Collaborative Learning with Affective Artificial Study Companions in a Virtual Learning Environment.

Timothy Barker, B.Sc. (Hons.), M.Sc.

Submitted in Accordance with the requirements for the degree of Ph.D.

The University of Leeds,
Computer Based Learning Unit,
School of Education.

October, 2003.

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to work of others.

Acknowledgements.

"...the further one travels
the less one knows."

Lao Tzu,
Tao te Ching.

This research has been partly carried out by a team that has included Dr. Rachel Pilkington, Professor Roger Hartley, Dr. Catherine Bennett, Mrs. Rosa Hall and Mrs. Patricia Greenwood. My contribution to the work has been to review appropriate literature and software, inform the development of and evaluate the resultant Virtual Learning Environment as well as design, develop and evaluate the Learning Companion.

I would also like to thank friends, family and colleagues (old and new) for their constant support and encouragement. Finally, I would like to thank the music.

Abstract.

This research has been carried out in conjunction with Chapeltown and Harehills Assisted Learning Computer School (CHALCS) and local schools. CHALCS is an 'out-of-hours' school in a deprived inner-city community where unemployment is high and many children are failing to meet their educational potential. As the name implies CHALCS provides students with access to computers to support their learning. CHALCS relies on many volunteer tutors and specialist tutors are in short supply. This is especially true for subjects such as Advanced Level Physics with low numbers of students. This research aimed to investigate the feasibility of providing online study-skills support to pupils at CHALCS and a local school. Research suggests that collaborative learning that prompts students to explain and justify their understanding can encourage deeper learning. As a potentially effective way of motivating deeper learning from hypertext course notes in a Virtual Learning Environment (VLE), this research investigates the feasibility of designing an artificial Agent capable of collaborating with the learner to jointly construct summary notes. Hypertext course notes covering a portion of the Advanced Level Physics curriculum were designed and uploaded into a WebCT based VLE. A specialist tutor validated the content of the course notes before the ease of use of the VLE was tested with target students. A study was then conducted to develop a model of the kinds of help students required in writing summary notes from the course-notes. Based on the derived process model of summarisation and an analysis of the content structure of the course notes, strategies for summarising the text were devised. An Animated Pedagogical Agent was designed incorporating these strategies. Two versions of the agent with opposing 'Affectations' (giving the appearance of different characters) were evaluated with users. It was therefore possible to test which artificial 'character' students preferred. From the evaluation study some conclusions are made concerning the effect of the two opposite characterisations on student perceptions of the agent and the degree to which it was helpful as a learning companion. Some recommendations for future work are then made.

Contents.

Index of Figures.

Index of Tables.

Chapter 1: The Research Context, Aims and Methodology.

1.1 Introduction: The Need for Virtual Learning Environments.

Widening participation in education is of global importance. The current knowledge-based society, fuelled by enormous technological advances, for example in Information and Communications Technology (ICT), Biotechnology and Nanotechnology, can not be maintained without a workforce capable of understanding the complexities of life in the Twenty First Century. However, there are still great deficiencies in the balance of the knowledge commodity both between nations and indeed within national communities. Thus, in the interests of participation and productivity world citizens must be empowered through education otherwise they run the risk of subjugation, worthlessness and ultimately poverty.

'Flexible learning' where students can, for example, study at a distance enables people to participate in new learning experiences who previously had been excluded, for example perhaps on the basis of unfavourable geographical or economic factors. This flexible learning must take place under a whole new set of personal circumstances. Potential students, particularly mature students, may be located in disparate countries or may only be available during 'unsociable' hours. There is therefore a need to provide facilities that accommodate this new Educational ethos and corresponding approaches. Virtual Learning Environments (VLEs) adopted in colleges, Universities and in the commercial training sector are used as tools that can help deliver flexible learning and so work towards this goal of widening educational access. VLEs provide course materials, usually accessible around the clock, integrated communication tools such as chat and bulletin boards and a host of features that support the learner.

New technological approaches to education require new associated pedagogical approaches. Collaborative Learning is based upon a view that knowledge is socially constructed (Cole & Wertsch, 1996; Crook, 1998; Vygotsky, 1978) rather than constructed in individual cognitive isolation. In this view students participate in learning communities through shared goals, artefacts and associated communications. Improvements in learning occur as a result of exchanging knowledge or experience with our peers, ideally in situations where we each have overlapping but not identical areas of expertise, knowledge or experience to share. However, to date, VLEs designed for this approach suffer from the Achilles heel of the need for peers or tutors with which to collaborate. Obviously such availability is dependant upon precisely the set of circumstances that the flexible learning approach seeks to address, such as differing time zones or unusual working hours. One possibility for overcoming this weakness is to provide *artificial* peers or tutors with which students can collaborate in the absence of human collaborators. Hence the work presented here has been greatly influenced by Chan's (1996) vision of the Global Social Learning Club in which the student is placed at the centre of a wide variety of learning resources including not only conventional texts and the tutor but both artificial and human companions, peers and outside experts. The need for such artificial Study Companions is driven by the need to widen access to collaborative styles of learning when human peers or tutors are simply not available. Furthermore, it is proposed that through the use of personalisation of these Companions, for example in terms of their 'affective' qualities as realised in behaviour and appearance, the resultant collaboration is enhanced through an increase in student motivation and reward.

Chapeltown and Harehills Assisted Learning Computer School (CHALCS), the context for this research, is a case in point. CHALCS was established in 1987 after a period of civil unrest in this inner city district of Leeds in the UK. Poverty, crime and poor housing, indicators of social exclusion, led to Educational problems (such as truancy and under-achievement) in the community. CHALCS hoped to address these social symptoms through enhanced educational provision. However, CHALCS' tutors, who are in the main volunteers, are in short supply resulting in a 'waiting list' for pupils to gain entry to CHALCS. Therefore, through discussions with staff at CHALCS it was felt that a VLE might be able to alleviate some of these access problems to CHALCS' facilities and thus widen access and increase the flexibility of provision in line with the arguments presented above.

Three of the key areas which CHALCS addresses are those of Science, IT and Literacy. Through consultations with staff it was decided to implement a VLE which would address all of these areas. An exemplar Advanced Level Physics course would be developed in a little supported area, Astronomy and Optics. Through the use of this VLE and suitable training of both pupils and staff those involved would also gain an understanding of IT and New Technologies like the Internet, particularly those relevant to Education. Finally, the project would champion the emerging policy of integrating key skills within pupil's learning, particularly those related to deeper reading for understanding and effective note-taking felt to enhance learning and hence employability. Reading for effective note-taking is considered to be a key component of literacy and is already the subject of a literacy programme at CHALCS. The specific aim was to encourage pupils to practice the key skill of summarising learning materials presented to them (in the specific example case the Physics course notes to be contained within the VLE). Summarisation was chosen as it is an exemplar key skill required during the initial 'Acquisition' stage of the proposed pedagogical approach (see section 2.1.2). To motivate students to engage constructively with the learning materials students would collaborate with an artificial Study Companion, employing affective qualities, and in this process create their own summary document of the Physics VLE material.

It is proposed that the affective qualities of the Study Companions do not have to rest on complex Artificial Intelligence (AI) techniques to be appealing. That is, the Companions will employ mostly 'design-based' affective behaviours or 'affectations' (Barker, 2003), enabling the use of readily available and supportable technologies that schools might realistically use. In addition the hope is that students motivated by the companion will engage more deeply with the reading for understanding and the summary note-taking task.

1.2 The Background and Context of the Research.

This research takes place within the kind of contexts alluded to above, that is, a global view of educational provision and a more local view of community reform. Inevitably both of these viewpoints, concerned with the necessity of Virtual Learning, attract attention from traditional political and social reformists. Federal European and National Governments have spearheaded campaigns to encourage school participation in the Networked Learning opportunities provided by ICT including such initiatives as the European Netdays and the U.K. National Grid for Learning. Such initiatives are aimed at widening school access to ICT and raising teacher awareness of positive uses of ICT to support learning.

1.2.1. The National Grid for Learning.

The National Grid for Learning (NGfL) :

> "is a Government initiative to help learners and educators in the UK to benefit from ICT. It is a vital part of the Government's commitment to the creation of a connected learning society in which learning is increasingly accessible and adapted to individual needs...The NGfL is intended to increase and widen access to learning opportunities for everyone in both formal and informal learning environments such as schools, colleges, libraries, homes and workplaces."
>
> (BECTA, 2001)

Indeed, over £1 billion has been earmarked by the Government to fund ICT in schools during the period 2001 to 2004 under the NGfL initiative. The NGfL is a strategy comprising of the three key elements of *infrastructure*, *content* and *practice.*

'Infrastucture' is concerned with providing the hardware, networks and associated services necessary to sustain the initiative such as the provision and management of ICT equipment, see www.managedservices.ngfl.gov.uk. Further initiatives in this key element include 'Computers for Teachers' which offers subsidies for teachers towards the purchase of a personal computer, 'Independent ICT Procurement Advisory Service' (IPAS) - an independent source of advice on purchasing equipment and 'Building the Grid', a source of help on planning, procurement and practice.

'Content' is concerned with encouraging the development of suitable resources and ensuring their availability. It is realised in the form of the NGfL portal. This portal covers all sectors of the educational infrastructure including Schools, Further and Higher education, Lifelong Learning, Career Development, Libraries and Museums, Community Grids, International Networks and Government Agencies (see www.ngfl.gov.uk). As well as providing links to other web sites the NGfL also provides resources and advice for both teachers (e.g. the Virtual Teacher Centre, www.vtc.ngfl.gov.uk) and pupils or students alike. Furthermore, all sites linked from the NGfL portal must pass strict criteria, the so-called Ground Rules, which maintain standards such as accuracy and legality and are enforced by the watchdog organisation GridWatch. Additionally, it is intended to implement an initiative, Curriculum Online, to provide digital resources for teachers and learners targeted specifically at the National Curriculum. A further award has also been established, the New Opportunities Fund (NOF) Digitisation award, to develop digital content suitable for Lifelong Learning. There are also further initiatives developing resources for teaching and learning targeted specifically at Northern Ireland, Scotland and Wales.

The final key element in the NGfL strategy, 'practice', is concerned with ensuring that the infrastructure and content can be used to their best advantage through proper training of the people for whom it is intended. The NOF programme targets teachers and school librarians (£230 million) and public librarians (£20 million) by providing training in the use of ICT in education.

The NGfL is manifesting in Leeds as the Leeds Learning Network (LLN), see Appendix B and www.leedslearning.net The LLN provides the following key features:

1. Content – access to local, regional and commercial content as well as schools' own internet and intranet sites and staff and pupil web space;
2. Security – web and email filtering plus secure authority-wide networks;
3. Integration – seamless integration of school and LEA networks;

4. Access – free, unlimited, high speed access within reasonable limits plus remote access for teachers and access to the authorities other networks, e.g. museums, libraries and administrative departments;
5. Support and Guidance – on-line technical support and single point of contact as well as access to experts in, for example, web design;
6. Low Cost – help in the form of initial grants with no connection cost during the day and low costs after 6pm;
7. Ease of Use – information and training on the LLN.

It can therefore be seen that this city-wide network has the potential to increase individual and group participation in the local, national and global learning communities through electronic means thus supporting the ideal posited above of widening of educational provision both within the city of Leeds and beyond. In this current climate it can be seen that since this research began, there has been increased interest in developing appropriate models of educational use of ICT in schools (particularly the use of the internet for learning and the development of IT literacy across the curriculum) which make the focus of this research particularly timely.

1.2.2. Enhancing Citizenship Through Key Skills.

The UK Government is keen for the working population to posses a number of key skills which will better equip them for life in the Twenty First Century.

> "Key skills are intended for everyone, from pupils in schools to Chief Executives in large companies. Key skills are the skills most commonly needed for success in a range of activities at work, in education and training and life in general. They focus candidates' attention on where and how they are using skills for the purpose of improving the quality of their learning, work and performance."
>
> (QCA, 2002)

Table 1.1 summarises the current categories of these skills. These skills are seen as essential to enable people to fully participate within the work force and society at large including participation in Lifelong Learning and hence widening provision of education. The key skills initiative was piloted across two hundred centres including schools, training organisations, workplaces and colleges of further and higher education.

Key Skills Units	Example
Communication	Contribute to discussions, summarise information
Application of Number	Tackling numerical problems and collecting, recording, interpreting and presenting results
Information Technology	Search for and select information, plan, interpret and present information from different sources
Working With Others	Organise tasks to…support co-operative working
Improving Own Learning and Performance	Help set targets and plan how these will be met
Problem Solving	Identify problems & come up with ways to solve them

Table 1.1. Key Skills Examples (DfEE, 1999)

From September 2000 a new key skills qualification addressing Communication, Application of Number and Information Technology was introduced at five different levels of proficiency. It is available, for example, as part of GNVQ's, A Levels, Modern Apprenticeships, New Deal participants, Degrees, Doctorates, NVQs etc. As the DfEE states :

> "Key skills enable people to be competent and confident in communicating both orally and in writing; applying number - tackling numerical problems and collecting, recording, interpreting and presenting data; using IT; working with other people; being able to improve on their own learning and performance; and problem solving skills."
>
> (DfEE, 1999)

As previously stated, summarisation is highlighted in the key skills Specifications under the 'Communication' module. The proficiency expected of students on courses varies depending on the level of their study. It was intended to initially instruct CHALCS students encountering this skill, possibly for the first time formally, to Level 2 of the key skills specifications. To elaborate, Level 2 of the Communication module states the following Learning Objective:

> "In assessing the candidate 's performance, evidence must show that she or he is be able to select relevant material, identify accurately lines of reasoning and the main points from both text and image, and summarise the information to suit her or his purpose."
>
> (QCA, 2002)

The aim in our exemplar case study is for the students to be able to construct their own notes for subsequent revision purposes as well as to encourage deeper engagement for understanding the Physics material.

1.2.3. The Research Context: Chapeltown and Harehills Assisted Learning Computer School.

As mentioned in the introduction CHALCS was founded in 1987 as a possible answer to some of the problems facing this inner-city area of Leeds. The area faced riots and civil unrest during this year leading to a government task force being established to address the problems. CHALCS aims to address the Educational problems born of the underlying social problems such as vandalism, drug abuse and dysfunctional families inherent in Chapeltown and Harehills. Table 1.2.shows the aims of CHALCS.

CHALCS Aims
1. Provide computer facilities and activities which will develop and enhance the educational potential and performance of young people (particularly in English, Mathematics and Information Technology) from the Chapeltown and Harehills districts of Leeds.
2. Encourage a supporting role for parents and parental groups to the CHALCS activities and to local schools.
3. Provide support for the work of local schools and a suitable environment for young people to do their homework and out of school assignments.
4. Help raise the expectations and aspirations of both parents and children for a better future.

Table 1.2 CHALCS Constitutional Aims

Category	Chapel Allerton	Leeds	England and Wales
Total population	18,206	715,402	52,041,916
Lone parents (with dependant children)	11.1%	7%	6.5%
Health 'not good'	11.4%	9.8%	9.2%
Unemployed	5.1%	3.3%	3.4%
White	67%	91.8%	91.3%
Pakistani	5.9%	2.1%	1.4%
Indian	5.0%	1.7%	2.1%
Black Caribbean	11.2%	0.9%	1.1%
No Qualifications	31.4%	30.9%	29.1%

Table 1.3. Demographic Comparison of Chapel Allerton Ward with Leeds City Authority District and England and Wales, National Census, 2001
(Source: National Statistics website: www.statistics.gov.uk Crown copyright material is reproduced with the permission of the Controller of HMSO)

Table 1.3 highlights some of the demographic factors affecting the Chapel Allerton ward of Leeds (where CHALCS is situated) in comparison with the Leeds Metropolitan district and England and Wales. As can be seen the area is notably multi-ethnic with a high rate of unemployment, poor health and single-parent families in comparison to the rest of the city and England and Wales. These factors contribute to an atmosphere leading to educational difficulties both at home and school.

An inspection carried out by (OFSTED, 1999) of 48 inner-city and rural schools sought to highlight the relative performances of pupils in both primary and secondary education in the four ethnic minorities of Bangladeshi, Black Caribbean, Pakistani and Gypsy Traveller. In addition the report analysed schools' strategies for raising attainment, their policies on race relations and the extent of LEA's assistance. A further 34 schools were visited who had had some success in minority ethnic pupils' education.

One conclusion of this inspection was that subsequent practice may take the form of extra-curricular activities such as IT clubs or sports clubs. The practice could also involve an increased pastoral element and head-on tackling of racial stereotyping. The report also highlights the need for improved links with parents, such as home visits and with the community at large, e.g. the local mosque or local businesses. In the light of these recommendations we can now examine the role of CHALCS in an inner-city community.

Students attend CHALCS on a voluntary basis during the evenings and at weekends. It is independent of the local education system in terms of funding but it's students obviously are very much a part of that system, coming from surrounding schools such as Chapel Allerton, Allerton Grange and Notre Dame.

Appendix A contains an overview of the CHALCS organisational structure. It's Director is responsible for the overall running of the organisation whilst the Deputy Director administers the day-to-day tasks. Apart from the secretarial contingent most of the staff, i.e. tutors, are appointed on a voluntary basis. They are experienced tutors nonetheless and are well motivated. For example, the A Level Physics tutor is in fact a Research Fellow at the University of Leeds and is dedicated to helping his students through their examinations and onto University or Employment.

CHALCS has a well-documented positive image in the community. It is housed in a building, TECHNORTH, which also contains various small technologically based businesses such as training and business support. It encourages the public and media to examine its practices, for example at Open Days. Also, given the nature of CHALCS' history and current national relevance, Members for Parliament are occasionally invited to inspect its achievements. In addition other initiatives have used CHALCS as a venue, for example, the European Internet awareness initiative, Netdays.

Mohammed (1996) carried out a case study of pupil's abilities in learning Science Topics at CHALCS. One of his key conclusions relating to this work was that:

> "CHALCS Science-like materials and instruction are effective in encouraging interest, attitudes and participation which get students involved in science. The problem of effectively linking this to improved performance in the school science curriculum, and classroom, remains. The use of electronic support, focused on school tasks under the community schools concept could help to achieve this objective..."
>
> (Mohammed, 1996)

A later study by Ravenscroft and Hartley (1998) supported this hypothesis:

> "...the links with schools may be strengthened. At the moment the pupils make the link via their learning activities, and this may be a useful basis (aided by electronic communication) to strengthen inter-communication with schools."
>
> (Ravenscroft & Hartley, 1998)

Thus both sets of researchers suggest that the communication with these more traditional schools could be improved with the introduction of ICT. This would imply that the technology helps to foster a more holistic learning community independent of racial or economic background and location in the city. As can be seen this notion has formed part of the rationale for the work presented in this thesis which introduces a VLE at CHALCS (Barker, 1999a) where students can communicate across city and indeed country boundaries and thus learn collaboratively. However, to enable true integration between schools and CHALCS obstacles in terms of organisation and infrastructure would need to be overcome to enable students to access school work from CHALCS and CHALCS work from school. In reality this is likely to be some years away. The focus of this work, therefore, has been the introduction of the VLE to CHALCS and the use of the VLE by individual CHALCS students. The VLE was also piloted with a local school but it was not possible to develop a collaborative programme between the local school and CHALCS within the time-scale of the project.

1.2.4. The Research Inspiration: The Global Social Learning Club.

A pivotal inspiration for this research is the vision of a Global Social Learning Club proposed by Tak Wai Chan, extending his former ideas of Social Learning Systems which are *"environments where multiple participants, either computer simulations or real human agents, work at the same computer or across connected machines, taking various roles via a wide range of activity protocols*" (Chan, 1996). The 'activity protocols' include cooperative learning, competition or peer tutoring, for instance. Hence the Global Social Learning Club widens the Social Learning Systems vision to include the possibilities provided by the Internet and other modern ICT opportunities.

This vision of Chan's builds upon his earlier doctoral research concerning building artificial or Virtual Learning Companions (VLCs). Chan states "the goal of the learning companion is to stimulate the student's learning through collaboration, competition and demonstration" (Chan & Baskin, 1988). Learning Companions are based upon a social constructivist view of cognitive development initially proposed by researchers such as Vygotsky (1978). Chan in particular refers to Vygotsky's *Zone of Proximal Development* or ZPD (see section 2.2.1) which briefly refers to the difference in performance of a learner (apparent development level) when they work with a more capable peer as opposed to working without such a companion. However, Chan expands Vygotsky's framework to include collaboration with a VLC which he calls the ZPD-VLC. That is, the difference in learning occurring when a student collaborates with an average human peer as opposed to with a VLC. Although Chan states the need for a weak Turing assumption such that given a sufficiently low bandwidth of communication a student cannot detect that their collaborator is a VLC. Furthermore and fundamental to this thesis Chan points out that "having a VLC in the network can enhance the availability of synchronous learning since the VLC is always present" (Chan, 1996). That is a VLC can lead to widening educational provision because it is available 24 hours a day, 7 days a week.

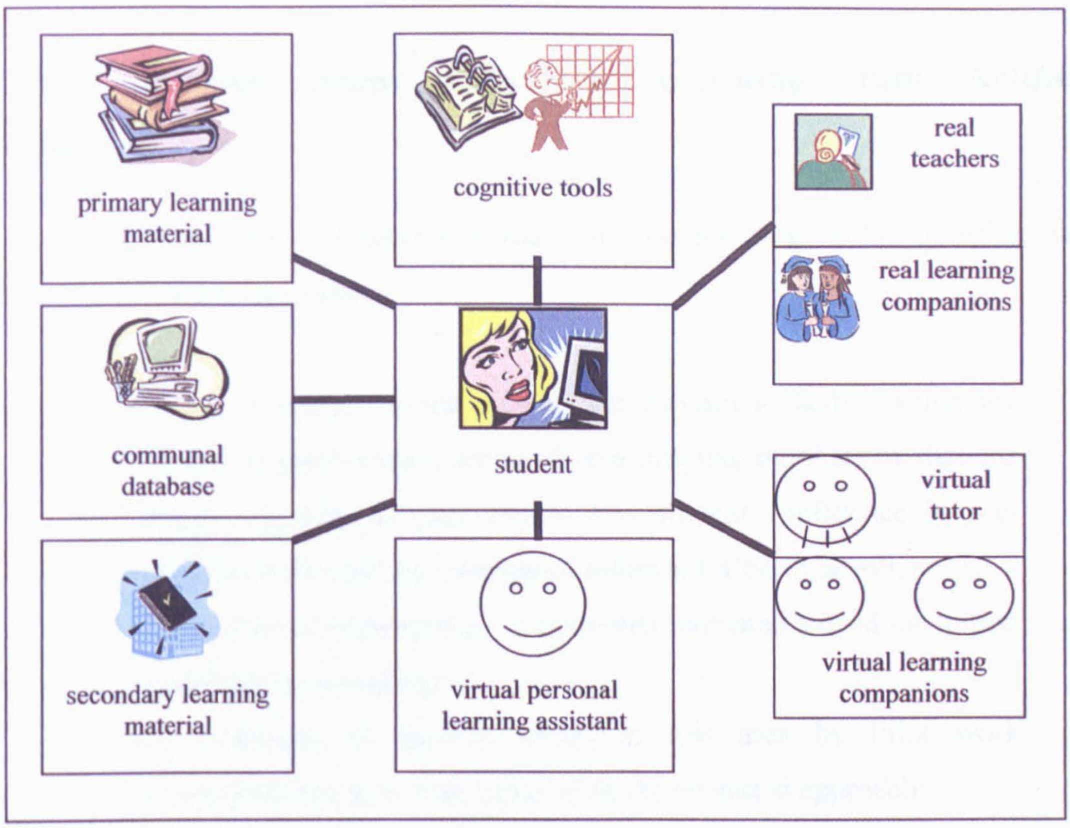


Figure 1.1 Student-Centred learning environment (adapted from Chan, 1996)

Figure 1.1 shows Chan's Student-Centred Learning Environment, a model which sees the student at the centre of a range of learning supports. These systems comprise of many elements that supplement the networked learning experiences of students. Extending this concept of learning environments the Global Social Learning Club provides a vision in which the student is part of a wider networked community existing to support all of their Educational needs. This is in keeping with the aims of the National Grid for Learning and more locally the Leeds Learning Network (although these initiatives have not yet embraced the notion of artificial learning companions as a component of learning networks).

To fully realise Chan's vision of a Global Social Learning Club, and in line with the objective of widening access we also require the Virtual Learning Companions that provide collaborative assistance when humans are not available. This will be discussed further in chapters 5 and 6.

1.3 The Research Aims: Effectively Learning from Artificial Companions.

As we shall see in Chapters 3 and 4 it emerged from initial questionnaires, observations and interviews that :

1. CHALCS wanted online support for Physics to help develop the subject (at present there are small and irregular numbers of students studying physics in part due to low student confidence in pre-requisite skills and the shortage of tutors at GCSE A level);
2. assessment of student help requirements indicated a need for tuition in effective note-taking;
3. an evaluation of specific needs in this area by Pilot work investigated the potential utility of an Agent-based approach;
4. based on these results it was decided to develop Agents to support note-taking in Astronomy and Optics.

Consequently, the aims of the research are set out in Table 1.4.

Research Aims
1. To design a pedagogical model to support both students and tutors working within a VLE in line with Chan's Global Social Learning Club
2. To evaluate existing VLE design tools and to implement the pedagogical model as far as possible with a particular selected VLE instantiated with materials and tasks to support an area of the curriculum appropriate to CHALCS
3. To evaluate the potential of this instantiated VLE for supporting the development of key skills within this area of the curriculum
4. To consider key skills in this curriculum area which might benefit from the addition of Agent-based support to the VLE and, focusing on one such skill, develop an exemplar Learning Companion to support it
5. To trial the Learning Companion to suggest Agent design characteristics which may prove motivating and effective in assisting students to learn from the VLE material

Table 1.4 The Research Aims

In the CHALCS context the key skill which emerged as the most suitable for the development of Agent-based support was that of summary note-taking for the reasons stated above. Summarisation is a useful first step in the 'acquisition' stage of the pedagogical framework (see section 2.1) in the form of revision notes for instance. Summarisation also complements the current literacy initiative at CHALCS by encouraging pupils to engage more deeply with the material they read (read for understanding) by selecting material, identifying lines of reasoning and progressing to form their own interpretation of material (QCA, 2002). Finally, if the summarisation task is successful in promoting a deeper engagement with the learning material it will also result in an increase in performance. This is likely to occur if the task is successful in promoting more active learning such as self-explanation and problem-solving behaviour typical of more collaborative styles of learning (Trafton & Trickett, 2001). The chosen curriculum area, preferred by CHALCS, is that of Advanced Level Physics. This choice was based on consultations with staff and observational studies carried out at CHALCS and a local school as discussed in Chapters 3 & 4.

In terms of CHALCS, the long-term aim of this research is for ICT to help promote inclusiveness within the local educational system, something which is an obviously prominent concern given CHALCS' history and raison d'être. In the medium and short-term the goal was to design a system that will support student's learning and tutor's instruction at CHALCS by exploiting ICT (specifically Internet technologies) to motivate students to engage deeply with study materials.

1.4 The Research Methodology: Qualitative and Quantitative Inquiry.

Educational research is divided into two camps. On the one hand there is the normative or positivist school with its roots in Aristotelian deductive reasoning and Baconian induction, i.e. start with the relevant data, generate hypotheses then validate them (Cohen & Manion, 1994). The desired result is a generic theory derived from observed phenomena. The alternative school however, emphasises the essential differences between the natural sciences and social sciences, notably that the object of attention in the latter is human and therefore not susceptible to the laboratory style experimentation as typified by the positivist approach. This post-positivist or interpretive approach is characterised by the involvement of the researcher in the process itself, the "human as instrument" (Lincoln & Guba, 1985, p.192) leading to a more inductive and holistic result.

The outcomes of these two schools are, very broadly speaking, two correspondingly distinct sets of methods. The normative school tends to prefer the collection of quantitatively-based data to test theories, for example employing statistical analysis, whereas the interpretive school lends itself more to qualitative investigation. An example of the latter would be the ethnographic approach derived from early research in Anthropology where the researcher integrates with the society under consideration so that a rich account can be derived and analysed.

A problem for the positivist tradition has often been that the manipulations required to control variables experimentally may distort what is studied to the extent that what is studied is not the natural phenomenon of interest e.g. by removing the phenomenon from its natural context to the laboratory. However, a problem for the interpretive school has often been one of subjectivity in interpretation and the difficulties in separating the influence of self or features of the local micro-context from real or general phenomena.

1.4.1. Methodological Triangulation

Given the educational objectives of this research the methodology adopted for this study, was mainly one of qualitative inquiry. As the studies were to be carried out in a working classroom environment it would not be appropriate to implement classical experimental methods, for example exploiting the use of control groups and random sampling. In such situations where the researcher must work alongside practitioners in authentic learning situations an experimental approach would not, in any case, be possible. Instead, the methodology adopted mostly takes a case-based approach and examines staff and students in their natural environment. This does not however obviate the need to adopt a variety of methods to be used in triangulation with qualitative observations. As Lincoln and Guba (1985) state triangulation is a necessary element of establishing 'trustworthiness' as examined in the next section.

Cohen and Manion (1994, p.236) provide an overview of Denzin's (1970) typology of triangulation. This overview consists of six separate techniques for achieving interrelation of results and thereby some validation of investigations. However they go on to say:

> "Of the six categories of triangulation in Denzin's typology, something like four have been used in Education. These are: time triangulation with its longitudinal and cross-sectional studies; space triangulation as on the occasions when a number of schools in an area or across the country are investigated in some way; investigator triangulation as when two observers independently rate the same classroom phenomena; and methodological triangulation. Of these four methodological triangulation is the one used most frequently and the one that possibly has the most to offer."
>
> (Cohen & Manion, 1994, p.236)

In the research reported here the first two methods have been employed where possible within the constraints of the context, such as investigating populations of students over time at the same institution and the use of more than one school population in the area of inner-city Leeds. However, *methodological triangulation* is the key technique employed. Both qualitative methods (such as interviews, classroom observations and video recordings) and quantitative methods (such as statistical analysis of web page 'hits' or other analysis of software or chat logs) are combined to provide as full an interpretive picture as possible. The sample size is necessarily small but the largely qualitative methods often give a clearer insight into the reasons why individuals respond in particular ways. Quantitative data, provided by the VLE for instance, is both more easily collected and analysed for larger numbers of interactions but although the results may have more generality their precise interpretation can be more problematic, and this is why a combination of approaches is preferred.

The practical details of the actual techniques used to produce this triangulation will be further elaborated within the subsequent chapters relating to the specific studies.

1.4.2. Generalisability of the Research Findings: Establishing Trustworthiness.

The notion of generalisability is it seems grounded in the reductionist, deterministic positivist approach. The notion that a grand rule can be deduced, or induced for that matter, runs contrary to currently prevalent notions of indeterminism which belie the foundation upon which reductionism and hence generalisability rests. Furthermore, generalisations are constructions formulated on the basis of limited samples of populations which are then said to represent the complete population. These constructions are "not found in nature; they are active creations of the mind" (Lincoln & Guba, 1985). The argument, is that any theory derived by induction is probabilistic leading to the belief that truth is relativistic and not absolute. It is also a fact that paradigms, especially those of Science, exist in a set time and context (Kuhn, 1962). This context is not a rigid entity but a shifting perspective, particularly in Social Science, and as such subject to the peculiarities of the time. Thus generalisations, to put it crudely, have a shelf life which undermine the utility of their formulation in the first instance.

The issue then which is central to the discussion of generalisations becomes one of trustworthiness. Table 1.5 shows the equivalent criteria used to establish trustworthiness in both the normative and interpretive research methodology schools of thought.

Normative Nomenclature	Interpretive Nomenclature
internal validity	truth value/credibility
external validity	applicability/transferability
reliability	consistency/dependability
objectivity	neutrality/confirmability

Table 1.5 Normative and Interpretive Concepts of 'Trustworthiness'

In the interpretive account 'truth value' is concerned with ensuring that the reconstructions of the observed reality are credible. Lincoln and Guba (1985, p.301) define various techniques to operationalise this such as "prolonged engagement, persistent observation and triangulation", peer review, resultant hypothesis refinement (or negative case analysis), referential adequacy (comparing derived theories with finer-grained data sources, such as video, as mentioned above) and reviewing constructions with members of the community being observed. Additionally, the researcher cannot hope to predict the novel context which will be introduced sometime in the future so in order to enable 'transferability' a '*thick description*' is required consisting of a thorough description of the context being observed and the actions or processes taking place within it. 'Consistency' is analogous to the positivist notion of replicability, that is the experiment can be repeated, under the same conditions, arriving at equivalent results. Consistency though in the interpretive school could at its simplest form rely on the notion, borrowed from normative methodologies, that validity is reliant on reliability and thus, mapping across the terms, credibility cannot exist without dependability (Lincoln & Guba, 1985, p.316). That is, if the kind of techniques exemplified above, such as triangulation, for establishing truth value (or credibility) are utilised then there is no reason to establish dependability independently. Triangulation can further be seen as a method of ensuring "confirmability" or neutrality which is analogous to the positivist notion of objectivity but instead, and importantly, shifts the emphasis onto the *data* rather than the researcher.

Each data source in the studies that follow in this thesis is listed in the relevant Chapter together with a rationale for their collection and a description of the particular analysis technique employed in the hope of furnishing a 'thick description'.

1.5 Summary and Outline of the Following Chapters.

The initial motivation for this research was a need for widening educational access and was established within the current global and local contexts. VLEs were suggested as a response to this emerging need. Of particular interest is Chan's vision of the Global Social Learning Club including provision of artificial Learning Companions based upon a social constructivist view of collaborative learning. Since beginning this research there has been a substantial push from Government to establish networks in schools which can exploit Virtual Learning. Additionally the UK Government is championing the notion of the necessity for Twenty First Century citizens to be proficient in key skills including IT literacy and the ability to learn from a variety of media and information resources. Furthermore, the specific context of this research was introduced, that is Chapeltown and Harehills Assisted Learning Computer School (CHALCS). Within this context, and in line with the goals of the organisation, the research programme aimed to:

1. design VLE support in an area of the curriculum with low student numbers and reliance on volunteer tutors;
2. test the potential of this VLE in this context;
3. identify residual support requirements which cannot easily be met by the VLE or the human tutor and which may be a candidate for Agent-based support (the chosen key skill was summary note-taking as justified in Chapter 3);
4. formatively design and test an exemplar case of such Agent-based support – in this case for summarisation;
5. evaluate factors affecting the effectiveness of the Agent-based approach primarily in engaging students with the VLE material and in guiding summary skills development.

The key skill of summarisation was introduced as an exemplar key skill providing a focus for this research within the subject area of Advanced Level Physics. Finally in this chapter, the methodological approach consisting of both qualitative and quantitative methods was outlined and justified particularly in terms of facilitating triangulation leading to establishing the generalisability or 'trustworthiness' of the results.

Chapter 2 will examine previous research into Pedagogical issues, the Collaborative Learning methodologies, Summarisation issues and Affective Computing and will outline the Pedagogical Model to be adopted. Chapter 3 begins by evaluating possible VLEs then proceeds to explain the reasons behind the selection of WebCT and the development of the Physics course implementation within WebCT. Chapter 4 concentrates on the specific studies into students' methods and problems in working with the WebCT course concluding with some detailed recommendations concerning their support requirements. This is followed, in Chapter 5, by a review of artificial Learning Companions then a study which sought to inform the design of the Learning Companion reported in this research. Chapter 6 describes the further design and implementation of this Learning Companion to support students taking notes within the WebCT Physics course. Chapter 7 then details the evaluation of the Learning Companion exploring issues such as affective communication, student motivation and the Companion's effect on student's note-taking behaviour. Chapter 8 illustrates the final system in the form of a 'walkthrough'. Finally, Chapter 9 evaluates the research as a whole and includes a summary, conclusions and directions for future research.

Chapter 2: A Review of Previous Related Research.

The intention of this chapter is to place the research in the context of four key areas which together form the foundation for the work. First a Pedagogical Framework for the successful execution of a VLE is developed based on a review of the Psychological and Educational literature. Secondly, the Collaborative Learning literature is briefly reviewed to inform the design of the Agent with which students may collaborate on the summarisation task. Thirdly, relevant research on Summarisation is reviewed in order to suggest strategies by which the computer-based Agent might aid summarisation. Finally, research on Affective Computing is introduced to suggest relevant characteristics which might motivate and encourage students to interact with the Agent on the summary note-taking.

2.1 The Pedagogical Framework: The Three A's.

A pedagogical framework with three key elements is being proposed for the effective implementation of the ICT-based educational programme at CHALCS. The framework has emerged gradually, is rooted in the literature and developed through interaction between members of the course design team, students and teachers. The three aspects of the framework: Acquisition, Argumentation and Application derive from learning theories and empirical studies reviewed in this chapter.

2.1.1. The Development of the Pedagogical Framework.

Learning has been classified according to many different typologies, perspectives and theories of learning which are inevitably susceptible to interpretation within historical and cultural contexts. Bigge and Shermins (1999) have attempted to classify these learning theories into three camps - *theistic*, *behaviourist* and *interactionist*.

A theory of learning prevalent *before* the Twentieth Century, considered learning a product of 'mental discipline'. In this theory the mind is considered a 'muscle' which has to be exercised. Corresponding pedagogies employed *Theistic* methods such as drill and practice with punishment for students who did not sufficiently learn the subject matter at hand.

However, during the Twentieth and Twenty First centuries two new types of learning theory emerged. The first are grouped into *Stimulus-Response* (S-R) theories of the behaviouristic school and the second can be grouped under *interactionist* theories which rest upon the assumption of underlying cognitive processes. S-R theories argue that "learning is a change in observable behaviour, which occurs through stimuli and responses becoming related according to mechanistic principles" (Bigge & Shermins, 1999). These theories suggest extrinsic reward can be used to manipulate motivation and hence learning. They utilise conditioning procedures such as stimulus substitution and reinforcement. In these procedures reward is contingent upon desired behaviour. The desired behaviour is broken down into small steps towards the goal and is shifted towards the goal so that whatever gained the reward yesterday is not sufficient to gain the reward today.

For instance, Skinner's Operant Conditioning (Skinner, 1953) consists of Positive Reinforcers – rewards that could include verbal praise, a good grade or a feeling of satisfaction as well as more obvious rewards such as food, tokens or money. These positive reinforcers strengthen behaviour – in other words positively reinforced behaviour will reoccur. Negative Reinforcers on the other hand result in the removal or avoidance of behaviour. Negative reinforcers punish behaviour and involve using an aversive stimulus such as criticism or smacking, the removal of rewards (response cost) can also be experienced as punishment. Skinner found that applying negative reinforcers when a desired behaviour was absent was ineffective, negative reinforcers are only effective in reducing unwanted behaviour that is part of the current repertoire of behaviours produced and then only if the negative reinforcer immediately follows the unwanted behaviour. If there is a delay then the negative reinforcer may only succeed in changing the last behaviour. For example, if you own up to something and receive criticism what may be reinforced is that owning up results in punishment not that the original bad behaviour results in punishment.

It is far more effective to reward accidental or naturally occurring instances of good behaviour than to withhold reward because no such behaviour has been produced. This is probably because it is unclear to the learner what the new required behaviour is or that such behaviour will be rewarded. It is unfortunate that in parenting and in education good behaviour is often ignored whilst bad behaviour is punished leading, overall, to a negative rather than a positive learning experience. As with punishment it is important that positive reinforcers are given immediately following the desired behaviour and that they continue once the behaviour is established (if after a while the good behaviour never receives reward then the good behaviour may drop off). The key to shaping behaviour in the direction desired and to developing positive and motivating learning experiences is to continue to give reward for good behaviour and to reward it as close in time to its occurring as possible.

In order to 'shape' behaviour toward a new goal it is necessary to reward the behaviour currently being produced that is closest to the goal. Gradually, the learner will produce more and more goal-like behaviour and some of it will be more goal-like than others. Reward for less goal-like behaviour amongst this new (generally more goal-like) behaviour can now be gradually withdrawn so that only even more goal-like behaviour now receives reward. By breaking down a new behaviour to be learned into a series of structured easy steps, the first of which is something the animal or child can already do and so can be rewarded, behaviour can be gradually 'shaped' toward the goal and the overall learning experience remains positive. These principles led to the notion in Computer Based Learning (CBL) that information should be presented in small amounts and in ascending order of difficulty with the starting point always well within the student's current grasp so that a correct response could be positively reinforced. Thus branching programs took children through at their own pace with the emphasis on each child, whatever their current level, always achieving a high success rate in answering questions. The immediacy of feedback (positive or negative) which the computer can provide through its one-to-one interaction remains one of the most important qualities that computer based learning has to offer.

Interactionist theories, on the other hand, are concerned not with behaviour but with mental processes and structures that are created in the mind of the learner. These theories seek to explain what the individual *understands* as a result of *interaction* with a teacher or *psychological environment*. The emphasis is on the construction of appropriate cognitive structures within the student's mind and is typical of the approach of constructivists such as Bruner (1966), Piaget (1970) and Vygotsky (1986). These writers suggest that knowledge is actively constructed out of experience through interaction in the physical and social worlds and that the child actively seeks interaction to provide a model for the development of new theories about how the world works that he or she can use to guide action in the world. These ideas are elaborated further in the section on collaborative learning (See section 2.2.2).

For the moment synthesising these stances: students acquire information and/or skills through interaction or direct experience and then seek to reinforce and refine this knowledge through discourse with others and by applying it in new settings and contexts to test its adequacy. New ideas are integrated with old ones, conceptual understanding is refined and appropriate cognitive structures are built, enabling application in new situations. Such new situations form a test-bed through which reflection and a new cycle of learning may occur, e.g. see Kolb (1984). In this way meta-cognitive skills and strategies emerge – children begin to learn how to learn and become increasingly autonomous and responsible for their own learning, developing stable and successful general approaches to learning. Such approaches and strategies may be termed transferable skills. Through acquiring such skills (an example of which might be note-taking) it is hoped that students will be able to learn to manage their own acquisition processes, being able to acquire new knowledge by bringing such skills to bear on new domains. Since such general skills and approaches require practice and opportunity for application across a range of contexts, for example see Gagne (1975), there is a need for appropriate tasks that help students develop them.

In the following sections a pedagogical framework for enabling learning through aperceptive, interactional and experiential processes is synthesised and elaborated based on a review of the literature.

2.1.2. Phase One: Acquisition.

"**acquire**…**1** come to have **2** learn or develop (a skill or quality)…
ORIGIN Latin *acquirere* 'get in addition'"

(Soanes, 2001)

This phase of the Pedagogical Framework is obviously the initial contact between the student and the learning material. As the above definition states we are concerned with the student learning new facts and figures but perhaps more importantly, skills which will not only furnish them with the new knowledge that we are attempting to impart but also provide a means of manipulating that knowledge. In the parlance of symbolism we are facilitating construction of mental models (Johnson-Laird, 1983) which students can then subsequently manipulate mentally, test through action in the world and upon reflection revise.

Gagne (1975) emphasises the 'domains of learning', that is the "primary categories that limit the generalizability of conclusions about the learning process". His five domains are:

1. **Motor Skills** "mediate organized motor performance" and are acquired through practice;
2. **Verbal Information** or factual 'knowledge' ideally presented in "an organized, meaningful context" if it is to be understood and not simply acquired by rote;
3. **Intellectual Skills** are the ability to make discriminations from concepts and generate and apply rules. These "require prior learning of prerequisite skills";
4. **Cognitive Strategies** are "skills that govern the individual's behaviour in learning, remembering and thinking";
5. **Attitudes** which are best learned vicariously and are cultured through participation in a learning community or community of practitioners.

Gagne (1975) states that recognition of these domains of learning and their different attributes are prerequisites to appropriate instructional design which will, importantly, vary between subjects. In the acquisition phase verbal information or factual knowledge may be "acquired" through listening to a teacher or reading course materials such as those presented in a VLE (provided that they are presented in a structured and coherent fashion). However, in order to develop intellectual skills and cognitive strategies other tasks in the argumentation and application phases are required. In these phases students articulate their ideas and suggestions, with each other and the tutor, and test the limits of these within supporting activities that serve to give a meaningful context in which to practice developing skills.

A key element of acquisition was highlighted by Bruner (1975) who recognised the trait of 'going beyond the information given' by coding the information we receive from the senses into some more readily recognisable form. That is, forming *associations* between that which we already know and that which presents itself as new. These coding systems may subsequently be combined, from quite different subject fields, to allow for predictions leading to novel discoveries. This suggests that the primary phase of acquisition forms a basis for the furtherance of novel associations leading to self-actualisation of the student. That is, through the interaction between prior knowledge and new information new and meaningful mental models are built leading to creative constructions in turn giving rise to opportunities for creative expression that can be intrinsically motivating and have a positive effect on self-esteem.

A word of warning is however sounded by Ausubel (1985) who discusses the difference between *rote* and *meaningful* learning noting that, in contrast to the former, meaningful learning implies the ability to transform or manipulate as well as to recount information. Moreover, if the student is unable to find meaningful overlap between existing prior-knowledge structures and new information to be assimilated there is a danger that the information may:

- Not be assimilated at all (or deleted from memory before a stable long-term structure is built);
- Not integrated correctly with prior-knowledge structures but isolated in long-term memory from them, and so not recalled appropriately in a range of appropriate contexts;
- Integrated with inappropriate structures (a 'forced fit') resulting in misconceptions. See also, for example, Kintsch & van Dijk (1978); Britton, Glynn et al. (1985); Mannes, & St. George, (1996).

This seems to reiterate that prior knowledge, meta-cognitive strategies and skills such as Gange's "intellectual skills" are just as important as raw facts if learners are to construct mental models, such as those described by Johnson-Laird (1983), from verbal information and successfully access and use these models later in a range of appropriate contexts. Ausubel continues to further decompose meaningful learning in particular into *concept* and *representational* learning and posits an 'assimilation theory' to explain how new information is integrated with existing cognitive structures in a hierarchical fashion. The distinction between *accommodation* and *assimilation* is an important one as the latter requires no critical thinking whereas the former requires deeper understanding. Relating this to Gagné's 'domains of learning', 'assimilation' relates simply to the acquisition of 'verbal information' whereas 'accomodation' necessitates the learner employing some basic intellectual skills as well as cognitive strategies.

It is possible to devise learning tasks that help students acquire the meta-cognitive knowledge needed to develop intellectual skills, cognitive strategies and the 'key skills' mentioned in section 1.2.2. The ability to manipulate and transform information to create new meanings is of necessity a creative process. Writing tasks can be viewed as involving such creative processes and can also be used to promote creative knowledge transformation with the right guidance (Berieter & Scardamalia, 1987).

To summarise, this section has highlighted some of the research on how new knowledge is assimilated and accommodated within existing knowledge structures if it is to be understood and used effectively. Moreover, for this to occur, new information must be presented in a meaningful context – a context which relates the new information to the learners' prior knowledge. This provides us with some guidance for instructional design. In addition, work on collaborative education (investigated in section 2.2) further highlights the potential benefits of sharing the acquisition task with a tutor or peer. With this in mind the next phase in the Pedagogical Framework, argumentation, can be seen to naturally follow by allowing a theatre of practice where students seek to reassure themselves of their acquired knowledge or are challenged to construct new meanings through interaction with peers or a significant other.

2.1.3. Phase Two: Argumentation.

"**argue**...1 exchange conflicting views heatedly.
2 give reasons or evidence in support of something...
ORIGIN Latin *arguere* 'accuse'"

(Soanes, 2001)

At this phase in the Pedagogical Framework it is necessary to corroborate internalised knowledge by seeking evidence, comparing or explaining, possibly in a social context (as in the case of collaborative group work – see section 2.2) thus validating the mental models acquired in the acquisition phase. It is proposed that argumentation is one important means by which this takes place. It is useful then to examine interpretations of the meaning of argumentation.

It is apparent from the above definition that in the formal sense 'arguing' is a form of reasoning (clearly not always a dispassionate form as "heatedly" and the word's origin imply but a form of reasoning nonetheless). In the Natural Sciences this usually implies deduction and induction. Johnson-Laird (1988, p. 220) proposes a taxonomy of thought which includes these forms of reasoning and others, notably 'daydream', 'calculation' and 'creation'.

However, taking a broader view of the term and applying it to the written form "argument" is used to denote a particular style of writing which seeks to explain, justify or persuade through the logical presentation or sequencing of ideas. Andrews (1995) proposes a variety of approaches to helping students argue such as teaching linguistic techniques, exploring the connection between narrative and argumentative composition in both speaking and writing, examining planning strategies and the provision of argumentative resources (eg. advertisements, letters of complaint, speeches, etc.). There is a notion here though that argumentation does not just produce a product but is an actual *process*. This process is of interest here in terms of supporting the kinds of intellectual skills that Gagne (1975) highlighted – see the previous section. Any work which seeks to provide this support must take into account arguing in its written form as a process as well as a product.

This distinction between process and product is seen as typifying the difference between the behaviourist and interactionist schools of thought highlighted in section 2.1. In the process view knowledge is selected, manipulated and presented to serve a goal and usually with a specific audience in mind. However, when we choose to play both sides of an argument (when reasoning for instance) we do so as a means of applying a systematicity, objectivity or rigour to our thinking that serves as an aid to problem-solving or decision-making. We thus test the completeness and consistency of our own thinking. In debate we rely on our interlocutor to perform a similar function, the joint product forming such a dialectic.

In terms of tractable solutions to support argumentation, work by Collins (1977) has elucidated 24 rules which comprise the Socratic method of teaching. In this method the student learns through systematic exploration of variables affecting cases and their interrelations. Reasoning skills such as how to form and test hypotheses, distinguishing between conditions and making uncertain predictions are encouraged. Collins' rules have been used to design an automated support system using a subset of the rules to model the Socratic dialogue (Stevens et al., 1982). This method is of interest in the context of this research as it attempts to decompose argumentation into manageable rules which could easily be employed by a tutor or, indeed, an artificial Agent.

A related, partially implemented framework is proposed by Ravenscroft (1997) in his "Learning as Knowledge Refinement" which proposes an Intelligent Tutoring System (ITS) dialogue based on Vygotskian principles aimed at helping students overcome misconceptions in Physics based on dialogue game theory work of Pilkington (1992) and Moore (1993). He uses his developed pedagogy as an "instructional method for achieving educational goals" and sees his knowledge refinement aim as the next paradigm progression in dialogic tutoring systems using both dialogue structure based strategies and a generalised domain model to test for completeness and inconsistency in reasoning. Learning as knowledge refinement is thus defined as a process which:

> "requires a constrained dialogue to be maintained in which the learner expresses and refines qualitative causal explanations, examining and revising their assertions and conceptions to form complete, consistent and explanatory models."
>
> Ravenscroft (1997)

There is a hope then that this argumentation phase will encourage the kinds of reasoning that will not only lead to evaluation and reflection of the results of a discussion but also cause some reflection on the processes and tactics that led to the results. This would be particularly pertinent in collaborative group work where, for instance, the tutor engages students in discussion or debate using a process of knowledge refinement then a transcript of a chat session in a VLE could be posted and reflected upon.

However, although this argumentation phase is an important validation phase it is notoriously difficult to achieve lasting conceptual change through such dialogues. The conceptual change in Physics programme (Hartley et al., 1991) found students may readily appear to change their conceptions when backed into a corner by the timely presentation of a counter-example only to revert to previous conceptions outside the experimental situation.

For genuine accommodation to occur there is a need for the student to actively seek evidence, compare and explain for themselves. This entails opportunity to test and apply knowledge in a valid psychological environment such as that proposed in the final phase of the Pedagogical Framework. Thus mental models which have been formed through acquisition then validated through argumentation are finally tested in a novel situation indicating their worth to the student.

