:** **Performance outcome:**	U23
15	**Question #:** 5.6, 6.1 **Description:** e^x, 2^x, 10^x displayed in graph- their respective lines are not labelled. **Cause:** lack of graph labels and colouring. **Location:** Page1, graphs. **Effect:** users have to remember while line is which when reading graphs, and distinguish lines which are all black. **Actions:** **Performance outcome:** may slow comprehension down, make user redisplay graphs, prolonging task time.	U1,2 C1
16	**Question #:** 1.9 **Description:** text changes too quickly to read, if not read at start and watched instead. **Cause:** text in graph **Location:** Page 1, text in graphs. **Effect:** user may not notice text first time and be confused as to which line is being drawn. **Actions:** may repeat graph displays several times. **Performance outcome:** slow task speed down.	U2,4
17	**Question #:** 2.3 **Description:** number line is not associated close enough to the text that refers to it. **Cause:** distance between text and number line. **Location:** Horizontal reflection leaflet, text. **Effect:** user may not notice it immediately or not read the text that follows this instruction once diverted. **Actions:** **Performance outcome:**	U8
18	**Question #:** 6.1 **Description:** labels are not near graph lines. **Cause:** labels at top of graph rather than at top/bottom depending on which line they refer to. **Location:** Horizontal reflection leaflet. **Effect:** may slow comprehension of graph down. **Actions:** **Performance outcome:** slow task completion down, reduces comprehension.	U11
19	**Question #:** 1.1, 1.6 **Description:** summary information is not highlighted and obvious. Navigation buttons take prominence on page. **Cause:** information in normal font. Should be highlighted somehow. **Location:** Page 2, leaflet. **Effect:** text explanation goes unnoticed. **Actions:** **Performance outcome:** may reduce comprehension.	U12, 23
20	**Question #:** 6.1 **Description:** graph lines are not labelled clearly enough so it is not easy to tell if the –x line is on the right or the left. **Cause:** **Location:** Vertical reflection graph leaflet, p. 2. **Effect:** difficult to identify which graph is for which equation. **Actions:** **Performance outcome:** reduces comprehension.	U16

No	Problem Description	User Problem Match
21	**Question #:** 1.1 **Description:** important words highlighted 'above' or 'below' when press 'show me' but the lines of graph aren't drawn. **Cause:** show me demo **Location:** Pages 2 & 3, demo/test **Effect:** incomplete demonstration that may confuse or is useless. **Actions:** **Performance outcome:** slow task performance down.	U13
22	**Question #:** 6.1, 6.2 **Description:** graph answers to test questions aren't labelled. **Cause:** label & text questions. **Location:** Page 3, graph. **Effect:** difficult to comprehend graph. **Actions:** redo test **Performance outcome:** slow comprehension.	C3,5
23	**Question #:** 6.1 **Description:** text question asks for scaling, when wrong it displays it in the graph. **Cause:** text box, graph display. **Location:** Page 3, graph. **Effect:** difficult to see connection if not noticed straight away, answer never appears in text box. **Actions:** redo task. **Performance outcome:** slow comprehension and task.	
24	**Question #:** 6.1 **Description:** the 2 lines displayed on graph are not labelled at all. The bold line does not relate to bold text, and this could easily be used to distinguish the 2 lines. **Cause:** lines in graphs. **Location:** Page 4 & 5, demo/test. **Effect:** graph is not that easily comprehensible. **Actions:** repeat graph display. **Performance outcome:** poor comprehension.	C12

Problems Predicted by Multimedia Designers Using Own Judgement

No	Problem Description	User Problem Match
1	**Question #:** **Description:** graph test question box obscures graph that questions are being asked about. **Cause:** question box location. **Location:** Pages 2 to 5. **Effect:** may annoy user and make them give up or stop the exercise prematurely. **Actions:** box has to be moved so question can be answered and full graphs can be seen. **Performance outcome:** may prevent comprehension.	
2	**Question #:** **Description:** icons at bottom of page unclear as to what they do. **Cause:** no text explanation or roll-over text. **Location:** bottom of every page. **Effect:** confusion. **Actions:** clicking to see what happened. **Performance outcome:**	U6