2.1.4. Phase Three: Application.

"**apply**…**1.** make a formal request for something…**2.** bring into operation or use **3.** be relevant **4.** put (a substance) on a surface **5.** **(apply oneself)** put all one's efforts into a task…

ORIGIN Latin *applicare* 'fold, fasten to'"

(Soanes, 2001)

For students to reinforce reasoning skills and domain knowledge it is proposed that they will engage in appropriately designed tasks. Aside from their argumentative skills they will also need to practice their crucial problem-solving skills for such skills are valuable in every-day life but also (in our case), as Physicists, these skills will be used throughout their academic career. As an example, the following problem is to be found in the General Certificate of Education (Physics), 1996 in the "Astronomy and Optics" module in relation to depth of focus and depth of field in a simple variable focus single lens camera :

"Calculate the ratio of light transmitted when different sized apertures are involved".

Such problems require the identification of a set of appropriate procedures and the correct application of these procedures to arrive at a suitable answer. The ability of the student to identify the category of problem and hence the appropriate procedural steps will depend crucially on effectiveness of acquiring appropriate facts, procedures and/or the cognitive strategies and intellectual skills needed to identify or match the appropriate approach to the characteristics and features evident in the new problem statement. In symbolic terms both domain rules and generic task operators that can identify and manipulate these correctly are needed.

Symbolic Artificial Intelligence (Newell & Simon, 1976) in particular has studied problem-solving in great detail as it was deemed central to the notion of 'intelligence'. However, it was typical of this early pioneering work to choose problems that were easily reduced to computational models, such as the Blocksworld or Chess domains, e.g. see Russell and Norvig (2003, Chapter 1) for a synopsis of the history of AI.

However, the symbolic approach did not take into account human biases in purely deductive reasoning (Wason & Johnson-Laird, 1973). As a result some AI researchers turned to nature to solve problems such as mimicking evolutionary techniques (Langton, 1997), using types of situated stimuli-response models (Brooks, 1991) or, indeed returning to sub-symbolic or connectionist models (Rumelhart & McClelland, 1986). Nevertheless, AI research into problem-solving provides a foundation with which to further investigate human problem-solving strategies and provides valuable insights into the design of computer-based aids such as artificial peers or tutors that would work with students on learning tasks or problems. Although it is beyond the scope of this thesis to develop such artificially intelligent agents capable of solving new physics problems or coaching students in developing problem-solving techniques. Furthermore, the complexity of developing such agents for the more open-ended problem-solving task of summary note-taking is even more problematic. Nevertheless this research can help inform less knowledge-rich approaches. This will be discussed further in Chapter 6.

Chi and Bassock adopted a case-based approach to the development of problem solving skills utilising examples to guide student's learning. Their focus was: "how the declarative knowledge encoded from text or from the teacher's instruction becomes proceduralized into a skill" (Chi & Bassock, 1989). They were further concerned with whether or not students correctly assimilated domain knowledge (facts and domain-specific reasoners) and the degree to which these are "understood". They are sceptical of generalised heuristics and prefer instead "general domain principles" such as "the sum of all the forces acting in the system is zero" on the basis that students do not easily develop meta-level problem-solving heuristics that can be readily transferred to new domains. Thus, Chi and Bassock (1989) conclude that whilst it is desirable to teach learners better learning strategies such as to more closely monitor their comprehension so that this triggers the formation of explanations this is not easy, especially for novice learners. However, Chi and Bassock concluded that students who were able to generate such explanations for themselves (self-explainers) were more successful in solving problems. Moreover, others have suggested that the role of the tutor in prompting timely generation of such explanations (or reflections on performance) by students may also serve to assist problem solvers and/or the scaffolding of problem-solving strategies. See, for example, Schwartz (1995), Bielaczyc et al. (1995), Pilkington & Parker-Jones (1996), Hume et al. (1996) and Katz (1997).

It is important to recognise the differences between novice and expert problem-solvers. Evidence from a range of problem-solving contexts suggests that novice problem-solvers are more likely to be distracted by shallow rather than deep attributes of the problem. Hence they are more likely to be distracted by surface differences in the ways in which problems are presented and less likely to identify deeper structural similarity with previously solved problems. In contrast, expert solvers more readily categorise problems on the basis of abstract, deep or structural features enabling them to generalise more readily to previously solved problems, suggesting a possible course of action, see e.g. Newell & Simon (1972).

Precisely how experts do procedularise problem-solving methods or notice common features of problems is not well understood. Undoubtedly, the advice to novices who wish to become experts would be 'practice makes perfect' hence the need to supply real authentic problems. Additionally, it would be important to recognise these important differences between experts and novices in a tool which aimed to support problem-solving, be it specific Physics problem-solving or a problem-based approach to effective summary note-taking. In particular, in contrast to novice summarisers, expert summarisers are more likely to produce shorter summaries than novices whilst also producing summaries that contain more information rated "important" and deleting more information rated "low-level detail" or less relevant or important (Johns, 1985; Tawalbeh, 1994). Supporting learners in identifying what is important to understanding and not being distracted by more shallow aspects of the presentation will involve working with real exemplar texts and adopting strategies for selection and condensation of material.

Collaborative problem-solving offers more than an economic opportunity to practice problem-solving as students share resources. Collaboration also exposes students to alternative viewpoints and corresponding explanations thereby challenging their own ideas. This in turn can facilitate the development of deeper and more meaningful conceptual understanding.

2.2 The Need for Collaborative Learning.

As Education moves from traditional settings such as schools and universities into more 'distance' or 'flexible' methods of delivery as championed, for instance, by the UK Government's drive for Lifelong Learning there is a corresponding need to develop more independent study skills. Many advocate that in order to develop these there is a need to move away from more prescriptive or didactic teaching methods towards more constructive educational practices where a learner is immersed in community of peers who learn collaboratively as a group rather than as individuals (Johnson, 2001). In this model the tutor is, by necessity, less of a 'sage on the stage' and more of a 'guide on side', meaning that the educationalists role also shifts from being one of prescriber to one of enabler of co-constructed knowledge.

These 'distance' learners do not necessarily have access to tutors or indeed peers exactly when they need assistance – they could be in completely different time zones or working unsociable hours. Thus if this new generation of flexible learners are to feel empowered to manage their own learning experience and therefore reclaim their learning independence, there is a need to provide methods for them to pursue learning activities when it suits them. One means of achieving this is by providing infrastructures such as the National Grid for Learning (reviewed in Section 1.2.1) thus facilitating networked learning (Barker & Pilkington, 2000). As we have proposed, a further means would be the provision of more advanced computational assistance, utilising, for instance, AI techniques in the design of artificial Learning Companions (Chan, 1996) which can be engaged in collaborative learning activities.

However, limiting the current discussion to the specific context of this research such innovations in Computer Supported Collaborative Learning provide a potential solution to two problems as CHALCS: the voluntary nature of staff and the small number of students, particularly in A level Physics. First, tutors would be able to offer one-on-one tutoring of students outside class hours and at a distance. This does not necessarily mean that students would receive less tutoring but that this tutoring would be available more often and "just in time". Additionally, if students were using the Internet to contact experts in the field, for instance, this would also allow a tutor to spend time with students who specifically require *their* attention. Secondly, again if students are using ICT to collaborate with a peer on a problem at a distance (for example, with students studying the same syllabus in different schools) then the limits imposed by a small class size, such as finding a 'suitable' collaborating partner need no longer apply. Furthermore, this peer need not be human but could be an artificial Learning Companion, as is proposed by this research. Thus collaboration partners are just a mouse click away.

These notions are no longer fanciful ideals with the advent of web-based Virtual Learning Environments (Barker, 1999a; Dillenbourg, 2000) that provide a communications framework for students' collaborative learning experience. An appropriate guiding Pedagogical Framework has already been introduced in section 2.1 for the introduction of such VLEs. It was suggested there that such environments need to enable students to move between phases of Acquisition, Argumentation and Application (not necessarily in this order but with opportunity for all three) if they are to effectively master all the domains of learning necessary to be competent in the subject. In part this was based on 'social constructivist' or 'experiential' theories of learning as characterised by the pioneering work of Vygotsky (1978) amongst others. These theories give central importance to the need for interaction both with the environment and with a peer or tutor if deep or meaningful learning is to take place. Such theories underpin models of Collaborative Learning and, in the next section, these models are expanded upon.

2.2.1. Defining Collaborative Learning

As previously mentioned in section 2.1.1 the driving rationale behind students working together on a task is that they are exposed to their peer's viewpoints on possible solutions thus forcing them to challenge their own (mis)conceptions and explain their own reasoning before arriving at a conclusion or decision. However, distinction needs to be made between Cooperative and Collaborative Learning although this is not an easy distinction to make as sometimes they are used synonymously in the literature, see for example Burton et al, (2000). Roschelle (1992) attempts to distinguish between the two methodologies: collaboration involves the students maintaining a shared conceptualisation of a problem whereas cooperation simply involves each student working on a sub-problem independently. However, Burton et al. (op. cit.) note this distinction may sometimes be a matter of level or granularity. In other words, at some finer level of description individual agents within a collaboration will be working to some extent independently and have their own differing perceptions of the problem at different times. Collaborative learning methods such as the jigsaw method (Aaronson et al., 1978) may explicitly mix phases of collaboration in which there is a sharing of the conceptualisation of the problem and negotiation of problem-solving strategies with phases in which individuals are working co-operatively on sub-parts of the problem, later coming together to re-share perceptions of progress and collaboratively negotiate or plan further phases of independent working. Therefore the key point to discriminate between collaborative and cooperative learning is the extent to which the task is shared or divided amongst learners, that is their symmetry or, indeed, asymmetry in both the division of labour and understanding of the current state of the task. As Dillenbourg (1999a, p.11) states "in cooperation, partners split the work, solve sub-tasks individually and then assemble the partial results into the final output. In collaboration, partners do the work "together"."

Dillenbourg (1999b) provides a good overview of Collaborative Learning issues where the distinction between cooperative and collaborative interactions are but one facet of what it means to be engaged in Collaborative Learning. Apart from this *situational* aspect of collaboration other aspects include the nature of the *interactions* such as the interactivity ("the degree of interweaving between interaction and action", p.12), the synchronicity (or asynchronicity) of the communications and the negotiability of shared views ("when two partners misunderstand, they have to build explanations, justify themselves, reformulate statements and so on – all of these activities that can lead to learning", p. 14). This latter point is something which any artificial Learning Companion must also attempt to capitalise upon if it is to claim all the benefits of collaborative learning for its interactions with students.

The other two defining aspects of collaborative learning posited by Dillenbourg include processes and effects where, *processes* are the adoption of mechanisms appropriate to collaborative learning, such as inductive correlation of joint-constructed representations, cost/benefit analysis of sharing cognitive load, self-explanations (particularly as extrapolated to group situations) and both self and social conflict. Other collaborative processes include internalisation and appropriation or grounding (reaching a common conceptual framework) necessitating a model of the collaboration partner possibly leading to "differential reasoning", i.e. highlighting discrepancies in one's own model by comparing it with that of the partners. This may be viewed as an argument for user or student models in artificial collaborating agents helping to work towards this 'grounding'. On the other hand, *effects* of collaboration concern not those learning effects of individuals but instead the achievement of the group. These could be measured a priori, i.e. an 'experimental' approach or a posteriori, i.e. a 'qualitative' approach such as dialogue analysis although Dillenbourg states that it is perhaps more appropriate to "zoom in" on the collaborative interactions – such as the detailed logs of interactions provided in CSCL environments (and utilised in this research).

Thus, for Dillenbourg (1999b, p.17), a theory of Collaborative learning must include "criteria for defining the *situation* (e.g. symmetry, degree of division of labour), the *interactions* (e.g. symmetry, negotiability), *processes* (e.g. grounding, mutual modelling) and *effects*." Additonally, Kanselaar et al (2001) would add *task design*, *mediating tools* and *collaborator's characteristics* to Dillenbourg's definition of collaboration. Hawkridge (2000) further highlights Dillenbourg's omission of task design in the definition. It is not possible to tackle all of these issues within the scope of the work reported here, nevertheless these issues need to be considered in designing VLEs in which agents are intended to act as Learning Companions and will be returned to in discussing the design of the artificial Learning Companion in Chapter 6.

2.2.2. The Theoretical Basis of Collaborative Learning.

The primary theoretical basis for Collaborative Learning comes from the "constructivist" learning theories of Piaget (1896-1980), Bruner (1915-) and Vygotsky (1896-1934). Common to these authors is the notion that learning is an active process in which learners construct new ideas or concepts out of experience. The learner selects and transforms information, constructs hypotheses, and makes decisions, relying on cognitive structures to do so. Vygotsky, in particular, emphasised the potential role of the instructor in helping students to discover principles for themselves by engaging the student in an active dialogue whilst Bruner further suggested that the task of the instructor is to translate information to be learned into a format appropriate to the learner's current state of understanding and that the curriculum should be organized in a spiral manner so that the student continually builds upon what they have already learned.

It has often been suggested that Piaget emphasised the way in which the child interacts with or manipulates the material environment to construct conceptualisations and theories out of experience and to test these whilst Vygotsky emphasised the importance of social interaction (particularly with a more mature individual) in the development of such concepts and theories and in the challenging and revising of these. However, Cole and Wertsch argue cogently that this difference between the two in these respects has often been exaggerated to make the point (Cole & Wertsch, 1996). It would be fairer to say that both recognised that both interaction with the material world and social interaction are means by which children formulate conceptions about how the real world operates and, as they attempt to apply these conceptions in practice or test them against the conceptions of others, inconsistencies force them to accommodate new theories and expectations based on these experiences.

Thus, Vygotsky agreed with Piaget on the role of the internal dialectic but extended Piaget's work in emphasising the importance of language and communication with others as a mediating factor in a child's development of concepts. This led to

Vygotsky experimenting with language-independent tasks with children. A conclusion of this work was a new emphasis on the role of internalised language in directing the child's own internal dialectic reasoning processes.

As the child develops their conceptions through this internalised mediation Vygotsky became interested in the role that a more mature collaborator can play through the use of dialogue. Vygotsky considered the difference between a child's performance alone and with a more mature partner such as a tutor which he called the **Zone of Proximal Development** (ZPD), defined as:

> "the distance between the actual development level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers"
>
> (Vygotsky, 1978, p.86)

Vygotsky goes on to eloquently say of the ZPD that it:

> "defines those functions that have not yet matured but are in the process of maturation, functions that will mature tomorrow but are currently in an embryonic state. These functions could be termed the "buds" or "flowers" of development rather than the "fruits" of development. The actual developmental level characterizes mental development retrospectively, while the zone of proximal development characterizes mental development prospectively."
>
> (Vygotsky, 1978, p.86)

Thus Vygotsky argues, social interaction with a more mature or experienced individual activates the ZPD and hence the potential development level, mostly through a process of imitation. Vygotsky further argues that learning must precisely aim for this advancement of development through an internalisation of the more advanced developmental processes occurring as a result of this interaction. This research intends to achieve these goals by employing an Agent or artificial Learning Companion which is "more capable" in terms of the summarisation task than the student using the system in the hope that the student will enter the ZPD, internalise some of these developmental processes, more precisely those related to the meta level study task of summarisation, and hence learning will occur as their developmental level approaches its potential. Additionally, the student will begin to 'acquire' the material presented in the VLE through this process of active engagement as they hopefully begin to imitate the VLC.

2.2.3. Models of Collaborative Dialogue.

An appropriate methodology must be employed to guide collaborative dialogue between peers and between peers and their tutor. A number of methods have been stated in the literature and will be examined here, notably reciprocal inquiry dialogue, Socratic dialogue and argument negotiation.

Reciprocal inquiry dialogue (Rosenshine & Meister, 1994) is actually not a method of collaborative learning as such but a method for enabling tutors to encourage students in question-formulation. The idea is that by encouraging students to ask appropriate questions of each other rather than answering the questions posed of them by the tutor, initiative for learning is handed back to students resulting in more active and constructive learning. The hope is that by providing a systematic method by which the tutor can gradually pass on control of discourse to students, students learn to formulate questions for themselves without help from the teacher which results in them becoming more independent learners.

As mentioned in Section 2.1.3 Socratic dialogue , see Collins (1977) provides a template for expert-guided inquiry dialogues by providing rules for interactions modelled on scientific inquiry. For example, an opening move might be 'hypothesis identification' where a student will state a hypothesis, claim or prediction to open up the debate. This could be followed up by a 'hypothesis evaluation' where the tutor asks the student to justify their hypothesis. Further possibilities include counter-examples, systematic variation of cases (i.e. introducing extreme examples) and even entrapment which would force the student to state a misconception before having to justify it. However, Socratic styles of dialogue could appear aggressive and require careful handling in a truly congenial collaborative inquiry, such as with a (artificial or human) peer.

Argument (or debate) has already been highlighted in section 2.1.3. Suffice it to say that argumentation in a collaborative setting should be constructive rather than destructive implying a degree of empathy between peers and with the tutor and shared will, i.e. commitment to a common objective. Providing it is constructive, argument can be more beneficial than simple agreement. For example, Burnett (1993) found that collaborative discussion that involves "substantive conflict" was more likely to improve writing quality than discussion which involved high levels of simple agreement. Mercer & Wegerif (1999) made a similar distinction concerning children's talk. They categorised talk as either exploratory, disputational or cumulative and concluded that "exploratory talk" involving neither simple disagreement nor simple agreement but the expression of alternative views together with explanations improved children's reasoning skills. Walker & Pilkington (2001) suggest some of the social and discourse behaviour that tutors need to encourage if an environment in which argument can be constructive is to be created. Such polite behaviours include encouraging students to validate each other, encouraging each other to take part in the conversation (turn sharing) and avoidance of personal criticism. Note that encouraging these behaviours involves giving immediate and direct positive feedback and avoiding unwanted adverse stimuli (as discussed in section 2.1.1) and are potentially important in 'shaping' or modelling future behaviour. Such qualities and behaviours and their implications for the design of a learning companion will be discussed in section 2.4 on Affective Computing and Chapter 5.

Finally, the work of Ploetzner and Kneser (1998) examined the collaborative interrelation of pairs of students taught either qualitative or quantitative knowledge in classical mechanics where each member of the pair had the different kinds of knowledge. In addition the problems to be solved were beyond the competence of each member of the pair thus necessitating collaboration. The students were asked to complete multi-component tests before and after the collaboration. These tests consisted of questions related to the qualitative and quantitative instructional units plus tests which required interrelation of quantitative and qualitative knowledge, i.e. should mostly only be solvable after the collaboration. In addition the collaborative problem-solving dialogue was later analysed. The study concludes by saying that interrelation does take place where necessitated, subjects move from exclusively quantitative representations to combinatory but that the qualitative subjects gained more from quantitative subjects rather than vice versa. Finally they say that "qualitative representations form not only a good starting point for subsequent construction of quantitative representations during problem-solving but also form a beneficial starting point for learning quantitative representations" (Ploetzner & Kneser, 1998). This suggests an order for the introduction of physics problems in an instructional programme as students move from qualitative 'representations' of real Physics problems to procedures for calculating quantitative solutions to such problems.

With some of the theory of collaborative learning in mind it is now timely to examine some key research into computer-based approaches to collaboration which has taken place over recent years.

2.2.4. Computer Supported Collaborative Learning.

It is proposed that Computer Supported Collaborative Learning (CSCL) is further sub-divided into categories of systems that support different types of collaboration such as human to human, human to Agent or Agent to Agent where 'Agent' is taken to mean an artificial (i.e. computer-based) companion. The latter category is a sub-field of Distributed Artificial Intelligence known as Multiple Agent Systems (Wooldridge, 2002) useful for modelling or simulating behaviour. For such a simulation model of collaboration in an educational context see Burton, Brna & Pilkington (2000) which utilises Dialogue Game Theory and explores the influence of roles in collaboration. Key systems used in human-to-human collaboration include Betterblether (Robertson et al., 1998), Chatterbox (Pilkington & Kuminec, 2002), Sensemaker (Bell, 1997), Belvedere (Suthers & Jones, 1997), SCI-WISE (White & Shimoda, 1999) and Clare (Wan & Johnson, 1994). Additionally, more recent work by Sheremetov and Arenas (2002) straddles both human-human and agent-human paradigms.

Some systems such as Betterblether and Chatterbox are aimed at developing childrens' discussion skills through augmenting the facilities of chat systems. Others such as Belvedere aim to support their users in part through the provision of a joint collaborative space in which to co-construct a representation, in the case of Belvedere argument structures are represented in a diagrammatic form helping students to 'see' the weight of evidence for and against a position. In such systems the joint problem-space gives a continual shared record of progress on the task. It is worth noting that during a summary construction task, particularly between an artificial Agent and a human student, the summary document itself could be viewed as one such joint space, an *external representation* of the problem. The summary is composed of phrases selected from the main text and inserted in the joint summary-construction space and any new organisational features added there (such as bullet points, underlining, etc.). The joint construction space thus becomes a cognitive tool (or mediating artefact) to aid the collaborative effort. This tool could then be used to prompt self-explanations (Chi et al., 1989) as well as provide a focus for the collaborative discussion.

Drawing on ideas from work on human-human interaction using chat tools and the notion of a joint collaborative space in which a learning product is under construction, the main focus of the work in this research is the design of an artificial study companion with which to engage in the construction of a representation of text-notes in a joint shared space – a companion in a CSCL environment in which the task is summary note-taking. The field of Learning Companions will be further reviewed in Chapter 5, section 5.1 . Suffice it to say at this stage that these systems are built upon notions of collaboration (in turn based upon notions of social constructivism). However, it is important to note at this stage that in collaboration between humans and artificial peers the interaction is usually overseen by either human or artificial tutors. Peers can be seen as equals (Chan & Baskin, 1988), more or less competent (Hietala & Niemirepo, 1998) or as "trouble-makers" whose aim is to challenge the learner's behaviour and thereby prompt changes in behaviour (Aimeur & Frasson, 1996). In the latter case students need to learn to be fairly robust in order to be able to defend their views when these are correct. Collaboration can then sometimes take the form of peers working together on a problem or it can take the form of the human student teaching the artificial Learning Companion.

As the artificial Learning Companions proposed for this research will collaborate with students on a summary note-taking task the next section examines some pertinent literature concerning summarisation both as a key skill which students must master and a creative writing process that may encourage deeper conceptual understanding.

2.3 Summarisation.

Although summary construction is of central concern here it is prudent to place summarisation within the context of wider writing study skills. The next section aims to do this by identifying three approaches to writing. This is followed by a discussion of the meaning of summarisation, two writing process models that might apply to effective summary note-taking are then explored. Some techniques tutors use to try to develop and improve summariser's skills are described next then finally work in automatic summarisation by computer is reported.

2.3.1. Approaches to Teaching Writing.

Jordan (1997) defines three schools of thought on the teaching of writing skills, notably **product**, **process** and **functional** approaches. The product approach is concerned with emphasising the end result as being of most importance, for example, the structure inherent in producing a thesis or other genres. However, the product approach is often combined with a functional approach to produce a functional-product approach. That is, as well as being concerned with structure, cohesion, grammar and style (the product) this hybrid approach is also concerned with the language functions or "the logical arrangement of discourse forms" (Jordan, 1997), e.g. description, narrative, explanation, definition, etc. Possible ways to teach this approach would consist of a model text to serve as an exemplar and exercises designed to draw attention to its features. This method can also incorporate structure and vocabulary aids and, for example, lists of connectives and referencing guidelines.

On the other hand the process approach employs a more learner-centred rationale in that it emphasises the composing process, i.e. it encourages discussion tasks, drafting, feedback and informed choices. Feedback takes the form of either peer evaluation or tutor-student dialogues or written comments thus facilitating the role of draft and revise in the creative writing process. It is thus hoped that by explicitly stating a composition process, e.g. plan, draft, revise (at a general level), students are engaged in the creative act of writing. This can then be applied to a variety of situations, i.e. they are not tied down to one written form as may happen in the product approach. Practically, the process approach can be taught by guiding students through the phases by means of suitably designed exercises such as categorising information from texts and discussing results leading to feedback and further drafts. In addition, as will be seen later, a computer-based aid could provide the peer or tutor feedback inherent in the process approach. If this is the case then it needs to support the phases that students negotiate, i.e. the plan, draft, revise phases. A study will later be described (see Chapter 4) which sought to investigate the processes that students negotiate as they attempt to produce a written summary of the Physics course material.

As mentioned, feedback is an important part of any writing exercise and a number of guidelines are presented by Jordan (1997).These include :

1. Be specific (this helps with revision);
2. Actively involve the students, i.e. don't just correct;
3. Use codes for grammatical errors;
4. Only flag frequent errors or those that interfere with communication i.e. don't overwhelm the student;
5. Use self-monitoring i.e. students indicate where they need help;
6. Employ peer-correction.

It is worth bearing these factors in mind in the construction of a tool aimed at collaborating with the student to guide them.

2.3.2. What is a summary ?

As it is proposed to provide additional computer-based support in summary writing it is necessary to first attempt to define a summary. Johns (1985) states that the summarisation task involves the "identification of main ideas and condensation of text while maintaining the focus of the original". Tawableh (1994) states that summarisation leads to deeper processing of text thus aiding comprehension. He goes on to say "it focuses the student's attention upon the ideational structure of text and increases reminding of and a reinterpretation of background knowledge which aids retention and recall". Summarisation can thus be seen as a process that complements the acquisition phase of the pedagogical framework. Tawableh assumes comprehension to be a pre-requisite to summarisation thus leading to the formation of higher order 'macros', a notion based upon Kintsch and Van Dijk (1978). Note that this notion of comprehension may be quite limited; it does not necessarily imply correct understanding it merely describes a process of integration with existing knowledge structures in ways which enable condensation of the larger part of the material to form a 'gist' of the text read. If the material is highly unfamiliar this may be inaccurate or large parts of the material may fail to be integrated at all and may simply be deleted from memory. In this case the reader may report making 'no sense' of the material read. Thus, summarisation is seen as an active process involving integration of real world knowledge. In the Kintsch and Van Dijk model schemata, representations of real world knowledge and summarisation strategies, operate at a micro level on propositions and at a macro level on the 'gist' of the text. The processing of schemata is carried out by the application of rules which govern the summarisation process.

Tawableh has reviewed the summary literature with a view to informing the design of a process model which he subsequently evaluates. A review of Tawbleh's and Kintsch and Van Dijk's summarisation process models follow.

2.3.3. Process Models of Summarisation.

A common approach to the design of Computer Assisted Learning and Artificial Intelligence programs is to examine the Psychological literature relating to the area of human expertise that they are proposing to simulate. The following two process models of summarisation have been reviewed with this aim in mind, the idea being that they might lend themselves to the design of additional computer-based support.

Kintsch and Van Dijk (1978) propose that summarisation takes place on two distinct levels which they term the 'micro' or "propositional representation of the semantic content" and 'macro' or "global meaning". The micro level consists of an ordered list of propositions, i.e. predicate/verb/adjective plus argument (noun phrase). These propositions relate to each other by referential coherence or inference as modelled in a hierarchical graph. Furthermore, the higher in the list the proposition is then the higher the probability of it being recalled. The macro level consists of the 'gist' of the text, i.e. the global meaning. A number of rules are identified which operate to reduce and transform the micro structure to the macro representation. These rules are as follows :

1. Deletion – removal of unessential propositions;
2. Selection – deletion of propositions which can be inferred from others;
3. Generalisation – substitution of specific terms by generic ones;
4. Construction – introduction of a new proposition allowing inference of deletions (or related material from the readers own knowledge).

Essential in Kintsch and Van Dijk's model is the utilisation of schemata which operate in conjunction with the macro rules and associated world and linguistic knowledge to create the macro 'gist' from the micro structure. The linguistic knowledge contains elements of phonology, syntax, semantics and pragmatics. The schemata encapsulate the world knowledge plus the text structure and can be activated in a bottom-up manner, i.e. are data driven, or a top-down manner, i.e. are conceptually driven. To conclude, Kintsch and Van Dijk propose that comprehension and associated summarisation of a text involves construction or activation of schema leading to a plausible and coherent model of text meaning.

Tawableh (1994) emphasises the importance of text structure in construction of summaries. He cites Meyer's (1975) classification of rhetorical structures of expository prose :

1. Collection – groups ideas by association or sequence over time;
2. Description – hierarchical, identifying elements in relation to constituents or parts;
3. Causation – causal relationships, i.e. cause / effect;
4. Problem / solution – problem and/or solution involving subordination;
5. Comparison / constraint – similarities or differences.

Dowd & Sinatra (1990) and Carrell (1985) combine explicit teaching of these kinds of text structure with, for example, graphing techniques which aim to illustrate inter-ideational relationships with a consequent increase in rates of recall of texts amongst students.

Tawableh seeks to identify a process model of summarisation by utilising Dialogue Analysis techniques to inform the design of a Summary Analysis. In a review of other Summary Analyses he again identifies that none of the schemes are based upon the underlying structure of the text and that none of the analyses encompass all phases of the summarisation process. Based upon a rhetorical predicate analysis (Meyer, 1975) Tawableh arrives at a model of the summarisation process consisting of the three phases of selection, condensation and organisation but which do not occur sequentially and are in fact indistinct processes. This model is further expanded upon in tabular form as shown in Appendix C. Tawableh states that comprehension of the text is a pre-requisite to effective summarisation which is then followed by selection to distinguish and identify important propositions. Condensation involves transforming and condensing propositions whilst maintaining the rhetorical function then organisation involves linking together propositions using cohesive ties and discourse markers. Tawableh formulates a scheme for summary analysis based upon this model which he validates using think-aloud protocols and then employs to examine differences between summaries produced by first and second language English speakers. The results of this analysis are conclusions regarding improving second language speaker's summarisation skills. These are then used to inform the design of a computer-based aid to summarising although this implementation is limited to the selection and condensation phases of a contrastive text.

These approaches to the process of summarisation together with the results of a study of CHALCS students' summary note-taking (described in Chapter 4) provided a framework for analysis and further development of the Virtual Learning Environment and artificial Learning Companion to support effective summary note-taking. In the next section Tawalbeh's model is explained further together with related work to suggest ways in which such a companion might seek to support summarisers.

2.3.4. Improving Summarisers' Performance.

Jordan (1997) presents a number of techniques and strategies which can be used to improve summariser's performances. These include such simple devices as tables of abbreviations or symbols, frameworks or structures of idealised summaries, discourse markers or cues in the original texts, use of colour, highlighting, layout, numbered sections, lists of connectives and even concept mapping. However, these apply more to the process of note-taking than actual summarisation which, as highlighted above, may be more akin to the task of abstracting in which there is a requirement to reduce the text in ways which faithfully reflect the author's intent in terms of content and organisation and which communicate these to an audience other than oneself. This may require a deeper level of cognitive application and is certainly a different process from representing the gist of text in ways suited to one's own needs and purposes. The task of effective note-taking is an aid to knowledge acquisition and is clearly a cognitively demanding task in its own right.

Tawableh reviews a number of recommendations from other researchers which include the kind of techniques espoused by Jordan (Garner & McCaleb, 1985; Johnson, 1981) through to inductive and deductive teaching of summarisation rules (Hare & Borchardt, 1984). However, these techniques often lack the important integration of strategies and rhetorical structure although Johns (1988) uses both in the limited problem-solution structure but omits condensation rules. Armbruster et al (1987) suggest using both text structure and summaristaion strategies to summarise a problem-solution text with encouraging results. Tawableh criticises this work though on the basis of it's limited applicability to other text structures and, again, lack of explicit condensation rules – the specific linguistic techniques that serve to reduce the text.

Tawableh's own suggestions for improving summarisers performance, at least in a second language context, took the form of a computer based aid. This CBL package was based upon empirical evidence emerging from a case study seeking to improve subjects' summarisation skills. That is, reorganising rhetorical schemata that identify the structure of the text and using these together with the reader's understanding of the text to guide the selection of important material, in keeping with the proposed process model, i.e. the three phases of selection, condensation and organisation. Therefore the first phase of the CBL system asks the user to identify the type of text structure then based upon this structure the system asks the user to enter text in boxes effectively performing a mark-up (by hand). For example, if the student selects "comparative text" the system prompts for propositions that are similar and propositions that are different, providing feedback as appropriate. The system then asks the user to select information they wish to include in their summary, again providing feedback as appropriate.

The relationships which the system prompts the user for (in the above instance "similarity" and "difference") and feedback offered by the system rely on mark-up of the text based on a rhetorical predicate analysis (Meyer, 1975) of the text. The process of mark-up is further explained in Chapter 6, section 6.6.2. Appendix H contains a section of Physics course notes that have been annotated to show the rhetorical hierarchical analysis. 'Rhetorical predicates' are "labels for the relationships between content words" (Meyer, 1975, p.41) which show the ideational relationship amongst propositions in a text. The eighteen rhetorical predicates which Meyer derived are shown in Table 2.1. Rhetorical predicates can thus link ideas across sentences or paragraphs either at the same level or a subordinate level in a hierarchy (as can be seen in Appendix H) showing their relative importance to the text. Pilkington (1988) expanded the list of rhetorical predicates by the inclusion of, for example, 'similarity' (linking two common concepts) and 'difference' (linking two uncommon concepts).

In the condensation phase the user is asked to select a condensation method (in accordance with Tawalbeh's model – Appendix C) then select appropriate propositions. The system subsequently guides the user through the process step-by-step based on the mark-up. Unfortunately, Tawableh did not implement the organisation phase in a CBL system but recommends that the system can provide facilities to delete redundant information, delete or add punctuation, use synonyms, use cohesive ties and discourse markers and join sentences. These facilities would be supplemented with menus of pre-stored exemplar techniques.

Tawalbeh's work provides much inspiration for the development of the tool described later as part of this research. As well as being implemented by a sound theoretical model the system is developed through a formative evaluation method with target users, i.e. the resultant system was proven to work as intended by helping English as a second language speakers with part of the summarisation process. However, the tool relied on a complicated mark-up of the text to be summarised and the user-interface would not be appropriate for younger age range of students. This 'user-friendlines' will be a major concern of this work. In addition Tawalbeh's tool was never completed. Whilst it is not intended to re-address all of the summarisation processes that Tawalbeh delineated in such depth, it is the intention to address singular processes in each phase as exemplars, i.e. selection, condensation and organisation strategies. Furthermore, the tool will be more 'user-friendly' as appropriate for the target group of younger and less independent learners. The resulting co-constructed (between the student and the Agent) summary, should be useful, e.g. for revision purposes, as well as helping students understand the key skill of effective summary note-taking. In addition it is hoped that by simplifying the mark-up process the system may be more easily extended at a later date through the construction of tools to assist web-authors in quickly applying mark-up to notes enabling the agent to "run" over a wider range of material.

Rhetorical Predicate	Description
1. Alternative	equal weighted alternative options
2. Response	equal weighted question/answer, remark/reply or problem/solution
3. Attribution	qualities of a proposition
4. Equivalent	restates same information in a different way
5. Specific	gives more specific information about something that was stated in a general manner
6. Explanation	previously stated information is explained in a more abstract/concrete manner
7. Evidence	evidence through perception of a situation to support some idea
8. Analogy	given to support an idea
9. Manner	way an event or event complex is performed
10. Adversative	relates what did not happen to what did happen
11. Setting Time	gives time of setting in which information being related occurs
12. Setting Location	gives location of setting in which information being related occurs
13. Setting Trajectory	gives changing background of location and time that occurs as characters in a narrative travel through various places
14. Representative Identification	singles out one element of a group and makes it stand for the whole group
15. Replacement Identification	one thing standing for something else
16. Constituency Identification	identifies a part in relation to some whole
17. Collection	list of elements related in some unspecified manner
18. Covariance	relation often referred to as condition, result or purpose with Antecedent/Consequent

Table 2.1 Meyer's Rhetorical Predicates (adapted from Meyer, 1975)

As well as reviewing work carried out on the summarisation process it is also relevant to examine research into the automatic creation of summaries by computer software. The next section looks at some of the work in this field.

2.3.5. Automatic Summarisation.

Interest has been renewed quite recently in the field of automatic text summarisation. As we are ever more bombarded by information (for example from the World Wide Web, Television and Audio communications) there is a need to extract salient points in order to prevent an 'information overload'. Agent-based software can help with this onerous task, e.g. Maes (1997). It is hoped that eventually important features of a 'document' (in it's broadest sense) can be distilled and presented to a user enabling them to gain an overview or digest the myriad of sources with minimal amount of effort. Today this is even more pertinent as we move from desktop computers to mobile wireless systems with inherently less storage resources.

Early work in automatic summarisation could be typified as an information-processing approach in that it would rely on a computer's ability to perform statistically-based analysis on large texts at a speed greater than a human operator. For example, Luhn (1999) (originally published in 1958) reports on pioneering work which took a text then would derive a list of significant words based on their frequency of occurrence. The software that he describes would take into account similarities between words and their relevance as calculated by a cut-off frequency both at a lower point (i.e. the word does not occur frequently enough to be significant) and an upper point (i.e. the word is a common word). The latter cut-off point which can be likened to 'noise' could be implemented by a look-up table of common words, e.g. pronouns, prepositions and articles or, as Luhn proposes, a statistically-based method. Sentences are then analysed for their inclusion of these high frequency words within a certain distance of each other thus arriving at a "significance factor" for each sentence in the text. Finally, a cut-off value is needed for the software to derive which significance factors indicate a sentence should be included in the derived summary.

This surface-level approach can produce interesting and valid approaches to automatic summarisation particularly with the addition of weightings corresponding to cue-phrases and a phrase's location in the original text. Cue phrases can include such elements as "to conclude", "this paper" or "finally" and, indeed, similar cue phrases are mentioned by Meyer (1975), Pilkington (1988), McKeown (1985), Mann and Thompson (1989) and in Pilkington's DISCOUNT scheme (Pilkington, 1999) as signals to rhetorical predicate structures. A cue phrase can also be given higher significance if, for example, it appears in the document's title or at the beginning of a paragraph.

A further addition to this kind of automatic summarisation is to derive measures of significance from corpora of text (Kupiec et al., 1999). This method would, for example, compute frequency of words and derive cut-off points and weightings. Aside from these "knowledge poor" (Hahn & Mani, 2000) approaches a different approach has been taken by researchers relying on natural language processing and dialogue analysis. These are termed "knowledge rich" techniques. For example, and relevant to this research, the approach proposed by Marcu (1999; 2000) uses a rhetorical structure tree (Mann & Thompson, 1988) to help indicate a proposition's importance within a text. Marcu has implemented software which will derive the 'discourse tree' of a text using cue phrases. This tree is then analysed in terms of promotion sets derived from a phrase's status as either a nuclear phrase (e.g. "mars experiences frigid weather conditions") or satellite (e.g. "with its distant orbit…and slim atmospheric blanket") and position as either a leaf or a node. A formula is then derived to compute a significance score based on a promotion phrase's first occurrence in the depth of the discourse tree.

Hahn and Mani (2000) report on various systems which currently exist to automatically summarise text including Microsoft's Word Autosummarize tool, IBM's Intelligent Text Miner, Oracle's Content and Inxight's Summarizer (part of Alta Vista). However, these tools can be viewed as employing extraction (i.e. condensation) as opposed to abstraction (i.e. paraphrasing). An example follows using Microsoft's tool (Figure 2.1) based on the content of this section on Automatic Summarisation (with the exception of the summary itself and the final paragraph).

Microsoft Word uses a simple summed word frequency-based method to select important sentences (Gore, 1997). As seen, the effect can be quite useful even though such a crude method is employed. In this example (see Figure 2.1) 835 words in 31 sentences contained in the original document were summarised as 219 words in 9 sentences by selecting a ratio of 25 %.

Interest has been renewed quite recently in the field of automatic text summarisation. For example, Luhn (1999) (originally published in 1958) reports on pioneering work which took a text then would derive a list of significant words based on their frequency of occurrence. Finally, a cut-off value is needed for the software to derive which significance factors indicate a sentence should be included in the derived summary.

This surface-level approach can produce interesting and valid approaches to automatic summarisation particularly with the addition of weightings corresponding to cue-phrases and a phrase's location in the original text. For example, and relevant to this research, the approach proposed by Marcu (1999; 2000) uses a rhetorical structure tree (Mann & Thompson, 1988) to help indicate a proposition's importance within a text. Marcu has implemented software which will derive the 'discourse tree' of a text using cue phrases.

Hahn and Mani (2000) report on various systems which currently exist to automatically summarise text including Microsoft's Word Autosummarize tool, IBM's Intelligent Text Miner, Oracle's Content and Inxight's Summarizer (part of Alta Vista).

Microsoft Word uses a simple summed word frequency-based method to select important sentences (Gore, 1997). Current research into automatic summarisation is beginning to abstract information from multiple documents (as found on the World Wide Web), multiple languages and multiple modalities (e.g. video and audio).

Figure 2.1 A Summary Produced Using Microsoft Word 2000.

In terms of the evaluation of automatic summaries, differences have been highlighted between formative evaluations where a summary document can be compared, for example, to an idealised summary produced by an expert and extrinsic evaluations, for example whether a summary allows set questions to be answered based on the original text. Current research into automatic summarisation is beginning to abstract information from multiple documents (as found on the World Wide Web), multiple languages and multiple modalities (e.g. video and audio). Although, this work presents huge new challenges and directions.

A Pedagogical Framework for the implementation of a VLE has so far been presented and contextualised. This research assumes one particular learning style, notably that of collaboration, as we will utilise an agent as a student partner in the construction of a summary. Hence this section also reviewed the literature on writing and summary note-taking in particular. The next section is an important one as it relates this research to previous work on creating believable, engaging artificial agents with which students feel they can effectively collaborate.

2.4 Affective Computing.

There has been much interest over recent years in an area of Computer Science which has come to be known as Affective Computing (Picard, 1997). Basically, it involves taking into account affect in computer systems, be it in terms of making computers aware of user's affective states (e..g in Human-Computer Interaction) or modelling affective states within a system (usually employing Agent-based approaches). By recognising user's affective states it is hoped that a system can be more empathic (Brna et al., 2001) and symbiotic. By modelling affective states within a computer it is hoped that artificial cognition will begin to more closely mimic human cognition. The latter could lead to, for example, faster searches of state-spaces by employing 'intuition'-like heuristics or 'somatic markers' (Damasio, 1994) or 'emotional interrupts' (Simon, 1979) and more relevant to this research, model affective states of artificial (animated) Agents thus creating a more 'believable' character (Bates, 1994).

This review will not cover much of the work on analysis of affective states in human users although, as will be demonstrated in Chapter 9, this may be of interest in future work. For a review of this area the reader is directed to Picard (1997) including the section on Affective Wearables which are an interesting area of personalised ubiquitous computing devices designed to gather psychophysiological data of their users. Instead, as already pointed out, this research is more interested in creating believable characters by employing affective methods. In fact the affective techniques employed have more accurately been termed *'affectations'* (Barker & Pilkington, 2001) as they do not strictly employ models of affect in an Agent. However, the work is placed within the Affective Computing context as previous work in this area has informed this research. It is also worth mentioning that 'emotion' and 'affect' (or indeed 'passion') will be used synonymously here, examples include 'love', 'hate', 'sadness' etc. First though it is necessary to examine the origin of the study of Affect.

2.4.1. The Philosophical Basis of Affect.

Columbetti (2001) and Solomon (Solomon, 1993) present an overview of the philiosophical development of notions of emotions and cognition. Columbetti begins by illustrating concepts pertinent to Stoicism which viewed emotions as some form of illness with no relevance in rationality. In fact, the view of Stoicism is that feelings should be conquered and controlled in an attempt to be objectively rational. She describes Descartes' position as less extreme saying that to him the main function of reason is to control the excesses of passion and that in moderation they are part of life's rich experience. This more liberal view can, she says, be traced back to Plato and she cites examples of Plato's thoughts on pain and pleasure being due to "the breaking and reestablishment of the harmony of the elements". In turn, Aristotle emphasised "the practice of character", i.e. feelings, in order to display true virtue. In fact, Columbetti (2001) states that Aristotle extols the use of emotions in the Rhetoric where they can be used to persuade an audience. This is an example of the strategic use of affect to influence perception. Aristotle additionally attempts to classify feelings such as fear and confidence in terms of psychological explanations. For Hobbes and Hume, Columbetti states, "human nature is fundamentally passionate". However, whilst on the one hand Hobbes views feelings negatively, to be controlled by a government, Hume, in contrast, sees them as additionally positive, for example determining social benevolence. This extra social dimension of affective reasoning is interesting to this research for, if feelings are essential in determining social actions, then computer systems which hope to successfully interact with humans must be 'aware' of the effect of affect on social intercourse.

Columbetti goes on to discuss work by Le Doux (1996) and Damasio (1994) which points towards the notion of cognitive or deliberative reasoning leading towards an "imagined" situation inducing an affective response, e.g. fear conditioning. This is what Damasio calls the somatic marker hypothesis. In addition Columbetti puts forward arguments for 'strong embodiment', that is the necessity to experience an emotion in terms of its body arousal states such as increased heart rate, or perspiration. This argument is evidenced in terms of representations of the body in the brain, particularly in the case of 'phantom limbs' patients. Damasio, in fact, sees emotion as distributed throughout the body. He posits the following three statements relating to embodiment :

1. "The human brain and the rest of the body constitute an indissociable organism, integrated by means of mutually interactive biochemical and neural regulatory circuits (including endocrine, immune and autonomic neural components);
2. The organism interacts with the environment as an ensemble: the interaction is neither of the body alone nor the brain alone;
3. The physiological operations that we call mind are derived from the structural and functional ensemble rather than from the brain alone: mental phenomena can be fully understood only in the context of an organism's interacting in an environment."

(Damasio, 1994)

Here, Damasio is arguing for the integration of mind and body in suitably derived theories of mental life. Leading on from this work, the argument can be taken en extremis to cover such ground as the necessity of notions of self in mental functioning (and ultimately consciousness), including the perception of emotions.

Damasio (1994) aims to disambiguate between primary and secondary emotions. To summarise, primary emotions are considered to be reactionary in nature, for example a 'startle' response if we hear a loud bang. They seem to occur predominantly in the limbic structure of the brain, which is a centre for emotion, attention and memory and appear to be more instinctual and less easily controlled. On the other hand, secondary emotions may be cognitively generated such as 'grief' which results from an understanding of the process of death. These seem to involve the limbic system in the brain but also with activity in the cortex or 'gray matter' (Picard, 1997). These results indicate that there is a level of emotions that are reactionary in nature involving little deliberative intervention. This has some implications for the design of an Agent employing notions of a hybrid reactive/deliberative architecture as will be discussed in a subsequent section and Chapter 6.

Whilst this research does not hope to find answers to such searching questions as the nature of mind-body interactions and ultimately the meaning of consciousness it is worth bearing such fundamental concepts in mind when attempting to construct affective artificial organisms for these are at the heart of the philosophical debate. The more immediate relevance of work on Affective Computing to this research will next be examined.

2.4.2. The Relevance of Affective Computing: Creating Believable Agent-Based Personalities.

> "*Personality* is the set of psychological traits that uniquely characterize an individual. Personality distinguishes each individual from all others and colors his or her behaviour in a pervasive and recognizable fashion. Personality makes people fascinating to one another."
>
> (Hayes-Roth et al., 1997)

Much work has been carried out on creating believable agents by modelling animated characters based on an underlying emotional model which it is hoped will lead to convincing agent-based 'personalities'. Indeed, personified Pedagogical Agents are emerging as a distinct field within Artificial Intelligence in Education as evidenced by, for example, work presented at the workshop on Animated and Personified Pedagogical Agents at AIED 99 in Le Mans, France.

For example a web-deployed Agent, Adele (Shaw et al., 1999), utilises an animated graphic interface together with a synthesised speech component to supplement a medical educational environment. An architecture for Adele was proposed as a typical client/server application. In the server component resides the case, student, persona, references and simulation parameters whilst the client contains the text-to-speech engine, a task planner and assessor, an animated persona, reasoning engine, simulation engine and Graphical User Interface (GUI). Problems in the believability of Adele amongst students are reported although Shaw et al. state that Adele's hints were generally considered useful. (A demonstration of Adele can be downloaded at www.isi.edu/isd/carte/, last accessed 9th February, 2002).

Related work includes that of Paiva et al (1999) and Abou-Jaoude and Frasson (1999) who are both considering the underlying emotional model which guides an agent's interaction. Abou-Jaoude et al seek to elicit emotional states from students then incorporate these into a "believability layer" of an Agent architecture. Paiva et al. are based the personifications of their Vincent Agent on the OCC model of emotions (described in Section 2.4.4). Interestingly, they propose a development methodology for the implementation of personified Agents consisting of :

1. choosing the character's personality (e.g. using bipolar personality traits);
2. extracting the behaviour features (the different ways a character will perform the same behaviour);
3. defining the relevant events in the outside world;
4. defining the characters goals (e.g. "aiming that the trainee will master the domain");
5. defining the Agent's emotional profile (i.e. emotional resistance and memory and a set of emotional reactions).

These steps provide a concrete approach to the development of an Agent's personifications, thus enabling the implementation of a personality, which is complimentary to the pedagogical aims.

However, this work will not attempt to create such emotional models although this branch of AI shares one goal of the research reported here since the role of affect in such models is to create a perception of personality within the mind of the user or student. Although our own work does not approach the apparent sophistication of such artificially intelligent models (and indeed it will be argued does not have to), lessons can be learnt from such work that can inform the design of the agent to be reported here. Further Affective Computing applications are described in section 2.4.5.

Lester et al. (1997) carried out empirical research on student's use of an animated Agent known as 'Herman the Bug' which inhabited a learning environment. They state that the aim of such Agents is to increase the motivation of student's interaction with a learning environment, as Lester et al. (1997) succinctly stated in describing the 'persona effect' : "*the presence of a lifelike character in an interactive learning environment can have a strong positive effect on students' perception of their learning experience*". They describe an empirical evaluation of five clones of an Agent with varying degrees of expressiveness concluding that the clone with the greatest number of modalities, i.e. principle-based, task-specific, animated and verbal advice produces the best student problem-solving performance.

Hietala and Niemirepo (1998) further explored the concept of Learning Companions displaying personae which took the form of 'strong' and 'weak' problem-solving capabilities. The 'weak' companions however gradually improved their performance as the human collaborator became more competent herself. Other persona-based cues took the form of appropriate Learning Companion dialogue, i.e. "commanding" language for 'strong' versus "hesitant" language for 'weak' companions and distinct names and graphic appearances. Of the two groups of students identified it emerged from studies that extrovert students preferred the weak companions whilst the introverts preferred the strong companions. On the other personality traits of the companions Hietala and Niemirepo (op. cit.) state that "although more research is needed these 'social' features are important in providing a socially rich environment for different kinds of learners".

2.4.2.1. What Are Personality Traits ?

The previous quote by Barbara Hayes-Roth is a concise summary of the views of personality theorists whose work relates to this research, notably those of the 'trait theory' school of thought. Aside from personality colouring our actions and portraying our individuality it can also make us, or indeed an artificial agent so embued, an interesting object of attention - something which is essential in a computer-based learning package vis à vis the 'persona effect' cited earlier. But what exactly are these traits which we neatly use to encapsulate an individuals personality? Furthermore, how do they relate to the design of artificial pedagogical agents?

The psychological literature on personality traits has a long history together with a correspondingly varied set of conclusions. However, it has become common to refer to a seemingly common set of traits which can be identified in the literature. This taxonomy of traits, defined by consensus amongst personality psychologists, is known as the Five Factor Model or the 'Big Five' (John, 1990; McCrae & John, 1992). Table 2.2 delineates the Five Factors together with typical adjectives and scales commonly used to define them.

Factor	Adjective	Scales
I. Extraversion (E)	Active	Warmth
	Assertive	Gregariousness
	Energetic	Assertiveness
	Enthusiastic	Activity
	Outgoing	Excitement Seeking
	Talkative	Positive Emotions
II. Agreeableness (A)	Appreciative	Trust
	Forgiving	Straightforwardness
	Generous	Altruism
	Kind	Compliance
	Sympathetic	Modesty
	Trusting	Tender-Mildness
III. Conscientiousness (C)	Efficient	Competence
	Organised	Order
	Planful	Dutifulness
	Reliable	Acheivement Striving
	Responsible	Self-Discipline
	Thorough	Deliberation
IV. Neuroticism (N)	Anxious	Anxiety
	Self-pitying	Hostility
	Tense	Depression
	Touchy	Self-Consciousness
	Unstable	Impulsiveness
	Worrying	Vulnerability
V. Openness (O)	Artistic	Fantasy
	Curious	Aesthetics
	Imaginative	Feelings
	Insightful	Actions
	Original	Ideas
	Wide interests	Values

Table 2.2 The Big Five Personality Traits, Common Adjectvies and Scales (from McCrae and John, 1992, p.178-179)

The Five Factors are based on previous psychological research which analysed natural language as well as previous questionnaires. Natural language was turned to in the analysis because researchers believed that if traits did exist then there must be corresponding trait terms used in everyday language. As McCrae and John (1992, p.184) state: "The lexical hypothesis holds that all important individual differences will have been noted by speakers of a natural language at some point in the evolution of the language and encoded in trait terms; by decoding these terms, we can discover the basic dimensions of personality". On the other hand, questionnaires utilising scales devised by personality psychologists existed in their hundreds, mostly based on sound psychological theory and indeed, when analysed by researchers, revealed a certain commonality. For instance, psychologists agreed that scales to measure chronic negative emotions and those used to measure interpersonal activity (the two dimensions being known as N and E) were "to be found in a wide variety of instruments" (McCrae & John, 1992). Thus it became possible to combine the various approaches to personality research through factor analysis resulting in the five common traits shown in Table 2.2. Note that the (roman) numerical and letter alternatives to the names given for the Big Five in Table 2.2 are sometimes used to avoid the ambiguous and possibly misleading names commonly given to them.

Of the Big Five personality traits (McCrae & John, 1992) McCrae and Costa (1989) have identified two, extraversion and agreeableness as being most relevant to interpersonal interaction. In fact in the interpersonal circumplex theory of personality (Wiggins, 1979) the extraversion trait (also known as power, status or control) is said to range from dominance to submissiveness and the agreeableness trait (also known as affiliation or warmth) is said to range from warm to cold or agreeable to quarrelsome. This research assumes the extraversion and agreeableness traits as they are seen as crucial to interpersonal interaction, something which the Agent to be designed in this research must be adept at. Later, in Chapter 6 (particularly see Table 6.2), the 'affectations' or behaviour-based affective qualities of the Companion will be compared to the interpersonal dimensions of personality proposed by Wiggins (1982) showing that the Companion's design is constructed so as to be effective interpersonally. This is obviously of primary concern in software which hopes to be 'sociable' in terms of creating a suitable environment for successful pedagogical interactions. This last statement leads onto a discussion of the role of personality in Human-Computer Interaction.

2.4.2.2. Evidence for the Role of Personality in Human-Computer Interaction.

Clifford Nass has been involved in research investigating the hypothesis of Computers As Social Actors (CASA). In The Media Equation, Reeves and Nass (1996) describe research undertaken at Stanford University Communications Department which aimed to refute or assert the Media Equation hypothesis, that is, that people respond to mediated communication in exactly the same way as they do to face-to-face communication: "People respond socially and naturally to media even though they believe it is not reasonable to do so, even though they don't think that these responses characterize themselves" (Reeves & Nass, 1996, p.7). This means that Reeves, Nass and colleagues were able to turn to the extensive literature on, for instance, the psychology of interpersonal relationships and look for evidence of equivalent behaviour in mediated communication.

The result is a series of experiments which *systematically* test the transferability of social science theories to mediated forms of communication. This systematic experimental methodology is typically carried out by Nass and colleagues as 'experimental' style research utilising, for example, randomised subject designs, dependant variables and statistical analysis.

An obvious benefit of this approach to studying mediated communication is the derivation of design guidelines and the ratification of the whole socially-oriented style of Human Computer Interaction prescribed by this PhD research and that of others in the field (particularly see the next section). As Reeves and Nass (1996, p.8) conclude from their work: "Humans are experts on social relationships, and they are experts on how the physical world works. Rules for using media as tools, on the contrary, are often arbitrary and must be learned. When media conform to social and natural rules, however, no instruction is necessary. People will automatically become experts in how computers, television, interfaces, and the new media work."

By way of illustration, the Media Equation work of Reeves and Nass on personality which is seen as demonstrating their approach and is jointly of most interest to this research will now be explored in more detail.

Reeves and Nass (1996, p.75) first describe a study they undertook to ascertain if children regarded mediated personalities (actually as portrayed on Television) in terms of the same personality traits in real life. Their results were among the first indications that media may indeed not be such a special case, that is, both social and personality psychology results could legitimately be applied to mediated communications. These insights were again apparent in a subsequent similar study only this time involving adults. Interestingly, this follow-up study utilised just basic line drawings but still created an impression of personality as perceived by the subjects thus indicating that full-blown photo-realistic characters may be unnecessary. As a result of these studies, Reeves and Nass emphasise the need for a personality to be consistent thus leading to a likeable reaction rather than one of dislike and confusion although personality flaws (such as those found in the real world), they state, may indeed add to believability.

Reeves and Nass also recount an experiment to "give a computer a personality" (Reeves & Nass, 1996, p.91) by solely using text to convey a dominant and a submissive 'personality'. That is, dominant personalities utilised assertions and commands whereas submissive personalities utilised questions and suggestions. Additionally, the dominant computer displayed higher confidence than the submissive computer in the form of certainty scores. Further manifestations of the personality traits included the sequence of interactions, i.e. the dominant computer always took the first turn in conversation whereas the submissive one would follow. Finally, they chose the names of the two computers to reflect their personality type. Their hope was that dominant text would imply a dominant personality and submissive text would imply a submissive personality, that dominant or submissive people would indicate that the equivalent computer 'personality' was more like them and that they would prefer this equivalent computer personality to the opposite – known as the "similarity-attraction hypothesis".

In the Reeves and Nass experiment the subjects were analysed as being dominant or submissive themselves then half of each group worked with each of the two computer personality types on a scenario type of problem. Results indicated that the two computer personalities could be easily distinguished plus the two human personality types did identify with their equivalent computer personality and, indeed, preferred their similar personality counterpart even though the content of the interactions was the same across the two computer types. This similarity-attraction of personalities also resulted in participants rating the computer's *ability* more highly and also led to increased enjoyment of the task. Reeves and Nass go on to state that:

> "the *creation of personality on a computer is not primarily an issue of artificial intelligence*"
>
> Reeves and Nass (1996, p.97)

This, as will be seen, is central to this thesis and the notion of *affectations* (Barker & Pilkington, 2001), which are, affective behaviours without recourse to underlying emotional models. Indeed, all of the above mentioned results concerning the two personality types are very relevant to this research described in Chapter 6.

Reeves and Nass (1996) also state that rich descriptions of personality, e.g. virtual reality renderings, are not a necessary precursor to the communication of personality thus they rely as much on social science findings on personality, such as those mentioned above, in the context of dominant and submissive personalities. They also state that a "short series of questions" could be used to identify potential user's personality which could then inform the choice of an appropriate computer personality. However, ultimately the choice of computer personality would depend on the task it was designed for. For example, Reeves and Nass (1996, p.98) state that "if the help is from a peer, the personality could be more friendly than dominant". They go on to say:

> "Making one personality fit all cases is hard. One way around this is to offer multiple personalities that can be selected by the user. The content of the interface wouldn't have to be rewritten; all that is needed is to slightly modify the style of interaction...If each personality is strong and well-defined, users will recognize the ones that are similar to their own. These are the personalities they would enjoy and the ones that would make them think better of the media product and themselves."
>
> (Reeves & Nass, 1996, p.99)

The above quote is included because it is particularly relevant to the approach taken in this research in that two personalities will be presented, notably consisting of one 'dominant'/'cold' and one 'submissive'/'warm' character. The intention then is that students will choose the personality with which they feel the greatest degree of attraction, i.e. the most similar personality if the similarity-attraction hypothesis is to remain valid.

A further study carried out by Reeves and Nass (1996) studied the "gain theory" of interpersonal relationships. Basically, this theory states that a personality which changes over time to be more like another's personality will please this other more than one that was similar all along. Reeves and Nass thus carried out an experiment similar to the one described previously using dominant and submissive personalities in both computers and participants. However, during one of the trials the personality changed for half of the subjects (of mixed personality types) from either submissive to dominant or vice versa. This manipulation of the experimental variables thus allowed Reeves and Nass to examine the results when the computer changed to be more similar in personality to the participants, less similar and consistently similar or dissimilar. They found that, as described in the previous study, similar personalities produced more favourable results than dissimilar. However, they also found that those computer personalities which changed to be more like the subject's personalities were liked more than those that stayed consistently similar, i.e. the 'gain theory' had held true for mediated communications. Addtionally Reeves and Nass (Reeves & Nass, 1996, p.104) state: "when the computer changed to match the participant's personality, participants thought it was more competent and they thought that the interaction was significantly more satisfying intellectually", even though the computers are conveying the same information.

One particular implication of these results is that "media should adapt to the personality of the user" (Reeves & Nass, 1996, p.106). This tenet will be returned to in Chapter 9 of this thesis. For their part, Reeves and Nass think that language can play a central role in the conveyance of this adaptation of personality. However, as previously claimed, they rightly point out that a candid elicitation of a user's personality through the utilisation of pre-tests prior to an interaction, for instance, would be detrimental. Instead Reeves and Nass propose alternatives such as monitoring a user's linguistic style, the use of "cautious claims" or "propensity to interrupt". There are doubtless many ways to ascertain a user's personality that are less invasive than a questionnaire approach. These will be further explored towards the end of this thesis when such an approach may present itself as the logical next step. Furthermore, this approach need not obviate the need for multiple personalities as espoused above, a computer personality could still adapt in degrees to be more similar to that of it's user.

Nass et al (2000) describe similar 'experimental' style work exploring issues concerning "Embodied Conversational Agents" (ECAs) as opposed to the simple renderings of personality described above. Their first study examines the ethnicity of

ECAs, concluding that it does in fact have an effect on user attitudes and behaviours. However, of more interest here is their work on ECAs and personality. Nass et al (2000) report on how they manipulated both verbal and nonverbal cues in an embodied agent to portray an extrovert 'personality'. Verbal manipulations included confident choice of words and phrasing and speaking "fluidly". Nonverbal manipulations included both expansive posture and gestures and the propensity to approach (reduce the perceived distance between user and ECA). Their prediction was that these kinds of manipulations used to present extrovert and introvert behaviour of ECAs would result in these intended traits being recognised in such mediated communications, similar to the dominance-submissive study outlined above. Hence, the intentions of the study were to firstly determine if these manipulations created impressions of intended personality traits with participants and secondly to see if inconsistent characters were disliked or perceived as neutral. Thus introverted or extroverted characters were assigned a number of conditions consisting of matching or mismatching variations of consistent or inconsistent verbal and nonverbal behaviour. The task was again a scenario type of problem which the participants undertook with the ECA.

Results of these experiments indicated that participants did indeed recognise both verbal and nonverbal cues of extroversion and introversion. Other results indicated that participants liked the ECA more, perceived the interaction as more useful and have more fun with a consistent personality plus they found it more useful than an inconsistent ECA even though, again, actual content did not vary. Additionally, participants had more fun when ECA personalities matched their own. Overall the study confirmed to Nass et al. that people interact in the same way with ECAs as they do with other people and that they preferred consistent characters in these interactions, just as they do in real life. Design implications thus include the need to create "attractive" characters as "perceptions of seemingly objective criteria, such as intelligence, can be influenced by attractiveness" (Nass et al., 2000, p.397) plus, due to the kind of consistency issues mentioned above "it may in fact be more important that the character sends a clear message about its personality than that it matches the user" (Nass et al., 2000, p.398).

Isbister and Nass (2000) further elaborate on the CASA work described above studying ECAs. As personality is seen as central to the creation of a character, drawing inspiration from a number of disciplinary perspectives such as film, television, comic books and written fiction (Isbister & Nass, 2000, p.252), this work studies an important aspect of personality notably the extroversion/introversion traits, paying particular attention to whether consistency or inconsistency is the best design practice. The intended traits are manifested as described above, i.e. though utilisation of both verbal and nonverbal behaviour. Drawing on the interpersonal Psychology literature Isbister and Nass (2000) hypothesise that two results may occur in their experimental work. The first is the similarity-attraction hypothesis stated above but the second is also to be found in the Psychology literature which, in contrast to the first hypothesis, states that people will seek out behaviour which compliments their own. Thus, extroverts prefer to interact with introverts and vice versa. However, due to previous studies in Human-Computer Interaction just confirming the similarity-attraction hypothesis this was the expected outcome for these studies. Additionally, the studies sought to find if verbal and nonverbal behaviour created impressions of personality.

Results indicated that consistency of character design was of paramount importance: "Inconsistent character cues may undermine a character's relationship with the computer user, not only leading to less liking, but also to less influence. *Designers of characters that are meant to play a tutoring, guiding or shopping role may want to be especially careful about consistency.*" (Isbister & Nass, 2000). This last point is of obvious import for the design of the Learning Companions reported in this thesis. Isbister and Nass also point out that inconsistency can occur in a character over time. For example, a computer personality could adapt from a complimentary one to a similar one as described previously. There is a possible contradiction here that *it may be preferable for a character to adapt it's personality so that it is similar to a user but this apparent lack of consistency would be contradictory to this empirical work on ECAs* and, indeed, other disciplines involved in the creation of a character's personality such as those previously mentioned. It may be that an adaptive personality along the lines of the similarity-attraction hypothesis takes precedence over the need for consistency. This is something which will need to be examined in future work. However, to further complicate matters this study by Isbister and Nass also found that participants actually preferred a complimentary personality thus contradicting the previous studies.

One possibility for this outcome proposed by Isbister and Nass includes the added effects of utilising a character as opposed to the text personifications utilised in the earlier study which they state is something they would like to study further. A further proposed explanation could be the differences being due to two different trait scales being manifested, i.e. the earlier study was concerned with dominant-submissive traits whereas this study is concerned with extrovert-introvert traits which are related but dealt with slightly differently in the interpersonal Psychology literature. Again, Isbister and Nass (2000) state that a further study is needed to investigate the differences between the two scales of traits. Finally, of particular relevance to this research, they say:

> "When designing for initial engagement with the interface and ongoing general social interaction, the intorversion-extroversion scale may be more relevant; when designing interactions in which joint control and control shifts become important to manage, dominance-submissiveness may turn out to be a more important factor in user experience"
>
> (Isbister & Nass, 2000)

That is, tutoring systems employing Learning Companions which share the responsibility for producing a joint product, such as a summary, may be best employing dominant-submissive personifications. However, "initial engagement" is obviously of paramount importance in capturing the attention of a potential collaborator and "ongoing social interaction" is necessary to prolong a collaboration. To this end, this research, as already mentioned, actually utilises 'affectations' in the design of believable agents relating to both these scales, see Chapter 6.

This section sought to emphasise the importance of affect in the development of animated pedagogical Agents, in particular in the development of appropriate personalities. The next section will demonstrate an approach to conveying the 'humanness' of an agent resting, not principally upon the necessity of an agents' 'personality', but instead upon the grounding of that personality in a 'body'.

2.4.3. An Adjunct of Affective Computing: Creating Embodied Agents.

An argument has already been put forward above to state the case of the interrelatedness of the mind and body. This argument mainly rested on the premise that emotive experiences are inextricably linked to their physical grounding, both in terms of bodily states of arousal and the submergence of that body in an environment. An additional argument for creating artificial bodies for artificial agents is proposed by researchers which basically rests on the assumption that human-agent (and possibly agent-agent) interactions are more meaningful in terms of social interactions if all possible modalities of these interactions are employed. In human terms this involves utilising Non Verbal Behaviour (NVB) as well as speech in conversation. As Cassell (2001, p.68) states:

> "people communicate with and to other people and not in a vacuum. Eyes gaze at other people and focus other people's attention on shared targets, hands gesture between people, faces express to other people. These behaviours are the external manifestations of social intelligence and trustworthiness...as well as a localization of the conversational processes of grounding information and a representation of information in their own right."

This would lead us to believe that in order to fully utilise the plethora of social communications which "come for 'free'" (Cassell et al., 2001) these social aspects of our artificial agents can be represented in terms of artificial bodies. However, it is crucial, as Cassell et al (2001) point out, that these other communication modalities are integrated functionally rather than simply performing 'additive' roles. That is, on occasions, these artificial bodies can better represent communicative functions than can their associated speech act, for instance. This point is, in fact, made by Cassell et al as a criticism of other research in this area of embodied agents and has been duly noted in this research. For example, using a head nod rather than a verbal "OK" to confirm understanding of a command. Furthermore, Cassell separates *interactional* or 'social' functions, i.e. social conventions for conveying meaning such as 'turn taking', and *propositional* or 'domain' functions, i.e. the world knowledge requiring an understanding of the discourse (Cassell, 2001). Cassell's work on embodied agents is out of necessity based upon research of human-human conversations, particularly those relating to NVB (Cassell, 2000).

Cassell et al (Cassell et al., 2000) describe the Functions, Modalities, Timing, Behaviours (FMTB) conversational framework which is used as a model for the generic design of Embodied Conversational Agents (ECAs). Briefly, 'Functions' refers to the distinct separation of propositional and interactional function already mentioned. 'Modalities' refers to sensitive use of multimodalities such as appropriately choosing speech if an agent is currently silent or a head nod otherwise to affirm. 'Timing' is important in terms of synchrony of, for example, animations and speech and is fundamental to the notion of *entrainment* or the synchronisation of speakers behaviours to one another. Finally, 'Behaviours' refers to conversational function existing separately from its form, for example a character could gain attention by applauding or shouting an expletive. This model, together with research in linguistics, sociology and human ethnography (Cassell, 2000) is then taken to arrive at design requirements for ECAs (Cassell et al., 2000, p.41). These design requirements are then, in turn, used to arrive at a generic architecture for ECAs (Cassell et al., 2000, p.43) consisting of an Input Manager (to convert input into a form suitable for processing), Hardwired Reactions (enabling instantaneous agent activity, such as user tracking), a Deliberative Module (consisting of an Understanding, Decision and Generation sub-module, the Decision module being further decomposed into Interactional and Propositional processing and a Response Planner) and finally an Action Scheduler which schedules and synchronises, where necessary, motor processes.

This FMTB architecture has been implemented as an agent named REA (Real Estate Agent) who acts as an agent engaging in both formal and informal conversation concerning property purchasing. This implementation utilises a number of imaginative technical solutions such as IBM ViaVoice for speech recognition, BT Festival for speech generation and STIVE for image processing. A full description of this system can be found in Cassell et al. (2000; 2001).

In terms of the evaluation of the FMTB model Cassell et al (2000) mention the lack of a theory to generate conversational behaviours from their functions, that is, there is a need to have a "morphology of conversational behaviours" (Cassell et al., 2000, p.59). Additionally, there are reported weaknesses in the synchrony of events in REA leading to a lack of the aforementioned entrainment, an essential quality, without which REA may appear to be malfunctioning. Finally, in terms of evaluation, Cassell et al have used a Wizard of Oz approach to compare user's reactions to agents with and without interactional behaviours finding that "users judged the version with interactional behaviours to be more collaborative and more cooperative and to exhibit better natural language (even though both versions had identical natural language abilities)" Cassell et al. (2000, p.59). Interestingly for this research, future work carried out by Cassell et al. will examine ways to enable REA to anticipate which conversational style is best suited to a user, e.g. whether or not to use 'small-talk' to personalise a conversation.

Andre and Rist (2001) describe an ongoing research effort to build presentation agents, that is, embodied agents which present information to users thus supplementing or completely obviating the need for more 'traditional' methods of presentation. The hope is, similar to the work of Cassell cited above, that the use of embodied agents will thus utilise the 'free' communication aspects of human-human communication.

Andre and Rist (2001) make the distinction between agent behaviour which is *scripted* and that which is *generated*. They define a script as "a temporally ordered sequence of actions including body gestures, facial expressions, verbal utterances, locomotion, and (quasi-) physical interactions with other entities of the character's immediate environment" (Andre & Rist, 2001, p.53). They point out that scripts can be written in specific languages or in the case of MS Agent (the system used in this research) Visual Basic or Java can be used to control the agent thus facilitating the programmer with the usual high level language facilities. Interestingly, Andre et al (Andre et al., 2000) also distinguish between a script and *self-behaviour*, i.e. "not only gestures that are necessary to execute the script but also navigation acts, idle time gestures, and immediate reactions to events occurring in the user interface". However, manually generating scripts in such a way can be an overbearing burden for the agent behaviour author so to address this problem Andre and Rist employ what they term a "generative mechanism", that is a method which composes scripts based on some higher level generative rules. Furthermore, these rules can be employed together with a plan-based approach to behaviour generation which can decompose high level goals to lower levels of detail, of course depending on the objectives of the agent interactions. It must be noted though, that the work reported here by Andre and Rist is concerned solely with the domain of *presentation agents* whereas the work described in this thesis is concerned with the behaviour of the agent which relates to the process of effective summarisation (studied in Chapter 4). However it is worth examining some of the issues surrounding presentation agents as they potentially relate equally to the summarisation agents.

Andre and Rist (2001) describe their work on the PPP Persona and the AiA system. The PPP Persona system was designed to provide an animated agent interface to a system to describe technical devices whereas the AiA system helped users utilise the web by ordering information and helping them navigate the 'hyperspace'. They describe the essence of their script generator, already mentioned, as a typical AI planning procedure, that is a plan operator represents a communication goal plus the body of the operator indicates which acts will achieve this goal. A hierarchical type of planning decomposition process thus ensues to reduce high level presentation goals to "elementary production-retrieval or presentation tasks" (Andre & Rist, 2001, p.57). Although, the complete script is not entirely produced in advance but at key interludes where the user is allowed to influence the navigation space through a hypermedia type of interface. However they do criticise this kind of scripting approach with its limited interactive capabilities. On the other hand they had some positive results during evaluation, stating that subjects who had used a persona "found the presentations themselves and the corresponding tests less difficult than subjects who had seen presentations without persona" (Andre et al., 2000, p.222). This is compared to the 'persona effect' of Lester et al (1997). They do, however, embellish this claim with the additional statement for the need to "model the agent more deeply – for instance, by giving it personality" and further elaborate by saying "a necessary requirement for the success of such presentations is that the agents come across as socially believable individuals with their own distinct personalities and emotions" (Andre et al., 2000, p.223).

Andre and Rist (2001) present a number of arguments for the use of presentation teams as opposed to the single agent systems described elsewhere. They say that such approaches enable different rhetorical relations to be anthropomorphosised (e.g. pros and cons), allow less monotonous repetition (by using multiple agents for the same information) and provide different perspectives on the same information. To this end they have investigated a car sales scenario, inhabited marketplace (IMP), and a tool to help visualise message exchange (MAGIC MONITOR). IMP takes into account a user's profile and the role and personalities of the characters giving the presentation can be chosen by the user, see Andre et al (2000, p.229) although there is no user interaction during run-time. To elucidate, the character can be a buyer or seller and can have an *extroversion* personality trait (with values of extravert, neutral, introvert) and an *agreeableness* personality trait (with values of agreeable, neutral and disagreeable). Andre et al go to great lengths to explain that they do not have an emotional model per se (as described in section 2.4.4) but instead use these internal character 'states' as constraints upon the selection of appropriate behaviours. This compares favourably to the approach taken in this research with the difference that individual characters have set internal 'states', that is, although distinct, they cannot by customised by the user. Instead, the onus is placed on the user to self-select their desired characteristics by choosing the character they prefer.

MAGIC MONITOR graphically demonstrates the originators of dialogue in the MAGIC LOUNGE virtual meeting space. The IMP system uses a similar generative planning system to that described above for PPP and AiA but, of course, multiple characters must be taken into account. This is achieved by extending the communicative acts by dialogue acts such as turn taking, etc. Similar to the work of Cassell cited above dialogue acts now not only contain propositional content but also the interactional content. Additionally, Andre and Rist point out "*characters have to be realised as distinguishable individuals with their own areas of expertise, interest profiles, personalities, emotions and audiovisual appearance*", i.e. they make a concerted effort to create distinct characters in the interactions. Hence, the character's distinct personality is used as a 'filter' for the dialogue. MAGIC LOUNGE utilises this same planning approach, acting first as a 'screenwriter' to convert the dialogue to character-based presentation then as a 'facilitator' of the presentation. Thus, in both IMP and MAGIC LOUNGE all information pertaining to the dialogue must be known in advance of the presentation so that behaviours can be authored by the planner. This led Andre and Rist to consider agents which did not have this kind of a priori knowledge requiring improvisational presentation such as soccer commentary.

ROCCO II (Andre et al., 2000) is a system to generate live commentaries on the simulated RoboCup soccer competition. Two animated agents sit on a sofa in front of the game, sipping beer and commentating on the match as it unfolds. The difference with this approach compared to the previous work such as AiA and IMP is that the agents are no longer able to be modelled utilising the planning approach, instead they are modelled as (semi-) autonomous, distributed agents: "we assign each agent its own reactive planner and code the agents' dialogue strategies as operators of single planners" (Andre et al., 2000). The resulting dialogue is thus a result of the interaction of the two reactive systems. Andre et al employ a notion of background information which is used when there is no live commentary available. Additionally, as mentioned above, a character's personality is again used as a filter. Furthermore, each character is assigned a soccer team. These embellishments thus help to coordinate the dialogic behaviour between characters. However, Andre and Rist point to problems inherent in this approach such as the resultant dialogue being incoherent as there is no prior agreement between agents leading to a lack of "global organisation", i.e. the dialogue lacks a goal and may simply 'ramble'.

Also relevant to this research is the use of language, both in terms of linguistic and acoustic realisation. Andre et al (Andre et al., 2000, p.242) classify utterances in terms of verbosity, specificity, force, formality, floridity ('flowery' language) and bias (an evaluation of an event). For example, positive bias dialogue templates will be selected if a character is in favour of a team and forceful language is used for extroverts whereas hesitant language is used for introverts. Furthermore, acoustic variation of speech can lead to the impression of affect, for example Andre et al represent arousal by using a higher talking speed and pitch range. They do say that they are limited by the choice of speech generation engine, TruTalk, which they are using in terms of the parameters available to program. They also state that MS Agent provides no means for "detailed intonational markups" (Andre et al., 2000, p.234). However, this is not quite the case for the research described in this thesis which utilises individual phoneme assignment of pitch through the use of the MASH 'sing' option (see Chapter 6 and Barker, 2003). As Andre et al (Andre et al., 2000, p.246) conclude from their evaluation studies "*the look and voice of a character are indeed important cues concerning it's personality and interest profile*". This issue will be addressed further in Chapter 6. They go on to state that the systems they are building are not meant to be complete models of personality but instead "test beds that allow for experiments with various personalities and roles" (Andre et al., 2000, p.251). These experiments, they say, would include investigations of which personalities and roles are appropriate to which situations and point towards future research to investigate the automation of such selection.

To increase user-participation in presentations Andre and Rist (2001) are exploring the use of 'helper' characters which can influence both an antagonist and a protagonist in a resultant dialogue, such as the discussion of selling and buying a car, respectively. The user can thus cause changes in the character's emotional state or present a new dialogue goal, such as providing evidence for a claim made by the seller. The adoption of this approach arose out of the need to constrain the dialogue utterances (that is, the user can only influence dialogue utterances indirectly) and the desire to allow the user to either assume active (i.e. influencing the dialogue) or passive (i.e. allowing the agents to interact unaided) participation. This latter direction in their work has led Andre and Rist to consider the unpredictability of human intervention in their presentation systems, something which had to be considered from the outset in this research and was mainly met by the imposition of an artificial dialogue 'language' as something of a compromise as described in Chapter 6.

Rist and Schmitt (2002) are modelling the effect of social relationships between embodied characters utilising socio-physiological theories of 'cognitive consistency'. That is, in their Avatar Arena demonstration system, multiple agents, who must interact to solve diary management functions, must balance their cognitive configurations. They illustrate the meaning of cognitive configuration by stating for Person P, another person O and object X "if P likes O and P perceives the objects X1 and X2 being similar and at the same time believes that O also perceives X1 and X2 being similar, P's perception of the situation (i.e. it's cognitive configuration) is balanced" (Rist and Schmitt, 2002, p. 82). If a character subsequently realises that its assumptions about another characters beliefs were unfounded, i.e. conflicted with it's own, then the cognitive configuration will become unbalanced resulting in an increase in the social distance between the two characters. Rist and Schmitt see a major application of this approach being in training people in social interaction although, to date, their system lacks a model of affect currently under development.

The important aim then of this type of embodiment in artificial agents is to emphasise the essential qualities of a 'natural' conversation, utilising the 'free' abilities employed in everyday face-to-face interactions. However, although some of this work is informative to our research it is not intended to fully implement a human-like conversation. In particular it is not intended to employ all of the typical NVBs, attempted by Cassell's work, as the ultimate goal is not to reproduce as human an interaction as possible. Instead, a more modest goal is to provide Edutainment (Barker, 2002) utilising artificial dialogue and, to some extent, artificial NVBs. That is, pedagogical and motivational value are of more importance than the creation of a totally 'natural' Human Computer Interaction. However, as can be seen, the work mentioned in this section on personality, in particular, has informed parts of this research, as will be demonstrated in Chapter 6.

As already mentioned, there has been some interdisciplinary work previously carried out on emotion synthesis such as the OCC model and others. The next section reviews some of the most notable of these models and the subsequent section will illustrate this work with implemented systems.

2.4.4. Examples of Affective Computing Models.

Emotion synthesis is a key element of imbuing computers with emotional behaviour and is of particular interest here in the development of personified Agents. Picard (1997) decomposes emotional systems into five components:

1. Emotional behaviour;
2. Fast primary emotions;
3. Cognitively generated emotions;
4. Emotional experience: cognitive awareness, physiological awareness and subjective feelings;
5. Body-mind interactions.

Picard (1997) states that "just as all animals do not need emotion systems as sophisticated as a human emotion system, neither do all computers". As will be seen later, it is intended to implement the first component of emotion synthesis within an artificial Agent as part of this research (the 'affectations' already referred to). However, it is worth reviewing key emotional models employed in the design of computer systems which simulate human emotions. Thus it is hoped that future software development may be informed by these approaches. In fact, most emotion models implemented as software only address the 'cognitively generated' category shown above.

The choice of the following two exemplars is based on the reviews of Picard (1997) and Hudlicka and Fellous (1996). Picard delineates approaches amongst three categories relating to the five components stated earlier, notably: emotion synthesis via cognitive mechanisms, emotion synthesis via multiple mechanisms (including low-level approaches) and synthesising emotions' influences (e.g. in decision-making processes). Hudlicka and Fellous (1996) in turn first delineate approaches to emotion synthesis amongst five characterisations :

1. *level of abstraction* (e.g. architecture/task/mechanism level);
2. *computational model* (e.g. symbolic/connectionist);
3. *type of emotional phenomenon* (e.g. intrapsychic, social processes, impact on cognitive processing and appraisal, role in monitoring/planning and dysfunctional);
4. *underlying theoretical framework* (e.g. emotion as secondary or central to cognition);
5. *research objective* (i.e. 'performance' versus 'process' models).

Hudlicka and Fellous (1996) then concentrate on the 'type of emotional phenomenon' characterisation by which to categorize theories and associated models. The frameworks highlighted thus include: cognitive appraisal, emotion as goal-monitoring, emotion and cognition interaction and emotion as social coordination. In terms of the two categorisation systems the examples that follow and any implementation details present themselves as the most pertinent to this research and illustrate the approach of emotion synthesis.

The OCC Model of Ortony et al (1990) is reputedly one of the most widely accepted and utilised models of emotion. In fact, although computer modelling was used in the validation of the model it was never intended to be used to synthesise emotions as such. Ortony et al (1990) say of their work : "the point of such an enterprise is not to create machines with emotions – we think that such an endeavour would be pointless and futile – but to create a computer model that can 'understand' what emotions people would be likely to experience under what conditions". However, they succeed in modelling a comprehensive range of emotive phenomena in a tractable form suitable for a computer-based implementation. This has led to its widespread popularity in emotion synthesis.

The model is based upon an appraisal system where the world is divided into situations consisting of events, agents and objects. Emotions, then, arise from valenced (i.e. positive or negative) reactions to these situations. The diagram in Appendix D shows the 22 emotion types that the model delineates. They emphasise that their choice of emotion words is not critical. Instead, they are more concerned with the place of the model in the structure with the words serving a role as technical terms rather than as ultimate definitions. In the diagram, labels in upper case represent structural elements whereas labels in lower case represent emotional, or potentially emotional, states. Panels at the bottom of boxes containing individual emotions contain the name given to the group. CONSEQUENCES OF EVENTS are "reactions of the experiencing person to the implications of events for the person experiencing the emotion". ACTIONS OF AGENTS represent "emotions having to do with people's reactions to the agency they attribute to agents". ASPECTS OF OBJECTS represent "undifferentiated affective and aesthetic reactions to objects". Other structural elements are similarly defined. Variables are introduced in the model which affect an emotions intensity. For example, events and their consequences have a *desirability*, actions of agents have a *praiseworthiness* and aspects of objects have an *appealingness*. The variables are then used in rules to derive the valence of a particular emotion's intensity. Furthermore, emotions can interact with each other or trigger themselves. Implementation details of the OCC model are sketchy. However this has not precluded it from being influential in computer-based emotion synthesis.

A further theory worthy of mention here is the CogAff architecture (Sloman, 1999) which proposes a schema consisting of a triple-layered approach to an affective architecture. The three layers consist of a reactive layer, a deliberative layer and a self-monitoring layer. The reactive layer relies on automatic 'visceral' responses to stimuli, for example, 'startle' or 'disgust'. For example, a response could result automatically on hearing an unexpected loud noise close by. A suitable emotional label for the behavioural reaction to the noise might be 'startle'. The deliberative layer, on the other hand, involves some evaluation of stimuli, for example constructing a plan to attain a behaviour or learning how to achieve actions. Any plans learned by this layer can subsequently be passed to the reactive layer. Emotional labels which this layer may generate include for example, 'pleasure' resulting from achieving a plan. The third layer involves the use of reflective meta-management processes such as preventing goals in the lower layers from interfering with each other or choosing more efficient strategies. Emotions which this layer may attain include shame and grief, for instance. The model appears to be based on the supporting psychological and neurological literature (Damasio, 1994), particularly the idea of primary and secondary emotions. However, its use in synthetic character work is limited, to date.

2.4.5. Examples of Affective Computing Applications.

The following examples serve to further illustrate the use of the OCC model and its derivatives in particular as, has been noted, this is a key approach used in affective computing.

The aim of the 'Oz' project at Carnegie Mellon is in "developing technology for artistically interesting simulated worlds" (Bates et al., 1992). To this end Agents are broadly designed to include goal-directed reactive behaviour, emotional state and behaviour, social knowledge and behaviour and natural language abilities. Bates et al. are concerned with creating believable characters inspired by the master animators of Walt Disney (Bates, 1994). They state that at any one time a clear emotional state must be displayed by animated characters thus creating the *illusion* of life in a character. They have developed an architecture called Tok which integrates various sub-components including those implementing perception (Sensory Routines and Integrated Sense Model), reactivity and goal-directed behaviour (Hap), emotion and social relationships (Em) and language analysis and generation (Gump and Glinda respectively). As the interest in this research is presently focussed on emotion synthesis Em will be examined in detail. Bates et al illustrate their Tok architecture by means of the implementation of a simulated cat, Lyotard, living in a simulated world. The perception and action components are described. Interestingly, Hap, the action component, bases the Agent's action on perception, current goals, emotional state and behavioural features amongst other components. It uses plan schema to derive behaviours based on context conditions and success tests. However, of primary interest here is the Em model.

The Em model utilised in the Emotion and Social Relationships component of Tok is, like Elliot's AR project to follow, based on the OCC model (Ortony et al., 1990). It is cognitively based in that external events are compared with goals, actions are compared with standards and objects (which can include other agents) are compared with attitudes. Emotions used in Tok include: hope, fear, happy, sad, pride, shame, admiration, reproach, gratification, remorse, gratitude, anger, love and hate. As Hap runs goals are created, succeed and fail. These results affect the emotional model Em. For example Happiness and Sadness occur when goals succeed or fail or Hope and Fear occur when there is a chance of a goal succeeding or failing. Standards are used to either approve or disapprove of actions resulting in Pride, Shame and Reproach. Anger, Gratitude, Remorse and Gratification result from combinations of other emotions. For example, Gratification is a result of Happiness and Pride. Some emotions always coexist, for example Pride-Gratification due to the choice of standards. Finally, positive or negative attitudes towards objects are used to derive the emotions Love and Hate, for example Lyotard will Love a human who makes him grateful. Em also models the decay of emotion over time. Em, then, adjusts behavioural *features* which "modulate the activity of Hap" (Bates et al., 1992) thus creating emotional behaviour. Reilly (1996) provides more details of the Tok and Em architectures.

The Affective Reasoner (AR) (Elliot, 1992) is also based on the OCC Model work described above by Ortony et al (1990). It is an appraisal-based computer model of emotive states. Elliot et al (1999) describe the AR's conception of emotions as "a by-product of goal-driven behaviour, principled (or unprincipled) behaviour, simple preferences, and relationships with other agents". In fact, they go to great lengths to dispel claims of Elliot's notions of emotions being related in any way to those grounded in a physical body. The comparisons between his approach and that of the OCC model are obvious. He describes "pseudo-personalities" modelled as:

- appraisal frames – representing an agent's goals with respect to events;
- principles – with respect to perceived intentional actions;
- preferences – with respect to objects;
- moods – temporary changes to the appraisal mechanism;
- channels – for the expression of emotions.

To elucidate, the appraisal frames, or *goal-based emotions*, for example joy and distress, are activated when the agent matches an event with its goal-based concerns. For example, a typical goal for Steve (Rickel & Johnson, 1999) – a pedagogical agent which inhabits a virtual environment for Naval training - might be "I will give explanations about the subject domain and discuss interesting details about the domain with the student" (Elliot et al., 1999). It is thus hoped that Steve's expression of happiness resulting from being able to give the student some useful information, for instance, will convey 'enthusiasm' to the student. However, these goal-based emotions may be further modified by Steve's relationship to the student. For example, if Steve and the student are 'friends' then Steve may 'feel sorry' for the student if they get something wrong otherwise he may 'gloat' which may be useful to encourage competition. To further elucidate, *principle-based emotions*, for example, irritation and reproach, are generated when an agent makes a kind of moral judgement about another agent's actions. For example, a typical principle for Steve may be "Student should attend to me when I am talking with them". Thus, if the student does not currently have Steve in their field of vision when Steve is attempting to demonstrate something, for instance, then Steve will manifest a kind of 'annoyance'. Finally *preference-based emotions* are activated, for instance, when an object matches with Steve's preferences. For example, he could 'like' or 'dislike' that object. Thus situations which the agent encounters match to twenty-six different emotion types, twenty-two of which are based on the OCC model (see Appendix D). Furthermore, *emotion intensity variables* determine the quality and intensity of each emotion.

Returning to the effects of this affective modelling on an agent's personality, Elliot et al (1999) state "it is thus necessary to describe the personality of the automated agents in terms of the situations that arise, characterized as sets of goals, principles and preferences in the content domain." In addition to Steve, Elliot's 'AR agents' utilise a number of expressions to convey emotional state or this 'personality'. Elliot &

Brzezinski (1998) state that they "have highly expressive schematic faces, speak with somewhat emotionally inflected voices, listen to users through speech recognition and use a rich set of musical selections to help reflect their current states". In terms of evaluation of these agents Elliot & Brzezinski (1998) state that users more correctly matched emotive content of the AR agents compared to that of videos of human actors when both were reading ambiguous news items effectively filtered through their respective 'personalities'. Future work proposed by Elliot will focus on affective modelling of user's states in an attempt to further 'humanise' the social relationship formed between computer and user which is particularly useful in the field of tutoring systems such as that described in this thesis. Methods to achieve this aim, or at least approximate it, include utilising such mechanisms as:

- inquiry – ask the user how they feel;
- stereotypes – make assumptions about user types, e.g. like to win, etc;
- context – e.g. a user who has repeatedly failed may feel bad;
- affective stereotypes – infer how most users would feel;
- self-inspection – base an assumption of a user's affective state on the agent's state.

Further work on affective perception will be described in Chapter 9 as the need for it in future research will then become more evident. Finally though, to round off this section on Affective Computing, some general issues pertaining to the field will be discussed.

2.4.6. Concerning Affective Agents.

Johnson and Rickel (2000) present an overview of concerns in the development of animated pedagogical agents. These include benefits such as navigational guidance, gaze and gesture as attentional guides, nonverbal feedback, conversational signals, conveyance and elicitation of emotion, virtual team-mates (c.p. 'Learning Companions') and adaptive pedagogical interactions. They constantly refer to implemented systems such as Steve (Rickel & Johnson, 1999), Adele (Shaw et al., 1999) and Herman the Bug (Lester et al., 1999). They go on to elucidate technical concerns in the development of animated pedagogical Agents. Of special concern here are the problems associated with 'conveyance and elicitation of emotion' as, as mentioned, these are seen as central to the animated pedagogical Agents. Of special concern here are the problems associated with 'conveyance and elicitation of emotion' as, as mentioned, these are seen as central to the creation of a believable character. Johnson and Rickel state that in order that Agents can display appropriate emotional behaviour emotional intent (as designed by the behaviour artist) has to be integrated with pedagogical intent. In fact due to the pedagogical implementation of these types of Agents it is not necessary to implement a broad range of emotional behaviour. For example, as Johnson and Rickel (2000) state: "they should be able to exhibit body language that expresses joy and excitement when learners do well, inquisitiveness for uncertain situations (such as when rhetorical questions are posed), and disappointment when problem-solving progress is less than optimal". They cite examples of how the Cosmo Agent (Johnson & Rickel, 2000) employs an emotive-kinaesthetic behaviour sequencing approach. Firstly, designers must create appropriate behaviour for Cosmo in the 'behaviour space'. Secondly, these must be indexed to their emotional content. Finally, the emotive behaviour must appropriately be integrated into the pedagogical behaviour to create longer cohesive behaviour fragments. They conclude by saying that research which integrates emotive computer models with interactive learning environments "has only just begun".

A further point is that there is a danger in such anthropomorphisation of agent software which is well documented in the literature (Preece et al., 2002, p.155), particularly in relation to embodied interface agents. In some instances it seems that attributing such human qualities to software is an unnecessary step in overcomplicating software design. For instance, Elliot et al (1999) describe a situation where Steve may have a goal of becoming 'angry' in order to obtain the attention of the student but then has a 'patience' principle which leads to feelings of 'remorse' or 'shame' for possibly losing it's 'temper'. This raises the question of the validity of the creation of the 'angry' goal in the first place. Without it Steve would not have to 'feel' shame as the action would not take place. Even more protracted affective reasoning may, indeed, be possible. Undoubtedly, modelling affective states of agents (either artificial or human) is a complicated business. However, it is worth not losing sight of the original aims of the inclusion of affective states, notably to increase the sociability of human-computer interaction with a resultant increase in student motivation (this is the more modest aim of this research). To this end, complex models of affect may prove to be unnecessary or indeed inappropriate.

2.5 Summary.

This chapter began by introducing a 'pedagogical framework' to be employed in the implementation of a successful Virtual Learning Environment, such as the one studied within the context of this research. It highlighted three phases of the framework, notably Acquisition, Argumentation and Application. Each of these three distinct but interweaved phases was then elaborated upon in terms of relevant literature. Subsequent chapters will demonstrate the effectiveness of this approach through empirical validation and report on techniques which can be utilised to improve and further refine it.

The second section of this chapter touched on some of the Educational and Psychological literature pertaining to collaborative learning with a view to providing a rationale for the future design of a tool with which students collaborate to perform the summarisation task. Additionally, Section 5.1 will discuss collaboration between a human learner and an artificial Agent in more detail.

The third section of this chapter reported on the task of summarisation. In particular the task was placed into the context of generic writing skills and approaches to teaching them such as the differences between writing as a process or as a product. An exploration of a definition of a summary was attempted together with an investigation of how pedagogical techniques can be employed to improve summarisers' performance. Finally in this section, the role of the computer in automatic summarisation was explored with a view to providing a tool which could aid students as they summarise.

The fourth and final section of this chapter was concerned with the role of Affect in Computing. The section began by placing the study of Affect into the historical context of its philosophical development before discussing it's relevance to this research. The next section discussed the design of exemplar suitably personified "believable" Animated Pedagogical Agents with hypothesised consequences for student motivation. In particular the section discussed issues of personality in the role of believability and reported on research which sought to investigate the relevance of interpersonal Psychology theories in the realm of Human-Computer Interaction. The subsequent section considered a particularly relevant adjunct of Affective Computing, notably the need for and means of creating Embodied Agents. The next section then highlighted some key models of Affective Computing then the penultimate section culminated in a discussion of some implemented computer systems which have taken these models into account, for example the use of the OCC model by Bates et al in the 'Oz' project. The final section in this Chapter then rounded off the reviews by considering some of the more general issues in animated pedagogical agent design such as the integration of emotion research and the dangers of anthropomorphisation.

With some of the key literature relevant to this research now explored subsequent sections will report on studies carried out in the field (at CHALCS and other local schools) which relate to the original Research Aims as stated in Chapter 1. In addition later chapters will report on the design, implementation and validation of a Learning Companion System which seeks to help students summarise course notes within the Virtual Learning Environment. First though is a report on the development of the web-based Physics course material to be embedded in the VLE.

Chapter 3: The Development of a Web-Based Physics Course.

3.1 Introduction.

In this chapter the development of the web-based Physics course at CHALCS is described. The selection of the tools and development of the course materials was based on the results of an initial questionnaire administered to CHALCS students to assess support requirements also described in this chapter. The development of the course materials and initial piloting of them within the Virtual Learning Environment (VLE) took place over a 12 month period. The necessary Internet infrastructure was put in place and novel material developed and uploaded into the chosen Virtual Learning Environment (VLE), WebCT. This chapter reports the lifecycle of the web-based course production. Firstly, the tools examined in the early days of the work are highlighted and compared followed by an overview of the main functionality of WebCT. Secondly, the development process itself is described together with an overview of the course. Thirdly, preliminary results of what may be termed the 'pilot' phase are presented. Finally, conclusions are reached as to the future direction of the research.

3.2 Review of Web-Based Learning Tools.

Before the course materials could be developed and piloted with CHALCS students a suitable set of tools with which to create them and the learning environment needed to be identified. A review of the available technologies at that time was undertaken at the beginning of the project in which a number of integrated VLEs (web-based classrooms) and potential 'component' solutions (based on a number of different stand-alone software systems) were examined. Based on the pedagogical rationale described in chapter 2 key features of such a Virtual Learning Environment included that there be opportunity to provide areas in which students had access to resources and course materials, space to articulate with each other and the tutor their understanding of that material and opportunity to practice or apply their knowledge within both self-paced and collaborative activities on which they would receive peer and or tutor feedback. Given these requirements it was important to select a VLE for the project which would incorporate both content development and management tools and communication and discussion tools together with tools to support collaborative group-working plus provide opportunity for self-assessment or reflection when working independently. Traditional VLEs have a 'structural' approach to courses as content, generally with asynchronous discussion and student activity tracking, and sometimes with Multiple Choice Quizzes. Some VLEs like Firstclass primarily support synchronous and asynchronous discussions but may be more limited in their content management tools (Bostock, 2000).

Table 3.1 compares the tools available at the time on the key attributes relating to these requirements. It is important to note that since the outset of the project several new VLE tools have been developed and the capabilities of some existing tools extended including the chosen tool WebCT - see, for example, CHEST's updated comparison grid (CHEST, 2002) for five popular VLEs involved in the recent JISC interoperability pilots.

The key point here is that for the requirements of the proposed course there was a need to balance adequate discussion tools with the need for some content management and student tracking facilities and the ability to provide opportunity for collaborative group-work and tutor feedback.

Features	Learning Space	Top Class	WebCT	Merlin	Event Ware
HTML knowledge	No	No	No	-	No
Security	Yes	Yes	Yes	Yes	No
Student space	No	Yes	Yes	Yes	No
Quizzes	Yes	Yes	Yes	-	No
Student management	No	Yes	Yes	-	No
Email	No	Yes	Yes	Yes	No
Bulletin board	No	Yes	Yes	Yes	No
Chat	No	No	Yes	Yes	Yes

Table 3.1 - A Comparison of Virtual Learning Environments.

In Table 3.1 'HTML knowledge' indicates that students or designers needed a knowledge of HTML to use the environment (i.e. to upload, edit, etc. course materials); 'Security' indicates that the environment was password protected (and, for example, that communication with tutors or peers can be confidential); 'Student space' indicates whether students are able to have their own area within a course (thus ensuring that students have a 'safe' space to work in, possibly in a group); 'Quizzes' indicates that the environment supported inclusion of either Multiple Choice or Long Answer quizzes (e.g. to support student self-assessment and/or provide the tutor with feedback on student progression); 'Student Management' indicates that the environment provided tools for the designer to manage students (e.g. help with maintaining a score database thus easing the administrative overheads); 'Email' refers to integral email facilities; 'Bulletin Board' refers to integral asynchronous threaded discussion forum facilities with tutor editing capabilities and 'Chat' refers to integral real-time synchronous chat facilities (these last three facilities providing the infrastructure for the collaborative learning, as we will see in a subsequent section). A '–' indicates that at the time support for a facility was unknown.

Of all the VLEs reviewed at the outset of the project, WebCT alone supported all of the necessary functionalities. In addition it was free to download and explore with test material and, importantly, had been well received by colleagues in other institutions. This led to the development of sample material in order to test its functionality.

Initially 'Study Skills' materials were developed around a theme of the American Civil Rights movement using the WebCT server at the University of British Columbia (the originators of WebCT). This investigation proved that WebCT was definitely capable of providing the facilities we required for the implementation of a course at CHALCS. Furthermore, the WebCT licensing package is such that courses can be developed at no cost. Instead, licenses are granted per number of students taking the course. This is certainly an incentive to trial the software thus enabling a full evaluation prior to purchasing. The implication then is that course costs can be passed onto the student so that for a minimal personal cost they have access to the numerous facilities available. This model was to be adopted for the CHALCS course. However, it eventually became possible to negotiate a free limited license with the vendors of WebCT.

In addition to these advantages, the WebCT server can be installed on both UNIX and Windows NT platforms. Clients access the system from a standard web browser such as Netscape Navigator or Internet Explorer. Therefore, once the server is installed it is possible for students to access the course from any number of platforms supporting web-based Internet connectivity. This is an obvious advantage for students connecting from CHALCS, school or possibly home and thus helps achieve this research's goal of widening provision for those wishing to study at CHALCS.

Finally, WebCT provided good support over the Internet in comparison with competitors at that time. This support took the form of email, discussion lists, web pages and, later, the e-learning hub. This was undoubtedly a deciding factor in the uptake of WebCT at CHALCS. In the next section the features of WebCT are illustrated in greater detail.

3.3 WebCT.

Later versions of WebCT require a student to have a global login ('MyWebCT') with which they are able to access all the courses to which they belong on a particular course server although this facility was not implemented in the WebCT version used during initial development. In the initial development the tutor created a login for the student notifying the student of their user id and password. From the MyWebCT page they can then follow links to the chosen course. By selecting a course link they are then taken to the front page for their course and, once inside, they can check for announcements, select any of their personal or public bookmarks or change their password. The course front page will contain a course title and a number of icons which serve to either take them to the course notes or to integral tools provided by the course designer. Tools available to the student include the integrated email facility, the bulletin board, synchronous chat facility and a shared whiteboard tool. Included amongst the facilities for email is the concept of folders in order to manage messages. In fact, most of the functionality is similar to that provided by regular email clients. The folder concept is echoed in the bulletin board tool in the form of forums where discussions take place related to a subject heading. The discussion within the bulletin board is 'threaded' so that replies to a posting will appear indented below the original message (see Figure 3.1).