Appendix 6.7.2: Problems Predicted by Subject Matter Experts using the EMMCW Method

No	Problem Description	User Problem Match
1	**Question #:** 1.1. **Description:** only the headlines of the text are highlighted. **Location:** all pages. **Learning difficulty:**	
2	**Question #:** 1.2, 1.3 **Description:** there are important words which are not highlighted. The student may fail to read these, because the whole text gives the impression of boring theory explanation. **Location:** all pages. **Learning difficulty:**	
3	**Question #:** 1.4, 1.5 **Description:** students' attention is not drawn to any part of the text. **Location:** all pages. **Learning difficulty:**	
4	**Question #:** 1.6 **Description:** the button on the right hand side of the page are too big. The student will probably stop reading the text on the left and will read the buttons' captions. **Location:** pages 1-5. **Learning difficulty:**	
5	**Question #:** 1.7, 1.8 **Description:** there is no animation or image that could interrupt the students while they are reading, but the fact that the program window is not a full screen window will certainly tempt students to look at the desktop's icons. **Location:** **Learning difficulty:**	
6	**Question #:** 1.9 **Description:** there is not any dynamically displayed text. **Location:** **Learning difficulty:**	
7	**Question #:** 2.1, 2.2 **Description:** there are no highlighting techniques used, although the software is coloured there are only 2 colours in the images. **Location:** pages with graphs. **Learning difficulty:** text is not enough to identify a curve, but you need colour when you have to look at 2 curves at the same time.	C1,2,3,6,12 C13
8	**Question #:** 2.3, 2.4 **Description:** there are buttons labelled with 'i' which provide more information. They are not highlighted at all, on the contrary the user may think that these buttons are disabled. **Location:** leaflet pages **Learning difficulty:**	U33
9	**Question #:** 2.5, 2.6 **Description:** the only highlighting technique is bold line in the graphs, which is not enough. These bold lines are not sufficient to attract student's attention. **Location:** **Learning difficulty:**	
10	**Question #:** 2.7, 2.8 **Description:** each object is revealed gradually but the program generates a window in which the user has to make an input, however it is displayed in front of the graph. **Location:** graph pages. **Learning difficulty:**	U40

No	Problem Description	User Problem Match
11	**Question #:** 4.1, 4.2 **Description:** the text is not sufficient. What follows is not very well presented in the text. **Location:** all pages. **Learning difficulty:** Students may have difficulties in understanding the images.	C7,8,11,12
12	**Question #:** 4.3, 4.4 **Description:** the level of detail presented in the text is not adequate. Even after looking at the graphs there should be some text explaining to students why a certain transformation has been done in this way. **Location:** all pages. **Learning difficulty:**	U35,36
13	**Question #:** 4.5, 4.6 **Description:** the text does not sufficiently direct student's attention to important parts of the image. **Location:** **Learning difficulty:**	
14	**Question #:** 4.7, 4.8 **Description:** I could not find these links. If there are any they have to be clearer. **Location:** **Learning difficulty:**	
15	**Question #:** 4.9, 4.10 **Description:** there are some links but they are not very clear and are not included in the text. When you finish reading, you can see an information button but it's too late. **Location:** **Learning difficulty:**	C2,3,12
16	**Question #:** 5.1 **Description:** text gives the impression that it is a summary of something which is not presented. It could be written in a better way. **Location:** **Learning difficulty:**	
17	**Question #:** 5.3 **Description:** concept descriptions assume that students are familiar with the subject. That does not happen all the time for all students. **Location:** **Learning difficulty:**	
18	**Question #:** 5.4, 5.5 **Description:** the detailed description is missing. There is a confusion presented without any explanation. **Location:** **Learning difficulty:**	
19	**Question #:** 5.6, 5.7 **Description:** the order of objects is clearly indicated but somehow they are mixed up. I could find 'horizontal translation' in the 'vertical translation' topics in the main page they are supposed to be different topics. **Location:** page 1 of vertical translation leaflet. **Learning difficulty:**	U25, 26
20	**Question #:** 6.1, 6.2 **Description:** objects in the images are not labelled appropriately, including the interface buttons. Everyone can easily get confused. **Location:** **Learning difficulty:**	U41 C1,2,3, 12
21	**Question #:** 6.3 **Description:** the design of the icons does not help the student to identify its function. Icons need more fantasy and colour. **Location:** **Learning difficulty:**	U41, 50