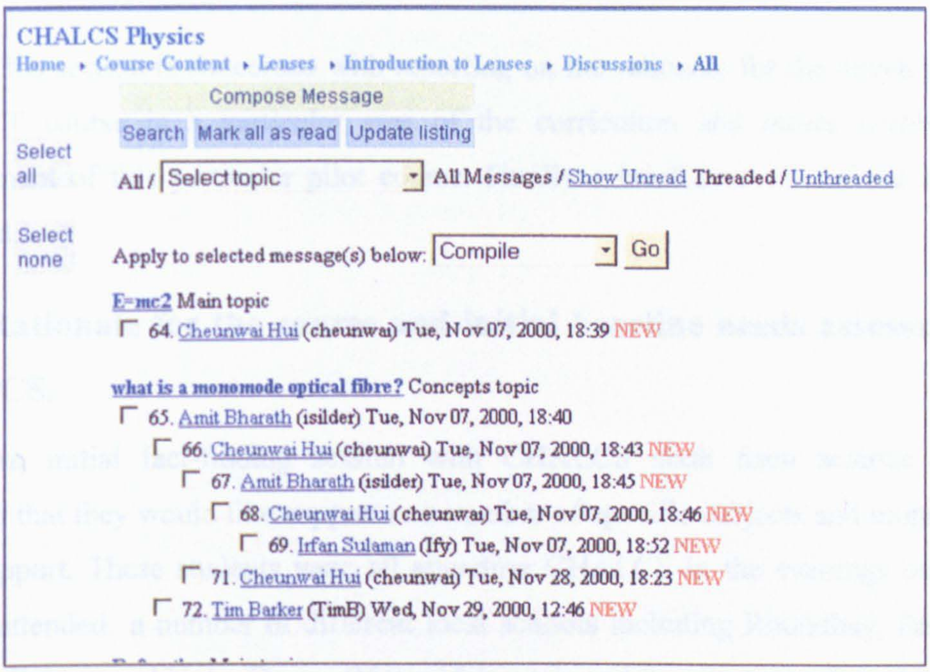


Figure 3.1 WebCT Bulletin Board Threads.

The result is a simple graphical representation of the 'threads' within the discussions. Whilst the email and bulletin board are asynchronous the chat, on the other hand, is synchronous in that all correspondence takes place in Real Time. The student chooses a 'room' in which to participate and can see all contributions from each of the students in that room (as listed) from the point at which they enter.

In addition to the communication tools WebCT also provides numerous other tools. One of the facilities central to this research (demonstrated in Chapter 4) is the 'my-notes' facility. This is a rather simple mechanism which allows the student to annotate the course notes with their own notes. There is also a 'resume' facility which will take the student to the page they were last working on, there is a 'compile' tool which concatenates notes (for example, to print out) and there is a simple 'search' tool. WebCT also contains a group 'calendar' where, for example, dates and times of chat sessions could be published. The facilities available in the actual course notes pages will be elaborated upon in the subsequent section. They include the glossary, navigation features and those particular to the Physics course itself. First, though, it is important to talk about the rationale underpinning the course and its consequent development.

3.4 Implementation of a Physics Course.

This section is concerned with reporting on the rationale for the development of a WebCT course in a particular area of the curriculum and issues involving the development of this particular pilot course. Finally a brief overview of the course is presented.

3.4.1. Rationale for the course and initial baseline needs assessment at CHALCS.

An initial fact-finding session with CHALCS sixth form science students indicated that they would like support in a number of specific subjects and more general skills support. These students were all attending CHALCS in the evenings but during the day attended a number of different local schools including Roundhay, Park Lane, Allerton Grange and Notre Dame. Of the 15 interviewed the most popular subject areas were Maths with 10 of the fifteen studying this subject, 7 studying Chemistry and 6 studying Physics. In terms of extra support required the most popular areas were reference material (7), relevant theory (6) and past papers and writing skills both having 4 students declaring an interest. Reassuringly, all students stated that they had a basic familiarity with Information Technology with a third saying they had Internet access. Finally, all students said that they would like to participate in this research.

Taking this profile into account and the availability of a tutor at CHALCS able to work on the project it was decided to implement a web-based course in Physics which would offer support in generic note-taking skills so that the model could be adapted for use in other subjects. It was, therefore, hoped that this would be the forerunner or exemplar case inspiring other CHALCS staff to author their own material for other subject areas. Astronomy and Optics was chosen in particular because, at the time, this was a new subject area on the N.E.A.B. syllabus being used at CHALCS with little provision in traditional text books, thus the content itself would be novel, filling a gap in the material already available. The volunteer tutor thus felt it was a particularly suitable topic for providing additional support materials via the web.

3.4.2. Development of a VLE course.

Prior to commencement of the project there existed at CHALCS a number of IBM PCs linked by a Local Area Network. However, there was no Internet connection so this was obviously the first priority in the early stages as connectivity was central to the research goals. Due to problems experienced by the installation contractors causing approximately 3 months delay, the implementation of a VLE had to be postponed until Internet connectivity was established. This connectivity consisted of a PC running RedHat Linux supporting the World Wide Web server as well as the WebCT service.

There emerged two distinct teams involved in the implementation of the on-line Physics course. These might be named the **development team** and the **delivery team**. This researcher took a role in coordinating activity between the two teams, assisting with course design activities and supporting the tutor in delivering the material at CHALCS. The development team consists of the author of the Physics content, two secretaries to enter the material into Word and the WebCT course designer who uploads the course notes, creates the course appearance, etc. The delivery team consists of two Physics tutors (the material author and the CHALCS tutor) and a WebCT technical consultant. There is obviously some overlap in these role definitions and they are dynamic.

Aside from the development team authoring and uploading course notes to the WebCT environment it is also necessary for the delivery team to plan lessons using the tools. The three pedagogical stages of Acquisition, Argumentation and Application (see Chapter 2) are mapped onto the tools provided by WebCT (with the addition of class-based face-to-face teaching and the use of Microsoft Word) as shown in Table 3.2. Note that following early trials with WebCT reported later in this chapter the WebCT 'My Notes' facility was not used and was replaced by Microsoft Word running as a separate application. This is the only non-WebCT tool used. The reasons for not using the 'My Notes' facility for note-taking are explained further in section 4.2.

The delivery team used this iteratively developed mapping of the learning objectives to tasks and tools to plan lessons using WebCT. Table 3.3 lists example abridged tasks taken from a CHALCS Physics lesson. These are arrived at by firmly grounding the appropriate tools for each stage in the Physics material. Note that each task is preceded by a demonstration of the facilities provided by WebCT and is complemented by a full set of Learning Objectives.

Stage	Tool	Function
1. Acquisition	Face-to-face	Summary & Introduction
	Quiz	Feedback
	Chat	Brainstorming
	Bulletin Board	Structuring ideas
	Compile	Selecting own notes
	Word (Not WebCT's 'My Notes')	Annotating own notes
2. Argumentation	Chat / Bulletin Board / Face-to-face	Group based explanations and judgements
3. Application	Chat / Face-to-face	Brainstorming qualitative solutions
	Email	Quantitative solutions
	Whiteboard/Chat	Group based model answer
	Chat transcripts / Email feedback & Whiteboard answer	Reflection

Table 3.2 Mapping tools to aspects of the pedagogy.

Example Lesson Tasks.
1. Start to compile the Physics course notes then cut and paste them into Micorsoft Word, adding your own thoughts as you see fit. These are your own notes. Complete this as homework.
2. Discuss solutions to the following problem using the bulletin board: a) Ideally, how thick and how smooth should a glass window in a house be ?
3. Use the whiteboard and chat to solve the following problem as a group: "A person stands at the end of a swimming pool 8 metres in length ...What is the depth of the pool ?..."

Table 3.3 CHALCS "Reflection and Refraction" Lesson Tasks.

Note that each of the three tasks can be seen to relate to the three consecutive phases of the Pedagogical Framework. Appendix E contains a full lesson plan which resulted from this kind of development process and was used at a real CHALCS session. It contains Learning Objectives, Performance Measures, Resources Required and the Lesson Tasks including Demonstrations.

3.4.3. The Physics Course Overview.

The 'Properties of Light' module was used for the purposes of this pilot. Figure 3.2 shows a course notes screen (from the original version of WebCT which actually underwent several revisions during the course of this research).

The top right window is a glossary definition of the term 'electromagnetic spectrum' obtained by clicking on the hyperlink. The left hand frame in the main browser window shows the course pathway and the top frame shows the tools and navigation icons. The main frame shows the course content for 'Properties of Light' including a diagram scanned from the course textbook. The 'gears' icon represents an activity whilst the 'URLs' icon is a link to external resources such as encyclopaedias. In addition a 'question mark' icon is sometimes used when students are asked a question which requires an elaborate answer, for example by searching the on-line encyclopaedias in the 'links'.

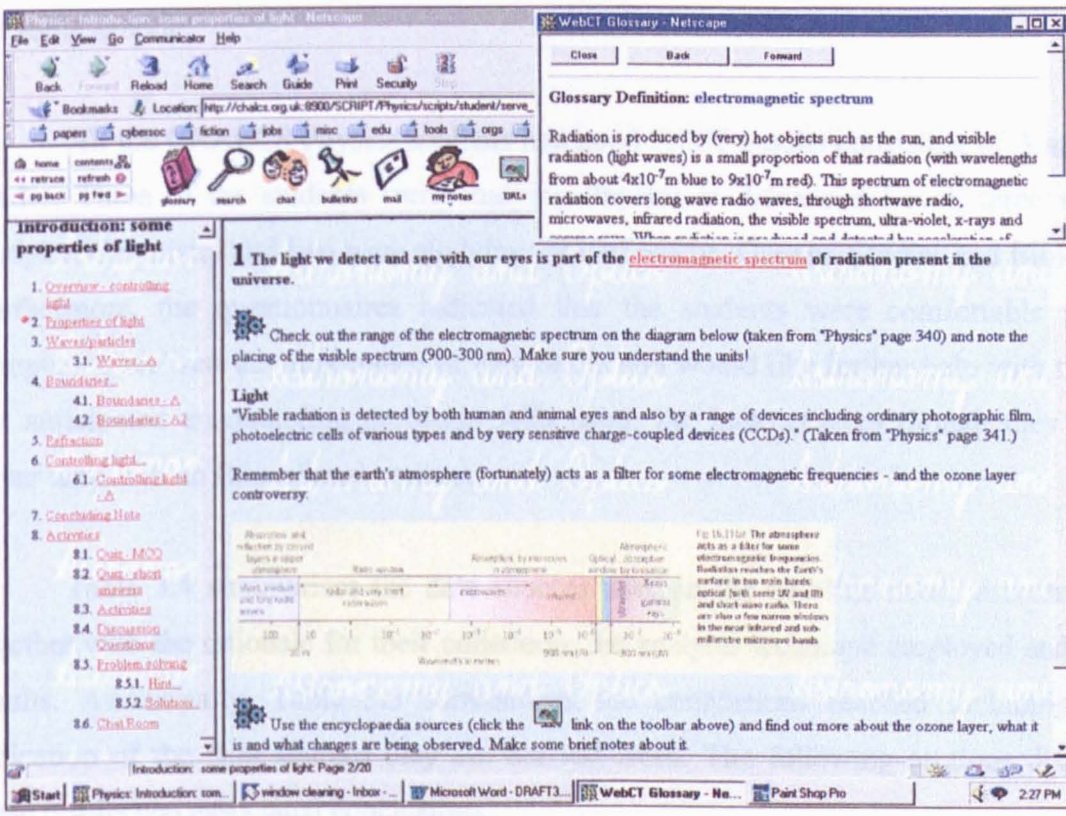


Figure 3.2 The 'Properties of Light' Introductory Physics Course Notes.

The purpose of including the icons mentioned above is to more deeply engage the student as they interact with the course materials and helping to shift the nature of the learning from passive to active. For example, by including a 'question mark' it is hoped that this will cause the student to reflect upon the concept being described and generate explanations before reading the appropriate answer rather than passively acquire an exposition. A similar rationale forms the basis for inclusion of the activities. The inclusion of the Learning Companion, as detailed in Chapter 6, would serve to further engage students in the acquisition of the course notes by providing a more interactive learning experience which the current version of the static WebCT notes does not provide.

3.5 Results of the Preliminary Evaluation.

The preliminary evaluation consisted of a baseline questionnaire/needs assessment plus a familiarisation session in using WebCT from which observational notes and online tracking were used to test the tools proposed. The baseline questionnaires were completed by the three students regularly attending CHALCS for Physics, one male (age 16) and two females (age 17). The two females were studying at the same school.

All the CHALCS Physics students had good GCSE results consisting of A and B grades. Three of the students were studying Physics at Advanced Level, three were studying Chemistry and two were studying each of Maths, General Studies and Biology. Furthermore, the questionnaires indicated that the students were comfortable with computers, realised the importance of Key Skills and would like further help with them but anticipated encountering problems with using the Internet even though they had never used it in an Educational context.

Table 3.4 summarises the data sources used in the baseline needs assessment together with the rationale for their collection, the analysis technique employed and the results. Additionally, Table 3.5 summarises the conclusions reached including an indication of the data sources they are derived from. The following sections discuss these results and associated conclusions.

DATA SOURCE	RATIONALE	ANALYSIS TECHNIQUE	RESULTS
1. Questionaires	obtain base-line data e.g. age, gender, etc.	gather individuals personal details + categorise and rank responses to more general questions	1 male (age 16), 2 female (age 17) – all had good GCSEs, comfortable with computers, think key skills are important & anticipate problems in using the internet
2. Video	enables closer inspection of WebCT use e.g. tools	look for problems or successes in tool use	- students demonstrated problems using 'my notes' e.g. including images - problems in identifying whiteboard contributors led to collaboration difficulties
3. Webct 3a. BB/email 3b. chat logs	to see if they are used effectively	look for evidence of correct use, e.g threads, forums (BB) or collaborative debate (chat)	- inappropriate posts made to BB - email understood - chat: arguments are easily developed but more time is needed to develop them
3c. page tracking	to evaluate tools and notes usage	look for patterns (e.g. increases) in statistics (e.g. means)	- usage increased between sessions
4. Observational notes	capture data not captured by other sources	look for problems using WebCT tools or in general	- 'acquisition' stage best used outside of face-to-face as apart from note-taking or external resources usage it requires little tutor intervention - substantial overheads in course management

Table 3.4 Summary of Results of Preliminary Evaluation.
(Key: 'BB' = Bulletin Board).

DATA SOURCE	CONCLUSION
3a + 3c	LC 'peer' for chat/BB collaboration
2 + 3a	LC 'coach' for whiteboard/BB use
2	LC 'guide' to effective note-taking strategies
4	LC 'tutor' to help reduce administrative/management load
3b	students soon became accustomed to using WebCT

Table 3.5 Summary of Conclusions of Preliminary Evaluation.
(Key: 'BB' = Bulletin Board, 'LC' = Learning Companion)

In addition to the questionnaires Video and audio footage was gathered of one student's on-screen interactions together with observational notes and on-line data (such as page tracking and chat logs supplied with WebCT). Thus data could be combined or triangulated in-keeping with the methodology described in Chapter 1.

Preliminary data analysis indicated that the CHALCS Physics students soon become accustomed to using the WebCT facilities. During the first session the students only had a mean hits score of 33 with 3 articles read or posted to the bulletin board. However, this increased to a mean hits score of 43 with 24 articles read or posted in the follow-up session. Chat, in particular, was used extremely effectively to develop arguments with the tutor and peers but time limitations cut short the session. However, inappropriately placed postings were made to the bulletin boards indicating that the concept of fora and threads were not properly understood. It became apparent that the Acquisition stage is best carried out outside of face-to-face contact time as it required little tutor intervention therefore freeing classroom based sessions for the more interactive Argumentation and Application phases. Tutor intervention was, however, periodically required during Acquisition to guide the students through the notes or to encourage them to use the on-line encyclopaedias. The 'my notes' facility provided by WebCT proved cumbersome to use to structure course note extracts and annotations, especially if images were required since these could not be cut and pasted into my notes and links to glossary items were also not maintained when pasted in (this is discussed in more detail in section 4.2). The WebCT whiteboard tool was also problematic in terms of identification of contributors which exacerbated problems inherent in collaboratively authoring diagrams. In other words it wasn't possible to distinguish who had contributed what to the whiteboard. Finally, it became apparent that the tutor's job was very time-intensive involving administration overheads, for example pruning the bulletin board or preparing diagrams for the whiteboard, as well as lesson-based work.

3.6 Emerging Support Requirements.

From these preliminary evaluations a number of possible areas that might benefit from additional online support emerged as potentially useful, particularly the online agent-based support or Learning Companions envisaged by Chan (1996) which would be capable of engaging students on a one-to-one basis for anytime, anywhere coaching. At this stage of the research, one possibility identified was to provide Agent-based support at the level of each of the tools, e.g. a *peer* to chat or use the bulletin board with, a *coach* for use of the whiteboard or the bulletin board, a *guide* to course notes navigation and 'my notes' construction and a *tutor* to assist the human tutor.

To elaborate, a peer Learning Companion would allow students to collaboratively develop arguments when no suitable real peers are logged on which would encourage access from outside of the classroom-based face-to-face session, i.e. home, school or CHALCS open access times. A coach could provide more technically-based help in the form of the proper use of threads and fora in the bulletin board or the best way to draw diagrams collaboratively in the whiteboard. The guide could support 'my notes' construction by acting as a Learning Companion. It could offer guidance by collaborating with a student in effective note-taking strategies. Finally, the tutor would be used across all three stages to decrease the workload of the real tutor by aiding course administration such as bulletin board editing and email filing as well as possibly answering Frequently Asked Questions. However, it is not possible to design and implement all of these kinds of Agent-based support so it will be necessary to focus on one of these contexts to evaluate the potential of the Agent-based approach.

3.7 Conclusions.

This early 'pilot' work indicated that the navigation of WebCT could be easily learned by students as evidenced by their enthusiastic uptake of the Properties of Light Introductory material and successful utilisation of the integral WebCT tools. This success together with initial evidence of what students found difficult motivated the development team to provide additional learning resources including both short answer and multiple choice *quizzes*, material to form the basis for chat *discussions* and *activities* such as hints on summarising the Properties of Lens course notes themselves (see Figure 3.3).

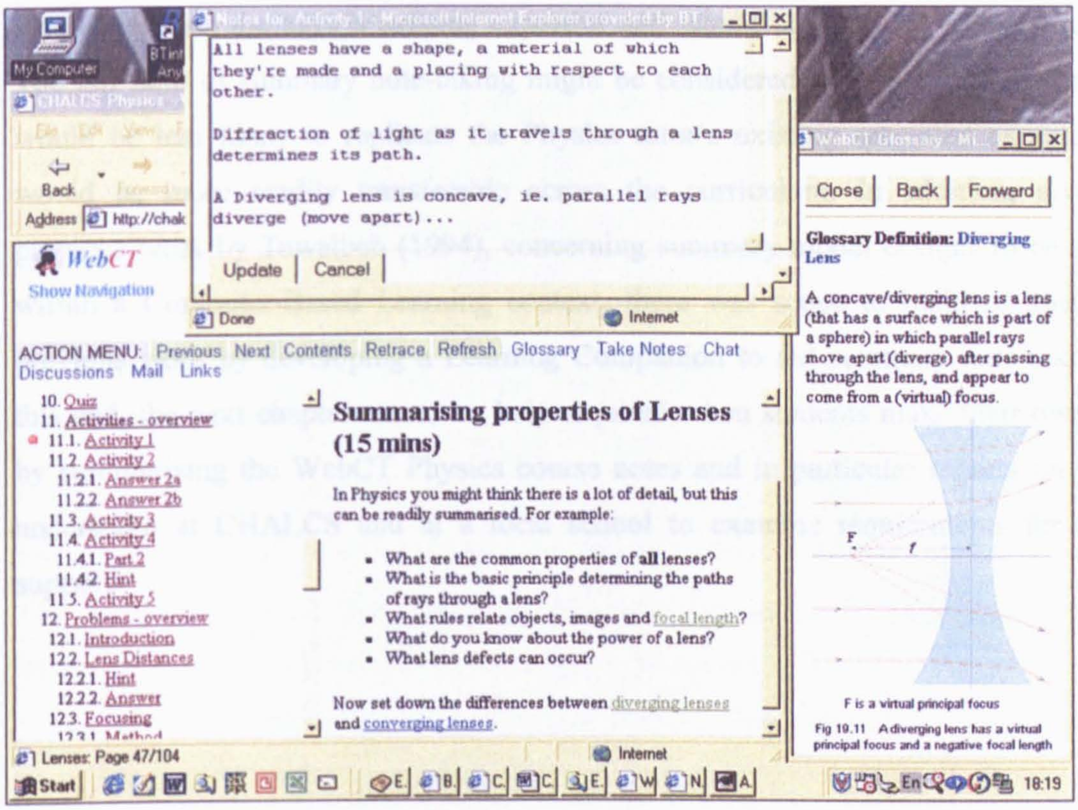


Figure 3.3 A Summarisation Activity in the Lenses Module.

To conclude, earlier in this chapter the results of a fact-finding session at CHALCS were reported indicating students own requests for support. The most popular requests, additional reference support and explanation of theory, were addressed through the development of on-line material and targeted use of the integral communication tools by the human tutor e.g. use of the whiteboard to support collaborative problem-solving. Four students each wanted additional support with revision and writing skills. It has been noted here and in the literature that the ability to take good course notes both to aid initial comprehension and as a revision technique is a key aspect of successful academic performance which many Universities recognise through their core skills / key skills programmes aimed at helping students master these skills at undergraduate level. Students who acquire these skills at GCSE level are likely to have an advantage both in their GCSEs and in their later studies (Kiewra et al., 1995; Slotte & Lonka, 1999). An online Learning Companion to support students in developing these independent study-skills through anywhere, anytime assistance has great potential and is likely to motivate and engage in ways that static notes cannot. Whilst support for problem-solving when the tutor is not present is also highly desirable it would represent a major undertaking in terms of developing an intelligent agent that would replicate the tutor's existing expertise and would be relatively domain-specific. The key skill of summary note-taking might be considered to be less domain-specific, would be less likely to replicate the Physics tutor's existing specialist expertise and would be more readily transferable across the curriculum. In addition, given the previous work by Tawalbeh (1994), concerning summary model designs to be applied within a Computer-Based Learning context, there was scope to build on this novel previous work by developing a Learning Companion to aid summary note-taking. To this end, the next chapter examines help required when students make their own notes by summarising the WebCT Physics course notes and in particular reports on studies undertaken at CHALCS and at a local school to examine requirements for such a support.

Chapter 4: The Summarisation Studies.

4.1 Introduction.

These studies aim to investigate more closely the support requirements of students as they take summary notes in the online WebCT environment with a view to informing the design of a note-taking Learning Companion. The first study took place at CHALCS and the second study took place at Notre Dame school. CHALCS has already been introduced in Chapter 1. Notre Dame is a Roman Catholic sixth-form college in inner-city Leeds. Some pupils from Notre Dame attend CHALCS' classes for supplementary education. The reason for including a study at Notre Dame was that the number of students enrolled in CHALCS who were studying physics at sixth form level was very small and although this meant it was possible to follow individuals closely it was not possible to be sure that their particular summary note-taking behaviour was typical. Notre dame offered the opportunity to take a larger class for a session working with WebCT. This would provide some additional evidence as to whether the strategies identified at CHALCS were typical, although it would not be possible to observe the summary note-taking process of children in the larger class in as much detail. A number of other CHALCS feeder schools were also contacted but felt unable to take part in any study.

From Chapter 3 we have argued for a number of potential support roles for a Learning Companion. However, we cannot hope to develop all of these roles within the lifecycle of this research project. Therefore, from these potential roles we selected that of note-taker as satisfying requirements of being "wanted" by students, not easy to provide 'just in time' by the tutor and potentially tractable as an exemplar case based on the previous work in the field (discussed in Chapter 2).

Specific Research Questions for this Chapter are listed below:

1. What help and support do students require in taking notes within the VLE ?
2. What do students understand to be the requirements of good summary note-taking ?
3. What assistance does the VLE offer the students in this task ?
4. What is the quality of student's summaries and are they aware of how their summaries could be improved ?
5. What types of residual help do students need to develop their summary skills that might inform the design of a Learning Companion ?

The key studies then described in this chapter are designed to find answers to these questions, especially Question 5, notably how to improve upon the functionality of WebCT. In particular, as emerged from the Critique of WebCT in Chapter 3 we wish to provide additional student support in the key skill of summarisation. The studies are therefore designed to address summarisation of the WebCT Physics notes. But first, we examine the features of the integrated WebCT summary note-taking aid. Note that the Rationale is common to both studies as are the final Conclusions and Summary. However, the Method, Analysis and Results are included for each separate study. The Method for each study differs slightly mainly due to the fact that the CHALCS study took place with a much smaller number of students and more time was spent on the task. However, both groups performed similar tasks.

4.2 Rationale.

As detailed in Chapter 3 preliminary evaluation indicated that there was a need to supplement the functionality provided by the 'My Notes' WebCT tool. Thus, in addition to conducting these studies in order to better understand the processes and strategies of students when taking notes using WebCT the aim of the studies was to gather further data on the ways in which note-taking support provided within WebCT could be improved through the provision of a Learning Companion and note-taking tool. The 'my-notes' facility is shown in Figure 4.1.

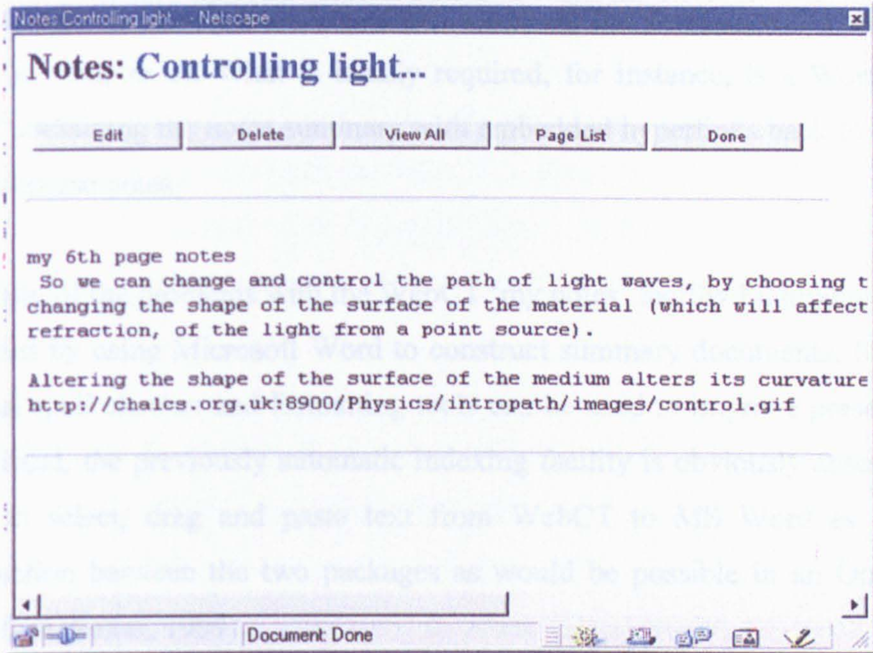


Figure 4.1 WebCT 'my notes' facility.

There are four options available, as indicated in the figure, consisting of :

1. **Edit** – a window pops up allowing the user to enter text for their annotation of the current page of course notes (N.B. There are no word processing type facilities e.g. spell check, formatting, etc. and hyperlinks are *not* active);
2. **Delete** – this operation simply removes *all* annotations from the current page, i.e. individual notes cannot be selected;
3. **View All** – this produces a list of all annotations across all modules;
4. **Page List** – this produces a list of hyperlinks consisting of course page titles, selecting one of these produces a list of all the annotations for that page (i.e. this is a simple indexing capability).

It can be seen that this 'my-notes' facility does little to support the *process* of summary construction. As the above list indicates, the facility lacks certain essential word processing features (e.g. bullet points, fonts, spell checking, etc.) as well as intuitive and complete editing capabilities. Also, URL's pasted into the annotations are not active so that links, for example, to glossary definitions appear as a meaningless URL instead of a hyperlink to the definition itself. Therefore the aim of the study detailed here was specifically to *identify student's needs in terms of ICT support when creating their own summary 'documents'*. These are exceedingly important in creating a solid foundation in the initial Acquisition stage of the Pedagogical Framework (as discussed in Chapter 2). What is ideally required, for instance, is a Word Processed document containing the notes summary with embedded hyperlinks back to the relevant text in the course notes.

Some of the problems with the WebCT 'my notes' facility highlighted above can be answered by using Microsoft Word to construct summary documents. For example, the integral spell checker and formatting tools can be used to improve presentation. On the other hand, the previously automatic indexing facility is obviously absent when the user has to select, drag and paste text from WebCT to MS Word as there is no communication between the two packages as would be possible in an Open Systems approach (see Barker, 1999).

The next two sections describe two studies, one undertaken at CHALCS and the other at Notre Dame Sixth Form College, to ascertain further support requirements for a note-taking learning companion. For each study the format of the report is to first talk about the study method (including its design – subjects and data artefact synopses – and its procedure), the analysis (in-keeping with the methodology described in Chapter 1 with full reference to data sources) and the results (particularly showing how they are grounded in the data). The subsequent conclusions which round-off this chapter are common to both studies and are thus derived from both studies' analysis and results discussions.

4.3 The CHALCS Summarisation Study.

The CHALCS summarisation study took place during an evening scheduled for the usual Physics lessons with the CHALCS tutor at the CHALCS premises in Chapeltown, Leeds.

4.3.1. Method.

As mentioned the method is divided into the design of the study, that is base-line details of the students and overviews of the data used, and a detailed report of the procedure.

4.3.1.1. Design.

Two students (females, age 17) were asked to collaborate in performing a summarisation task of a portion of the WebCT 'Lenses' module within the Astronomy and Optics course. They both study A Level Physics at the same school, Allerton High in Leeds and have been attending CHALCS for a number of years.

After a short demonstration, both students were asked to cut and paste from the WebCT notes into Word then swap documents and comment on each other's work finally emailing their summaries to the Instructor.

Artefacts collected for the purpose of analysis include observational notes, video/audio tapes both of tutor-led conversations and summarising (later transcribed), Word 'summary' documents including the collaborators annotations (later annotated by the instructor) and WebCT data (e.g. page tracking logs). See Appendix F for these documents. In addition, a questionnaire was utilised which aimed to gather student's background data such as age, gender, school and qualifications and answers to more directed questions such as attitudes to key skills and use of ICT for learning. The same questionnaire was used for both the CHALCS and the Notre Dame studies in order to maintain consistency. These data will allow the student's strategies when summarising and their preconceptions of summarisation to be examined as well as the quality of their resulting summary, i.e. they will relate to the Research Questions for this Chapter stated in section 4.1.

4.3.1.2. Procedure.

The video camera and microphone were first set up in the room directed at S1's computer screen, having already secured her permission. Aside from being aware of the video data, the students were also aware that the WebCT on-line data was also being gathered.

Data had already been gathered for the CHALCS students in the form of questionnaires in the previous 'pilot' session (see Chapter 3). The students were at this stage quite familiar with the WebCT tools, such as chat, email and bulletin boards but had not seen the 'Lenses' course material upon which this study was based.

This researcher began by demonstrating how to have two windows open, one containing WebCT, the other containing MS Word then how to cut and paste text from the former to the latter. The 'compile' tool within WebCT (which allows the student to form a concatenated list of selected course notes, for example for the purposes of producing a hard copy) was also demonstrated. Next it was necessary to demonstrate how to copy images from WebCT to Word by first saving them. Subsequently, we demonstrated the use of Word's annotation facility by inserting a comment in an example document.

At this point we held a discussion with the students concerning their thoughts on effective summarisation techniques. The pedagogical philosophy employed here and in subsequent sessions was for the tutor to facilitate rather than dominate the discussion much akin to an early stage of Inquiry style of tutoring (Palinscar & Brown, 1984). Table 4.1 contains an excerpt from this discussion.

In the initial discussion the tutor aims to focus the students on the task ahead – to get them to reflect on what a good summary is, eliciting what they already know about how to take good notes and why note-taking is important, for example questions at turn 1 and 8. At times the tutor summarises the outcome of the discussion too, for example at turn 6.

The students were asked to summarise section 1 "Introduction to Lenses" which comprises of 6 sub-sections including a summary. They were reminded of the navigation facilities within WebCT then proceeded to create their summary in Word. The video/audio were checked regularly to ensure the desired data was being captured. In addition, classroom based observations were recorded.

Towards the end of the summarising session the students were warned that they would have to end soon. After 43 minutes they were told to stop summarising. At this point the tutor intervened to save the summary Word document which, due to network restrictions, was achieved by attaching the Word document to an email within WebCT. Having effectively swapped documents the students were asked to comment on each other's work by adding annotations. First though the tutor intervened to initiate a discussion of student's impressions of their own and their collaborator's summary document including attempting to elicit strategies and decisions (see the excerpt in Table 4.2). The annotation activity was given just 11 minutes.

Following the annotation phase the tutor initiated another discussion concerning the differences between the two student's documents leading to a discussion of summarisation method (see Table 4.3 for an excerpt). This was an attempt by the tutor to cause the students to reflect on their summarisation experiences, in particular by noting differences between their two strategies, thus prompting self-explanation and collaborative problem-solving (in terms of identifying the best strategies with the help of each other and the tutor) with a resultant re-examination of their own summary note-taking techniques.

At this point the lesson was ended, floppy disks collected containing the summary documents and the video camera switched off.

1.	[T1] Yeah, the idea is that we wanna try and (in this session) think quite hard about what makes a good summary....what is it about doing a summary that helps you to learn ?
2.	[S2] Its basically...if you have...
3.	[S1] Find out what you've learnt from what you've written
4.	[S2] Bullet points or something
5.	[S1] Like a shortened version of what you've picked up from what you've
6.	[T1] OK, so one of the key things about a summary is that its shorter than the original text.
7.	[S2 & S1] yeah
8.	[T1] Why is that helpful for revision ?
9.	[S1] That you can be bothered to look at it.
10.	[S2] Its easier to get information from if its like
11.	[S1] Its just
12.	[S1] chop things off
13.	[S2] if you recap the bits you need
14.	[S1] you've picked up everything but you just do a recap of the main points
15.	[T1] right, right, right

Table 4.1 CHALCS Summarisation Study, Transcript 1 Excerpt – Pre-Summarisation.

(Key: S1, S2 = students, T1 = tutor 1)

1.	[S1] Yeah, read them all first, hmm…I've just condensed them, taken out like, the bits that you don't need.
2.	[T1] Right, so you did a, you selected the things you really wanted and just put them into Word.
3.	[S1] Hmmm…yeah, I've taken out some of like, its more like, mines more like bullet points in note form rather than…
4.	[T1] Right, so you spent a lot of time deleting stuff
5.	[S1] Yeah
6.	[T1] Once you got it in
7.	[S1] Yeah. Its only about one and a bit pages long, {there's a} bit on the page, cause I've taken out everything.
8.	[T1] Right, and did you do much re-writing or once you deleted anything did you do any re-writing after that ? Or did you just…
9.	[S1] I did a bit of sentence rearranging.
10	[T1] Yeah, yeah, did you do a similar sort of thing ? {to S2}
11	[S2] Well, I didn't delete as much as she did…
12	[T1] right
13	[S2] some bits I did but not as much as her
14	[T1] You've ended up with nearly much the same

Table 4.2 CHALCS Summarisation Study, Transcript 2 Excerpt - Post-Swapping.
(Key: S1, S2 = students, T1 = tutor 1)

1.	[T1] OK, what was the, what was the main difference then between what you both have put ?
2.	[S1] She's just, she's written the same, she hasn't read it, she's written the same thing three times on this.
3.	[T1] OK, so there's some repetition, yeah ?
4.	[S1] Yeah.
5.	[T1] But are there, are there, things that have been picked the same…as the ones that you'd pick.
6.	[S1] No, no, she's put a lot of information that I haven't even added on mine 'cause I didn't think it was relevant
7.	[T1] Right, so what, so what sort of things did you pick and why ?
8.	[S1] I've picked like the bare bones, like, just exactly like what I thought I needed and then she's put things like hmmm…about lenses, they're used to see near objects and then she's put molecules and bacteria and then distant ones and you know that, like, you know already that they're used in microscopes and telescopes. Do you see what I mean ?
9.	[T1] Yeah, so you've put in some extra information. Why, why did you think that ?
10	[S2] 'Cause, I mean we already had like whatever it wants, it's not as if you could pick it up, pick it up that quickly, do you know what I mean ?
11	[T1] Yeah.
12	[S2] So, I left it there for detail.

Table 4.3 CHALCS Summarisation Study, Transcript 3 Excerpt - Post-Annotation.
(Key: S1, S2 = students, T1 = tutor1)

4.3.2. Analysis.

Both students responded well to the initial 'training' session with just the annotation demonstration requiring repetition. At this point they are quite proficient in WebCT, Windows and Word and cutting and pasting between the two packages presents no real problems. However, it did later become apparent (See Appendix F, S1's summary) that pasting WebCT hyperlinks into the Word document was ineffective as this resulted in an error and ultimately missing words ,e.g. "Outer rays have a larger angle of incidence and hence a greater Error! Bookmark not defined." This led to confusion as picked up by S2's annotation#4: "the error has replaced many important words and therefore nearly all of the text doesn't make sense". It is worth pointing out that S2 avoided this possible error by making less use of cut and paste and actually typing her own notes or even writing and drawing diagrams on a separate piece of paper.

The initial tutor-led discussion elicited the student's preconceived notions of the content and process required to produce a good summary. These included :

- *Shorter than the original text*: "Like a shortened version of what you've picked up from what you've..." (S1)
- *Recap of main points:* "it's easier to get information from if its like...if you recap the bits you need." (S2)
- *Need to decide what to keep* : "You keep what you thinks important and relevant..." (S2), S1 saying "that's the hardest bit actually"
- In Physics *formulae, diagrams and definitions are important*
- Sometimes *diagrams are clearer than written descriptions* : "...some things, they're not obvious if you write out, like, a description you need to draw the diagram to show what you mean..." (S2)
- Use *note form*, in particular "trigger points" (S2)

It was observed during the actual summarisation phase that both students read the course notes first before commencing construction of the summary document (see Table 4.4).

	S1	S2
1	Introduction to Lenses	Introduction to Lenses
2	What is a lens?	What is a lens?
3	Convex lenses	Controlling light
4	Positive lenses	Convex lenses
5	Converging question	Positive lenses
6	Converging answer	Introduction to Lenses
7	Q1 Answer	Converging question
8	Q2 Answer	Converging answer
9	Concave lenses	Concave lenses
10	Concave - diverging	Concave - diverging
11	More concave lenses	More concave lenses
12	Summary	Summary

Table 4.4 Pre-Summarisation Student Tracking

It is interesting to see that S1 more or less navigated the section sequentially (with the exception of missing the "controlling light" introductory text) whereas S2 at one point went back to the first course page "introduction to lenses" then on return elects to omit the answers to the two supplementary questions on converging lenses. The two strategies employed elucidate different techniques to navigating the hypertext course notes. Whereas S1 employs traditional sequential reading techniques reminiscent of static texts, S2 employs techniques which take advantage of hypertext structure, that is, will navigate the text taking advantage of hyperlinks. Subsequent tracking indicates that during the actual summary construction both students navigate the course notes in exactly the same sequence, i.e. the sequential pre-defined course path.

A similar number of page hits were registered by WebCT in the study for both students (32 for S1 and 38 for S2). It is clear from the WebCT Hits Distribution, however, that S2 made greater use of the glossary facility whilst negotiating the summary process (see Figure 4.2). In addition S2 also read an article on the Bulletin Board.

From the video it was noted that S1 employs a 'paste and complete' technique, that is, sections of text are cut from the WebCT course notes, pasted then formed into complete sentences by typing. This technique was used extensively together with a separate 'editing' technique which was used more as completion approached. Upon a first draft S1 viewed the whole document, particularly looking at overall layout. As mentioned, the glossary hyperlinks did not work when copied then pasted into Word leaving error messages in the summary document. This resulted in S1 having to go through the entire document at the end and replace the errors with the actual words. Images were cut and pasted effortlessly to great effect (as can be seen in the excerpt in Figure 4.3, Appendix F contains the full summary). The Expert commented that both summaries were "very good" and that S1's was similar to their own.

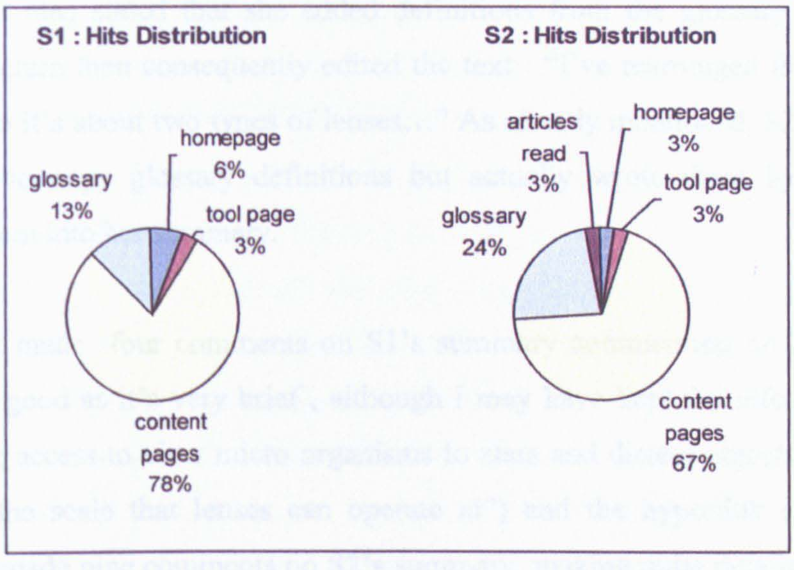


Figure 4.2 Comparison of Hits During CHALCS Summarisation Study.

At the first face the light enters a denser medium and will bend towards the normal. It will diverge. And at the second face it will diverge more because the light ray enters a less dense medium (air) and is refracted away from the normal as the light waves pick-up speed. So the image will appear to come from behind the lens i.e. the same side as the object (project the final ray back to where it crosses the axis) and so is a **virtual** image.

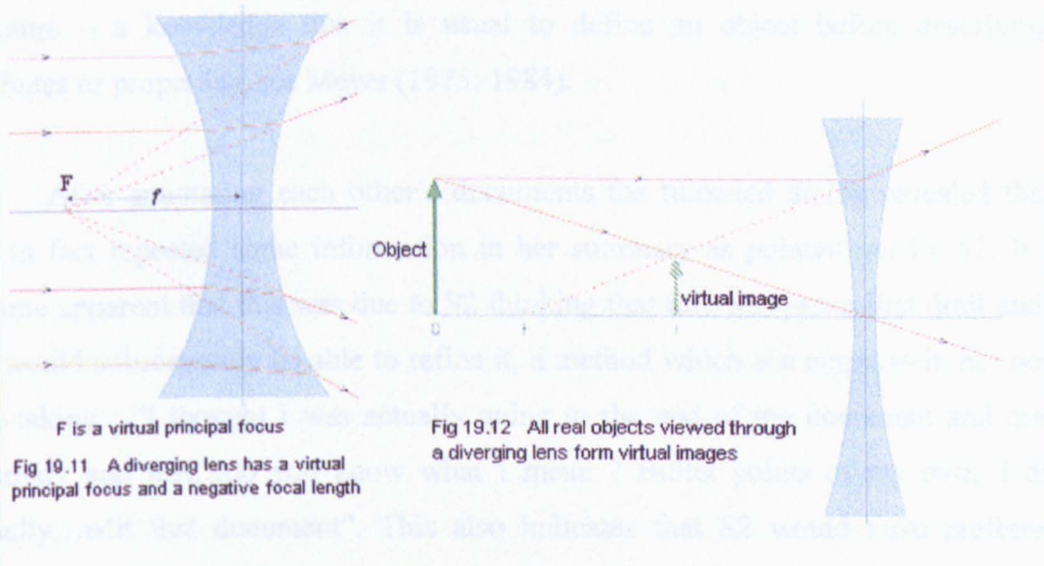


Figure 4.3 Excerpt of CHALCS Student (S1) Summary Document.

Upon completion of the summary documents the ensuing tutor-led dialog revealed that S1 not only selected and edited on a sentence level but also maintained a coherence between sentences by moving them around : "I've rearranged the sentences a bit because like you know like you read things in a different order". On the other hand S2 maintained the original sequence of notes : "I kept to the order it was in...". Interestingly S1 also stated that she added definitions from the glossary next to the appropriate diagram then consequently edited the text : "I've rearranged the writing to go with them so it's about two types of lenses..." As already mentioned, S2 did not use cut and paste to copy glossary definitions but actually wrote them by hand then incorporated them into her summary.

S2 only made four comments on S1's summary commenting on S1's brevity ("This is quite good as it's very brief, although I may have kept the info given about lenses allowing access to view micro organisms to stars and distant objects, it gives an estimation of the scale that lenses can operate at") and the hyperlink error already mentioned. S1 made nine comments on S2's summary, making quite detailed criticisms of her document. In particular she is concerned with S2's repetition ("this is the third time you've said this"), straight copying of text without explanation ("just straight copied out, no own notes made") and the general coherence of the document ("You've just started yacking on about concave lenses, you haven't told me what they are, come to think of it you haven't even told me what convex lenses are yet... this sentence has just been stuck in the middle of nowhere"). This shows an appreciation of rhetorical structure – a knowledge that it is usual to define an object before describing its attributes or properties, see Meyer (1975; 1984).

After annotating each other's documents the tutor-led dialog revealed that S2 had in fact repeated some information in her summary as pointed out by S1. It later became apparent that this was due to S2 thinking that this was just a first draft and that she would subsequently be able to refine it, a method which she employs in her normal note-taking : "I thought I was actually going to the end of the document and make a summary and just...do you know what I mean ? Bullet points of my own, I didn't actually...edit that document". This also indicates that S2 would have preferred to spend more time on the note-taking activity and that this may mean she needs to develop her strategies to work quickly when under time-pressure. However, the time-pressure for this task was high and it may also be reasonable to expect to be able to take more time taking notes. At the other end of the scale S1 relates "I just read it off, I never make notes" and is keen to point out the fact that she just utilises "trigger points" for revision forcing S2 to concede that "I don't pick up stuff that quickly". Thus, S2 consciously included more detail in her summary notes because she felt she needed them for later understanding.

4.3.3. CHALCS Summarisation Results.

In common with Chapter 3, Table 4.5 summarises the data sources used in the CHALCS summarisation evaluations together with the rationale for their collection, the analysis technique employed and the associated results. Table 4.6 summarises the conclusions reached including an indication of the relevant data sources which have been triangulated. The following sections discuss these results and associated conclusions. Finally, section 4.5 triangulates the results from both the CHALCS and Notre Dame studies to be described.

The two CHALCS students both reported feeling fairly competent with ICT on their initial needs questionnaire but anticipated problems with finding suitable material before the sessions began. In answers to a question about "searching for the right information or taking the right sort of notes" S2 reported "I would benefit from note taking and note making to make appropriate notes" and S1 reported "I always get confused. I often have to do things several times because I don't have the right notes." This is interesting in light of the earlier comments S1 makes about needing only the briefest of notes. It may be that whilst S2 might not have structured her notes optimally and may have copied too much, at the same time S1 may err in the other direction, copying across less than she will find she needs later.

The need for S2 in the CHALCS study to have to write and draw on supplementary pieces of paper is indicative of the inadequacies of the cutting and pasting methodology and subsequently would be an ideal area for further support, for example a '*scratch pad*' for jotting down sections of notes. Editing of the pasted text is obviously an important part of summary writing, the students taking care to add appropriate words to form complete sentences and indeed carefully selecting the text in the first instance although less rigorously in S2's case.

DATA SOURCE	RATIONALE	ANALYSIS TECHNIQUE	RESULTS
1. Video	To obtain a visual record of summarisation techniques/strategies	Noting summarisation strategy and time	S1 uses 'paste and complete' + some final editing S2 copies and pastes with very limited final editing
2. Word summaries	to obtain a record of the final summary and annotations (to be examined by an expert)	- looking for points of interest, eg. errors (+ the expert compared them to their own) - looking for comments about partners summary	- webct hyperlinks do not transfer to word (S1) - expert noted both summaries were good + S1s was similar to theirs - S1 liked S2s summary but would have liked more detail - S2 concerned with with repetition, copying & coherence
3. WebCT Data 3a Tracking 3b Page hits	- captures the sequence of course notes reading - to ascertain statistics on tools/page access	- examine logs for navigation strategies - statistical analysis of hits distribution	- sequential (S1) vs. hypertext (S2) contrasting styles - S2 makes more use of glossary & BB
4. Questionnaire (as used in Chapter 3)	to obtain base-line data	extract pertinent information	- 2 females, age 17 from Allerton High, good GCSE results, competent with ICT, both could benefit from note-taking help
5. Audio transcripts	- to elicit (pre)conceptions of summarisation & post-test reasoning + interrelationships	- highlight comments relating to summarisation - look for personality types, language use, etc.	- S1 rearranges sentences, S2 keeps to original order + S1 included edited glossary definitions - S1 is 'dominant', S2 is 'passive' but there is mutual respect - S1 is more economical, S2 admits being a slower learner - both use colloquialisms
6. Observational notes	to obtain misc. points of interest not covered by other data	- look for factors amongst notes affecting summarisation	- S2 utilised a separate notepad

Table 4.5 Summary of CHALCS Summarisation Study Results
(Key: Sx = subject x, BB = WebCT Bulletin Board)

FINAL CONCLUSION	DATA SOURCE	INTERMEDIATE CONCLUSION
C1	6	S2s notepage use may indicate the necessity of having a 'scratchpad'
C2	1 & 5	text editing is important although S2 does it less
C3	3a & 2	S1 followed a sequential pre-read but reordered her summary whereas S2 followed a 'hypertext' pre-read but kept the final order the same
C4	5	S1 is 'dominant', S2 is 'passive' but there is mutual respect
C5	5	both students use colloquial language

Table 4.6 Summary of Conclusions of CHALCS Summarisation Study
(Key: Sx = Subject x, Cx = a reference for the final conclusion, see Table 4.12)

It was interesting to see that although S1 followed the sequential path through the course notes when she constructed her own summary it did not comply with the original order but was adapted to be, in her eyes, more consistent. This is in contrast to S2 who employed a more hypertext kind of navigation yet constructed a summary in-keeping with the original order. Apart from the two extremes of summarisation techniques espoused by the students an interesting emergent factor of the dialogues is the dominating personality of S1 and almost submissive concessions of S2. There is, however, a mutual respect much in evidence between the two students, they realise their differences but whilst S1 may seem quite critical she is less judgemental than would first appear and S2 is very non-judgmental. These contrasting yet complimentary personality traits may be useful to model in a summarisation support (cp. Reeves and Nass, 1996 and Isbister and Nass, 2000 – described in Section 2.4.2). Also, in terms of personality traits, colloquialisms such as "like", "gonna", "sort of thing", "'cause" and "yeah" are much in evidence in the dialogues. Incorporation of this language may prove advantageous in establishing rapport with students but requires formative evaluation (see Chapter 5, the Wizard of Oz study).

Note that the joint Conclusions section, for both studies presented in 4.5, make further comments on the results presented here and presents a model of the two contrasting styles of note-taking as presented by the two CHALCS students. Preliminary conclusions here note that individual students may differ in their support requirements both in what they need to retain in their summaries for comprehension and revision and also in their preferred collaboration style.

4.4 The Notre Dame Summarisation Study.

The Notre Dame summarisation study took place during normal school hours at the college premises just opposite the University of Leeds in the inner city. A normal Physics lesson was utilised for the study in a computer-equipped laboratory.

4.4.1. Method.

Again, the method details the design of the study, the base-line student data and synopses of the data collected. Also included in the method is a detailed description of the experimental procedure.

4.4.1.1. *Design.*

The design of this study follows the design of the CHALCS study as closely as possible, given the different context. This context consists of a group of 15 students aged 17 and 18 with 4 being female and 11 male. A classroom was used consisting of 13 computers thus leading to 2 pairs of students at one computer, the remaining having a personal computer (note that the pairs are labelled S6 and S3). All students were in their second year of studying Advanced Level Physics following the NEAB syllabus.

The session was held in two halves. The first half concentrated on demonstrating the basic facilities of WebCT to the students and allowing them to practice using some of it's pertinent features. The second half was the main thrust of the summarisation study and involved the students in constructing a summary of the WebCT Phsyics course notes on Lenses which was something they were then studying in their usual classes.

Artefacts collected for the purposes of analysis include:

- permission slips;
- observational notes (used to document the summarisation process);
- questionnaires (to collect background data and e.g. attitudes to key skills);
- chat logs (used to analyse discussions concerning summarisation styles and general utility of WebCT/on-line learning);
- summary documents (with annotations, used to compare summarisation strategies);
- miscellaneous documents (e.g. scraps of paper, evidence of supplementary aids during summarisation).

These artefacts relate to the Research Questions stated in section 4.1 concerning informing the design of a Learning Companion.

4.4.1.2. Procedure.

The first task was to hand out the questionnaires to be completed in the student's own time and obtain completed permission slips. As already mentioned the aim of Part

1 was to familiarise the students with WebCT. To this end features of WebCT were first demonstrated including the course content i.e. glossary hyperlinks, icons (e.g. gears, light-bulb), links to external resources and an interactive java applet with mention of the Multiple Choice Quiz. Additionally, tools that were demonstrated included : email, chat and the bulletin board with mention of the whiteboard. The students were then asked to obtain their password and create their own 'MyWebCT' global login to enable access to the Physics course. Their first task was to prepare a definition of a term from the notes for inclusion in the glossary then email this to a partner, sending a copy to the instructor. Their second and final task of Part 1 was to use the chat to answer a question concerning the recognition of infra-red and X-ray electromagnetic radiation.

Part 2, the actual summarisation study, began with a demonstration of Microsoft Word which was to be used to construct the summary. The main focus though was on the use of the annotation facility within Word to enable peer commenting. In addition students were shown how to save an image to a local disk drive (thus enabling transference from the browser to MS Word), the use of bullet points (to neatly identify separate points) and cut and paste (between browser and summary document). All of these features were deemed particularly relevant to the task. The task itself was further decomposed into five parts as summarised below (shown also are the suggested times to be spent on activities which were enforced by the teacher):

1. discuss "what makes an effective summary ?" using chat [10 minutes];
2. summarise section 1 to 1.3 of 'Lenses' using Word [30 minutes];
3. email your summary to a partner and instructor then use chat to discuss differences [5 minutes];
4. annotate your partners summary using Word then email to instructor [10 minutes];
5. discuss your impressions of on-line summarisation using chat [5 minutes].

The final step in the procedure was to obtain all of the data (e.g. summaries, emails, etc.) from within WebCT for the purposes of subsequent analyses. In addition the completed questionnaires were collected about a week later.

4.4.2. Analysis.

Due to initial network problems time was short in the first part of the study which may explain why the email task was responded to badly. However, students seemed to have grasped the mechanism sufficiently for the study to proceed. In addition these same technical problems meant that the second chat task was unmoderated leading to students 'playing' with the chat rather than keeping within the topic in the lesson tasks, but again they mastered the skills necessary to use the facility. This mastery was also echoed in the general WebCT proficiency displayed, i.e. course navigation was sufficient for the study to proceed. It was noted, however, that the new 'MyWebCT' mechanism was confusing for the students resulting in a little lost time as they attempted to create global logins.

The summary task itself was responded to with mixed success. At one end of the spectrum one of the pairs of students (S3) composed a summary consisting of nine numbered succinct points with a clear title utilising a separate piece of paper to write out their summary before entering it into Word. At the other end of the spectrum the other pair (S6) simply cut and pasted the entire contents of the Lenses notes with no annotation or additions. Most students utilised a 'cut and paste' technique. Of the eight summaries which were emailed to the Instructor five utilised a title to structure their documents. The annotation activity was not carried out as expected with only two students carrying out the activity as desired although one of these did place comments within the document without the use of the annotation facility (S1). The remaining students did not attempt annotation. A selection of summary documents can be found in Appendix G (i.e. one good, one intermediate and one poor summary).

Table 4.7 shows information obtained from the WebCT Physics course 'tracking' facility relating to the pages which students examined as they went about the summarisation task. It shows the page 'hits' broken down for each of the 13 students (including the two pairs) into the main areas of the WebCT course. These areas are the course homepage, the tool pages (which contains, for example the 'compile' tool), the content (the actual course notes), the 'take notes' facility (the built-in annotation facility) and the WebCT glossary. Note that the students are matched with those referred to in Table 4.10 which follows. The 'sequence' column refers to one of five categories of course notes navigation styles (explained in the Key) derived from the WebCT logs and observation in the classroom.

student	total hits	homepage	tool page	content	take note	glossary	sequence
S1	21	12	3	6	0	0	B
S2	34	9	5	18	0	2	C
S3	2	5	5	12	0	0	C
S4	25	14	2	6	0	3	A
S5	17	4	4	9	0	0	C
S6	13	7	1	5	0	0	B
S7	28	15	6	7	0	0	B
S8	12	6	1	5	0	0	C
S9	36	16	7	11	0	2	B
S10	35	16	5	12	0	2	D
S11	13	10	1	2	0	0	E
S12	27	7	3	16	0	1	C
S13	33	7	3	20	3	0	B
Mean	**22.8**	**9.8**	**3.5**	**9.9**	**0.2**	**0.8**	
S.D.	**10.7**	**4.3**	**2.0**	**5.5**	**0.8**	**1.1**	

Table 4.7 Notre Dame Summarisation Study Tracking
(Key: **A.** Sequential navigation - following the pre-defined course path,
B. 'Hypertext' navigation - typically moving back and forth at will,
C. Summary use - skipping forward to the section summary,
D. Incorrectly utilising 'Properties of Light' course notes,
E. Just reading the Introduction to 'Lenses'.)

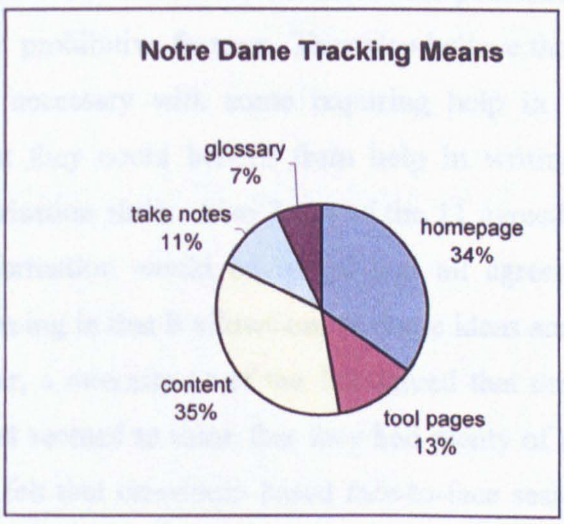


Figure 4.4 Notre Dame Summarisation Study Tracking Means.

Additionally, Figure 4.4 illustrates the percentage of time (i.e. the percentage of the total means) that the students spent on each WebCT tool.

No students used the Bulletin Board during the session although this should not be surprising as it was not included in the lesson tasks. As shown in Table 4.7, the mean total of hits was 22.8 with a Standard Deviation of 10.7, the range being from 2 to 36. The 'content' pages and 'homepage' were accessed a similar number of times (means 9.9 and 9.8, respectively) with the tool page being accessed the next most frequently (mean 3.5). Five of the thirteen students used the glossary between one and three times with an overall mean of 0.8. In addition, one student used the WebCT 'take notes' facility (equivalent to the 'my notes' facility described in section 4.2) three times.

4.4.3. Notre Dame Summarisation Results.

As with previous sections, Table 4.8 summarises the data sources used in the Notre Dame summarisation evaluations together their rationale, analysis technique and associated results. Table 4.9 then summarises the conclusions reached including the data sources they are based on. The following sections now discuss these results and associated conclusions. Finally, as already mentioned, section 4.5 triangulates the results from both the CHALCS and this study.

Of those who returned a questionnaire at Notre Dame (11), all but one state that they use computers a lot, at home and at school, and that they are "an essential part of the future" although cost, technical difficulties and problems searching for the right information may be prohibitive features. They also believe that key skills (DfEE, 1999) are important and necessary with some requiring help in this area. In addition, a minority stated that they could benefit from help in writing skills, one particularly mentioning summarisation skills. Also 7 out of the 11 agreed that help with searching for appropriate information would be useful and all agreed that group work is an essential part of learning in that it allows one to share ideas and prepares for work in the real world. However, a minority (4 of the 11) agreed that their problem-solving skills needed honing, most seemed to think that they had plenty of practice in other subjects. A similar minority felt that classroom-based face-to-face sessions were better in terms of developing social skills. Finally, only one student reported that their teacher had used the World Wide Web for learning purposes previous to this session.

DATA SOURCE	RATIONALE	ANALYSIS TECHNIQUE	RESULTS
1. Questionnaires (as used in the baseline assessment described in chapter 3)	to obtain base-line data	extract pertinent information	- age 17/18, 4 female, 11 male, all studying NEAB A Level Physics, experience problems searching for information on the internet, key skills important, require help with writing & summarisation skills
2. WebCT data 2a. Chat logs 2b tracking	- to collect attitudes towards webct & increase familiarity with tool - to examine reading strategies	- derive then illustrate response categories - derive navigation style categories & statistically analyse tool use	e.g. categories include 'off-task' chat + tutor's attempt to return to 'on-task' chat - evidence of 5 navigation style categories (see Table 4.7)
3. Word summaries	- to record final summaries to ascertain quality & areas for further support	- looking for points of interest - calculating % of original, to give an idea of quality	- 8 summaries were emailed to the instructor: one student produced a model summary whilst one student cut & paste everything, 5 used a title, 2 students annotated - a range of summaries were produced (see Appendix G)
4. Miscellaneous documents	-any relevant additional artefacts should be analysed	- look for evidence of summarisation strategies/aids	- one student used a separate piece of paper to draft the summary
5. Observational notes	- miscellaneous points not covered by other data + immediate feedback	- look for signs of tool proficiency + points of interest - look for summarisation strategies	- students soon mastered basic webct proficiency - one student pair took the study rather light-heartedly (S6) - students used various styles to pre-read - there was little time for reflective review

Table 4.8 Summary of Notre Dame Summarisation Study Results.

FINAL CONCLUSION	DATA SOURCE	INTERMEDIATE CONCLUSION
ND1	4 & 3	summaries written 'off-line' lead to better summaries when formatively evaluated (e.g. see S3 in this chapter and section 7.5.8) possibly showing a deeper engagement/reflection.
ND2	4 & 5	copying of the whole text to the summary is indicative of unmotivated students (possibly showing a shallow engagement)
ND3	3, 5 & 2b	this study generally corroborates the CHALCS process model – see Figure 4.4.
ND4	2a	affective tutor intervention may provide a model for the Learning Companion intervention, ie. use of sarcasm, encouragement, etc.

Table 4.9 Summary of Conclusions of CHALCS Summarisation Study
(Key: NDx = a reference for the final conclusion, see Table 4.12)

Due to the additional problems of coordinating a large class the results from the Notre Dame study are not quite so detailed as found in the first study. It was, however, apparent that a number of techniques are used by the students to arrive at their summary documents. For example, one conclusion of the utilisation of separate pieces of paper by S3 (also as evidenced by the CHALCS student) would be that this 'off-line' activity gives the students chance to reflect on the problem and obviously prevents copying verbatim thus potentially resulting in a deeper 'processing' and subsequent understanding – eventually producing a good summary. The summaries varied in length from simple three line documents containing very little useful information to massive documents, as mentioned, containing the entire course notes. Obviously, this latter method also contains little useful information in terms of a summary as this would result in information overload effect for revision purposes and shallow processing/comprehension of the text itself, likely to mean conceptual understanding and initial acquisition will be poor. This, however, was the contribution of a particular participant, S6, who took the whole study rather light-heartedly.

Appendix G contains a range of summaries. The first (S3), is a concise distillation of the key information written as a numbered list with a title, the second (S4), mid-range summary contains the key text which has been copied and pasted from the original course materials and the third summary (S7) almost completely replicates the original text. Table 4.10 contains information indicating the 'quality' of the summaries produced based on a simple word count. Of course 'quality' is not just dependant on summary length as information needs to be relevant and to the point. However, there appears to be some relationship between length and the ranking of the summaries for their (qualitatively assessed) quality. That is, summaries clustering around the mean of 204, i.e. 40%, (with the standard deviation of 111) were judged qualitatively 'better' summaries by an independent study skills expert than those at the extremes of the range which tend towards either brevity or verbosity. The total number of words used in the original course materials was 509. (Note that Chapter 7, section 7.5.8 further evaluates the quality of these summaries when they are compared to summaries produced with the aid of the developed Learning Companion.)

Student	Summary Length	% of original
S1	39	8%
S2	123	24%
S3	129	25%
S4	162	32%
S5	253	50%
S6	265	52%
S7	274	54%
S8	388	76%

Table 4.10 'Quantitative Assessment' of Notre Dame Summaries

The chat logs for the first part of the Notre Dame study revealed mostly dialogue disassociated with the task. For example, one student (S12) said: "anti neutrinos suit me better…they go with my shoes" and general chat about their plans for the evening ! However, one student (S6) did announce "does anyone think this is a seriously good piece of software ?" seeming to indicate that they were impressed with WebCT. There was a prolonged intervention by the instructor in an attempt to get the chat more on-task. Example interventions include (abridged):

1. can we chat about summaries please ?
2. what differences do you notice ?
3. what's better about your partners summary ?
4. how did you summarise compared to your partner ?
5. what did you/they leave out ?
6. does anyone actually need help in summarising ?

The instructor eventually became frustrated, took a step back and in fact made quite a cutting remark (point 6 above). It should be noted here that facilitating an online chat session differs markedly from facilitating a face-to-face discussion for a number of reasons. Importantly, since normal turn-taking can be disrupted due to the phenomenon of each person composing and sending replies or new topic openers simultaneously "multiple parallel threads" develop in chat which can make the discussion more difficult to follow and certainly more difficult to direct or command attention. Consequently the tutor needs to develop new strategies for handling discussion in this medium and there is emerging evidence that they cannot do this alone but must first scaffold the students in adopting certain protocols for online discussion if they are to be effective (Pilkington & Walker, in press; Salmon, 2000). Moreover, there is evidence that groups online need to follow the usual processes of 'forming, norming and storming' (Tuckman, 1965) in a similar manor to normal group formation even if the participants have been a face-to-face group for some time (Walker, 2003). Note that the proposed agent should not have such a difficult task in gaining attention in as much as the kind of discussion envisaged is a collaborative one-to-one rather than one-to-many discussion, although mechanisms for the control of turn-taking and gaining attention will need to be modelled. However, the tutor's final question does result in an ensuing on-task conversation thread albeit disjointed (see Table 4.11). To elaborate, one student (S6) reported that a summary is "a single piece of paper which teaches a whole module in a single bus journey" and another (S11) stated (as shown in Table 4.11) that WebCT is "Very good...its' like Yahoo...but more physics orientated." This comment further suggests that at least some of the students were already familiar with using chat environments in their leisure activities and this may have contributed to the sense that this was a 'fun activity'. Students may, however, also need to learn a more focused register for using chat in school (and later the work-place) – something which only practice in using chat in learning or other authentic working contexts will provide (see also Walker, op. cit.). The example cited above of S12 joking about neutrinos is a case in point. At this point a number of other students begin to engage in the conversation stating that they could benefit from more notes (Table 4.11, points 8 and 10) and one in particular (S1) criticising the folder facilities in the bulletin board (Table 4.11, point 11), thinking that "it would be a pain checking every single discussion section (from the pull down menu) to see if there was a query you could help with." However, as commented on by S10, one student does feel that the task is more fun than their usual lessons (Table 4.11, point 9).

These data do concur with the process model as derived in the CHALCS study consisting of decision-making (where students select the text thought appropriate as evidenced in their summary documents and utilising the two types of pre-read course note navigation techniques) and transfer (such as cutting and pasting the whole document as evidenced in observations and the final summary document). However, the reflective review phase was less in-evidence due to time constraints. In addition the action of first hand-writing a summary by the one student pair concurs with this model as exemplified by the S2 CHALCS student. Additionally, the use of the summary by the five students (S2, S3, S5, S8 and S12) when navigating the course notes points to the fact that they do find summaries useful in the 'acquisition' phase of learning. These five students can be viewed as utilising a sequential navigation strategy (exempting the viewing of the summary) whereas there were an equal number who used a more varied 'hypertext' style of navigation. Points of similarity are thus found with the two types of navigation demonstrated by the CHALCS students.

No.	Contributor	Chat contribution
1	T	so what do you all think ? good or bad ?
2	S11	Very good
3	S1	In what respect?
4	S11	It's like Yahoo....but more physics orientated
5	S1	I think the technology is a little bit overused - there's no need to have several user chat rooms with compartively few users.
6	T	I'm sorry we were pretty rushed...
7	S6	hello and welcome
8	S6	more notes
9	S10	What do you mean "This is far better than our usual lessons with Mr Coen"?
10	S6	need more notes
11	S1	same with the discussion board; i don't think segregating is the best way to do it. if someone knows physics, they, er, know it.

Table 4.11 Notre Dame 'Chat' Excerpt.
(Key: Sx = student x, T = Tutor)

In terms of the Learning Companion design the dialogue of the tutor intervention as described could inform an affective component in that the Companion will only reformulate a leading question for so long before resorting to techniques such as sarcasm to take control and/or refocus the attention of the collaborating student. In the example given above this technique seems to work well although a fine balance must be negotiated between constructive and detrimental intervention and the perception of this and/or preference for a particular style may vary with the personal traits of the student. The precise nature of the dialogic component of the companion will be revisited in Chapters 5 and 6.

4.5 Conclusions.

Table 4.12 summarises the overall results derived from triangulating results from both the CHALCS and the Notre Dame Summarisation Studies. The remainder of this section discusses these findings, reaching conclusions for the further note-taking support aid.

CHALCS/Notre Dame Associated Conclusion	Overall Conclusion
C1 + ND1	the further support should include a 'scratchpad' for jotting ideas
C2 + ND3	the further support should include a 'decision aid' to help with text selection
C2	the further support should include an 'organiser' to help with coherence, etc.
C3 + ND3	the further support should include a flexible 'process model' of summarisation
ND4 + C4 + C5	the further support should include an appropriate 'personality'
ND2 + ND4	there is a need to increase the motivation of some students, possibly through the use of a 'personality'

Table 4.12 Overall Conclusions of the Summarisation Studies.
(Note: see Table 4.6 and 4.9 for the 'associated conclusion' codes)

As has been argued, further support is needed to aid students when taking their own notes within WebCT due to the inadequacies of the 'my-notes' facility. From the studies described in this and the previous chapters and previous work in summarisation (described in Chapter 2) it has emerged that this support should include the following:

1. A scratch-pad for jotting ideas and diagrams (as used by S2 in the CHALCS study and S3 in the Notre Dame study) possibly leading to deeper processing 'off line' and resultant better summaries;
2. A decision-aid to selecting pertinent text thus helping students who simply select all the original text or who do not note the rhetorical structure (as evidenced by the CHALCS discussions – see 'decision-making' in Figure 4.4);
3. A cohesion aid to help students, for example, with cohesive ties or the use of lists, etc. to create a more readable summary (aiding students in "sentence rearranging" as evidenced by students in the 'reflective review' stage of their summary writing – see Figure 4.4);
4. A flexible underlying process model (supporting the varied stages of summarisation such as varying pre-reading and revision strategies);

5. A suitable 'personality' to motivate students (based upon the peer interactions reported above such as the use of humour, colloquialisms, and a choice of dominant/submissive personalities).

Therefore, to summarise, further support could serve to support the summarisation process in three stages :

1. in the initial selection **decision-making** (e.g. aiding selection of pertinent information and supporting different styles of course notes navigation);
2. during the **transfer** from one document to another (e.g. the possibility of a *scatchpad*);
3. the subsequent **reflective review** process (e.g. employing notions of organisational, syntactic and semantic correctness both at the level of sentences and complete documents, especially checking cohesion).

This support would have to be **flexible** enough to allow students to utilise differing strategies in navigating the source material as evidenced in the above studies and shown below in Figure 4.4. One possibility is to expand upon this work by synthesising a number of relevant interventionist strategies as further support mechanisms. To this end it is proposed that a Wizard of Oz technique could be employed to 'dry run' additional support designs before they are coded in software summarisation support aids (see Chapter 5).

The WebCT Physics course was well received at both CHALCS and Notre Dame both by pupils and tutors. However, technical difficulties do need to be overcome if the on-line classroom is to work seamlessly. The process model has been corroborated to a certain extent by both the CHALCS and Notre Dame studies. In addition the role of affect in terms of the personality of the tutor intervention has been highlighted as a valuable contribution to effective educational dialogue, helping to increase student motivation and subsequent engagement in summarisation and the Physics course notes. As highlighted in previous chapters, it is intended that use of affect will also be tested in the subsequent Wizard of Oz study and modelled in the artificial Learning Companion.

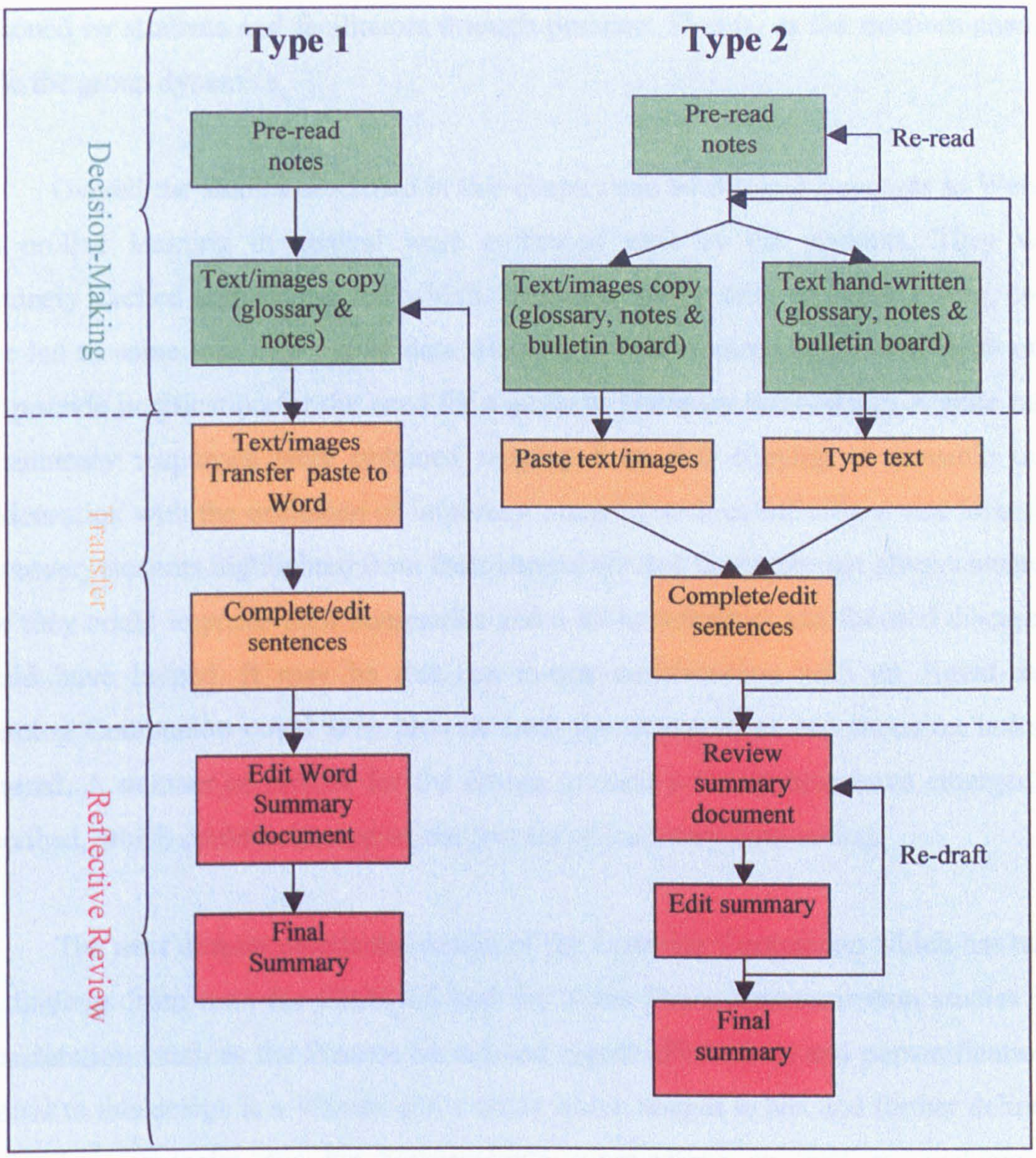


Figure 4.4 The Summarisation Process Model.
(Key: boxes = processes, arrows = process flow.)

4.6 Summary.

One of the major differences between the conclusions of the Notre Dame study and the CHALCS study was that in order to facilitate the dialogues on effective summarisation, etc. WebCT chat was used. With hindsight this may not be the best way to elicit information as to new users the allure of chat and possibly the novelty of its use in the classroom tend to make them respond off-task. This could have been avoided in a normal face-to-face session where tutor intervention would inhibit unwanted contributions with more authority than is seemingly attributable on-line. To elaborate, a new method of communication calls for new models of social interaction which need to be honed by students and facilitators through practice. That is, as the medium changes so do the group dynamics.

Overall the studies described in this chapter can be deemed a success as WebCT and on-line learning in general were embraced well by the students. They were genuinely excited and having fun which, ironically in the case of Notre Dame, could have led to some loss of accurate data. However, both studies complement each other and provide justification for the need for a guide to summary note-taking. A wide range of summary responses were obtained ranging from full copying to possible over-condensation with the exclusion of important material such as definitions and formulae. Moreover, students highlighted from their comments that they were not always aware of how they could improve their summaries and a more reflective and focused discussion would have helped. It may be that one-to-one collaboration with an Agent-based Learning Companion could help provide both the engagement and focus on tasks as required. A number of criteria for the design of such a Companion have emerged, as described, which could help support the process of summary note-taking.

The next chapter details the design of the Learning Companion which has taken the findings from both the CHALCS and the Notre Dame summarisation studies into consideration (such as the Process Model and Agent affectations and personifications). Central to this design is a Wizard of Oz study which sought to test and further delineate appropriate strategies for the Companion's intervention as well as evaluate the effectiveness of an animated Learning Companion.

Chapter 5: User-centred design of an Affective Learning Companion: Utilising the Wizard of Oz Technique.

5.1 Background to Learning Companions.

A Learning Companion is an artificial Agent whose goal "is to stimulate the student's learning through collaboration, competition and demonstration" (Chan & Baskin, 1988). Educational systems which utilise this technology are reviewed in a later section. For the moment we note that the name Learning Companion may be misleading as a Learning Companion (Chan, 1996) suggests the Agent acts as a peer of the student although there are in fact a number of roles that the Companion can fulfil. For example, the companion may act more like a student so that the human student must teach it how to solve a particular problem and thus learn by taking the role of the tutor. Another alternative is that the companion solves the problem itself whilst the human student 'watches' thus they learn vicariously.

It is worth summarising the possible roles for a Learning Companion within the WebCT Physics course (Barker, 1999b). These roles include:

1. a *peer* with which to collaborate on the task and provide encouragement and support;
2. a *guide* to show students around the VLE, alert students to what is new or relevant to them and assist in navigating course notes;
3. a *coach* to give instruction in the use of VLE tools such as the bulletin board or white board;
4. a *tutor* to monitor performance and give timely explanation and feedback during all three pedagogical stages of acquisition, application and argumentation and additionally relieve some of the tutor's administrative burden.

For the proof of concept that forms part of this research we have already argued for both pragmatic and pedagogic reasons that a Learning Companion to support students' note-taking would be a good exemplar case to test the utility of the approach.

It is proposed that the Learning Companion and the human peer collaborate (along the lines discussed in Section 2.2) to solve concrete problems involving extracting information from the Physics notes and jointly constructing an appropriate summary document.

The role of the peer in this context will actually be that of "more capable peer" (Vygotsky, 1978). Since note-taking is seen as central to the acquisition stage it is hoped that this 'more capable peer' will act as a 'role model' for students through demonstration of summarisation strategies and techniques, i.e. the student will be encouraged to develop their note-taking skills by entering the Zone of Proximal Development (as argued in Section 2.2.1).

Chan & Baskin (1988) proposed an overall strategy for such knowledge negotiation which involved discrete turn-taking between artificial and human peers with the problem initially being defined by a tutor (possibly also artificial) who is always on hand in case either student reaches an impasse. At CHALCS, however, it is hoped that the Learning Companion can be utilised when there is no teacher present. Therefore, an alternative needs to be considered since stopping mid-summary whilst a human tutor reads an email, maybe the next day, then replies with the aim of putting the peers back on track would result in a loss of impetus and possibly a resultant loss of motivation.

In the next section work in this field relating to the design of Learning Companion Systems will be reviewed in order to explore the range of possible alternatives within Social Learning Systems (Chan, 1996). This will be followed by an account of a preliminary formative 'Wizard of Oz' study designed to elicit perceptions and help specify characteristics of a note-taking Learning Companion prior its full implementation.

5.1.1. Social Learning Systems.

Central to the rationale for the integration of a Learning Companion within WebCT as mentioned in Chapter 1, is the notion of Social Learning Systems (Chan, 1996). As stated, an essential part of a Social Learning System is an artificial Learning Companion. A substantial amount of the functionality of Social Learning Systems is already being provided, for example: primary and secondary resources, human Learning

Companions and human tutors, within the the WebCT Physics course. However, to realise Chan's vision we also require artificial companions with which the students can collaborate. They are by no means intended to be a panacea to the problems of on-line learning but when viewed holistically, i.e. within the context of the Social Learning System, they help provide a complete educational support.

Interestingly, Chan's notion of s Social Learning System is very much akin to the British Government's notion of a National Grid for Learning mentioned in Chapter 1. Also mentioned in Chapter 1, more locally, we can see similarities between Social Learning Systems and the local networks beginning to emerge in British cities (see Appendix B). These networks exist to link public resources such as libraries and museums to local schools, colleges and universities thus creating an educational infrastructure with the common aim of knowledge sharing and communication. Thus, Chan's Social Learning Systems have never been more relevant.

5.1.2. Previous Pedagogical Agents and Considerations.

In addition to the work reviewed in Chapter 2 concerning personifications of Animated Pedagogical Agents there are some notable implementations that relate more closely to the goals of designing a collaborating agent in a learning context or a 'Learning Companion'. These implementations include the work of Aimeur and Frasson (1996), Chan and Baskin (1988; 1990), Goodman et al (1998), Dillenbourg and Self (1992), Uresti (2000) and Brna et al (2001). Each shall now be reviewed in turn. This is not intended to be an exhaustive review of Learning Companion implementations but does suggest the kind of issues to be considered in their development. A good review can also be found in Chou et al. (2003).

Aimeur and Frasson's (1996) work draws conclusions concerning the need to model a learner's knowledge level and affective state in order to select a "good learning strategy". They describe three classes of such learning strategies, notably 'one-on-one' (like traditional ITS), 'learning with a co-learner' (the idea being that the learner cooperates with a peer where "the knowledge level of the co-learner is slightly higher than the learner", cp. this work) and 'learning by teaching' (where the learner teaches a companion). They further propose and also test a new learning strategy "Learning by Disturbing" where a 'troublemaker' can give a wrong answer to a problem "in order to force the learner to react and propose the right solution" or wait for the student's answer then suggest a wrong solution or counter-example. By empirically testing their new strategy in comparison with co-learner and learning by teaching strategies the authors found that it is more suitable for people knowledgeable in the subject matter but could be dangerous for people with "unconfirmed knowledge". Thus indicating the need to know the knowledge level of the individual before applying this strategy.

The work by Aimeur and Frasson is interesting in terms of Learning Companion strategies but their strategy of learning by disturbing will not be adopted here – principally because the aim is to test the effect of different persona on the motivation and performance of students. For this reason there will be no attempt to vary the likelihood of an erroneous suggestion with either a 'passive' or 'dominant' character. In this way any differential affect of the 'character' with which the student studied should be due to perceptions of 'passivity' or 'dominance' and not the actual quality of advice.

The work of Chan has already been mentioned in relation to Learning Companion design. In fact Chan coined the term in a 1988 paper presented at ITS '88 (Chan & Baskin, 1988). This paper begins by quoting a Chinese proverb which recounts a Prince's companion during his lessons with a royal teacher leading to more effective education, hence "studying with the prince". His notion of a Learning Companion system is grounded in a socially situated, constructivist notion of learning based on the work of Vygotsky (1978) and Piaget (1970) amongst others (as reported in Section 2.2). Central to this foundation is that children learn more effectively when they have a peer with which to collaborate. With this notion in mind Chan sets about developing his own notions of how a computer support can capitalise upon this social advantage to Education citing the need for substitutes for unavailable peers in distributed learning environments.

The design of the Learning Companion Systems (Chan & Baskin, 1988) is based upon three interacting Agents. The first is the human student, the second is their artificial Learning Companion and the third is an artificial tutor. As Chan & Baskin state:

> "The role of the *computer teacher* is to offer examples, guidance and comments to both the student and the learning companion. The goal of the *learning companion* is to stimulate the student's learning through collaboration, competition and demonstration."
>
> (Chan & Baskin, 1988)

The teacher's role is seen as providing insightful comments on the lesson, justifications or advice on negotiations. This means that the student and companion are left to work together on the problem presented by the tutor. Chan & Baskin further define three possible roles for the companion. The first is '*companion as competitor*' i.e. where both companion and student work independently on the problem then "compare solutions, discover mistakes and self-correct mistakes in their own solutions" (Chan & Baskin, 1988). The second is '*working collaboratively with one working while the other watching*'. This method is self-explanatory, one Agent attempts a solution whilst the other watches then gives advice if needed or if neither can progress then a teacher will interrupt. Finally, the third role is '*working collaboratively on the same problem via responsibility sharing*'. This role type consists of one Agent being responsible for negotiation and the other for execution. First the decision-maker suggests a solution, next the executer suggests an alternative, this is followed by the decision-maker deciding which solution to use resulting in the executer carrying out the solution. The roles are then alternated. The second role is preferred for this research as it provides a simple model for collaborative learning in keeping with the aims of the research to investigate the effect of different persona rather than different collaborative styles. Moreover, in addition to be being (relatively) simpler to implement, the second role is one which is often spontaneously adopted by children in collaborative working and should be fairly natural rather than needing to be taught explicitly - see, for example, Crook (1994).

Chan & Baskin suggest two methods for the implementation of such mechanisms within an artificial Agent consisting of simulation and machine learning. The simulation route consists of selecting knowledge bases of varying degrees of completeness as a student's proficiency evolves thus simulating learning. The machine learning route would use Artificial Intelligence techniques so that the companion would learn and impart those learning strategies it used to the student. However, the former method is preferred by Chan & Baskin due to its simpler implementation. For this research, the idea is to fix the ability of the Agent so that comparisons can be made across students and the role of the Affectations (see Chapters 6 and 7) in the collaborations evaluated.

Wang and Chan (2000) have further developed Chan's ideas along a number of lines. First, they have developed a language with which to build Learning Companion Systems which are distinct from Learning Companions in that they may contain a whole multitude of artificial or human peers or tutors. To this end Wang and Chan propose an Agent-based architecture where agents communicate by event messages and are to some degree therefore autonomous. Secondly, Chan and his students (Chou et al., 1999) are beginning to investigate the utility of a Learning Companion model and propose General Companion Modelling which provides an architecture for companions which model their own beliefs, capabilities and behaviours. A further line of research presented by Chou at the AI-ED 99 Workshop on Animated and Personified Pedagogical Agents is that of 'disney-fying' companions, i.e. providing them with cartoon-like user interfaces and characteristics which make them more believable along the lines proposed in this research.

Goodman et al (1998) have developed a Learning Companion as part of an Intelligent Tutoring System which aims to teach Explanatory Analysis which is "the process by which an analyst formulates explanations for past or predicted events". The particular domain they concentrate on is that of satellite activity although they state the method could be used in e.g. legal reasoning and financial planning. As well as containing a coach which provides direct answers to student's questions the environment also contains a Learning Companion which is intended to promote reflection and articulation amongst students through the use of collaborative dialogue. As Goodman et al state:

> "A learning companion can help a student reflect on his thinking by critiquing, questioning or evaluating particular steps. Similarly, the student and learning companion may articulate those steps through further explanation or elaboration."
>
> (Goodman et al., 1998)

With their emphasis on articulation and reflection, which they see as central to the collaborative learning ethos upon which they justify their learning companion approach, Goodman et al concentrate on drawing out an involved dialogue between learning companion and student. This results in interaction with their companion, LuCy, being dialogue-oriented whereas the interactions with the computer-based tutor are not. For example, when elucidating the differences between tutor and companion they say that the nature of the interaction with the student is "authoritarian" in the former whereas the latter utilises "disputes, opinions and evaluations". LuCy can, in fact, provide partial or misleading information utilising mal rules to provide "buggy responses" which are fired depending on probabilities for varying student levels. This means that LuCY does not always provide a correct suggestion so it would be up to the student to notice this and challenge LuCy accordingly.

In contrast to the approach posited by Chan and Baskin (1988, op. cit.) where the companion is seen as a peer, LuCy is able to access the expert domain model and student model thus allowing recognition of student misconceptions. However, Goodman et al compare their work to that of Hietala and Niemirepo (1998), described in Chapter 2, saying that although LuCy adapts according to the student's knowledge "LuCy does not provide any personality parameters that can be tuned to exhibit particular personality traits". Whilst models of provoking reflection through sophisticated dialogue models that access user models is not the central focus of this research it is necessary to consider the ways in which the companion will articulate suggestions, challenges or requests for explanations and respond to those of the student whilst collaborating on co-writing a summary. For example, asking students to give reasons for inclusion of certain unknown or less important phrases in their summary, may cause them to reflect on their value as they attempt to articulate an answer.

The pioneering work of Dillenbourg and Self (1992) can be conveniently accommodated in this section although they make no claims to be designing Learning Companions as such. Instead they are interested in the collaboration which takes place in 'socially distributed cognition'. This collaboration can take place between two computer-based agents or between a human and a computer-based agent. This latter combination they call human-computer collaborative learning (HCCL). Their application domain, in which they test their ideas on HCCL, consists of a micro-world simulation of electoral processes which can consist of a "computerized 'co-learner'" amongst other components. Their view of agency is based on notions put forward by Minsky (1987), that is, the term 'agent' is used to refer to a cognitive process and the term 'device' represents "where agents are 'implemented' (a human brain or an artificial system)". They then postulate three axioms as follows :

1. "An individual is a society of agents that communicate. A pair is also a society, variably partitioned into devices.
2. The device border determines two levels of communication: agent-agent communication and device-device communication. Inter-agent and inter-device communications are isomorphic.
3. Inter-device communication is observable by each device. Therefore, inter-device communication patterns generate intra-device communication."

(Dillenbourg & Self, 1992)

Thus their research objective is that "mutual regulation processes, encompassed in inter-device communication, create self-regulation processes". Therefore their objective for their HCCL system, called PEOPLEPOWER is to apply rules on argumentation either between learners to derive a dialogue or with the same learner (i.e. reasoning) to derive a monologue. The procedure at the heart of the system comprises a theorem prover which utilises a depth-first search of a tree of rules (i.e. arguments). Links between these rules consist of 'refute links' and 'continue-links' which have an associated strength modified on the basis of dialogue and the results of elections. This modifcation is, Dillenbourg and Self say 'learning by experimentation'.

The main disadvantage of the system appears to have been in the cost associated with the interface which required a user's familiarity with the arguments as, basically, this dissuaded users from interacting with the system. Without a rich dialogue history for the HCCL to learn from the subsequent suggestions made by the co-learner were poor. Although the HCCL described above is a simply effective model of collaborative learning from which lessons can be learned in Companion design, little research effort went into creating a friendly interface as this was not the primary focus of the research.

A friendlier interface which allows a more naturalistic interaction would however be required if it were to be used by CHALCS students.

The design of such an interface for a companion is a more central goal for this research than the richness of the agent's reasoning. In addition Dillenbourg and Self state that a further limitation of their work is that it only allows for disagreement and agreement utterances whilst other work in collaborative dialogue analysis suggests that a whole host of utterances are necessary. The utterances allowed in SILA-student collaboration shall be described in Chapter 6, section 6.4.3.5. First though more recent work in Learning Companion design is described.

Uresti (2000) describes empirical work he carried out with a Learning Companion designed to be situated in an environment to teach Boolean Algebra. The hypothesis is that students teach a companion with less expertise than themselves how to perform the task and thus benefit their own learning. Uresti believes that in teaching the companion the student "will have to revise, clarify, organize and reflect on her own knowledge" (Uresti, 2000). The empirical work utilised two companions, one with a strong expertise and the other with a weak expertise. In addition, two further conditions were introduced in the form of strong versus mild (or 'free') encouragement to teach the companion. A score is displayed to the student which is related not just to their own performance but also to that of the companion. The companion's score improves mainly by teaching it whereas the student's score increases based on their own performance and their involvement in helping the companion. The mechanism by which the student teaches the companion is to modify the rules, for example their priority. Uresti found that the condition consisting of a weak companion with a motivated student showed a trend of being most beneficial to student's learning whereas the condition consisting of a strong companion with the free student showed a trend of being the worst condition for learning.

In addition to the conditions described, Uresti's companions also displayed human-like traits. The first of these consisted of the companion showing a greater reluctance to accept a new algebraic rule the more it was certain that it knew the correct answer (unless it was the same rule). This was an attempt to model a human's tendency to increasingly reject a concept the more they were sure they were correct. However, Uresti reports student's disbelief at the rejection of the companion, not expecting such a behaviour from a computer. As Uresti reports, this kind of student response has been termed 'The Plausibility Problem' as reported and explored by du Boulay et al (1999). A further human-like trait consisted of the weaker companion requiring to be taught the same concept more than once in a simulation of its inferior knowledge leading to a slower learning process. Students (especially in the weak companion, 'free' student condition) responded by decreasing the quality of their teaching to the companion as the lesson progressed, that is their reluctance to teach increased. In summary, Uresti concludes that "subjects who are strongly motivated to teach the weak LC [Learning Companion] benefit most from the teaching interaction" but that "users of the system were not prepared to deal with a LC which reacted as a human peer might do".

As Uresti's work utilises the 'learning by teaching' paradigm of companion design it is felt that the preference of strong students to teach weaker companions cannot automatically be generalised to a more collaborative style of interaction along the lines discussed in Section 2.2. i.e. where the role is not to teach but to complete a task together, maintaining a shared conceptualisation of the problem and a symmetry in division and understanding of the task. However, a similar finding could emerge in this work that strong students prefer to interact with weaker companions. This is something which will have to be born in mind as our Companion is evaluated.

The plausibility problem reported by Uresti is also of concern to this work as human-like traits in the form of 'affectations' are being adopted in the design of the companion (as discussed in Chapter 6). It is worth bearing in mind student's preconceptions of how software should behave in the design of companions if they are to be accepted, particularly as equals. Although it is hoped that by employing embodied, or animated, agents the hypothesised increase in believability will lead to a decrease in the student's conceptualisation of the companions as 'software' thus the plausibility problem may be less of an issue.

The work of Brna et al (2001) does not strictly belong to the field of Learning Companions as their Pedagogical Agent takes on more of a tutor's role. They have designed an Agent, Louisa, which is situated in an experimental narrative environment with which students construct stories based around pictures and text. It is, however, an important piece of work and of concern here because the role of *empathy* in the Human-Agent Interaction (HAI) is emphasised.

Rogers (1975), an influential psychotherapist, revises a previously attempted definition of empathy as a process rather than a state. This revised definition emphasises the need to enter "the private perceptual world of the other", being sensitive to and uncovering feelings, temporarily living another's life, communicating "sensings" of another's world which should be checked with the other and used to guide subsequent dialogues. Rogers states of empathy that:

> "to be with another in this way means that for the time being you lay aside the views and values you hold for yourself in order to enter another's world without prejudice".
>
> (Rogers, 1975)

A further definition of empathy is provided by Barret-Lennard (1962):

> "it is an experiencing of the consciousness 'behind' another's outward communication but with continuous awareness that this consciousness is originating in the other."
>
> (Barret-Lennard, 1962)

This last definition hints at the importance of maintaining one's own perspective in simulating the world of another, the 'as if' alluded to in Roger's original definition. For, without this degree of abstraction the objective analysis of the situation, feelings, etc. is lost. This may mean that the dialogue would not proceed as the guidance has been removed resulting, in terms of psychoanalysis in a lack of learning of the other's self or, in terms of a teacher, a lack of student learning of the subject matter. One final definition of empathy, particularly in relation to classroom experiences, is provided by Cooper and Brna (2002):

> "Central to communication and understanding is the quality of empathy, which enables people to accept, be open to and understand the perspective of others, whilst simultaneously developing their own perspective. Empathy supports and enables interaction, and creates the climate for both affective and cognitive support."
>
> (Cooper & Brna, 2002)

Brna et al (2001) state that the capability of their Agent, Louisa to demonstrate empathic characteristics is important even though it has no underlying model of affect. As they state:

> "Though an artificial tutor coupled with a believable agent cannot really empathise with or understand the students to whom it responds, it can demonstrate empathic characteristics which improve the learning climate and help to meet the individual learning needs of students".
>
> (Brna et al., 2001)

Of course this is a realistic statement in that the artificial tutor cannot experience empathy in a strictly human sense, i.e. it has no consciousness (alluded to above) or mind-body interaction - in the sense of Damasio's (1994) and other neuroscientists findings of interrelated responses to affective stimulation between mind and body leading to more effective interactions. However, this is in-keeping with the notion of Affectations (see Chapter 6, section 6.4.3.5) espoused in this research as the assumption is that "characteristics" of an Agent can be *interpreted* as empathic behaviour by the student. Furthermore, Brna et al (2001) state that they are more concerned with managing the emotions of the students rather than modelling the full range of human emotional behaviour which is of secondary importance to them. In fact they admit that currently the model is quite "shallow" but that they will build a "deeper model" taking into account such aspects as interaction patterns and levels of student expertise in story writing. These aspects together with the child's name, reading age and other work done could, they argue, lead to an 'empathic' relationship between the student and the Agent.

Brna et al (2001) conclude their work with an elucidation of desired empathic characteristics in teachers or, indeed, artificial tutors. These are:

1. Attend - "consider the actions, thoughts and feelings of the learner";
2. Engage - the teacher aligns "their actions, thoughts and feelings" with those of the child;
3. Value – express the value to the teacher of the child's work;
4. Encourage – encourage the child to go further utilising emotional, physical and cognitive aspects;
5. Parting – as the teacher switches her attention away from the child future availability is indicated and 'closure' avoided;
6. Available – can easily be called upon when required.

Cooper and Brna (2002) further define both empathic and unempathic characteristics of teachers (summarised in Table 5.1) based on a longitudinal, qualitative study of teacher/pupil relationships in UK schools.

These characteristics are of interest here because empathy in Pedagogical Agent design is certainly central to the Human-Agent relationship. The success or failure of this relationship could make or break student motivation which is often cited as of major concern in Pedagogical Agent design. By modelling some of the empathic characteristics shown in Table 5.1 in our Pedagogical Agents it may be possible to create a more empathic educational environment. To elaborate, based on their longitudinal study Cooper and Brna conclude that:

> "Empathy was central to high quality, effective teaching and learning, enabling greater understanding, better assessment, better academic and emotional support and consequently more appropriate teaching provision and more appropriate differentiation."
>
> (Cooper & Brna, 2002)

Characteristic Class	Empathic Characteristics E.g.	Unempathic Characteristics E.g.
Attitudes	open, warm, relaxed, fair	sees class as a group, impatient, intolerant
Facial Characteristics	frequent smiles, generally positive demeanour, expressive face	unexpressive, do not show emotions, facial expressions not in tune with words
Voice	positive, encouraging, expressive	unemotional, tone not in tune with words, matter-of-fact
Body Language	uses gesture, animated, tactile	wooden, unapproachable, distant, not animated, formal
Positioning	gets closer to child, less distance	distanced from children, front of class, more formal
Responses	uses students name frequently, listens carefully, gives a positive response but asks them to elaborate	not individualised, overrides, ignores pupil's comments, negative or unhelpful responses, does not value
Content of Teaching	relates directly to child's own experience, personal interest	sticks to curriculum, little differentiation, does not relate to child's interests
Method of Teaching	changes of pace and style, adaptable and flexible, differentiation	rigid, lacks variety, lacks interpersonal, elicits less from pupils, less reflection /discussion, interaction
Other Features	uses humour, 'not like a teacher', form personal relationships	behaves like a teacher, lacking in humour, shows false emotion, ignore emotion

Table 5.1 Example Empathic Teacher Characteristics (Cooper and Brna, 2002).

Finally, and importantly, Brna et al (2001) highlight ethical concerns for Agents (human or artificial), that is that they should adopt actions, thinking and emotional behaviour that they would like to see in their students.

This work on empathy is of great importance in the next generation of pedagogical agents (Kapoor et al., 2001) which will display this kind of 'social intelligence' utilising models of student affect (and hence raising the possibility of beginning to "enter another's world"), for instance, to affect their interactions leading to more conducive educational environments. As Rogers states:

> "Just as the client in psychotherapy finds that empathy provides a climate for learning more of himself, so the student in the classroom finds himself in a climate for learning subject matter, when he is in the presence of an understanding teacher."
>
> (Rogers, 1975)

With an idea in mind of previous work in Pedagogical Agents it is timely to examine the study at the centre of the User-Centred design of the artificial Learning Companion which forms part of this research.

5.2 Design of a Learning Companion: Wizard of Oz Requirements Analysis.

This section consists of an account of a study undertaken to arrive at the requirements for the Learning Companion and further reports on investigations which took place into the Design Specification, particularly those relating to the User Interface and Knowledge Representation. The study utilised an approach known as the Wizard of Oz (Preece et al., 2002, p.245). The rationale underpinning this approach is discussed, the study reported and conclusions and recommendations for the design of the Learning Companion made.

5.2.1. Introduction.

The Wizard of Oz study derives its name from the Wizard in the famous children's story in which Dorothy travels to a magical land by means of a tornado to defeat the Wicked Witch of the West. In the tale the wizard is actually a man from Kansas who operates a puppet from behind a curtain in an attempt to make his apparently lacking demeanour grander. This is in essence the technique adopted in Wizard of Oz studies in which the user or student interacts with a machine and the experimenter plays the role of Wizard manipulating the strings as it were behind the scenes and (usually) unknown to the user to simulate intelligent behaviour by the machine. The idea is to test the acceptability and plausibility of potential machine behaviours before they are fully implemented. In this study the task is to simulate the 'intelligent' behaviour of an animated pedagogical agent by remotely manipulating the graphical and textual/speech aspects of an on-screen puppet.

This section provides an overview of how the study was carried out as well as a full results analysis ending with conclusions and recommendations for future work.

5.2.2. Rationale.

The Wizard of Oz method is grounded in previous Human-Computer Interface work as well as more relevant user-centred designs for intelligent Agents. Maulsby et al. (1993) describe a prototype of an instructible Intelligent Agent called Turvy which utilises the Wizard of Oz methodology to evaluate its intelligence in the domain of text editing before it is actually coded. The EUROHELP project (Breuker, 1990) also used this technique to empirically investigate a 'coach' in an Intelligent Help System. More recently Kemp (Kemp, 1997) used the Wizard of Oz to prototype a Scenario-Based Simulation tutor, in the domain of social disputes, before any coding was again actually carried out. Additionally, Wooffitt et al. (1997) posit a design framework consisting of three stages. The first is 'design by inspiration', the second 'design by observation' and the third 'design by simulation'. They applied this framework in the design of a natural speech understanding system. As part of this design, during their third stage they employed the Wizard of Oz technique.