No	Problem Description	User Problem Match
22	**Question #:** 6.5 **Description:** any animation is displayed correctly the first time. If you want to watch it again the image does not get refreshed, so the old lines are there and you cannot see any animation. **Location:** **Learning difficulty:**	U14
23	**Question #:** 6.6 **Description:** changes of direction are not represented at all. That forces that student to guess, especially in the graphs where any kind of translation is presented. **Location:** **Learning difficulty:**	C4,6, 12
24	**Question #:** 1.1, 1.2, 1.3 **Description:** important words, phrases and concepts are not properly highlighted. Bold font is a better way to underline some words rather than italics. **Location:** all pages. **Learning difficulty:**	
25	**Question #:** 1.4 **Description:** "At first sight this seems unnatural so consult …", this is a suggestion but is put as a rule. It needs to be highlighted in a different way. **Location:** Page 5. **Learning difficulty:**	C12
26	**Question #:** 1.6, 1.7, 4.5 **Description:** it is not particularly clear if a proper order in showing the images or in making some tests has been suggested by the software. **Location:** all pages. **Learning difficulty:** the freedom in moving around the pages could confuse the students without giving them a proper method to approach learning the concepts.	
27	**Question #:** 2.1 **Description:** the graphs are not properly displayed, the lack of labels close to the graphs makes them unclear. There are no arrows pointing to any objects this would help to understand the different functions displayed. **Location:** all pages. **Learning difficulty:**	C1,2,3,12
28	**Question #:** 2.3 **Description:** in the tests both reflections of the function are displayed without properly describing them. This could confuse the student. Only the final answer is given, which is not enough. **Location:** page 2. **Learning difficulty:**	C12
29	**Question #:** 2.9, 2.10 **Description:** when the answer of a test has to be given, there is no sufficient time and no proper display of the several images to have a clear show of the answer. A more dynamic display of the text and graphs in sequential order may help. **Location:** Pages 2 to 5. **Learning difficulty:**	C2,12
30	**Question #:** 1.9, 2.10 **Description:** the text is dynamic only in the window where the graphs are displayed. The rate of changing the text is too fast. Displaying first the text and then the graph with enough time in between them would help their comprehension. **Location:** Page 1. **Learning difficulty:**	C1
31	**Question #:** 1.4 **Description:** the suggestion "you can alter the scale…" is repeated several times and it should be put in a different way. Sometimes it's displayed with other comments, and the attention of the reader could be lost. **Location:** all pages. **Learning difficulty:**	U1

No	Problem Description	User Problem Match
32	**Question #:** 4.1 **Description:** the presentation of the exponential function is not sufficiently clear to have a proper description of the function itself. There is a glossary, but it's not properly highlighted how it could be accessed. **Location:** page 1. **Learning difficulty:**	U5, 34
33	**Question #:** 4.5 **Description:** there are no visual links between the text and the graphs. This doesn't direct students' attention to the particular concepts. **Location:** all pages. **Learning difficulty:**	
34	**Question #:** 4.7 **Description:** in the tests, both reflections are displayed to find which is the correct answer. If not properly distinguished, the way to put a question and its answer could confuse the student. **Location:** Page 2. **Learning difficulty:**	C2
35	**Question #:** 4.9, 4.10, 5.1 **Description:** the final comment in the left window is not clear without a proper image. **Location:** Page 4. **Learning difficulty:**	C13
36	**Question #:** 5.1 **Description:** the terminology is really basic, and it should be familiar to the majority of the students. **Location:** **Learning difficulty:**	U2
37	**Question #:** 5.6 **Description:** there is no suggested order or structure. No steps are suggested, which could confuse a first-time user. **Location:** **Learning difficulty:**	
38	**Question #:** 6.1, 6.2 **Description:** there are no labels next to the functions. Some labels to show a different trend when some parameters have been changed could clarify the concepts. **Location:** **Learning difficulty:**	
39	**Question #:** 6.3 **Description:** it's not clear how to find the glossary. **Location:** all pages. **Learning difficulty:**	U41