In the study to be reported here the Wizard of Oz technique was aimed at illuminating particular requirements of such an animated pedagogic agent in the particular educational context of collaborative summary note-taking. Using the wizard of oz method it was hoped that a potentially acceptable design for such an Agent could be arrived at before the component was implemented through interaction with students. By simulating such an Agent before coding takes places necessary aspects of behaviour and details of preferred interaction styles are highlighted leading to a sharper understanding of the specific functionality required. In this way it is hoped that a more user-centred design with a higher acceptance will result. In addition the method offers insights into the degree of intelligence required by the agent in the interaction with the student in order to be convincing.

5.2.3. Design.

Figure 5.1 shows the environment presented to each of the 9 participants. The Agent shown (The 'Genie') is implemented using Microsoft Agent as described in Chapter 6. Microsoft Netmeeting is used to share both Internet Explorer (containing the

WebCT Astronomy and Optics course) and Microsoft Word (used for the construction of the summary document) between the participant and the Wizard. As shown in the diagram, the task is for the participant to construct a summary of the introductory material "What is a Lens ?". As well as the functionality of WebCT being available to both parties, such as glossary hyperlinks and external resources ("URLs"), the participant also has the total functionality of MS Word such as bullet points, spelling checker and more general text formatting options.

Figure 5.2 shows the Wizard's view of the environment. MS Netmeeting is shown for sake of completeness although when running this window can be minimised. The two shared windows containing Internet Explorer and MS Word are shown to the left. At the upper right side of the screen can be seen the puppet's control panel. This contains all of the possible moves allowed by the software as detailed in Table 5.2. The five categories shown naturally group the commands together. These moves are based on the work of Maulsby (1993) mentioned earlier. However, they are adapted for this particular context, for example by the inclusion of the 'Well Done' move so that the Agent can encourage a student. At this stage a minimum number of obviously needed moves were selected. It was expected that one of the outcomes of the study would be an idea of how limiting this set was for students and what additional moves they felt they needed.

By pressing the appropriate button the Wizard initiates a sequence of puppet actions together with an associated speech act. For example, pressing the 'Look' button results in the Genie first moving to the centre of the screen then moving to the right, gesturing left and saying "Look here ! This is important !" The commands "Say", "Same" and "Because" result in the Genie carrying out actions specific to these commands but also saying text as entered by the Wizard in the supplementary text box at the upper left corner of the Wizard control panel (see Figure 5.2). The participant had an equivalent set of commands at their disposal which could be selected from a menu or spoken to the Genie using a microphone and speech recognition. This would result in a dialogue box containing the command appearing on the Wizard's remote computer.

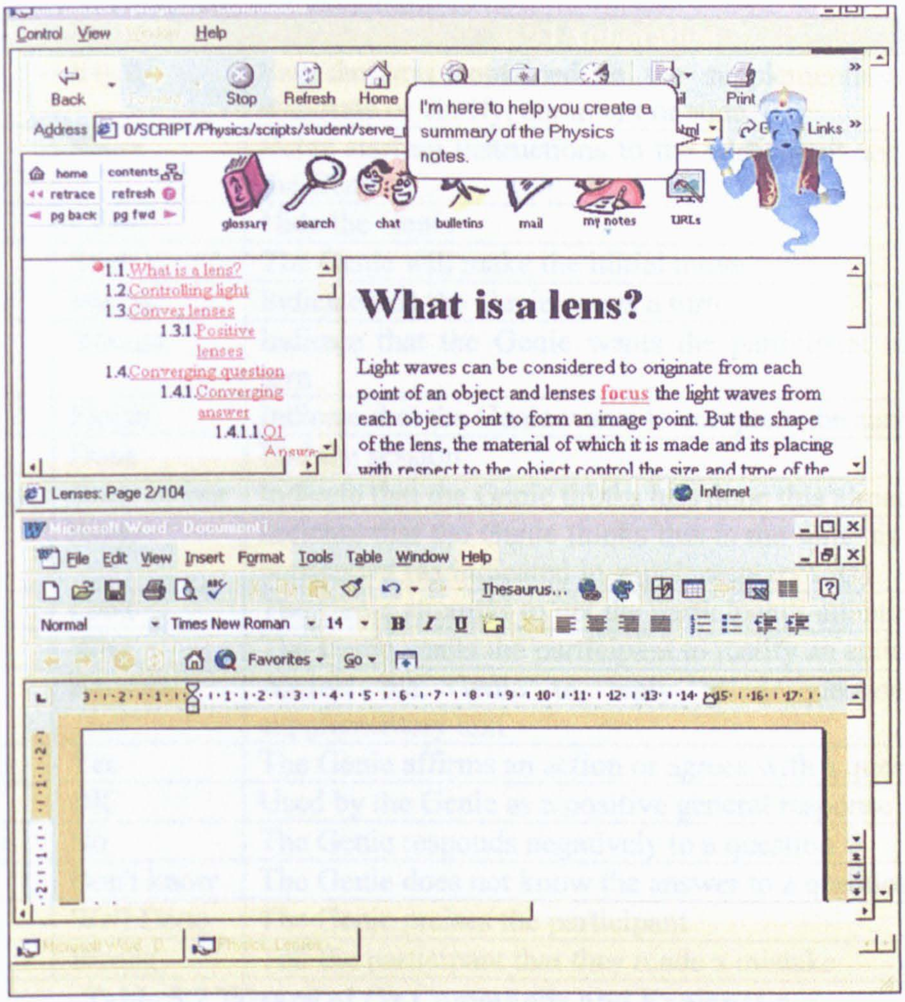


Figure 5.1 Participant's View of the Wizard of Oz Environment at Start-Up.

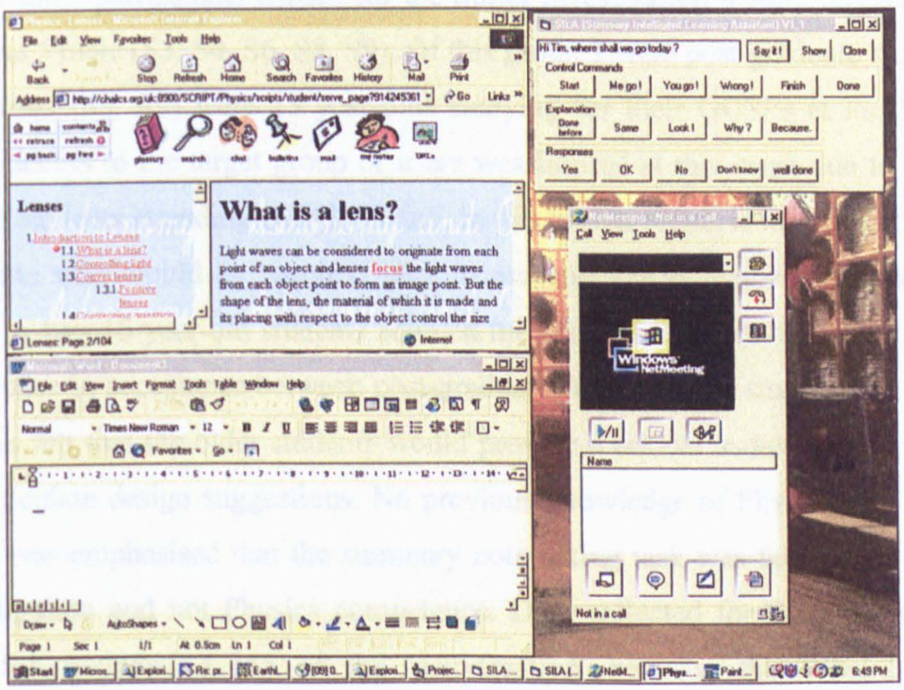


Figure 5.2 Wizard's View of the Wizard of Oz Experiment at Start-Up.

Category	Command	Explanation
General	Say it	Say the text contained in the supplementary box to elaborate or clarify, comment or state
	Show	Relay starting instructions to the participant and display the Genie
	Close	Hide the Genie
Control	Start	The Genie will make the initial move
	Me go	Indicate that the Genie wants a turn
	You go	Indicate that the Genie wants the participant to take a turn
	Finish	Indicate that the Genie wants to complete the summary
	Done	End the session
Explanation	Done before	Indicate that the Genie thinks he's done this already
	Same	Indicate that the Genie thinks this is the same as another participant (as indicated in supplementary text)
	Look	The Genie attempts to get the participant's attention
	Why	The Genie wants the participant to justify an action
	Because	Allows the Genie to reply to a question using supplementary text
Responses	Yes	The Genie affirms an action or agrees with a question
	OK	Used by the Genie as a positive general response
	No	The Genie responds negatively to a question
	Don't know	The Genie does not know the answer to a question
Feedback	Well Done	The Genie praises the participant
	Wrong	Tell the participant that they made a mistake

Table 5.2 Wizard of Oz Commands and Explanations.

Of the 9 participants chosen for the initial investigation 4 were female (S1, S2, S5, S7) and 5 male (S3, S4, S6, S8, S9). Of this group 7 were post-graduate students but 2 participants (S8, S9) were 15 year-olds studying for their GCSEs at local schools. Note that access to the target group of users was limited at this stage due to the small number of students attending CHALCS and, in this design phase, it was important not to use students who would be required later to evaluate the implemented system. The inclusion of two 15 year-old students enabled the researcher to provide some check on any difference in acceptance between post-graduates and younger students. At the same time it was felt that the older students would provide a critical audience who would be able to articulate design suggestions. No previous knowledge of Physics was assumed. Instead it was emphasised that the summary note-taking task was the central focus of the investigation and not Physics competence. Data collected for the purposes of the investigation included a log of the 'conversation' between the Wizard and the participant, the summary document, observational notes and a post-summary writing semi-structured interview. Table 5.10 summarises these sources, the rationale for their collection, the analysis carried out and the results.

5.2.4. Procedure.

Two computers were used connected by a Local Area Network, although, theoretically, any two computers connected to the Internet could be used as the software utilises the User Datagram Protocol (UDP). The utilisation of this protocol means that the first task to undertake is to determine the IP address of the computer (whether running the client or server) and enter this into the corresponding dialog box available on start-up together with the ports for transmission and receiving. In addition, the server also requires the user to enter a filename for the communication log upon start-up. The server is used by the Wizard, as described above, whilst the client is used by the participant (see Figure 5.3). As well as the computers requiring the client/server software the target participant's machine must also have all of the necessary Microsoft Agent components installed. These can be found on the MS Agent homepage (Microsoft, 2003).

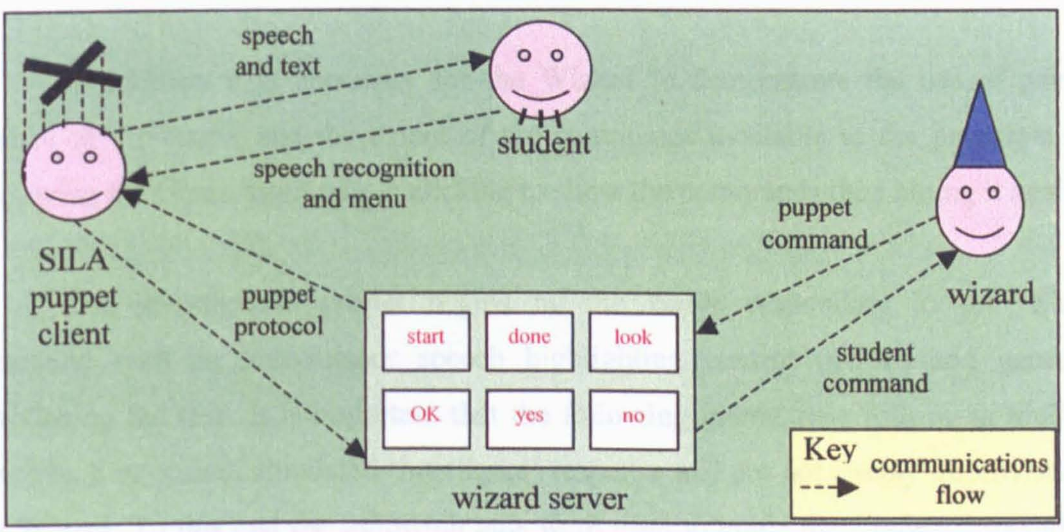


Figure 5.3 Wizard-Puppet Client-Server.

Once the software is installed and initialised it is necessary to set up MS Netmeeting. The Wizard's computer should place a call to the participant's computer using the data only option so that the speech output on the participants' computer is not overridden. Once a connection is established then the Wizard should enable sharing of

MS Word and Internet Explorer then point the browser to the Astronomy and Optics course.

With the environment in place on each participant's computer the instructions shown in Figure 5.4 were read to the participant.

> "Thanks for agreeing to help me in my study. The object of the study is to test the design of an animated Genie 'agent'. I will ask you to perform a task consisting of summarising a section of some on-line course notes (**Section 1.1 of 'Lenses'**) using MS Word. The Genie will attempt to help you by offering advice and demonstrating techniques for effective summarisation. You can communicate with the Genie by pressing the right mouse button over him to obtain a menu or you can speak the menu commands into the microphone. To gain control of the cursor you will need to 'request control' from the 'control' menu at the top of the window.
>
> On completion of the summarisation task I will ask you a few simple questions relating to your experiences. The whole study should last no more than 30 minutes."

Figure 5.4 Instructions for the Wizard of Oz Study.

In addition it is necessary for the Wizard to demonstrate the use of gaining control of the cursor and the extent of the commands available to the participant by displaying the Genie, right mouse clicking to show the commands then hiding it again.

The investigation proper begins by the Genie responding to the 'Show' command with an introductory speech highlighting control options and generally introducing the task. It is important that the following interactions follow, as much as possible, a structured simulated 'intelligent' response and are not simply improvised by the Wizard. To this end the following aids were used to guide the Genie's responses to participants actions:

1. An anticipated Agent Architecture;
2. An Instantiated 'Lens' schema;
3. A Rhetorical Predicate Analysis of 'What Is A Lens ?' See Appendix H;

4. A Theoretical Summarisation Process Model (Tawalbeh, 1994). See Appendix C;
5. An Empirical Summarisation Process Model (Barker & Pilkington, 2000). See Figure 4.4 of Chapter 4;
6. An elementary Dialogue Model (see Figure 6.9 of Chapter 6).

The personification philosophy and associated Genie actions are described in Chapter 6. Suffice it to say that the Genie was portrayed as a dominant, over-bearing helper who would at times interject when maybe not required to do so by the participant. The instructional philosophy was that the Genie demonstrates how to create a good summary rather than offer peripheral advice such as answering specific Physics questions. Additionally, the summary under construction is as much a focus of the interaction as the Genie itself. These philosophies as well as the above aids combine to guide the Wizard in the choice of responses to the participants actions.

On completion of the task the Wizard then interviewed the participants to obtain further insights into their impressions of the Genie's various aspects. Questions asked of the participants were :

1. What impressions (if any) did you have of the Genie's personality?
2. How did you feel when using the Genie ? Did you find it frustrating or helpful?
3. Were you able to express yourself sufficiently using the possible dialogue moves? What additional moves would you like to see implemented?
4. Did you find the Genie's animated behaviour over-the-top or was it entertaining and motivating?
5. Did you find the Genie's language appropriate (particularly for younger students)?
6. What did you think of the speech output and input?
7. How do you rate the advice the Genie gave on constructing your summary?
8. How else would you improve the Genie?

9. Did you have any problems with the Wizard of Oz study? How would you improve it?
10. Any other comments?

The resulting discussion was recorded and later transcribed. This transcription together with the Wizard-participant communication log, the observational notes and the resultant summary document were then used for the purposes of analysis.

5.2.5. Analysis.

The analysis can be further divided into six categories of interest notably strategies pertaining to the Genie's behaviour, User Interface/Personifications, the Dialogue Model, the Wizard of Oz study, Observed Summarisation strategies and more general observations. These categories emerged as the various data were examined for points of interest. Each of these categories shall now be examined in turn. This analysis again follows the methodology described in Chapter 1. A subsequent section (5.2.6) summarises this analysis. Finally section 5.2.7 draws key conclusions from the triangulation of these data which are summarised in Table 5.5.

5.2.5.1. Strategies of Genie's Behaviour.

The first category, Strategies for the Genie's behaviour, is also closely related to other categories such as the Dialogue Model and User Interface/Personifications. One of the most interesting observations was the participant's perceptions that the behaviour of the Genie was adaptive. That is, the 'domineering' or critical and 'autocratic' behaviour is prevalent during initial interactions but gives way to a more 'friendly', less interventionist behaviour towards the end of the task. Whether this phenomenon is actually taking place (that is, the Wizard is simulating this behaviour either consciously or subconsciously) or if it is simply a perception would require further study. However, it does have implications for the design of the Agent. As S4 pointed out: "At the beginning I don't think it's helpful, I got the feeling he wants to critique me or something like that, he wants to be cleverer than me…at last…I think he's helpful". S3 further said that this adaptive capability is a necessity: "when you get used to the summarising, you just need help, quick help or something, at the beginning yeah, it was good, after a while I prefer as a user to have more control of the process or the task." A long term goal may be to include some sort of adaptation of this kind although this will be outside the scope of this current project.

Another major observation brought up by a number of participants was that it would be useful for the Genie to offer hints or further specific help on the art of summarisation. For example S6 said: "It just jumped in and did it." This is an observation of the Genie's pedagogical philosophy of 'teaching by demonstration' rather than holding a conversation about summary construction. Obviously, different people have different ways of working but it is apparent that hints, for example on the use of bullet points and tables (as evidenced by S1's use for comparisons) and other effective summarisation techniques would explicitly help some participants. However, for some participants the philosophy is pertinent as evidenced by S2's comment: "I think what he does here is quite good really, he does structure it for you cause I'm crap at structuring stuff so it would be quite helpful for me to have somebody doing that". It would also be necessary to ensure that help is given only when deemed essential. S8 believes that the Genie, "only gives you help when you need it not all the time." It would be necessary to ensure that this faith in the level of intervention is maintained in future implementations.

It was mentioned by one participant, S5, that praise should not be used unnecessarily, they felt that sometimes the Genie could be "patronising". However, praise when used appropriately, i.e. when a proposition prominent in the rhetorical predicate hierarchy is selected, could result in much needed positive reinforcement. A further unwanted behaviour was that of the Genie interrupting unnecessarily. As S1 said: "Sometimes when you're in the midst of thinking then he will come in, you know, you feel 'where was I' then I have to go request control again...cause sometimes you might require a bit of thinking." This is mostly due to the 'domineering' personality of the Genie which will be discussed next. However, it is also related to the turn-taking mechanisms as controlled by the Dialogue Model. A mechanism would be required to ensure that the Genie is not interjecting unnecessarily and detrimentally, for example if the student it typing.

Further miscellaneous strategies that the participants would have liked to have seen in the Genie include was an ability to *justify* actions. This would also be reflected in the Genie's strategies for responding to user actions, such as a 'justify' dialogue request following the participant (such as S1) inserting a table, for instance. Another useful action which the Genie could perform would be to follow hyperlinks such as glossary definitions or external URL's such as looking up "photocopier" in an on-line encyclopaedia. As most participants pre-read the course notes before commencing their note-taking it would be prudent for the Genie to wait before performing any action. Finally, it became apparent that participants would benefit from a simple demonstration of moving and cutting and pasting text from the course notes to the summary document.

5.2.5.2. Personifications.

The second category of analysis concerns the User Interface and associated Personifications of the Genie. As already mentioned and as intended participants saw the Genie as having a 'domineering' personality. This was a main intention in the Genie's design (as described in Chapter 6) including choice of character animations, appropriate language and intervention strategies. However, different participants responded in different ways to this personality. By far the strongest dissent was evidenced by S2 who said: " I tend to react badly to domineering. [laugh] I feel like hitting him around the face. [laugh] So I'd go for one of your other characters normally I think." And goes on to say "That doesn't give you any power at that point, instead of saying 'do you need some help' or something like that he's saying 'I want to take over.' Sod you ! Whether you want it or not !" This participant was in fact vehemently opposed to the Genie's characteristics recounting an incident of encountering such a person in the past with contempt. This same participant also described the lighting of the face and "muscley" nature of the character reinforcing their conception of the "domineering" personality. On the other hand S7 did not interpret any damaging personality traits emanating from the Genie: "I thought that it was with a positive attitude towards me...so it wasn't meant to make me feel stupid". S1 seemed to completely anthropomorphisise the Genie saying that they "saw him more as a person" and "friendly and encouraging".

Most participants were happy with the Genie's language, S7 saying "it simulates sort of community language when people in a group debate or something like that, they use these sort of phrases yeah... to get familiar so if you want to be one of the group then you should follow the language." However, one participant (S6) went to great lengths to posit a more conversational style of interaction. For example they said: " If it was to say to you 'I've looked at your summary, you seem to have missed some points, would you like to put them in ?' and then if you go back and say 'well I don't know how' it could say 'would you like me to put in one for you ?' and then it could add in something and it still leaves some further points, it could say 'would you like to put in the other points ?' This is returning to the earlier mentioned contrast made between a conversational style of interface versus a 'teaching by demonstrating' philosophy. It is, however, worth bearing in mind that not all students will find the Genie's style of teaching complimentary with their preferred learning style. In addition most of the participants found the speech output acceptable with the exception of S8, one of the younger participants who commented: " he didn't sound very much like a Genie but.... I think it sort of lacks a bit of expression in the voice but it's alright." Although great care was taken to appropriately manipulate the Genie's speech intonation and inflections it seems that for this participant there is still room for improvement.

Another major aspect of the User Interface of the Genie are the animation sequences as described in detail in Chapter 6. These met with a mixed reception. One of the younger participants, S8, said of the animation that it was "quite suitable for the age range probably, not too disney-ish." Unfortunately, for some participants the animation had a negative effect on their summarisation experience. S5 said: "I think it distracted me a bit, moving around so much and being noisy whilst moving around... I didn't really know if I was looking at the text or the Genie moving around." This is an important criticism. It seems that although some participants will find that the animation supplements the perceived 'affect' to some participants animation sequences were perceived as over-the-top and had an unintended negative effect. It may be that an answer lies in an adaptive behaviour as suggested by S3 who thinks that the animation is appropriate at first but could be faster and later wants to concentrate on the task, not get distracted by unnecessary movement of the Genie.

The speech input did not prove at all popular with the participants. Some cited problems with their accent (e.g. S1) which was liable to confuse the software whereas other participants simply said that they were familiar with menus and so preferred to right mouse click over the Genie and use this option. For example S2 said "you do that because its visual". It was also pointed out by S3 that by grouping the commands in the menu they could be more easily understood. Finally, one of the younger participants, S9, seemed to enjoy the experience of working with the Genie. They said it was "funny", "well animated" and "very good".

5.2.5.3. Dialogue Models.

The third category Dialogue Models is concerned with the model of permissible moves allowed for the Wizard. A tentative model had been developed based on the experiences of Maulsby et al. (1993) together with the designer's insights following the design framework as mentioned by Woofit et al., (1997). One of the main outcomes was that three stages seem to naturally emerge from the model as observed in the Wizard-participant logs. To begin with the Genie and participant first negotiate who is going to commence the actual summarisation. As S7 commented, if the participant is unable to start summarising because they are perhaps unsure of their first action, the Genie can lead by example: "I didn't know how to start and the Genie started and I said 'OK, stop now' because I want to work." The second stage is the focus of the interaction as here turn-taking is taken care of as the summary construction takes place. As S7 again pointed out: "I like the way of saying 'I'll take over' or 'it's my turn' which is well, stay away until I do something, and it gave me a proper hint that the genie was thinking about something so it wasn't necessary for me to type at the moment so I could concentrate on reading or re-reading paragraphs." This stage is obviously central to the exchanges and seems to have been accepted by most of the participants as natural and simple. The final stage involves the negotiation of the end of the summary interactions. Either the Genie or the participant decide that they have completed the summary or that the other can finish. Sometimes, however, it is not always so straight forward as one participant in fact invited the Genie to complete the summary then as the Genie was doing so, interrupted it to take another turn themselves. Table 5.3 shows a log of typical Wizard-participant interactions for participant S3 as created by the Genie server. It illustrates the processes just described. Moves 2 to 4 show the commencement negotiation, moves 5 to 36 show the middle turn-taking phase (including several unsuccessful attempts to end the process) and moves 37 to 41 show the final end phase.

Move#	Participant	Move	Move#	Participant	Move
1.	genie :	show	22.	genie :	finish
2.	genie :	start	23.	student:	Me next
3.	student:	Start	24.	genie :	OK
4.	genie :	OK	25.	student:	You next
5.	student:	You next	26.	genie :	OK
6.	genie :	OK	27.	student:	Me next
7.	student:	Stop	28.	genie :	OK
8.	genie :	OK	29.	genie :	you go
9.	genie :	you go	30.	student:	You next
10.	genie :	look	31.	genie :	finish
11.	genie :	OK	32.	genie :	why
12.	student:	Start	33.	student:	You next
13.	genie :	you go	34.	genie :	OK
14.	student:	You next	35.	student:	Me next
15.	genie :	OK	36.	genie :	OK
16.	student:	Stop	37.	genie :	finish
17.	genie :	OK	38.	student:	Me next
18.	genie :	you go	39.	genie :	no
19.	genie :	well done	40.	genie :	done
20.	student:	Finish	41.	genie :	hide
21.	genie :	OK			

Table 5.3 Wizard-Participant Log of Interactions for S3.

Aside from these three phases of negotiation that typify the Wizard-participant interactions there were other kinds of interaction that appeared in the dialogue logs. One potentially important type of interaction was enquiry-explanatory in nature. It was noticed for example that the insertion of a table into the summary by one participant, S1, resulted in a 'why' move from the Wizard leading the former to 'justify' as mentioned above. Such explanatory phases (as discussed in Chapter 2) are likely to be important in that they indicate or can prompt a deeper level of engagement and can lead to improvements in reasoning or the final quality of a collaborative product.

Some moves by the Genie were simply ignored by the participant. There is an inherent reliance upon the participant to take part in the discussion and respond accordingly to the Genie's speech acts and actions. An uncooperative participant could easily greatly reduce the effectiveness of the Genie's contribution. This uncooperative stance was particularly prevalent in the logs of the younger participants (S8 , S9) who would continue with the summarisation process and ignore the Genie. Also, it wasn't always obvious which was the best move to use to convey the communicative intention.

Initially there did not exist a 'well done' move – something which the Wizard missed - but one was soon added in order to praise the participant appropriately. Students used some moves inappropriately for example S7 at one point used 'Don't know' instead of 'You go' and found themselves using 'Finish' instead of a more appropriate 'Say: what shall I do next?'

Finally, this last point leads onto a fundamental problem with the use of a restricted set of permissible moves in the Dialogue Model as opposed to a more open 'Natural Language' style of interface. For example, S7 mentioned that "at a certain point I wanted to say something and I wanted to mention something...I was looking for a move and it wasn't there or something like that then I forgot it". Obviously, not every possible move can be foreseen but if appropriate use is made of the 'say' move then not knowing which move to use for such instances can hopefully be circumvented. Having what goes into 'say' properly understood and responded to by the Genie is a different matter and a major Artificial Intelligence undertaking. Similarly, by perhaps replacing the moves with MS Netmeeting voice communications between the Wizard and participant, a Natural Language interface could be simulated. However, it is outside the scope of this study to implement a full Natural Language interface implementation for the Genie.

5.2.5.4. Effectiveness of the Wizard of Oz Approach.

The fourth category of analysis concerns the nature of the Wizard of Oz study itself. One of the main outcomes affecting the study resulted from the difficulty of performing the role of Wizard. The Dialogue Model, which took the form of a graphical network, is extremely difficult to follow, particularly when the Wizard has a number of other documents to be working from, as detailed above. Attempting to take these other factors into consideration e.g. looking at the rhetorical predicate hierarchy as well as the theoretical summarisation process model required too much cognitive effort resulting in the Wizard delaying their decision on the next best move. As S6 pointed out: " this is the problem with the wizard of oz thing that whilst you were busy doing it I didn't know whether it was still active or not, perhaps if he was to do something that's very recognisable when he's finished." As suggested one solution to this problem would be to initiate some kind of key repetitive animation in the Genie whilst the Wizard is considering their next action. One action easy to implement would be the 'processing' sequence which when executed results in the Genie spinning around on a fixed screen location. This would then communicate to the participant, given a suitable brief, that they must wait a short while for a response.

A further key outcome of using the Wizard of Oz approach became apparent during the post-test interviews. Of the total number of participants only S6 seemed to be aware that it was in fact a simulated intelligence which drove the Genie. Even this participant only realised this by physically turning around to see the Wizard operating another computer. Most people seemed willing to attribute the intelligence of the Genie to the software itself and were willing to attribute it with independent actions as already mentioned in terms of anthropomorphisations. S5, an Artificial Intelligence researcher, even went as far as to say: " I don't know how the knowledge base is built, maybe you could improve that a bit more", assuming that the intelligent 'knowledge base' had been coded. Some of these beliefs were further reinforced by participant's 'testing' of the Genie, for example by inserting a key word such as 'brain' into the summary text in order to see if the Genie would elaborate upon it. It is worth noting that for the majority of the studies the Wizard operated the server from behind the participant so that they were out of view with the exception of one case where the Wizard was sat in another room entirely. By holding the study in separate rooms the natural tendency for participant's to 'talk aloud' was suppressed. It would make for an interesting study to have the participants utilise a 'think aloud' protocol as they interact with the Genie. This was not, however, attempted here. Participants were requested to restrict communication to the dialogue moves in order to help the Wizard focus on the task which was quite demanding.

A number of additional outcomes of the Wizard of Oz include:

1. the necessity of precise task instructions (for example, explaining the interaction mechanisms clearly);
2. the delay and 'blurring' effects due to slow communication between client and server in MS Netmeeting causing frustration and confusion;
3. the desire to have Genie user commands visible at all times;
4. the elimination of MS Word auto-corrections;

5. a non-distracting desktop.

It was also apparent in some circumstances that, as the cursor became used for 'pointing' purposes, it's appropriation by the Wizard to operate the button console was also frustrating and disorientating.

5.2.5.5. Summarisation Strategies.

The fifth analysis category relates to Summarisation Strategies witnessed during the study. It was noted that seven of the participants did pre-read the Astronomy and Optics course notes before commencing the summarisation with two (S5, S9) launching straight into construction of their summary document. This is mainly in-keeping with the Summarisation Process Model from Chapter 4, Figure 4.4 (Barker & Pilkington, 2000) which forms the basis of the Genie's stages of intervention. There was also a mixture of editing techniques used by the participants ranging from cutting and pasting (e.g. S9) to typing whole sentences (e.g. S1) and all points along this range with the addition of the one participant (S6) pasting the whole text to MS Word then using Word's auto-summarise facility. Most participants preferred precise and terse points to represent the key points in the text but only one participant gave their summary document a title. It was noted that one participant (S8) seemed to be paraphrasing the course notes rather than extracting important points from them and condensing the knowledge contained in them. This raises the question of deep versus shallow processing. We are obviously trying to encourage understanding in the summarisation task so this kind of technique should be discouraged by any kind of support, e.g. the Genie. S8 said of the Genie's own summary points that they were "to the point, concise."

By way of illustration an example summary from S9 is shown in Figure 5.5. As can be seen the summary is precise yet contains enough detail to capture the gist of the course notes. In addition, bullet points are used to great effect to delineate statements. It is worth noting that the participant was one of the younger participants (age 15) and that English is not their first language.

- Lenses focus the light waves.
- The shape, material and the size of a lens controls the image.
- The eye is a lens.
- An object point focuses to an image point in the retina and all this points are interpreted by the brain. The eye has some differences and similarities to a camera, e.g. they both use lenses, the eye uses the brain to interpret the image.
- A whole array of cylindrical lenses are responsible for imaging a portion of text/picture in a photocopier.

Figure 5.5 Example Summary Document (S9).

5.2.5.6. Miscellaneous Outcomes.

The final category of analysis contains a number of observations which do not easily fit into any of the previous categories. Firstly, as has been noted the cursor became a valuable tool for both participant and Wizard to 'point' to referred text. S9 in fact used the cursor to show the Wizard which text they were reading at any given time. An interesting observation arose as one participant (S4) asked a question of the Genie/Wizard by placing it in the actual Word summary document. In the case of this study the Genie was able to correctly interpret and answer the questions although this will not be the case in the final software implementation. It does however seem to demonstrate to the participant that the environment is fully integrated as opposed to harbouring separate and distinct functionalities as they perceive that the Genie will understand their text in the main document window. There was a problem with the participants not being totally au fait with the WebCT environment as they did not realise that the hyperlinks within the course text would result in a glossary entry appearing upon selection. There was a correspondingly similar problem with the 'URL's' links to external resources such as on-line encyclopaedia as at one point a competent student would know to look up "photocopier" in order to obtain definitions, etc. Simple start-up instructions by the Genie would have made these options more explicit.

As already mentioned, S6 used the auto-summarise capability of Word to construct a summary. However, the Wizard was not entirely happy with the result although the participant seemed to be. The Wizard went on to make a number of small corrections but the finished document was far from the best summary produced. This is an interesting aspect of summarisation aids as automatic summarisation (as reviewed in Chapter 2) is a continually developing field. This result shows, however, that it still has a long way to go before it can match its human counterpart. It is hoped that the collaborative nature of the approach espoused here will help eliminate some of these concerns. It would be prudent to deny participants the auto-summarise facility in future studies. A final observation concerns the supplementary text boxes accompanying three of the possible dialogue moves. It became obvious at an early stage that they needed to appear on top of all other windows on the participant's screen otherwise they would go unnoticed and hold up the investigation. This was simply achieved by giving the boxes priority in the Windows Z-order.

It is now necessary to distil the concrete results of the above analysis before arriving at conclusions and recommendations.

5.2.6. Summary of Results.

Table 5.4 summarises the four data sources used to arrive at the results and recommendations. These main findings are summarised and explored in this section.

DATA SOURCE	RATIONALE	ANALYSIS TECHNIQUE
1. SUMMARY DOCUMENT	to record final summaries to help ascertain strategies and quality, etc.	look for points of interest e.g use of titles, bullet points, etc.
2. LOGS	to provide a record of agent-participant interactions and hence help to examine the dialog model's effectiveness, etc.	look for phrases and/or patterns of interactions, common inter-subject strategies, interesting emerging strategies
3. OBSERVATIONAL NOTES	to record any misc. observations not covered by other data such as summarisation strategies, participant utterances, problems with WoZ etc.	look for common inter-subject summarisation strategies and/or more general experiences such as WoZ-related problems
4. INTERVIEW	to elicit participants impressions of using genie, feelings, appropriateness of dialog model, genie language and animation appropriateness, speech input/output, genie improvements, problems with WoZ, etc.	extract typical quotes pertaining to the rationale (i.e. questions) + any other interesting participant observations/ insights

Table 5.4 WoZ Data Sources and Corresponding Analysis.

The perception of Genie adaptive behaviour by participants is one of the most interesting results. It has a bearing on the implementation of the software agent as it is possibly a projection of desired behaviour on the part of the participants whether it actually took place or not. It would require further studies and corresponding analysis. In addition hints and/or help are also of great importance to the participants, particularly in the area of effective summarisation techniques but there was also evidence that students did not want the Genie to just take over and do too much of the work. The final result relating to strategies is the need for the Genie to make relevant, necessary and timely responses, for example to praise, justify and wait before commencing any interaction as the participants pre-read the notes.

Of those results relating to the User Interface it became evident that participant's responses to the "domineering" attitude varied. It can only be hypothesised that this was due to past experiences, their own personality and preferred learning styles. It would be interesting to test these kinds of hypotheses with a further study. Most participants thought that the language and speech output were generally acceptable although, if time allows, more work could be carried out on speech intonation. Whereas most participants thought that the animation was acceptable others would have preferred less as they became more involved in the task. Finally, participants could have benefited from a separate list of voice commands which may have encouraged them to utilise this aspect of the Genie.

As for the Dialogue Model it became apparent that a three phase model was being followed on each interaction consisting of start negotiation, middle turn-taking and end negotiation. Additions to the model which emerged from the study include a 'why', 'justify' move and praise as 'well done'. (These moves were added before the next evalaution cycle). It was noted that the Genie relied on co-operation of the participants to maintain the 'flow' of conversation. This cooperation could be encouraged through the use of motivating, believable personalities (Lester et al., 1997). Finally, it became apparent that in some circumstances there was a little confusion over which dialogue move was the best suited to the intended meaning.

In terms of the Wizard of Oz study itself one of the most damaging factors to the effectiveness of the study were caused by the great cognitive load placed on the Wizard in examining multiple documents to select a move. This resulted in the participant left wondering what was happening or (in one case) actually realising that they were interacting with a human. However, in the main, the illusion created in the minds of most participant's was that they were in fact interacting with an 'intelligent' Agent.

The Summarisation Process Model seemed to be justified by this study as people mostly seemed to follow it's three stages although less emphasis was placed on the final Reflective Review phase. This could, however, have been due to the fact that the task only required a small summary to be produced which could be held in memory. It was also apparent that participants used a range of editing techniques to produce their summary.

Finally, an interesting outcome of the study was that participants seemed to perceive the various tools as part of an integrated environment. This hints at the success of the overall Wizard of Oz experiment and would indicate that, provided communication between the tools worked adequately, the final system should also be perceived as an integrated environment.

5.2.7. Conclusions and Future Work.

The discussions in the previous section which relate to the design of a potential agent can now be summarised in terms of key conclusions together with the corresponding data source from which they are derived, see Table 5.5. The next section then presents these as recommendations for future work.

The Wizard of Oz study described above exposed insights into the development of an Agent-based support for summarisation before undertaking major coding of an 'intelligent' component. A number of areas have been addressed, including: strategies pertaining to the Genie's behaviour, User Interface/Personifications, the Dialogue Model, the Wizard of Oz study, Observed Summarisation strategies and more general observations. In addition to the recommendations made for future Agent development the transferable technologies required to implement the User Interface have been tried and tested, ready to be integrated into future components.

Future work (described in Chapter 6) develops these components, building on the implementation considerations of this study's outcomes. As intended, the outcome will be the user-centred development of an Intelligent Agent to aid Summarisation in the WebCT Astronomy and Optics course.

CONCLUSION	DATA SOURCE	DATA EXAMPLE
1. adaptivity – interactional and propositional (Cassell, 2001)	4	S3: "at the beginning, yeah, it was good, after a while I prefer as a user to have more control of the process or task"
2. empathically praise, justify, wait	2 & 3	S5: unnecessary praise may be "patronising"
3. improve dialog model e.g. 3 phases + differentiate moves clearly	4 & 2	3 phases as illustrated in Table 5.3
4. incorporate summarisation process model (Fig.4.4) + rhetorical analysis, proposed agent architecture, Tawalbeh process model	3	7 of 11 participants pre-read the course notes
5. increase chances of human-agent cooperation by motivating with appropriate agent personality	4 & 3	S7: "...a positive attitude towards me...so it wasn't meant to make me feel stupid"
6. support a range of text editing techniques	3	different styles: cut and paste (S9) vs. typing whole sentence (S1)
7. include help and hints on summarisation	1 & 4	S6: "it just jumped in and did it"
8. improve speech input/output e.g. intonation	4	S8: "it sort of lacks a bit of expression in the voice"
9. a range of personalities may be required e.g. dominant *and* submissive	4	S2: "I feel like hitting him around the face" vs. S1: "friendly and encouraging"
10. employ a 'processing' animation if agent is 'thinking'	4	S6: "whilst you were busy doing it I didn't know whether it was still active or not"

Table 5.5 WoZ Key Conclusions Summary.
(Key to Data Sources: 1 = "Summary Document", 2 = "Logs", 3 = "Observational Notes", 4 = "Interview", see Table 5.4).

5.2.8. Recommendations.

The following ten points summarise the key aspects of future software Agent implementation which have emerged from this study.

1. Characters should adapt to the student e.g. they could display more help and/or animation at first then automatically gradually decrease (or possibly respond to a student request) or another alternative is for the student to choose from a number of characters which display different properties e.g. 'passive' or weak versus 'domineering' or strong characters. Selection of these may then vary with the student's style and experience. This is felt to be largely beyond the scope of the initial prototype to be built here but is considered important for future developments in the field and will be discussed further in Chapter 9.
2. Praise, justification and 'patience' could be more empathically implemented along the lines discussed in section 5.1.2.
3. The Dialogue Model could be improved to include a 'why/justify' move and the three phase start, middle and end negotiation model. This can be taken into account in the current implementation cycle.
4. Incorporate the Summarisation Process Model (as well as the other documents e.g. Rhetorical Predicate Analysis, Theoretical Summarisation Process Model) to guide the selection and condensation of sentences on the Genie's turns and enable explanations of choices.
5. A student can be encouraged to collaborate with the agent by employing a motivating and appropriate 'personality'.
6. Support a range of text editing techniques (and offer appropriate assistance).
7. Include hints/help on effective summarisation based on Tawalbeh (1994) and Meyer (1975), etc. and WebCT/MS Word use, e.g. the use of :
    a. Titles;
    b. Precise statements;
    c. Bullet points;
    d. The 'gist' of a text (as described in Chapter 2);
    e. Following WebCT links (e.g. on-line encyclopaedias), navigation, tools, etc.

f. MS Word e.g. do not use auto-summarise, use bullet points, formatting, etc.

8. Speech input and output could be improved by the addition of a voice command synopsis window and better intonation design.
9. A range of personalities may be required e.g. dominant and submissive.
10. Employ a 'processing' animation whilst the Agent is thinking.

The results of the Wizard of Oz study shown here indicate an approach to software design which by its very nature is designed to increase user acceptance of the final system, the Affective Study Companion, provided that the initial users are representative of the target population in the way they interact with the system and provided that the behaviour of the wizard can be sufficiently rendered in the final implementation. The detailed design and implementation of the companion is the subject of the next Chapter.

Chapter 6: The Design and Implementation of an Affective Learning Companion.

6.1 Introduction.

The Wizard of Oz methodology has proved useful in generating a number of recommendations which otherwise may not have informed system design. In particular the method allows a software developer to become more involved with potential users to achieve a more user-centred system development.

This chapter is concerned with a discussion of the design and implementation issues relating to the development of the software, Summary Intelligent Learning Assistant (SILA). SILA is to be situated within an environment consisting of the WebCT Physics notes which were initially developed as an exemplar of on-line learning at CHALCS. As hypothesised SILA should help motivate students as they engage with the Physics materials contained within WebCT, constructing summaries collaboratively and thus helping to improve their study skills. Based on previous research in collaborative learning (reviewed in Chapter 2) and pilot studies testing the potential acceptance of the agent-based approach (reported in Chapter 5) it is hoped that this engagement can be enhanced by including an affective component within the system manifesting as an animated pedagogical Agent displaying appropriate 'Affectations'.

This chapter discusses issues relating specifically to 'design and implementation', that is, techniques and technologies which enabled the development of SILA together with a detailed description of the current system prototype. Finally, issues surrounding this implementation phase are discussed and recommendations are reached.

6.1.1. SILA Design Rationale.

In considering the design of the system it was first necessary to establish the requirements of the system based on the needs analyses and empirical work conducted at CHALCS (see chapter 3 and 4). The main outcome of this work was the development of a case for building Learning Companions to help 6th Form students (i.e. ages 16 to 18) with their note-taking within the WebCT Physics course (as part of the

Acquisition phase reported in Chapter 2) and that these Companions should employ ideas from the research on Animated Pedagogical Agents in order to test the hypothesis that Learning Companions can help motivate students and can engage learners in ways consistent with Collaborative Learning theory. Moreover, the aim was to develop and test the effectiveness of alternative artificial 'personae' (behaviours which may be perceived as consistent with alternative personalities or styles) of Learning Companions on the engagement and performance of students when summarising.

The design rationale was conducted in the form of a gIBIS diagram (Conklin & Burgess Yakemovic, 1991) – see Figure 6.1. Its purpose is to explore the issues surrounding the design of SILA, particularly evaluating the pros and cons of the approach. This process helped to externalise the main design choices before the design process itself actually took place.

6.2 The Design Methodology: Evolutionary Prototyping.

For the purposes of this research the traditional software development methodology of the 'waterfall model' is not appropriate. The waterfall model consists of a sequential progress through a number of well-defined steps including: requirements specification, architectural design, detailed design, coding and unit testing, integration and testing and operation and maintenance. However, this approach is unsuitable for the development of highly interactive systems if the design criteria are not well established during the initial stages of the software development. As Dix et al (1997) candidly point out:

> "The only way to be sure about some features of the potential design is to build them and test them out on real users. The design can then be modified to correct any false assumptions that were revealed in the testing."
>
> Dix et al (1997)

This *iterative design* process has been adopted for the development of SILA due to the highly interactive nature of its central functionality and the perceived need to involve students by adopting a user-centred approach to development.

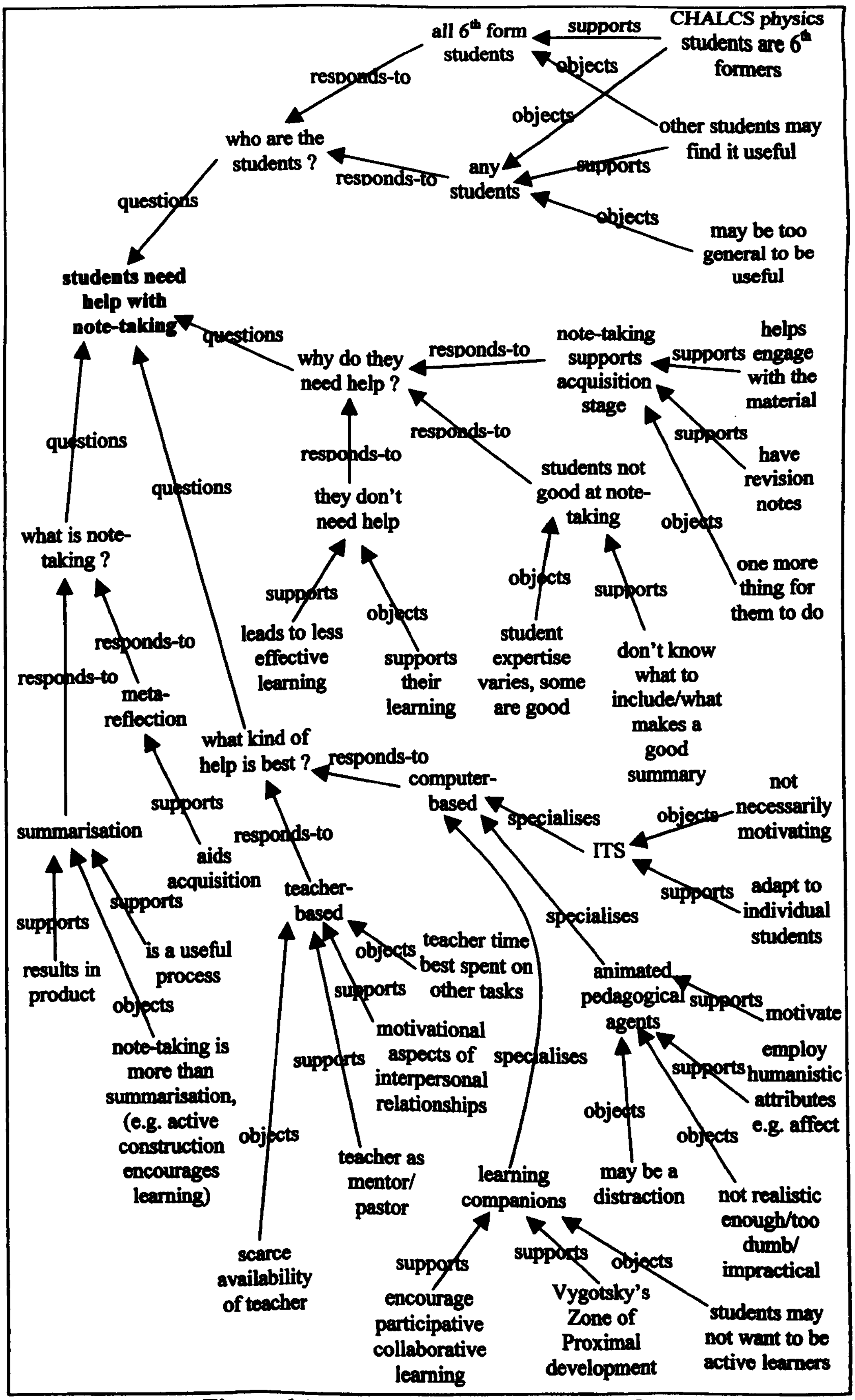


Figure 6.1. gIBIS SILA Design Rationale.

Dix et al (1997) go on to describe three main approaches to prototyping involving **throw-away** prototypes (where the iteratively developed system is discarded in favour of design requirements of the final system), **incremental** (where individual components are first delineated then gradually released to the user one at a time until the final system is arrived at) and **evolutionary** prototyping (where the system is seen to evolve from an early, less well-defined solution to a more thorough implementation as a result of subsequent user testing, i.e. the system is never discarded). There are, however, some problems inherent in this kind of rapid prototyping such as initial design decisions not being rectified in subsequent revisions or symptoms of bad design being treated in favour of re-thinks. One way to circumvent such problems (as followed in this research) is by utilising appropriate design documentation such as models of the user (e.g. the summarisation process model of Chapter 4) or modelling the dialogue of the human-computer interaction (e.g. those discussed later in this chapter in section 6.4.3.3).

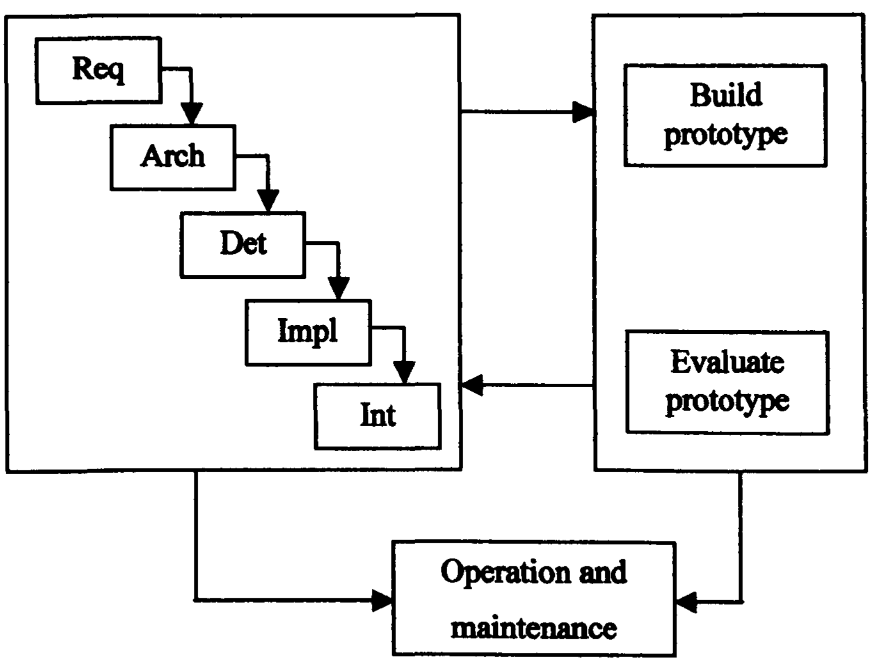


Figure 6.2 Evolutionary Prototyping (Dix et al., 1997).

Figure 6.2 provides an overview of a typical evolutionary prototyping approach to software design. This consists of the following five prototype design steps but importantly recognises that these five steps will undergo iterative build-evaluative cycles until the final software artefact is produced:

1. requirements analysis;
2. architecture design;
3. detailed specification;
4. implementation;
5. integration.

Each of the steps taken during the design and implementation of the five SILA prototypes, including documentation produced are detailed in section 6.4, 6.5 and 6.6. First though, the design development is outlined.

6.3 The Design Development.

Based on the design rationale for the system it was clear that the system would need three main components:

- Animated agents with affectations that could be varied to generate alternative 'personae' (affective behaviours perceived by students to be consistent with having a particular personality) which could then be tested for their ability to engage and motivate the student to complete the task;

- An interaction model or dialogue model through which the animated agent could engage the student in the joint collaborative task (the model would consist of dialogue mechanisms (moves) for managing the interaction with the student, including mechanisms for turn-taking and generating appropriate turns given the current state of the summary task);

- A model or mechanisms by which the system could summarise text itself within the collaborative space (joint summary under construction). This was needed not only in order to offer suggestions/participate in the summarisation process but also to comment on the suggestions of the student;

With an idea in mind of the design rationale, including the pros and cons of various decisions relating to the main issues, the next stage was to start development of the prototypes. The five staged prototypes are:

1. The Stand-Alone Prototype;
2. The Wizard of Oz Prototype;
3. The Pre-Evaluation Prototype;
4. The Evaluation Prototype;
5. The Post-Evaluation Prototype.

Each prototype will be discussed in detail in subsequent sections. For now we note that the stand-alone prototype was built to test the feasibility of technologies required to produce the animated agents and provide a basic mechanism for their manipulation in terms of the 'moves'. The Wizard of Oz prototype was next developed to test the capability of the 'pencil and paper' dialogue model when two collaborating humans, a tutor and student, worked together to produce a joint summary and the effectiveness of the animated agents in terms of the 'affectations'. This prototype was also intended to give further insights into the requirements of the summarisation model. The evaluation prototype was next developed incorporating the results of the Wizard of Oz study and implementing the dialogue and summarisation models as far as possible. Finally, the Post-Evaluation prototype made changes in line with the results of the evaluation study and to improve SILAs robustness for use with students, particularly in terms of allowing it to be deployed from the World Wide Web.

The next section describes this proof-of-concept in terms of the first four prototypes produced following the design methodology highlighted earlier in this Chapter. Chapter 7 provides a description of the final system implementation.

6.4 Incremental Development of the Summary Intelligent Learning Assistant (SILA).

This and subsequent sections describe the first four milestones eventually decided upon for the prototypes developed as part of this research including any problems encountered and the solutions arrived at to overcome them. The first three prototypes are discussed as part of this section then the prototype used in the evaluation is discussed in the next section. In-keeping with the process described in Figure 6.2 each prototype will be described in terms of design techniques relating to each of the first four prototyping steps of requirements analysis, architectural design, detailed specification and implementation. These techniques are:

1. Storyboarding – representing the 'look and feel' of the interface;
2. Data Flow Diagrams – representing a high level visualisation of the prototype architecture;
3. Pseudo Code, Hierarchical State-Transition Diagrams – representing the detailed solutions and the dialogue models, respectively;
4. Implementation Notes.

6.4.1. The Stand-Alone Prototype.

The first version of SILA was simply a stand-alone version, i.e. operating entirely on one computer. The stand-alone prototype was built to test the feasibility of MS Agent to convey 'affectations' and the suitability of Visual BASIC (VB) to manage the interaction. It consisted of the implementation of a control panel consisting of various buttons which, when selected by the User, initiated a 'move' in the MS Agent Genie character.

6.4.1.1. Requirements Analysis.

As can be seen in the storyboard (Figure 6.3) the requirements were simple in that just one agent was required which was controlled (like a puppet) from a Visual Basic form operating as a control panel. At this stage the mechanisms to jointly construct a summary were not considered, i.e. the next design cycle, the Wizard of Oz prototype would consider this.

6.4.1.2. Architectural Specification.

As can be seen in Figure 6.4 the 'puppet' console generates a command string which is passed to the main form. This form then interprets the command as 'speak' and 'play' MS Agent commands which make Genie speak text (i.e. using text-to-speech and a speech bubble) depending on the command and play an animation sequence e.g. 'announce' will make Genie look alert and raise a finger. In turn if an agent command is selected by right mouse clicking over the agent (or possibly using speech if this facility is enabled) then the command is passed back to the main form which displays it in a pop-up 'message box'. Although this architecture is quite simple it does allow some of the fundamental technologies to be tested.

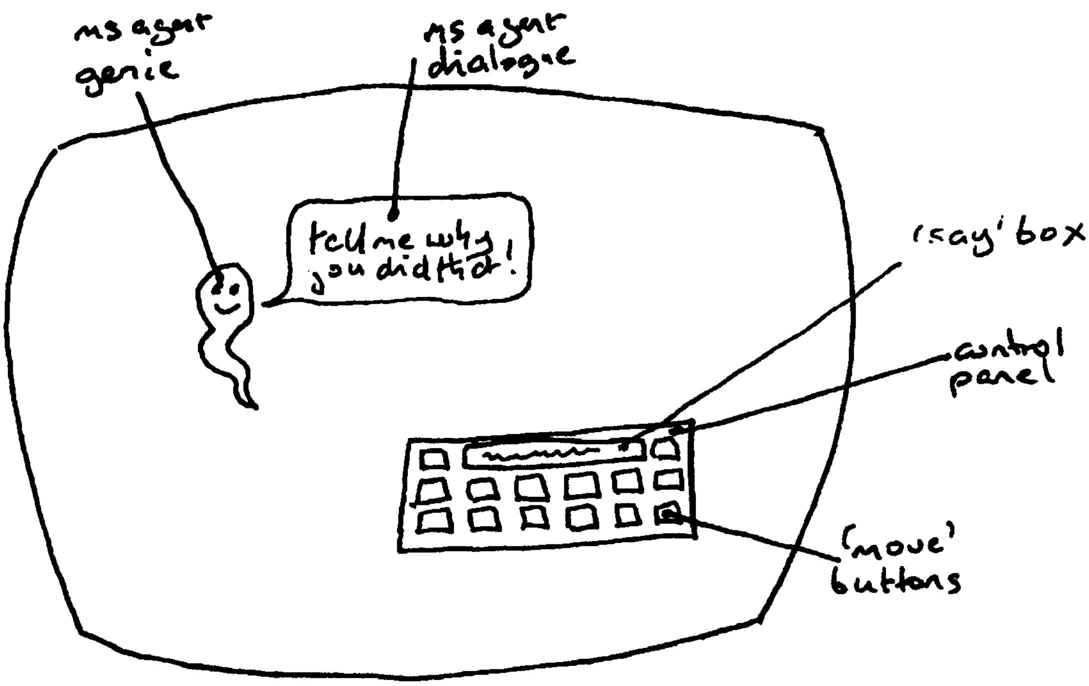


Figure 6.3. Stand-Alone Prototype Storyboard.

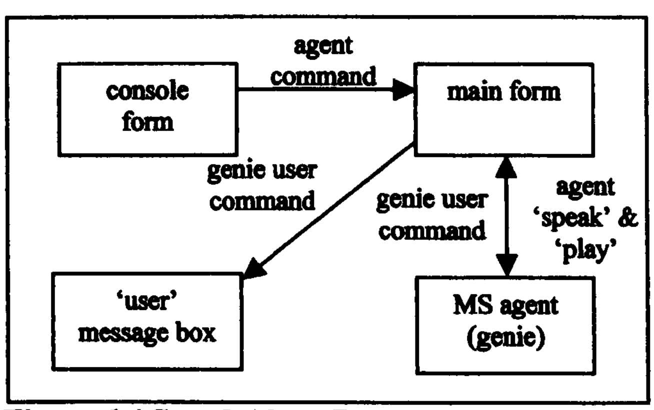


Figure 6.4 Stand-Alone Prototype Architecture.

6.4.1.3. Detailed Specification.

As can be see from the pseudo code in Figure 6.5 the stand-alone prototype code is fairly straightforward. One form interprets the button presses by converting them to a simple 'protocol' i.e. a string representing a command. This is then passed onto the main form. On loading, the main form will load the basic characters into MS Agent then add the user commands. The 'agent command' procedure then parses the 'console form' command then acts on it accordingly, i.e. shows or hides genie, says the special text or calls one of the 'moves'.

```
console form
        if button X pressed and X not = 'say' then
                command = "com" + X
                call main form ("com" + X)
        else if button X then
                call main form ("say" + textbox.text)
        end if

main form
        on form load
                load genie
                add commands to genie
        end
        proc agent_command (command)
                parse (command)
                if command = "say" then
                        genie speak text
                endif
                if command = "sho" then
                        show genie
                endif
                if commend = "end" then
                        end genie
                endif
                case command
                        "com01"
                                genie play("announce")
                                genie speak("I'll start")
                        "com02"
                                .
                                .
                                .
                case else
                        genie play("uncertain")
                        genie speak("I don't know what I'm...doing")
        end proc
```

Figure 6.5 Stand-Alone Prototype Pseudo Code.

6.4.1.4. Implementation.

The 'moves' for Genie were designed using MASH (a tool for authoring MS Agent scripts, see section 6.4.3.5) and were based on the design as reported in the previous chapter. At this stage the actual animations were sometimes, in hindsight, maybe too over-the-top and would later be refined. However, this prototype allowed the feasibility of the project to be tested, in particular the effectiveness of MS Agent to convey the 'affectations' and the suitability of Visual BASIC (VB) for the software development. This prototyping stage was successful and led to the implementation of software in anticipation of the Wizard of Oz study.

6.4.2. The Wizard of Oz Prototype.

The Wizard of Oz prototype was next developed to test the capability of the 'pencil and paper' dialogue model (shown below in Figure 6.9) when two collaborating humans, a tutor and student, worked together to produce a joint summary. This prototype was also intended to give further insights into the requirements of the summarisation model by enabling the tutor and student to share a document under construction. Essentially, it was envisaged that the Wizard of Oz (WoZ) prototype would be an extension of the stand-alone prototype employing a client-server architecture to enable the 'console' to operate the MS Agent 'puppet' remotely. This provides the opportunity for the Wizard to sit out-of-view of the participant, even in another room, thus creating the illusion of the 'puppet' behaving autonomously.

6.4.2.1. Requirements Analysis.

As can be seen from the WoZ Prototype storyboard (Figure 6.6) the WoZ participant would view WebCT and MS Word plus the Genie on one computer. The 'wizard' would share the view of WebCT and Word on another computer and would additionally have the 'console' form at their disposal plus the occasional pop-up message box informing them of participants' commands. These requirements, while progressively more complex than the Stand-Alone Prototype (for example, having to employ a network communications protocol) can be seen as a logical development of the Stand-Alone prototype requirements.

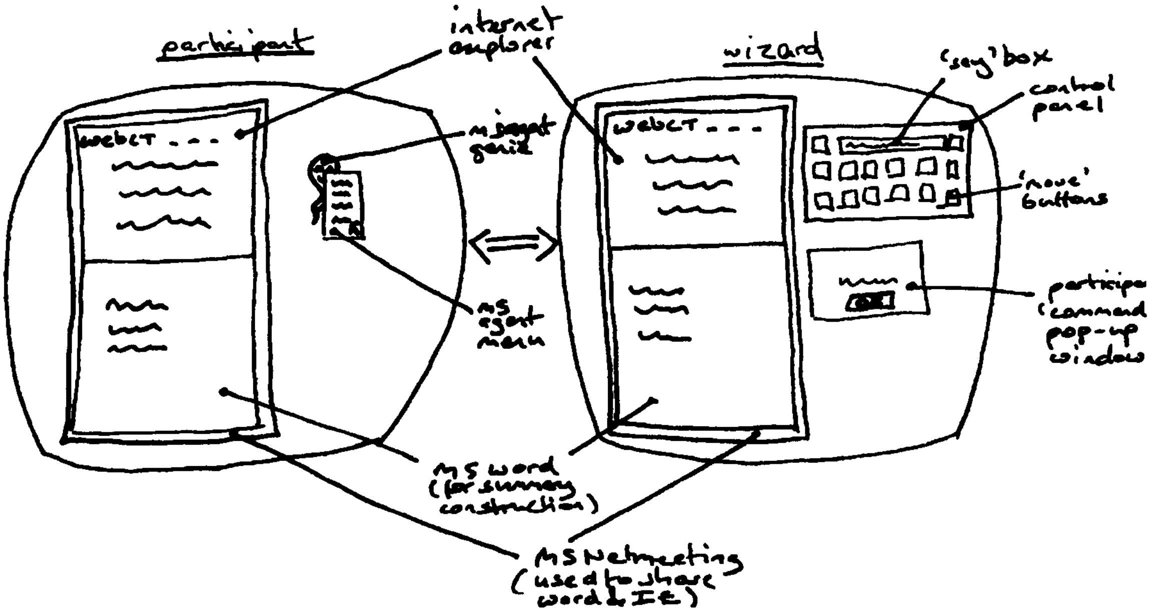


Figure 6.6. Wizard of Oz Prototype Storyboard.

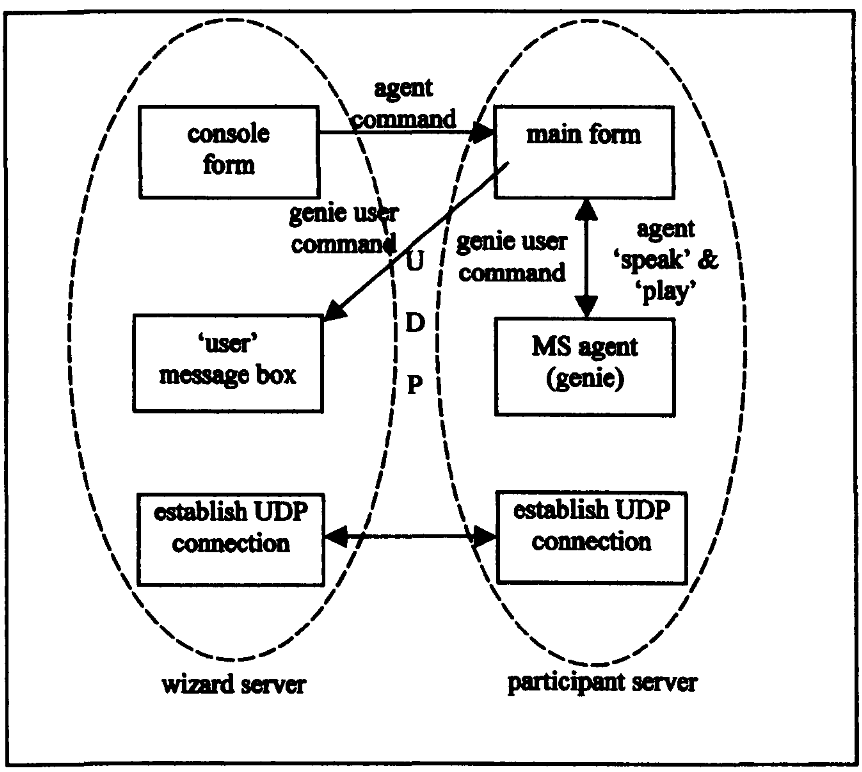


Figure 6.7 Wizard of Oz Prototype Architecture.

6.4.2.2. Architectural Specification.

Figure 6.7 shows that the architecture of the WoZ prototype (at least in terms of the Visual Basic code) is very similar to that of the Stand-Alone prototype shown in Figure 6.4. The only additional components are the simple routines to establish the

Universal Datagram Protocol (UDP) connection. Once this is established then the system operates in much the same way as described in Section 6.4.1 except that now all commands are just over the network using UDP as the 'Wizard server' and the 'participant client' are operating on remote computers.

6.4.2.3. Detailed Specification.

The Pseudo code described in section 6.4.1.3 adequately describes the specification used for the WoZ prototype. In fact, the code from the Stand-Alone version was mostly reused. Additionally though the pseudo code shown in Figure 6.8 shows the steps necessary to establish the UDP connection.

Client	Server
set IP address of server	set IP address of client
set remote host port (server)	set remote host port (client)
set local port (client)	set local port (server)
'bind' connection	'bind' connection

Figure 6.8 Wizard of Oz Pseudo Code.

6.4.2.4. The Specification of the Dialogue Model.

At this stage a 'pencil and paper' model of the SILA dialogue at a specification level was also developed, to be used by the Wizard, i.e. it is not implemented in software. It consists of State Transition Diagrams where nodes represent SILA dialogue 'moves' and links between nodes represent student 'moves'. The dialogue model is shown in Figure 6.9. It can be seen that SILA takes the initiative and starts the dialogue leaving the student with the possibility to decline this initiative, i.e. replying with 'no', or accept the initiative, replying with 'yes/ok'. The former student decision would assume that the student wants to make a contribution, i.e. choosing 'look', in which case SILA may ask for an explanation.

If the student accepts SILAs start initiative then SILA will make a contribution ('look') which gives the student the option of declining the contribution ('no') ask 'why' SILA did that (resulting in a 'because' explanation) or asking SILA to go next, etc. At this stage a SILA 'similar' move was anticipated when SILA would point out that a contribution was similar to a previous one.

Furthermore, it can be seen that a 'turn-taking' philosophy is assumed based on the cooperation of SILA and the student. This is essential to the success of SILA at this stage.

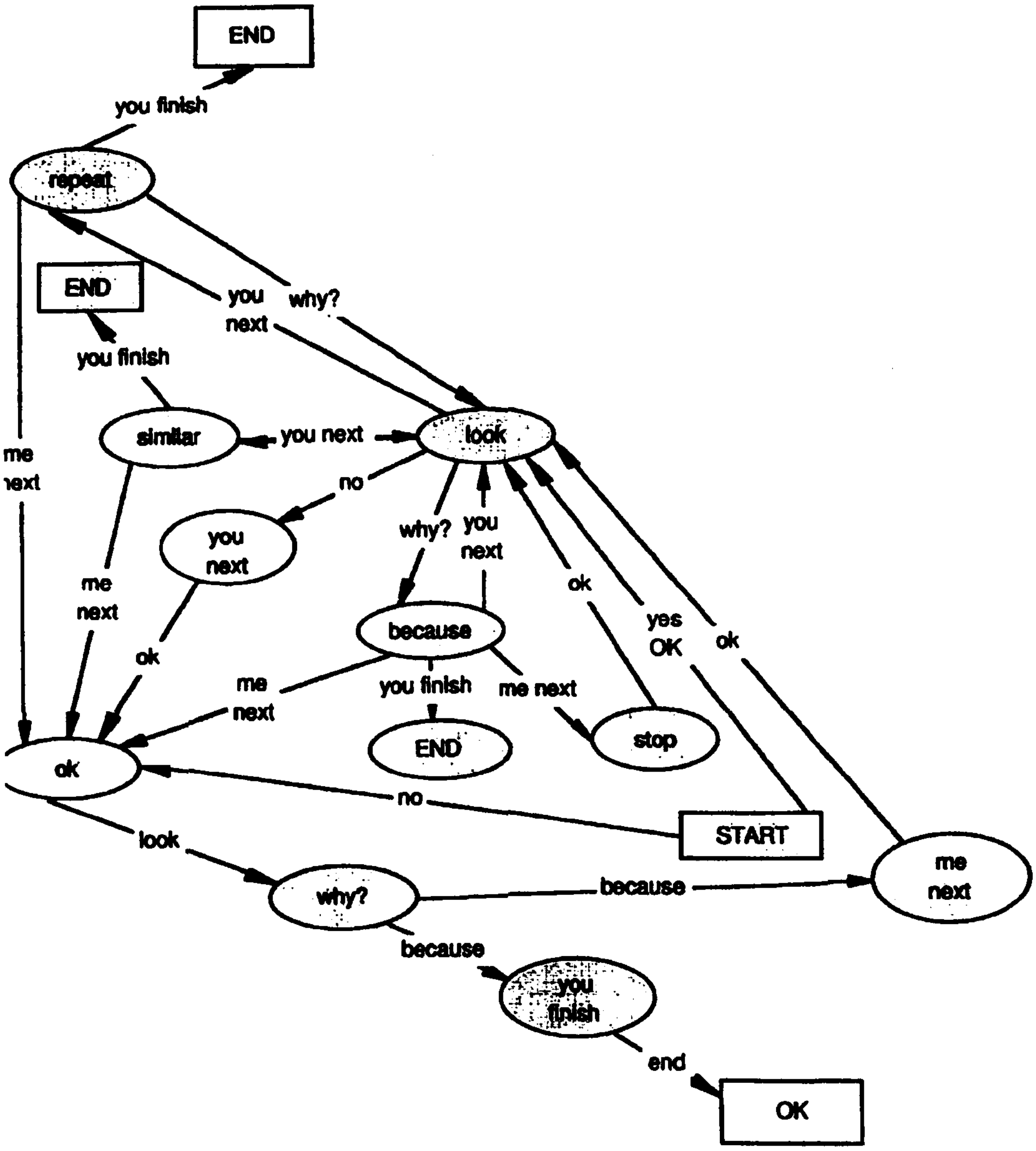


Figure 6.9 Wizard of Oz Dialogue Model.

(Key: nodes = SILA moves, links = student moves)

6.4.2.5. Implementation.

The initial thinking concerning the Wizard of Oz study was that software would have to be written which provided a collaborative text editor and web browser with facilities to remotely control the Agent from the Wizard's computer over a network. However, problems were continually encountered in producing software to enable the sharing of the summary document. These problems were mainly related to the kinds of client-server architectures supported by Visual Basic and the nature of communications allowed and the incompatibility of these with the best text editor solution technology. The eventual solution was to avoid the implementation issues surrounding sharing documents by utilising Microsoft Netmeeting to share MS Word. In addition, Netmeeting could also be used to share Internet Explorer and thus the Wizard and student could share both browser and word processor maintaining the illusion in the student's mind (as later attested to by students) that MS Agent was in control of these packages.

It thus remained to implement the remote calling of the Genie's moves from the Wizard's computer. The control panel of buttons was used from the previous prototype, the panel generating a command peculiar to that button. These commands were used as part of a simple communications protocol which was then transmitted over the Internet to the student's computer. A Universal Datagram Protocol was used for the purposes of the peer-peer communication which was achieved (after experimentation) with Visual BASIC. When the software is first initialised on both the Wizard and the student's computer the experimenter has to enter both the remote and local computer's IP address thus facilitating this communication.

The results of the study which used this prototype were described in the previous chapter. These results helped inform the design of the next prototype, termed the 'pre-evalutation' prototype as it was validated with post-graduate students and colleagues.

6.4.3. The Pre-Evaluation Prototype.

The Pre-Evaluation prototype began to more closely resemble the fully functioning system as envisaged at the initial design stage. Until this point only Genie had been implemented as a potential assistant. This phase added another character, Peedy the Parrot, and developed the modules (based on the architecture in Figure 6.11, below) which would provide the necessary functionality for SILA, namely this prototype implemented the reactive, deliberative and world model (which would be used to manage interventions in the summary task).

6.4.3.1. Requirements Analysis.

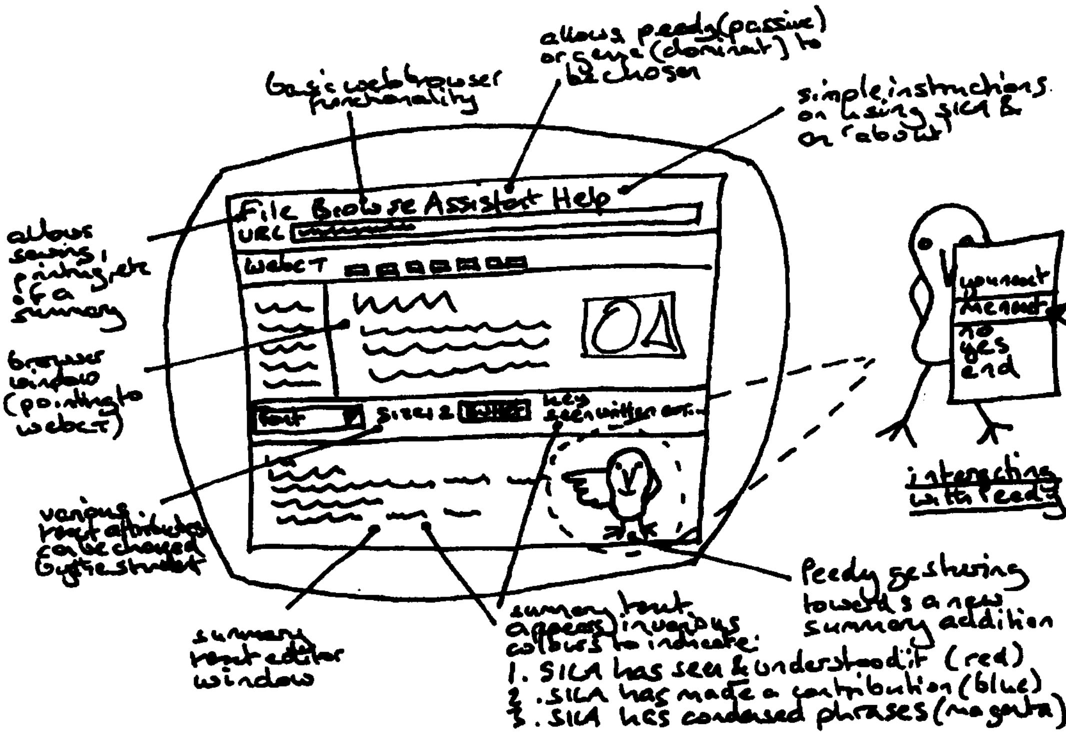


Figure 6.10. Pre-Evaluation Prototype Storyboard.

As can be seen in the storyboard (Figure 6.10) the software now executes on one computer again. That is, the network functionality necessary for the WoZ prototype is now redundant as the ultimate aim is to replace the Wizard through the programmatic control of the agents which will reside on the same computer as the previous agent client. The software contains a number of menus to provide the various necessary functionalities and two main windows to be used in summarising the WebCT notes. Basic text editing facilities are provided (in the form of an editable Rich Text Format box Visual Basic component) and a key used to show the function of the text. The facilities provided at this stage are still rather basic but the prototype allows the next level of technologies to be tested.

6.4.3.2. Architectural Specification.

The behaviour of SILA is modelled as a hybrid reactive/deliberative agent architecture (Aylett & Barnes, 2000; Sloman & Logan, 1999) see Figure 6.11. The permissible dialogue moves are conceptualised as a finite state graph (shown in Section

6.4.3.3) but actions at nodes make use of other computational mechanisms (described in section 6.4.3.3.8), particularly taking advantage of process models derived from the previous empirical studies (see Chapter 4) and those, reviewed earlier, in the literature on summarisation. This phase is obviously gaining much inspiration from the Wizard of Oz study as well as intentions of further behaviour-based affective personifications as described later (see table 6.2). Each component will now be described in turn.

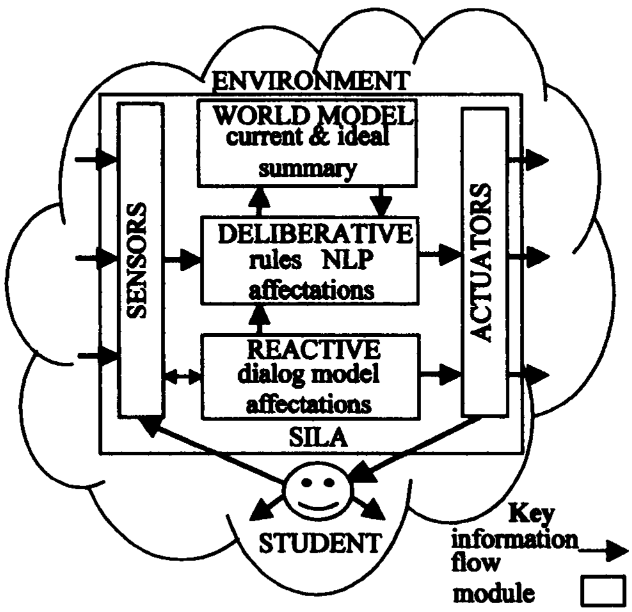


Figure 6.11 SILA Architecture.

6.4.3.2.1. The Environment.

SILA is situated in an environment basically consisting of one window containing World Wide Web browser functionality and another containing simple text editing capabilities. In addition a menu system is employed to provide text editing functionality (new, open, save, cut, copy, paste, etc.), the browser functionality (e.g. refresh, back, home, etc.), artificial assistant selection (Genie or Peedy the parrot), 'organisation' help, a scratchpad, facilities to open and close a log and help. Further text editing capabilities are provided directly above the editing window, i.e. font, text size and bullet points (see Figure 6.12). In addition a key indicates the meaning of the colours used for the summary text: *red* means that the text has been shown to SILA, *blue* indicates that the text has been written by SILA, *magenta* shows text that has been 'condensed', i.e. phrases combined into one phrase and *black* indicates that this is text that the student has entered but not shown to SILA.

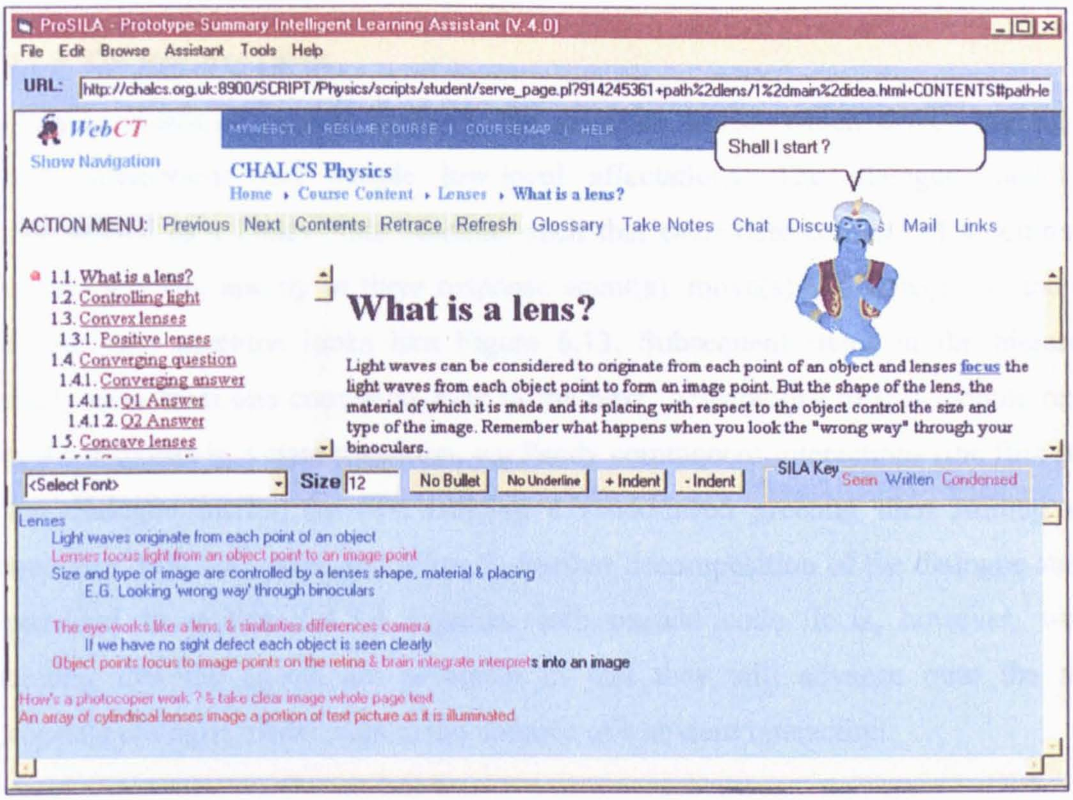


Figure 6.12 The SILA Evaluation Prototype Environment.[1]

6.4.3.2.2. Sensors.

The sensors provide SILA with all information relating to the environment. This is achieved through the menu or voice (i.e. speech recognition) commands services within MS Agent. The commands presented to the Student are adaptive in that those offered are dependant upon the current context, i.e. the current state of the dialogue model in the Reactive module. This allows the student to choose from a limited subset of the commands (some options now don't appear in the menu) thus alleviating the need for a long list of confusing possibilities. Importantly, this also constrains the difficulty of the Natural Language Understanding problem and also allows the implementation of the Finite State Device used in the dialogue model (see Section 6.4.3.2.3). The sensors also provide input from the text editor, in the form of raw text, to the Deliberative module which then processes this information.

---

[1] Note that this screeenshot is actually of the version of SILA used in the Evaluation studies (i.e. with some additional functionalities) as described in Section 6.5.

6.4.3.2.3. The Reactive Layer.

The Reactive module contains the dialogue model which drives the SILA-student interactions and simple low-level affectations. The dialogue model is conceptualised as a Finite State Machine such that each state consists of a command agent and move(s) and up to three response agent(s), move(s) and state(s). At the top level the state machine looks like Figure 6.13. Subsequent states in the hierarchy similarly map from one command state to the next. Affectations in this module result from the response to a state transition, e.g Peedy commences interactions (the first state in the dialogue model) by first uttering a randomised greeting then smiling and announcing "OK, let's have some fun !". Further decomposition of the dialogue stages is provided in section 6.4.3.3 together with pseudo code. It is, however, worth reiterating that the agents are proactive in that they will advance onto the next appropriate dialogue model state in the absence of a student interaction.

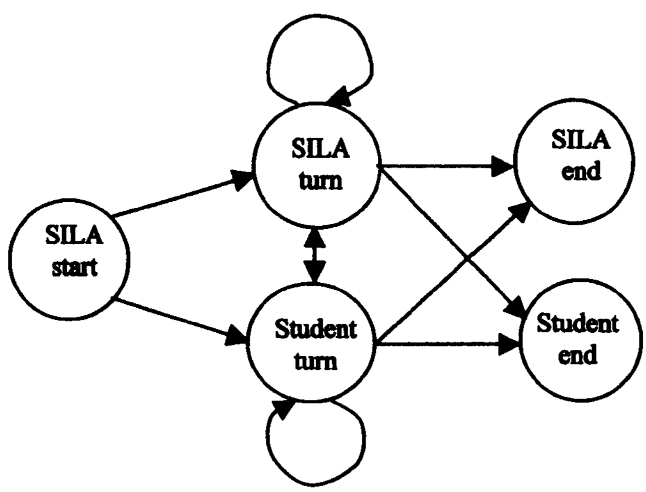


Figure 6.13 Top level Dialogue Model State Diagram.

6.4.3.2.4. The Deliberative Layer.

The deliberative module would contain the traditional reasoning mechanisms typical of AI systems. However, in the present prototype implementation these mechanisms are currently simplified. Notably, this module includes the summarisation rules derived from empirical results (e.g. the summarisation process model) and the literature on summarisation (Tawalbeh, 1994). For example, a SELECTION rule obtains the position of a rhetorical predicate in the rhetorical structure hierarchy from the ideal summary contained in the World Model, see Figure 6.29. In a future version it is envisaged that the SELECTION rule could then use this information to judge a predicates' worthiness to be included in the summary document although at present all phrase summaries are included. However, this rhetorical hierarchical position is used by a deliberative affectation rule which gives students feedback ranging from "well done !" to "why did you do that ?" depending on the perceived importance of the predicate they enter, see section 6.4.3.3.4. Additionally, SILA will indent all phrases according to how important (utilising the rhetorical positions) they are indicated to be, see the pseudo code in Figure 6.28 later. The CONDENSATION rule (which is based on Tawlabeh, 1994) examines phrases in the summary for common keywords then proceeds to match and condense them on the basis of their commonality. For the moment, matching phrases are concatenated through the use of an ampersand symbol.

The NLP (Natural Language Processing) sub-module takes the raw text input from the sensors and converts it to machine-understandable predicates, i.e. it matches a phrase a student has entered with one in the World Model. In the current prototype this is simply achieved through the use of the keyword matching, see the pseudo code in Figure 6.29 later. As well as communicating with the World Model (for both updating and consulting the current summary document and consulting the ideal summary) the deliberative layer also outputs to the actuators both to update the summary document in the Environment and of course achieve some of SILAs 'moves' (as shown in tables 6.3 and 6.4 below).

6.4.3.2.5. The World Model.

The World Model contains the current summary consisting of a list of predicates currently used and is informed by the deliberative layer as mentioned. It also contains the model of the ideal summary, that is, it holds a list of the predicates for the entire summary together with their rhetorical structure position, a justification (to be used by SILA's explanation facility) and keywords used by the NLP mechanism in the deliberative layer as mentioned in the subsequent section on SILA 'knowledge representation'. These data are then used by the deliberative layer, for example, to inform the choice of phrases to be used by the CONDENSATION mechanism.

6.4.3.2.6. Actuators.

The actuators are triggered by both the deliberative and reactive layers. They perform modifications to the summary document i.e. additions of predicates and all of SILA's 'moves' in the form of the animations and speech output e.g. feedback (both echoing input and yes/no/OK), prompts, affectations, justifications, initiating and responding, etc.

The architecture of the Evaluation version implementation of SILA has already been described which shows the system design at a conceptual level. The next section details specification and implementation which may be of interest to Users and Developers.

6.4.3.3. Detailed Specification.

In this section the specification of the dialogue and deliberative models are described in more detail.

6.4.3.3.1. The Dialogue Model.

The main 'driving force' behind the operation of SILA is the reactive dialogue model. It was planned in terms of permissible states along the lines mentioned in section 6.4.2.4 which describes the dialogue model for the WoZ prototype. In order to implement this model, based on the observations in the WoZ studies of discrete stages in the dialogue (see section 5.2.5.3 of Chapter 5), episodes of interaction were broken down into five stages as shown in Table 6.1. This table also shows which State Transition Network (STN) and Pseudo Code represent each dialogue stage.

Dialogue Stage	STN	Pseudo Code
1. Agent starts	Figure 6.14	Figure 6.15
2. Agent's turn	Figure 6.16	Figure 6.17
3. Student's turn	Figure 6.18	Figure 6.19
4. Student ends	Figure 6.20	Figure 6.21
5. Agent ends	Figure 6.22	Figure 6.23

Table 6.1 The Post Wizard of Oz Dialogue Model Five Stages

Note that *agent-command* as detailed in the pseudo code of section 6.4.1.3 is now separated into 'moves' specific to the two assistants, Genie and Peedy – becoming *genie-interpret-command* and *peedy-interpret-command* – but basically remains the same. However, the console form and the WoZ pseudo code of Figure 6.8 are no longer required. Additionally, the pseudo code for *get-user-response* (used in the dialogue model) and *add-command* (used by *get-user-response*) are provided in Figures 6.24 and 6.25. Also note that pseudo code for some of the more advanced features such as *condense* (i.e. the functionality to concatenate similar meaning phrases) and *analyse-contribution* (i.e. functionality to judge the value of a contribution and have the agent respond accordingly) is provided in section 6.4.3.3.8 in its final form. There follows a brief description of each dialogue model stage.

6.4.3.3.2. Agent Starts Dialogue Model.

First the agent will take the initiative and say it is starting. This is followed by presenting the student with an option to affirm or deny this 'move'. If they deny the agent the initiative then they may elect to start the summary. This will result in the genie denying them this opportunity or Peedy allowing it, i.e. in-keeping with the 'dominant' character of Genie and 'submissive' character of Peedy. If the student does not affirm or deny the start move then the agent will carry on regardless. However, if they confirm the agent 'start' move then they get the chance to tell the agent to 'go next'. In this case either selecting 'go next' or actually no action by the student (i.e. the default, autonomous agent behaviour) will result in a jump to the 'agent turn' dialogue model. Figure 6.14 shows the STN for Agent Starts and is followed by the pseudo code in Figure 6.15.

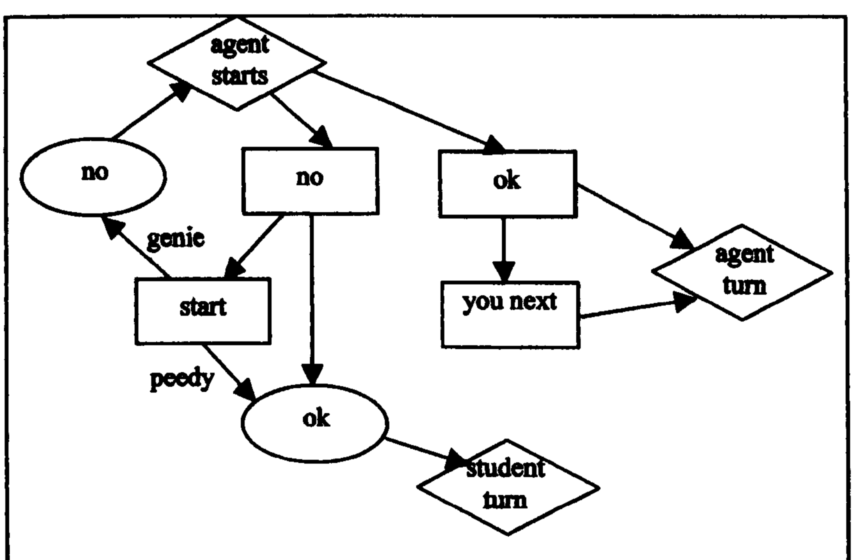


Figure 6.14 Agent Starts Dialogue Model STN.

(Key: ellipse = agent, rectangle = student, diamond = further diagram)

```
show agent
start:
        agent(start)
        get_user_responses(no,ok)
        if user_response = "no" then
                get_user_response(start)
                if user_response = "start" then
                        if genie then agent(no)
                        goto start
                        if peedy then agent(ok)
                        goto student_turn
                else
                if user_response = "null" then
                        agent(ok)
                        goto student_turn
                end if
        end if
        if user_response = "ok" then
                get_user_response(you_next)
                if user_response = "you_next"
                        or user_response = "null" then
                        goto agent_turn
                end if
        end if
```

Figure 6.15 Agent Starts Dialogue Model Pseudo Code.

6.4.3.3.3. Agent's Turn Dialogue Model.

The Agent Turn dialogue model commences with the agent stating that they will go next. The student is then given a chance to deny this move. If they do not then assuming a summary is not finished then the agent will gesture towards the summary and make a contribution to it (if the summary is finished then the dialogue model will jump to 'agent ends'). The student then has a choice of replying with 'why', 'ok' and 'well done'. If they respond with 'why' then the agent will justify its addition to the summary. In the case of the other responses Genie will immediately take another turn whereas Peedy will assume that the student wants to go next. This is a reactive 'affectation' as reported in the next section, i.e. Genie is more forthright whereas Peedy is more passive.

If the student did choose to deny the original statement by the agent to go next then the agent will affirm this choice and allow the student to go next, i.e. jump to

'Student Turn'. Figure 6.16 shows the STN for Agent's Turn and is followed by the pseudo code in Figure 6.17.

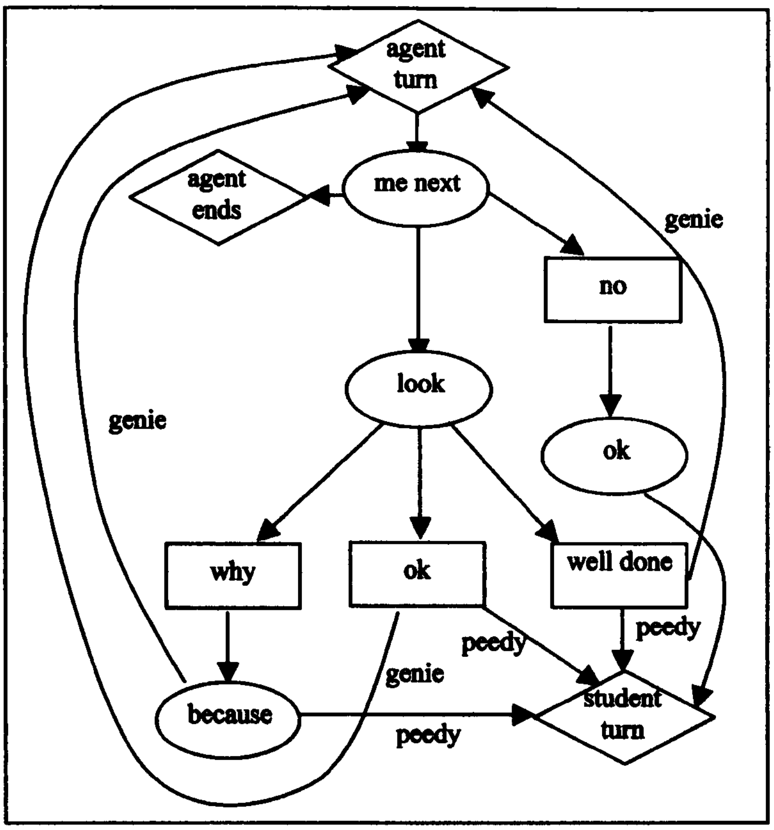


Figure 6.16 Agent's Turn Dialogue Model STN.

(Key: ellipse = agent, rectangle = student, diamond = further diagram)

```
agent_turn
agent(me next)
get_user_response(no)
        if user_response = "null" then
                if not(check_end) then
                        agent(look)
                        agent_action
                        get_user_response(why,ok,well done)
                        if user_response = "why" then
                                agent(because)
                                        if genie then goto agent_turn
                                        else goto student_turn
                                        end if
                        end if
                        if user_response = "ok" then
                                if genie goto agent_turn
                                else goto student_turn
                        end if
                        if user_response = "well done" then
                                if genie goto agent_turn
                                else goto student_turn
```

```
                end if
        else
                goto agent_ends
else
        if user_response = "no" then
                agent(ok)
                goto student_turn
        end if
end if
```

Figure 6.17 Agent's Turn Dialogue Model Pseudo Code.

6.4.3.3.4. Student's Turn Dialogue Model.

Initially the student is given the chance to either go next or end the summary. If they choose to go next then they are given a further chance to show the agent their summary addition (by highlighting it and selecting 'look'). This results in a value judgement by the agent based on the rhetorical hierarchical rating of the phrase, after it has matched it to a phrase in its World Model. That is, a phrase's importance rating which resides in the World Model together with keywords for it to be matched against (see section 6.4.3.3.8) is used to determine the agent's response. If the rating is poor then the agent will ask the student for an explanation of its inclusion, i.e. the student can respond with a 'because' move. If the rating is good then the agent will congratulate the student i.e. use a 'well done' move. If the rating is just average then the agent will simply reply with an 'OK' move. Each of these three rating responses will then jump to another agent's turn if the Genie is being used or if Peedy is being used will jump to the Student's Turn (the rationale for this was explained in 6.4.3.3.3, Agent Turn). Additionally, if the student does not reply to the chance to use 'look' then the whole Agent's Turn stage will commence from the beginning, i.e. that reported in section 6.4.3.3.3.

Furthermore, if the student does not initially choose to go next or the Genie is being used and a random number is above a threshold value (to add unpredictability to Genie's behaviour) then the agent may respond with a 'no', disallow the student's turn and go to Agent's Turn. Finally, if the student initially chooses to end the summary then Student Ends is jumped to. Figure 6.18 shows the STN for Student's Turn and is followed by the pseudo code in Figure 6.19.

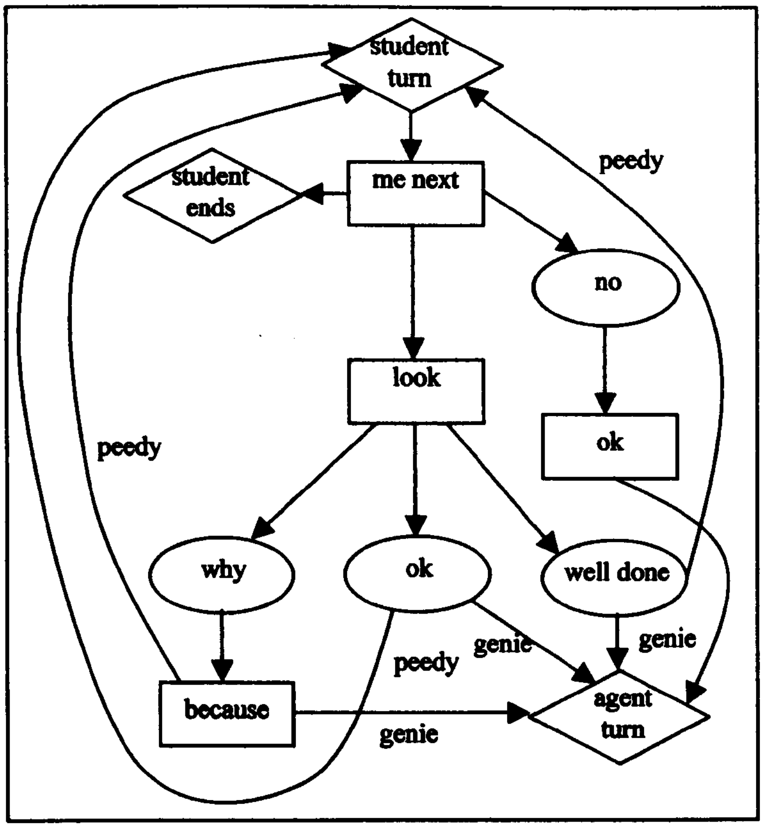


Figure 6.18 Student's Turn Dialogue Model STN.

(Key: ellipse = agent, rectangle = student, diamond = further diagram)

```
agents_turn
get_user_response(me_next, me_end)
if user_response = "me next" then
        get_user_response(look)
        if user_response = "look" then
        analyse_contribution(rating)
                if rating = question_selection then
                        agent(why)
                        get_user_response(because)
                        if genie then goto agents_turn
                        else goto students_turn
                        end if
                end if
                if rating = good_selection
                        agent(well done)
                        if genie then goto agents_turn
                        else goto students_turn
                        end if
                end if
                if rating = average_selection
                        agent(ok)
                        if genie then goto agents_turn
```

```
                else goto students_turn
                end if
            end if
        end if
        if user_response = "null" then
            goto agents_turn
        end if
    end if
    if user_response = "null" or (genie and rnd > 0.9) then
        agent(no)
        get_user_response(ok)
        goto agents_turn
    end if
    if user_response = "me end" then
        goto student_ends
    end if
```

Figure 6.19 Student's Turn Dialogue Model Pseudo Code.

6.4.3.3.5. Student Ends Dialogue Model.

First this model checks to see if the summary is finished. If it is then the agent will use the 'ok' move followed by presenting the student with the 'done' move which the student can affirm to end the session or deny to go back to Student Turn. If the summary is not finished then the agent will deny the student the chance to end and jump to Agents Turn. Figure 6.20 shows the STN for Student Ends and is followed by the pseudo code in Figure 6.21.

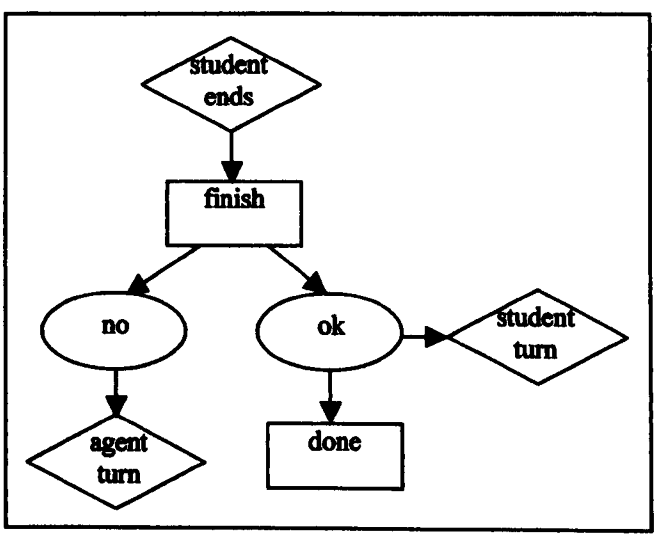


Figure 6.20 Student Ends Dialogue Model STN.
(Key: ellipse = agent, rectangle = student, diamond = further diagram)

```
student_ends
    if check_end then
```

```
        agent(ok)
        get_user_response(done)
        if user_response = "done" then
                goto end
        else
                goto student_turn
        end if
end if
if not (check_end) then
        agent(no)
        goto agents_turn
end if
```

Figure 6.21 Pseudo Code for Student Ends Dialogue Model.

6.4.3.3.6. Agent Ends Dialogue Model.

The agent will first announce that it will finish the summary, giving the student a chance to confirm this action or deny it. If they choose to deny it then the model will jump to Student Turn. However, if it is confirmed then the agent will announce that they have finished the session and will disappear. Figure 6.22 shows the STN for Agent Ends and is followed by the pseudo code in Figure 6.23.

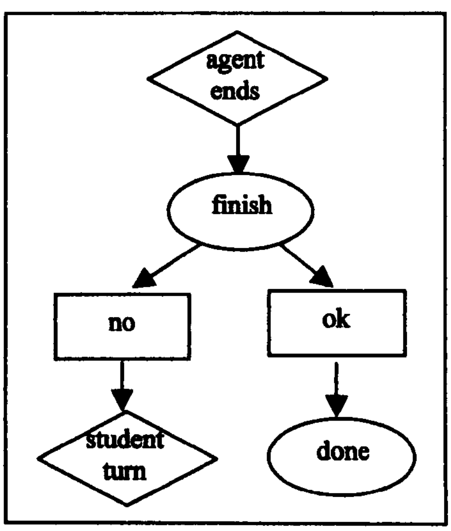


Figure 6.22 Agent Ends Dialogue Model STN.

(Key: ellipse = agent, rectangle = student, diamond = further diagram)

```
agent_ends
agent(I'll finish)
get_user_response(no,ok)
if user_response = "no" then
        goto student_turn
else
```

```
        agent(done)
end if
goto end
```

Figure 6.23 Agent Ends Dialogue Model Pseudo Code.

6.4.3.3.7. Dialogue Model Support.

As mentioned, Figures 6.24 and 6.25 demonstrate pseudo code which is utilised in the dialogue model pseudo code, notably the routine for obtaining a user's response *get_user_response* and a routine to support this by dynamically adding a series of permissable 'moves' to the MS Agent command options. These will not be explained further as they are quite self-explanatory except to say that "ACO" adds 'advanced character options' to the agents e.g. to set speech input parameters.

```
get_user_response (response1, response2, response3)
if genie then
        genie.commands.removeall
else
        if peedy then peedy.commands.removeall
        end if
end if
if response1 <> "" then
        add_command(response1)
end if
if response2 <> "" then
        add_command(response2)
end if
if response3 <> "" then
        add_command(response3)
end if
if genie then
        genie.add.commands "ACO"...
else
        if peedy then
                peedy.add.commands "ACO"
        end if
end if
if peedy then
        peedy.play (listening)
else
        if genie then
                genie.play (listening)
        end if
end if
set timer
do
        check for user_input
```

```
loop until timer timed out or user_input
if user_input then
        command_str = user_input
else
        command_str = null
end if
```

Figure 6.24 Get_User_Response Pseudo Code.

```
add_command(command)
        if peedy then
                case command
                        "say": peedy.commands.add "say"...
                                        .
                                        .
                        "well done": peedy.commands.add "well done"...
                end case
        end if
        if genie then
                case command
                        "say": genie.commands.add "say"...
                                        .
                                        .
                end case
        end if
```

Figure 6.25 Add Command Pseudo Code.

6.4.3.3.8. The Deliberative Module.

Pseudo code is also provided here for the *agent_action* (Figure 6.26), *condense* (Figure 6.27), *organise* (Figure 6.28) and *analyse_contribution* (Figure 6.29) which form part of the deliberative module of Figure 6.11. *Agent_action* is the main routine providing the functionality for an agent's addition to the summary which in turn calls *condense* and *organise* which respectively concatenate phrases and indent them according to hierarchical level. As can be seen from the dialogue model *analyse_contribution* is the routine called when a student makes an addition to a summary by highlighting it and selecting the 'look' move. These routines will not be described further as every attempt has been made to make the pseudo code as easily understood as possible.

```
agent_action
wait for gesturing animation to complete
i = -1
```

```
do
        i = i + 1
loop until i = number_of_predicates or predicate_selected(i) = false
if predicate_selected(i) = false then
        add predicate to end of summary
        indent predicate according to rhetorical level
        change text to blue
        predicate_selected(i) = true
        current_predicate = i
        actual_predicate(i) = predicate(i)
end if
if i = number_of_predicates then check_end = true
call condense
call organise
```

Figure 6.26 Agent_Action Pseudo Code.

```
condense
i = 0
do
        i = i + 1
        if predicate_selected(i) and not (condensed(i)) then
                call condensation (i)
        end if
loop until i = number_of_predicates

condensation
find the current predicate being condensed
step through each of its keywords
if the current keyword is not null
        and the current keyword matches another phrases keyword
        and the two phrases are not the same phrase
        and the current phrase is after the matched phrase
then
        a match has been found
```

```
end if
if a match has been found
        and the current predicate hasn't already been condensed
then
        take each of the current_predicates keywords in turn
                and match it with each keyword of the matched phrase
        if the current_predicate keyword is not the same then
                add current_predicate keyword to the phrase
        end if
delete the matched phrase
remember which phrase was condensed
```

Figure 6.27 Condense Pseudo Code.

```
organise
for each predicate
        if it is included in the summary and
                it has not been condensed then
                indent it to the rhetorical hierarchical position *3 spaces
```

Figure 6.28 Organise Pseudo Code.

(Note: 0 = most important, 3 = least important)

```
analyse_contribution
        change selection's colour to red
        if selection <> "" then
                predicate = -1
                do
                        predicate = predicate + 1
                        if not predicate_selected(predicate) then
                        loop
                                take each keywords(predicate)
                                match selection with it
                        until there's a match or no more keywords
```

```
                    end if
                    if user selection matches a phrase's keywords then
                            analyse_contribution =
                            rhetorical_hierarchical_place (matched_phrase)
                    actual_predicate(predicate) = selection
                    predicate_selected(predicate) = true
                    else if not match
                            analyse_contribution = 3
                    end if
            loop until predicate = number_of_predicates or
                    there was a match
    end if
```

Figure 6.29 Analyse_Contribution Pseudo Code.

(Note: predicate_selected is a boolean, true = the phrase is in the summary)

6.4.3.4. Implementation.

In terms of the environment, this prototype implemented the text editor which was to be the focus of the collaboration between SILA and the student as this is where the actual summary would be jointly constructed. At this stage it included basic text formatting capabilities (such as a font type drop-down menu and bullet point options) plus the coloured text and associated key as explained below. In addition, the World Wide Web browser functionality was also added at this point. This implementation was straightforward as VB contains a component to implement basic browser facilities similar to those provided by MS Internet Explorer.

The dialogue model in the Reactive Layer was implemented as a hard-coded sequence of if-then statements. This allowed the rapid development of this code based on the dialogue model. There is some modularity in terms of chunking of moves, notably into the sections outlined figure 6.13. In addition, statements were added which provided some of the more deliberative Affectations, such as faster response rates of Genie when compared to Peedy and the random choices of options within the dialogue.

The Deliberative Module was perhaps the most involved of the modules developed in terms of programming effort required as it contains the most computationally intensive solutions to problems, i.e. the implementation of the summarisation 'condensation' strategy and the 'Natural Language Processing' (NLP) capability. As explained in the 'knowledge representation' design section below (section 6.4.3.4.1) and the pseudo code in section 6.4.3.3.8, the condensation strategy uses keyword matching to concatenate phrases.

The World Model was designed to enable the agent both to decide upon and make suggestions or additions and changes to the summary on its turns and to enable the agent to comment on the suggestions/additions or changes to the summary made by the student on their turns. The World Model was implemented as a sequence of statements. That is, a summary of a phrase, its rhetorical hierarchy position, its justification clause and its keywords were all represented as assignment statements. In addition the model of the current summary was implemented as a number in an array representing a particular phrase. Section 6.4.3.4.1 further describes this 'knowledge representation' approach of SILA.

The system is roughly implemented as two main modules, one for User Interface related code and the other containing the 'Intelligent' code. In addition four forms are used (a 'form' is a concept used within Visual Basic to describe a user-defined window containing any number of interface components). There is one form for the main SILA window which can be viewed on starting the system, there is another form for display of the 'organisation' HTML help file, another for user input of supplementary text, e.g. to add a reason for the 'because' move, and the final form provides the sketchpad functionality already mentioned.

The 'interface module' providing the interface functionality contains procedures which interpret the student's input in a form suitable for the dialogue model which is contained within the 'intelligent module'. In addition and as mentioned procedures named *Genie_interpret_command* and *Peedy_interpret_command*, for Genie and Peedy respectively, convert commands from the 'intelligent' module to the sequences of animations and dialogue pertaining to each 'move'. In addition, the document editor functionality is provided in this module as is the browser functionality together with all menu options. This module is also responsible for loading the MS Agent characters upon start-up.

The 'intelligent' module is the main driving force behind the behaviour of SILA as its contains the dialogue model as identified in the Reactive Module in Figure 6.11 and described in section 6.4.3.2.3 and 6.4.3.3. The procedure named simply 'action' contains all of the code which performs the essential functions and is decomposed into the five sections of '*start*', '*student_turn*', '*agent_turn*', '*student_ends*' and '*agent_ends*'. As already mentioned, it consists of nested if-then statements and additionally case statements with further embellishments to handle some of the deliberative Affectations. Code within this 'action' procedure then calls either *Peedy_interpret_command* or *Genie_interpret_command* as appropriate.

This 'intelligent' module also importantly handles the user input via MS Agent, that is, it establishes the set of commands to be used (dynamically) then communicates all student input to the dialogue model. Further procedures, as mentioned above, defined within this module take care of the Natural Language Processing ('*analyse_contribution*'), selection of a predicate ('*agent_action*'), condensation of predicates ('*condense*') and the automatic indentation which forms part of the Organisation phase ('*organise*').

Finally, the system uses two files, one containing the HTML course notes taken directly form the WebCT course to be used in the evaluation and testing of SILA and the other containing the HTML Organisation help.

6.4.3.4.1. The SILA Knowledge Representation Approach.

The approach to knowledge representation adopted for the prototype stage of SILA may be considered by comparison a 'knowledge-poor' approach to summarisation (Hahn & Mani, 2000). This means that there is little structural or semantic information about the text to be summarised contained within it. It was decided that the system should use criteria based on the Wizard of Oz studies for the selection and condensation of material, etc. To summarise, information which the system uses should include:

1. a proposition's key words;
2. a proposition's rhetorical hierarchical level;
3. an expert's summary of the proposition;
4. an expert's justification/explanation of a proposition.

Thus the task becomes one of negotiating with the student for turn-taking then, when SILA has the focus, retrieval of an appropriate proposition's summary. Here 'appropriate' means that the proposition has not been used before, is in the right sequence and is relevant according to its relative importance in the rhetorical hierarchy. Keywords are used as a simple Natural Language Processing technique when processing a student's contribution to the summary. As already mentioned, they are also to be used by the condensation strategy which matches propositions based on equivalence of keywords then concatenates them. Finally, the justifications/explanations are to be used by the SILA explanation facility, for example, to provide a reason to the student when they enquire why SILA added a proposition to the summary.

Although this approach requires little reasoning on the part of SILA it is still, nonetheless, a basic knowledge representation. The approach can, in fact, be seen to belong to the *reactive* school of thought rather than the *deliberative*, a philosophy which was also adhered to in the design of the 'affectations'. Whilst the deliberative camp would propose modelling, for instance perhaps, beliefs and desires of an Agent, the reactive camp would be more concerned with the outcomes of the Agent, usually 'hard-wiring' input to output with little computation in between (at least of the sort usually seen in more traditional Artificial Intelligence circles). Brooks has been a proponent of this approach for a number of year, see for instance Brooks (1991), and it is now beginning to be quite commonplace within Agent circles, particularly more recently used in conjunction with deliberative mechanisms. For example, in terms of Affect, fast 'startle' responses could be modelled reactively (thus reducing the computational effort required) whereas more cognitively derived emotions such as 'sympathy' could be modelled deliberatively (requiring more computational resources and hence, potentially, a longer time to process). For example, Cassell reports work on her REA agent which utilises a 'hardwired' response system:

> "*Hard-wired reaction* handles rapid (under 200 milliseconds) reaction to stimuli, such as the appearance of the user. These stimuli can then directly affect the agent's behaviour without much delay, which means that, for example, the agent's gaze can keep up with tracking the user's movement without first processing the meaning of the user's appearance."
>
> (Cassell, 2001)

The knowledge representation approach highlighted may be seen to have a foot in both camps as does the theoretical basis of SILA's Affect, to be described subsequently.

6.4.3.5. Further Implementation: SILA User Interface Development.

As can be seen from previous sections the User Interface of SILA is a central concern of this research in particular the acceptance of the system by students. To this end, this section will explain the personification philosophy (i.e. the 'affectations') which are aimed at motivating students to use the system. This will be followed by the method of realising these personifications, that is the animation (including tools used and an account of the philosophy and actions of the animation sequences) and the dialogue (including the use of appropriate, colloquial language, register and the design of the speech synthesis utterances). The final section will then summarise the 'personality' design process.

6.4.3.5.1. The Personification Philosophy.

As stated earlier Lester et al (1997) noticed that the incorporation of personifications in Animated Pedagogical Agents leads to an increase in student motivation, something which they named the 'persona effect'. As will be elucidated, this research concentrates on the generation of behaviour-based affect, something which Picard (1997) calls "emergent emotions". That is, although there is no internal model of the Agent's Affective states it is the intention that emotive behaviour is used to simulate what have been called 'Affectations', i.e. "artificial manner" (Oxford Dictionary of Current English, 1989). This may, indeed, not be enough in itself to create a totally convincing 'artificial agent' since it will lack a deep-rooted emotional basis for behaviour. However, it is believed that careful manipulation of certain key affectations can create the illusion of character (Barker, 2003). In order that the effect of the behaviour of the Learning Companion could be studied it was decided to implement two distinct characters employing extremes of the Affectations, i.e. 'dominant' and 'passive' traits.

To recap, in recent years psychologists began to reach consensus on common personality trait scales. The currently prevalent version of which is the Big Five personality traits. These are known as Extraversion, Agreeableness, Conscientiousness, Neuroticism and Openness (McCrae & John, 1992). McCrae and Costa (1989) have identified two of these trait scales, Extraversion and Agreeableness, as being most relevant to interpersonal interaction. Furtermore, in the interpersonal circumplex theory of personality (Wiggins, 1979) the *extraversion* trait (also known as power, status or control) is said to range from dominance to submissiveness and the *agreeableness* trait (also known as affiliation or warmth) is said to range from warm to cold or agreeable to quarrelsome. Therefore, this research assumes the extraversion and agreeableness traits as they are seen as crucial in interpersonal interaction in current personality psychology. However, for the sake of convenience the two agents are labelled 'dominant' and 'passive' although they, in fact, are related to both of these traits.

Taking each of these traits as polar opposites, in terms of the two agent designs, the following list (Table 6.2) shows the two extremes manifested in a range of adjectives for two Agents ('Genie' and 'Peedy', the names given to two Agents packaged with Microsoft Agent) together with the method of realisation of these '*Affectations*'.

Genie		Peedy		
affectation	**realisation**	**affectation**	**realisation**	**trait**
hasty (dominant)	less time waiting for user input	patient (forceless)	waits reasonable time for user input	D-S
sarcastic (disrespectful)	says "clever you" & applauds	encouraging (charitable)	says "ok", "well done", etc.	Q-A
stubborn (domineering)	7/10 chance he'll ignore 'user next' request	amenable (unaggressive)	always gives way to users turn	D-S
macho (domineering)	deep voice & muscular	wimpy (meek)	feathered, rounded 'cute' parrot !	D-S
scary (warmthless)	upward lighting & arms folded	friendly (kind)	'light' voice tones & asks rather than tells	C-W
abrupt (firm)	terse language	casual (unauthoritative)	uses colloquialisms	D-S
demanding (discourteous)	e.g. "I'm going next"	polite (well mannered)	e.g. "can I go next please"	Q-A

Table 6.2 Affectations of Genie and Peedy, their circumplex correlates and the corresponding trait (Wiggins, 1979)
(Key: D-S = dominant-submissive, Q-A = quarrelsome-agreeableness, C-W = cold-warm)

It can be seen that by using a combination of reactive and deliberative behaviours including those based on response time, randomly-influenced decisions, appearance, language and voice tone it is possible to simply adjust the Affective qualities of the Agents (Barker, 2003). The greatest effect in terms of these first three qualities is achieved by what are loosely termed the *animation* whereas the *dialogue* model contains the mechanisms such as the last two behaviours mentioned. These two key components will be examined next.

6.4.3.5.2. Animation

The chosen package for implementation of the User Interface aspects of the Agents is Microsoft Agent (Microsoft, 2003). It is the commercial by-product of the Persona project at Microsoft whose aim was to "undertake the construction of a lifelike computer assistant, a character within the PC which interacts with the user in a natural spoken dialogue and has an expressive visual presence" (Ball et al., 1997). MS Agent provides a whole range of services for the designer/programmer which can be accessed from a variety of languages such as Visual BASIC or JAVA. Services include a text-to-speech engine, voice recognition, pre-defined animation sequences and menu controls.

Great care was taken to ensure realistic yet economical animation whilst conveying the desired Affectations for each character. Tables 6.3 and 6.4 detail the sequence of animations for each 'command' that the animation parsing routines should recognise in SILA. They are composite animations in that they consist of the core animations provided for each character in Microsoft Agent concatenated to form an intuitive sequence which it is hoped convey the intended Affect for that particular command. In addition, most commands have an associated speech utterance which could be supplemented by machine-generated utterances, for example, a 'because' move followed by the justification for the inclusion of a phrase within the summary document. Additionally, some of the animations provided with the characters have associated sound effects, such as the squawking as Peedy shrugs his shoulders.

The commands for each Agent were based on the seminal paper by Maulsby et al., (1993) who used the Wizard of Oz approach to prototype an intelligent interface Agent, Turvy, as mentioned in Chapter 5. By taking each of their proposed commands in turn and adjusting it to the intended context, the set of commands were derived as shown. Also, some commands were added which were obviously unique to SILA.

Command	Animations	Speech
show	appears from a cloud	-
hide	waves goodbye	"goodbye, my friend !"
say	points finger in air	[supplied text]
me start	salutation then one hand on chin, other on hip	"shall I start ?"
me take over	move to centre of screen, cloud magically appears (+ cymbal noise), rubs hands together, raises eyebrows, returns to original position	"I'm going next !"
you take over	move to middle of screen, pulls lightbulb from inside pocket, hand to ear	"you take over !"
me finish	scratch head, one hand on hips, other on chin, take out tablet and scrawl (+ whistle & scrawling sounds effects)	"I can do better than this ….hmmmmm….I'll finish this summary"
done	clapping (+ sound effect)	"We're done !"
look	move to bottom right corner of screen, wave arms, gesture (Genies) right then move back to original position	"look at this !"
why ?	hands shoot out to sides, mouth open aghast, one hand on hip, other on side of face with eyebrows raised	"tell me why you did that !"
because…	scratch head then take binoculars out and spin around	"It's obvious, because…" [+ justification]
yes	raised eyebrows, smile, clutching hands	"yes, definitely !"
ok	thumb up	"OK !"
no	hands out to side and shake head	"no way !"
don't know	hands out to side followed by one hand on hip, other on side of face and one rasied eyebrow	"I really don't know !"
well done	move to centre of screen, applaud then return to orginal position	"Clever you !" "Well done !"
unknown	one hand on hip, other on side of face and 'worried' expression	"I don't know what I'm meant to be doing" [whispered]

Table 6.3 Animation and Speech Acts of 'dominant' Genie.

In addition to these animation sequences for each command MS Agent supplies 'idling' behaviour for each character. That is, when the Agent is not being used they will start exhibiting random behaviour. For example, Genie will fall asleep and begin snoring (including the sound effect) and Peedy may take out a biscuit and chew on it or don some sunglasses. These animations help to relieve the monotony of a lack of interaction and are a welcome adjunct to the core services of MS Agent.

These animations were tested in the Wizard of Oz study as described in Chapter 5. The recommendation for the final implementation is to reduce the amount of animations for each command, for example removing the 'somersaulting' that the Agents perform around the screen. The tables shown are the improved designs for the final implementation.

Command	Animations	Speech
show	flies onto screen from nowhere	-
hide	waves wing then flies away (disappears)	"seeya later !"
say	paper aeroplane flies onto screen, grabs it, unfolds paper and reads it then crumples it up and throws it over his shoulder	[supplied text]
me start	bows, wing to beak and looks up pensively, gets calculator out and taps away (with sound effect) then puts wing to ear	"I'd like to start"
me take over	move to centre of screen, waves magic wand then returns to original screen position	"can I go next please ?"
you take over	moves to right of screen, switches a light bulb on then off, puts wing to ear then returns to original screen position	"I'd like you to take over"
me finish	wing to beak and looks up pensively, paper aeroplane flies onto screen, grabs it, unfolds paper and reads it then crumples it up and throws it over his shoulder, takes out pencil and paper and writes	"If you don't mind" "I'd like to finish the summary"
done	pulls out rosette with the number 1 on it (+ drum roll sound effect)	"Hooray ! We've finished !"
look	move to middle of screen, flap wings and jump up, gesture to (peedy's) right then return to original screen position	"Hey ! Take a look at this !"
why ?	wing to head, squawks (sound effect) then shrugs shoulders	"Wow ! Please tell me why you did that !"
because…	scratches beak and eyes rotate then takes out a telescope and looks through it while turning around (+ sound effect)	"I did it because…" [+ justification]
yes	smiles, raises eyebrows and teeth sparkle	"Yeah ! No problem !"
ok	stands erect and raises eyebrows	"OK !"
no	shakes head and eyebrows drop (affecting sadness)	"I'm so sorry, no !"
don't know	outstretches wings, shrugs 'shoulders' and eyebrows droop	"I'm sorry, I have no idea !"
well done	moves to centre of screen, wing to head, squawks (sound effect), smiles, raises eyebrows and teeth sparkle then returns to original screen position	"Well done ! You're so clever !"
unknown	shrugs 'shoulders' and eyebrows droop	"I don't know what I'm meant to be doing" [whispered]

Table 6.4 Animation and Speech Acts of 'passive' Peedy.

The design of animations is more of an art than a science and as such can be a long-winded process of create and test with subsequent revisions finally leading to the desired result. To this end a tool was utilised for this purpose named Microsoft Agent Scripting Helper (MASH). It cannot be praised too highly. It allows each core animation sequence to be initiated at the selection of an entry in a drop-down menu containing a list of all permissible animations for the currently selected character. It was also used extensively in the design of the speech utterance (see the next section). MASH then allows animation and speech commands to be exported in a form readily integrated in a

Visual BASIC program. Although MASH was basically used for the purposes described it contains, in addition, a host of other useful features including the ability to have MS Agent run various Windows services like opening email clients.

6.4.3.5.3. Dialogue.

As mentioned, MASH was also used for the implementation of spoken dialogue utterances, shown above in Tables 6.3 and 6.4. MASH allows the designer to map individual words onto a frequency, that is for each word (or phoneme for that matter) an octave and a musical note can be selected. In addition, MS Agent provides the ability, via the speech synthesis engine, to 'whisper' text or to speak in a 'monotone' plus the individual speed of words can be selected. Also, using the supplied text-to-speech engine (Lernout & Hauspie Truvoice), a number of voices can be selected from eight male and two female options. It was thus possible to provide not only unique voices for each of the two Agents but also unique speech inflections. That is, an attempt was made to make Peedy sound 'friendlier' than Genie who, in turn, was made to sound 'domineering'. To elaborate, Peedy used higher tones than Genie and, in fact, utilised quite a comical caricature-type of voice whereas Genie utilised lower tones and a more serious male kind of voice. A further example of the manifestation of their character is that Genie is likely to end an utterance on a lower note whilst Peedy will have a higher inflection at the end of his utterance.

In addition to the actual subtleties of the physical characteristics of the voice great care was obviously taken over the actual language of each character, in-keeping with the domineering versus passive qualities of Genie and Peedy. As can be seen in Tables 6.3 and 6.4. Genie will say "you could have done better !" whereas Peedy would say "I think you were wrong", that is Peedy is certainly more 'polite' than Genie who can be a little 'sarcastic' at times. Peedy's 'manners' are certainly better than Genie's too, being more likely to say "sorry" or "please". In addition, it was noticed at an early stage of the research (in fact, during the initial studies of summarisation at CHALCS) that the students used a particular colloquial language in their classroom interactions. As the intention is for SILA to be considered a peer by the students it was felt that these colloquialisms had a place in the speech utterances of the Agents. In fact, few colloquialisms were actually used (although a number were identified during the summarisation studies, see Chapter 4) which is something that could perhaps be built on in future versions. For example, Peedy will say "no problem !" and Genie will say "no way !"

6.4.3.5.4. The SILA 'Personality' Design Process.

To summarise the SILA 'personality' design process, the subtle expressions used in the implementation of the agents in this research consisted of carefully crafted variations of *language*, *voice*, *appearance*, *timing* and *randomisation*.

- *Language* was modelled after that used by the target students, that is, informal and colloquial (see Figure 6.30). Additionally, to appear dominant, statements may be used or to appear submissive, questions may be used.
- *Voice* was carefully crafted to be in-keeping with the characterisation of the agents, for example the character Genie has a deep register to mark a 'macho' dominant character whilst Peedy has a high pitched register to reflect his whimsical submissive character. Figure 6.31 shows the use of the Microsoft Agent Scripting Helper (MASH)[2] to carefully assign an individual pitch to each word thus creating the desired inflection and tonal register.
- *Appearance* conveyed the 'character' of the assistant. For example, the macho Genie character was strengthened through the use of a moustachioed, muscular male whereas the whimsical Peedy character was emphasised through the choice of a Disney-type of parrot. Furthermore, the appearance of the character must match the communicative intent. For example, Robby the Robot (who reads the instructions) will shrug his shoulders if making an uncertain statement, see Figure 6.32.
- *Timing* reinforced the impatient nature of the dominant genie character as it would soon interrupt whilst the submissive parrot would appear more patient in comparison. Timing was also very important in synchronising actions with dialog. For example, Figure 6.33 shows Genie gesturing towards an addition to the summary at the moment that it takes place.
- *Randomisation* had the effect of reducing the monotonous, predictable behaviour both in terms of idling animation as mentioned previously and shown in Figure 6.34 (where Peedy is slyly glancing from side to side whilst awaiting a student's

---

[2] http://www.bellcraft.com/mash/

interaction) and in terms of agent responses to student requests e.g. the possibility of Genie ignoring a student's request to go next as reported in section 6.4.4.3.3.

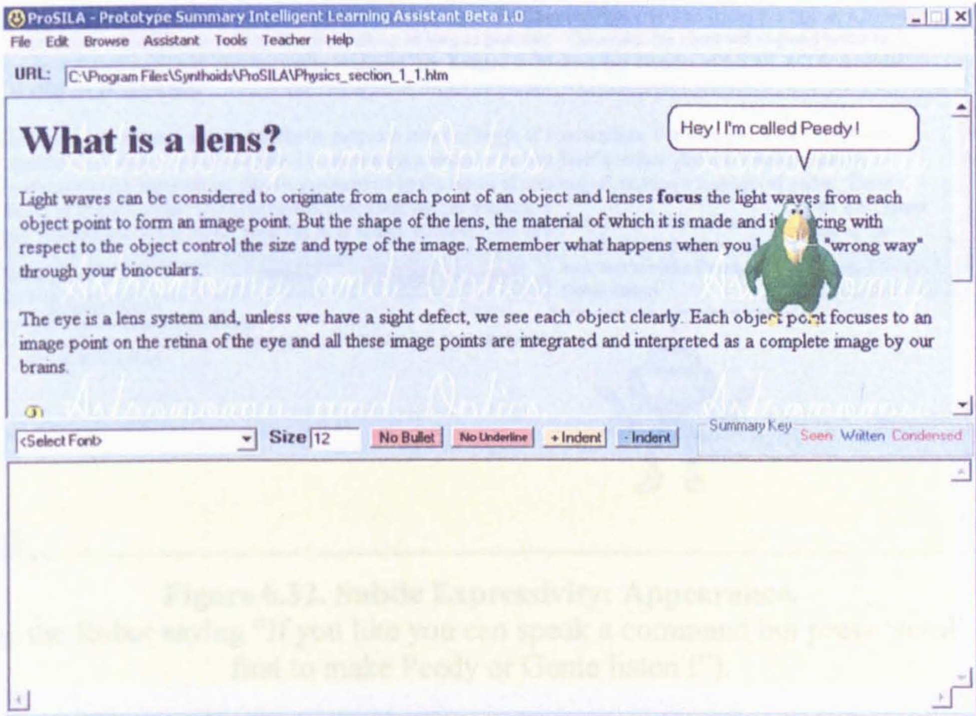


Figure 6.30. Subtle Expressivity: Language.
(Peedy saying "Hey ! I'm called Peedy").

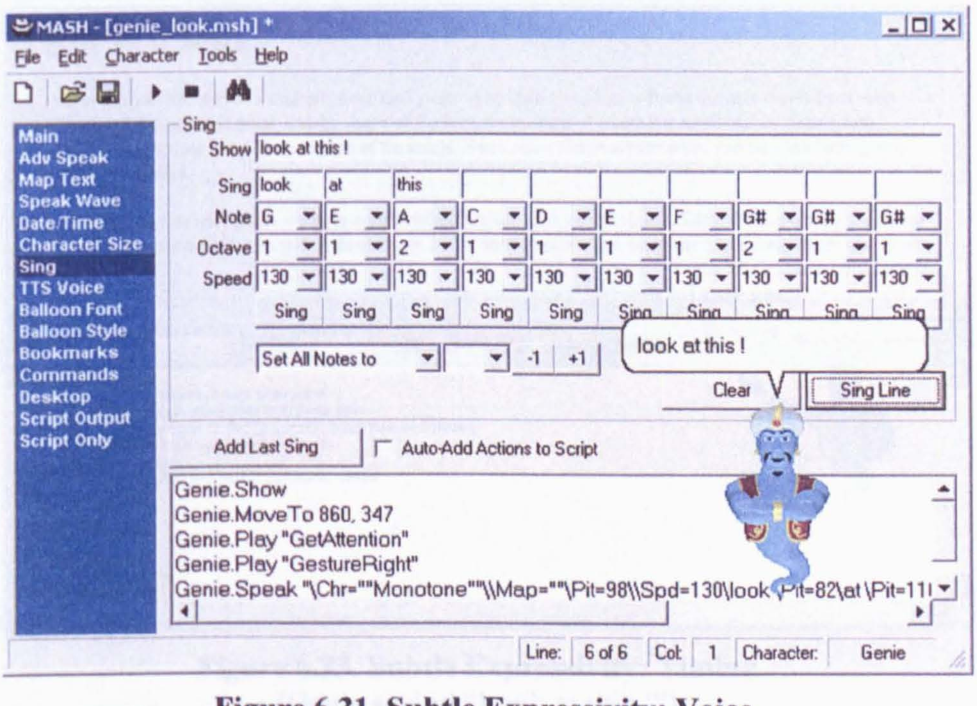


Figure 6.31. Subtle Expressivity: Voice.
(Genie 'singing' "Look at this !")

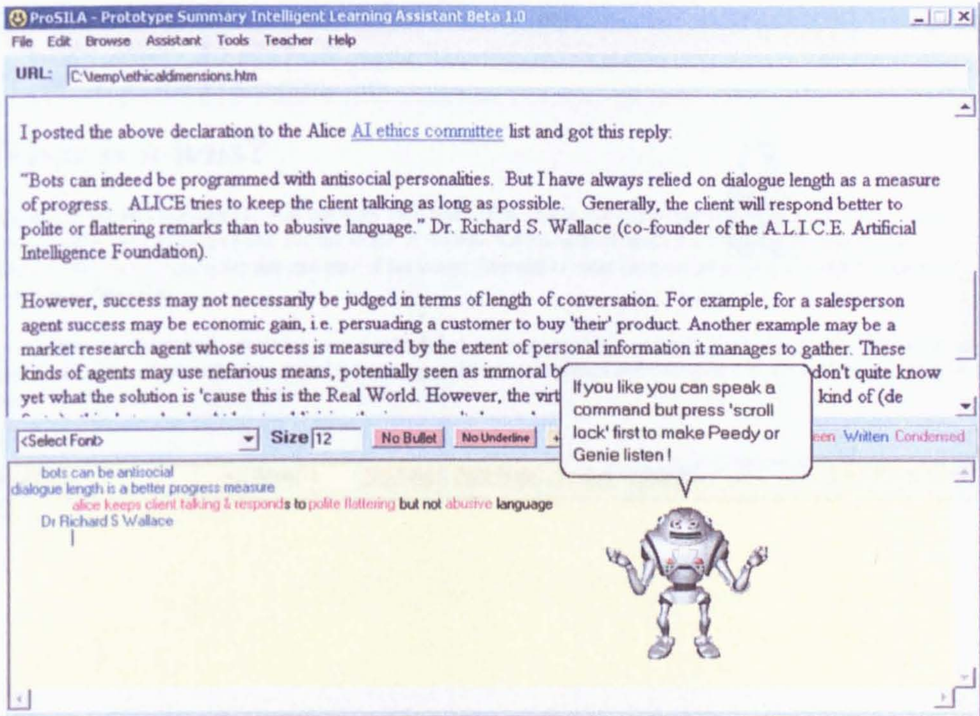


Figure 6.32. Subtle Expressivity: Appearance.
(Robby the Robot saying "If you like you can speak a command but press 'scroll lock' first to make Peedy or Genie listen !").

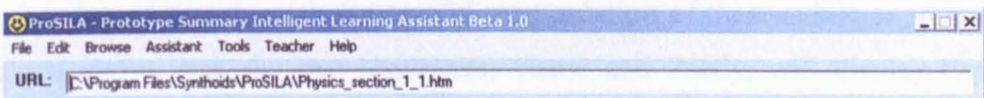


Figure 6.33. Subtle Expressivity: Timing.
(Genie saying "Look at this !").

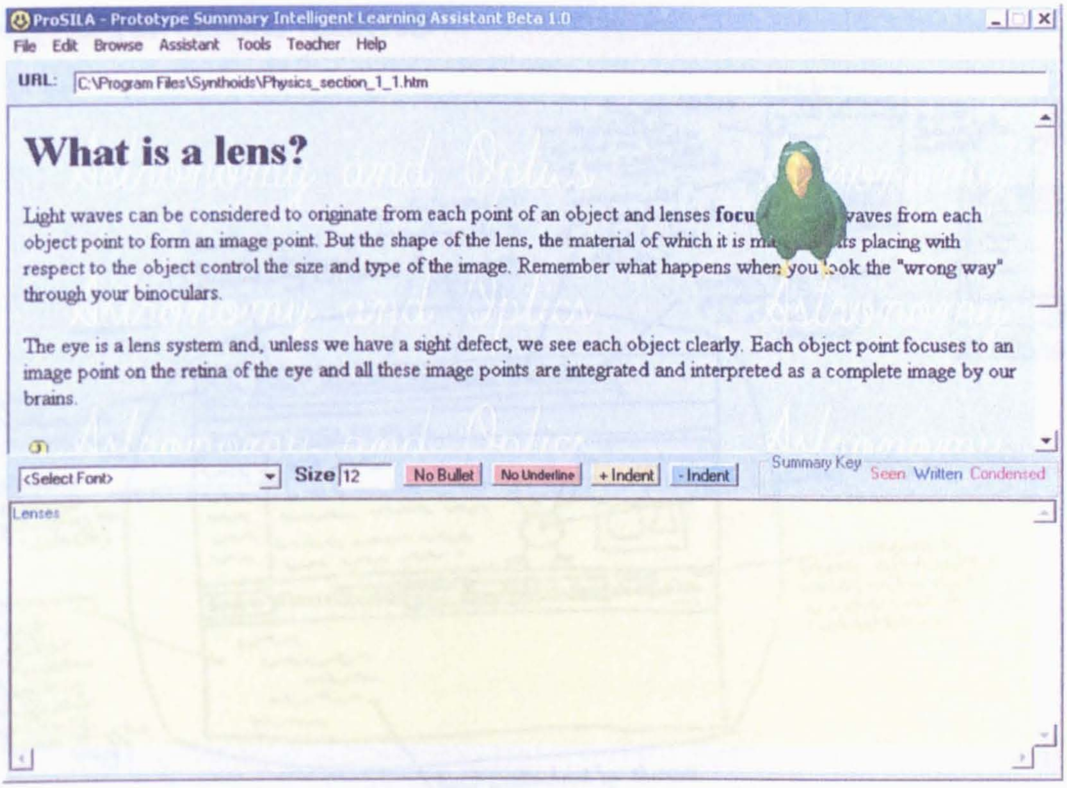


Figure 6.34 Subtle Expressivity: Randomisation.

It is hoped that through the careful manipulation of these subtle expressions, as highlighted above, that the two characters, Genie and Peedy, will display two quite different personas in-keeping with the general rationale of them displaying 'dominant' and 'passive' personas respectively. This kind of attention to detail does take time and much experimentation, as mentioned, however it is seen as being crucial in meeting the objective of displaying affective behaviour.

This pre-evaluation system was evaluated preliminarily with colleagues from the Computer Based Learning Unit at the University of Leeds (this preliminary evaluation is reported in the next chapter). As a result the Evaluation prototype was developed, as follows.

6.5 The Evaluation Prototype.

This section first describes the requirements analysis of the prototype used in the Evaluative studies described in Chapter 7 then reiterates its specification and architecture. This is followed by a description of this system's implementation.

6.5.1. Requirements Analysis

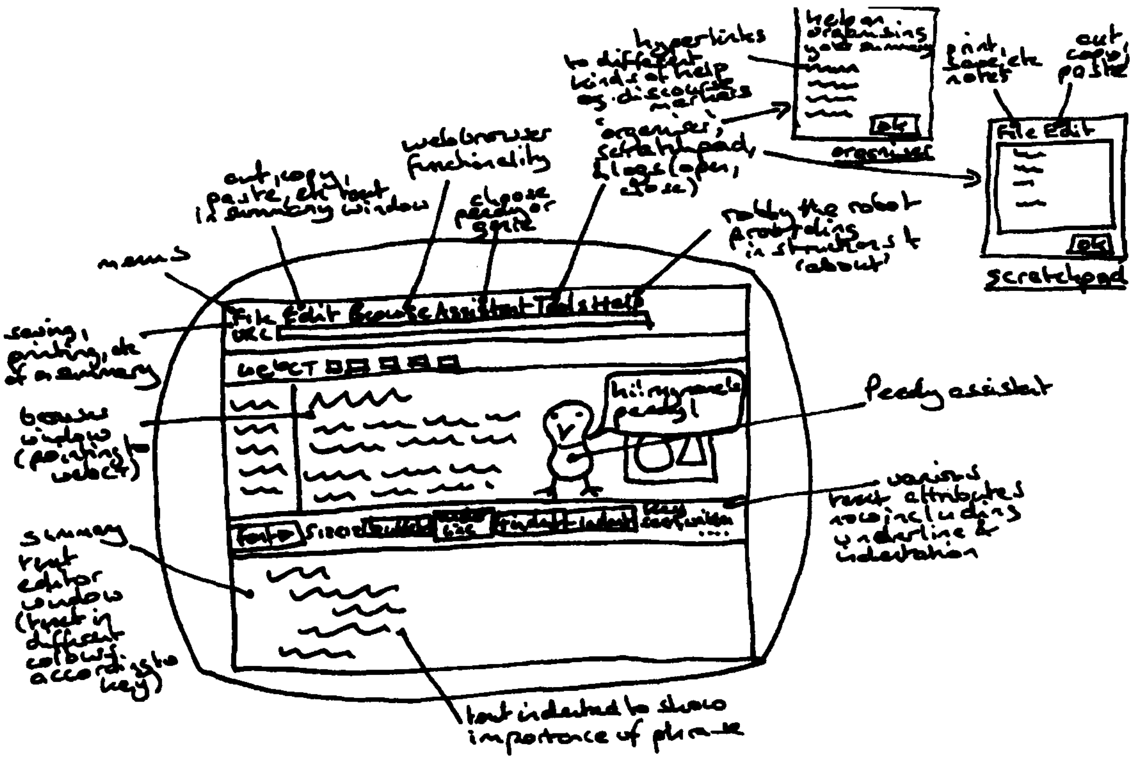


Figure 6.35. Evaluation Prototype Storyboard.

As can be seen from Figure 6.35 the Evaluation Prototype introduces an 'edit' menu to provide cutting and pasting functionality and a 'tools' menu providing "help on organising a summary" and a 'sketchpad' derived from observations of studies reported in Chapter 3. Most of the other functionality of this prototype is derived from the Pre-Evaluation prototype as described in the previous section. However, subtle improvements were made and are reported in section 6.5.4, Implementation.

6.5.2. Architectural Specification.

The architectural specification for this evaluation prototype version is identical to the pre-evaluation version as described in section 6.4.3.2 and so will not be elaborated further here except to say that the hybrid reactive/deliberative conceptualisation is retained in this version.

6.5.3. Detailed Specification.

Again, in-keeping with the evolutionary prototyping methodology a lot of the code developed for the Pre-Evaluation Prototype was reused and updated where appropriate. Hence the pseudo code described in sections 6.4.2.3 and 6.4.3.3 is still basically relevant.

6.5.4. Implementation.

The evaluation prototype made a number of improvements upon the pre-evaluation prototype basically consisting of those related to the dialogue model, the animation, the document editor and miscellaneous, less crucial developments.

Firstly, the dialogue model was modified to make the flow appear smoother. For example, in the previous prototype implementation, Genie would say "OK" after the student had selected a command from the MS Agent menu. Although, this may appear a sensible thing to do on the surface, as it provides immediate feedback, in practice it became monotonous, even annoying, leading to anticipated hostility towards the system. The solution was to simply have the Agent nod, utilising the visual channel in terms of appropriate non-verbal indications instead of the already saturated auditory channel. Underlying assumptions inherent in a smooth-flowing conversation were also implemented as *expectations*. For instance, if a student tells SILA that they don't want it to take the next turn then SILA can assume that the student wants to take the turn thus skipping the intermediate stage of the student *asking* to take the turn.

Secondly, the animation both of the Agent and the summary document were improved. Of the former, perhaps the most crucial addition was the synchronisation of SILA gesturing towards the summary document at the same time the new phrase was added. In previous versions these two actions were occurring at different times thus shattering the illusion that it was the Agent who is adding the phrase, something which is central to the 'believability' of SILA. Furthermore, the timing of highlighting of phrases as SILA 'considers' each one for condensation was improved so that it did not linger too long on one phrase. Finally, the amount of unnecessary behaviour of the Agent was reduced, mainly by taking away moves around the screen (creating a 'somersaulting' effect) thus economising on the animation with the effect of creating less distracting behaviour as had been reported in the Wizard of Oz studies.

Thirdly, minor modifications were made to the document editor. Indentation functionality was added so that the student could choose indentations of their own and editing capabilities were added in the form of cut and paste, etc. which could be activated in the menus at the top of the main SILA screen or by right mouse clicking over the summary.

Finally, miscellaneous improvements made during this prototype included adding a log of commands, times, etc. for evaluative purposes, a 'sketchpad' to be used as a separate jotter by students (as noted during the summarisation studies in Chapter 4), improved SILA explanations which reported the function of a phrase in a summary as well as its meaning and the disabling of student input during condensation thus minimising the chance of program "crash". Note that the final prototype listing contains a large number of comments to ease further development of the software.

6.6 Conclusions.

The implementation of an Affective Learning Companion is not an easy task. It requires a huge amount of time spent designing the system and an even longer time building it. The prototype described above is the end-product of a continually evolving specification which has resulted from the user-centred approach to design espoused by experts in the development of interactive systems. Various technologies and architectures were considered for this implementation before any coding actually took place. Of course, this is an essential part of any implementation phase as the choice of what transpires to be an unsuitable development decision will have lasting negative impact throughout the lifecycle of the project.

The result of this implementation phase, SILA, is an environment which supports and encourages students to write their own summary notes based on the Physics notes developed earlier in WebCT. They collaborate with one of two animated characters, Peedy the Parrot or Genie, in order to co-construct their summary document. It is thus hoped, in keeping with Chan's notion of a Learning Companion (Chan & Baskin, 1988), that they will vicariously learn the art of summarisation from this companion and its associated environment as described above.

As a user-centred prototyping methodology was used throughout the development of the SILA system distinctions between design and implementation inevitably fade and what may be considered an implementation by an outside observer may well be considered integral to the design evolution by the developer. However, two main design explorations currently hold promise for future developments. The first of these relates to continued development of Application Programmmers Interface (API) solutions enabling seamless integration of the functionality of the system proposed here with WebCT. The second relates to a more knowledge rich approach to the summarisation model together with authoring tools that support content mark-up in a time-efficient manner. These ideas are further outlined below.

6.6.1. WebCT API.

If taking a knowledge-rich approach to SILA implementation it would be prudent to integrate WebCT with the Learning Companion so that the Companion could form a model of the student's learning. Thus, the knowledge of which pages of the course notes the student has examined would enable the Companion to know the basis of their summary construction. Other tracking information such as number of articles posted to the bulletin board and glossary usage might also help the Companion ascertain the student's level of motivation regarding the course so that it could adjust it's persona accordingly.

There are two ways to provide this communication between the Companion and WebCT. The first concerns accessing the log file, resident on the WebCT server, directly perhaps over a Local Area Network. Whilst inelegant in that it is reliant on a Local Area Network (LAN) and not a Wide Area Network (WAN), such as the World Wide Web, it will provide all of the necessary information. The second method to access WebCT data was made possibile by an Application Programmers Interface (API) included in version (3.x) of WebCT (see Appendix N). There are three possibilities to call the API : use the command line interface to call webctdb followed by a number of parameters (such as student id, password and command etc.), use a standard URL followed by the parameters as key/value pairs separated by ampersands or use GET or POST methods to submit requests with parameters passed as data (for example, from standard HTML forms).

Unfortunately though, the API only allowed access to the student database which contains such information as login id's, names and quiz results. This information could be used to a limited effect to provide some of the required information but would require supplementing with the LAN solution data in order for the Learning Companion to obtain the complete Student Model. It is envisaged that a suitably designed quiz may be useful in gleaning a student's performance on the course or even the extent of their previous summarisation knowledge. It was envisaged early in the design phase of SILA that the Companion could limit itself to the API solution and, as the API is expanded upon by the WebCT designers in the future, for example, through the use of an "IMS Enterprise API" conforming to IMS Enterprise 1.01 Specification in the WebCT Campus Edition (WebCT, 2003), the Companion can be updated accordingly. However, it was decided to leave this possibility to a future prototype when a more knowledge-rich approach may be utilised. Thus linking with WebCT is not achieved on a system level in the current implementation although it would seem to be integrated into the SILA environment to the student as full functionality is available through the integral browser.

6.6.2. Mark-Up.

Tawalbeh's (1994) Computer-Based Summarisation Program (reviewed in Chapter 2) uses text *mark-up* techniques in the creation of a summary. This mark-up is based on the Standard Generalised Markup Language (SGML).

> "SGML is an international standard for the definition of device-independent, system-independent methods of representing texts in electronic form"
>
> Text Encoding Initiative (www.tei-c.org)

Tawalbeh's idea is that a text is annotated with 'mark-up', or embedded codes which represent the structure of a text, based on rhetorical analysis (Meyer, 1975; Pilkington, 1992). Rhetorical Predicates are relations such as comparison, problem-solution, instrument, cause-consequence, etc. These serve to link content sentences to form a coherent discourse and are staged (sequenced) according to the emphasis and importance the author places upon them. This mark-up can then be utilised by the computer-based system for example, selecting or condensing material according to its relative sequence and importance in the text as a whole. Tawalbeh delineates nine elements which are needed to mark-up an exemplar 'contrastive structure':

1. **topics being contrasted**: propositions should be exclusively linked to just one topic plus condensation strategies apply to topics;
2. **sequencing**: each proposition has a number to ensure correct ordering;
3. **attributes (headers)**: allow contrasting of topics on an attribute by highlighting differences plus the allocation of a proposition to its attribute further delineates it, e.g. the arctic and Antarctic are contrasted on the attributes of flora and fauna in the following sentence: "the arctic has a rich flora and fauna whilst the Antarctic doesn't";
4. **values**: an attribute's content further highlights how topics differ plus these may be deleted during selection, e.g. in the previous example "rich" is a value applied to attributes flora and fauna;
5. **hierarchical level**: marking rhetorical predicates, discourse markers and lexical items with a hierarchical level shows a proposition's importance in the text which is used by the selection and condensation strategies - see Meyer (1975) and Pilkington, (1992) for a description of the process of deciding level;
6. **importance**: indicates emphasis and significance for a proposition – this is added by the author through signals such as "most importantly" or by highlighting techniques such as the use of bold or "in particular";
7. **dependency**: shows dependencies between propositions aiming to differentiate between main and dependent attributes and used in organisation, e.g. to keep causally dependant items together in the sequence;
8. **collectivity**: used in organisation of the text "shows the authors emphasis on the quantity of attributes (one, few, many) on which the topics may differ" e.g. the author might say "there are 3 important variables affecting climate" and number these 1, 2, 3. In this case it would be important to select all three and keep them together;
9. **synonyms**: used in style variation.

(Tawalbeh, 1994, p.224)

Appendix O shows how Tawalbeh marked-up one text using the above scheme. Included is the marked text, the codes used to represent proposition number, values and the rhetorical predicates, etc.

As can be seen from the Appendix, the marking-up of a text presents quite a challenge for even a competent analyst as a variety of criteria such as rhetorical function, collectivity, synonyms, etc. need to be considered consecutively. This necessarily limits the utility of the approach since a sophisticated authoring tool to assist the material writer by automating some of this process would be required if the system were to be used by course writers. Although such a tool could be built to complement a future system, at this point in time, it was decided to adopt a much simpler technique. However, if the system is to develop beyond the proof-of-concept stage it would be prudent to once again consider a comprehensive mark-up approach to coding the Physics course notes and tools to maybe assist even along the lines of the modern-day successor to SGML, Extensible Mark-up Language (XML). This latter idea of course fits in quite well with the notions of web delivery of a course although, at present, XML is not supported by WebCT. The end result would be the possibility of developing a more knowledge-rich approach to summarisation than the present implementation described in this Chapter allows.

These and other investigations formed part of an early Feasability Study which investigated the possibility of implementing SILA. As mentioned, the technologies reported in this section were deemed outside the scope of this limited proof-of-concept design. Instead a more robust and achievable solution was envisaged.

The current prototype system, SILA, requires testing with students who would typically find such a tool useful. To this end, the next Chapter will report on studies undertaken at CHALCS, local schools and elsewhere, aiming to evaluate the system in terms of the aims of the software and, more fundamentally, the aims of this research programme as whole.

Chapter 7: Validation of the Learning Companion.

7.1 Rationale of the SILA Evaluation.

The purpose of the studies described in this chapter is to evaluate the effectiveness of the prototype Summary Intelligent Learning Assistant (SILA). In order to do this it is necessary to revisit the original Research Aims (as outlined in Chapter 1) which relate to the computer-based learning environment. Subsequent analysis of a variety of data will lead to conclusions regarding this effectiveness.

The primary relevant Research Aim to be considered here is that of point 5 in table 1.4 of Chapter 1, notably:

> "To trial the Learning Companion to suggest Agent design characteristics which may prove motivating and effective in assisting students to learn from the VLE material."

Decomposition of this Research Aim as it relates to this evaluation would lead us to examine the software on the basis of three **SILA Evaluation Objectives**:

1. *Validating* that a proof of concept system has been designed using the agent based approach which gives a reasonable approximation to the behaviour of the wizard in the Wizard of Oz study;
2. *Evaluating* whether the proof-of-concept system presents motivating and believable affectations and the extent of any differences between "personae" (characters with distinct affectations) in this regard;
3. *Assessing* whether the agent (or one of its "personae") provides reasonable support for summary note-taking of the VLE material.

The above suggests criteria against which we wanted to evaluate the system "top-down". In addition, a preliminary examination of the data also suggested some further criteria of analysis "bottom-up". The result, as shown later, will be interrelated themed comments grouped according to categories arising from this evaluation. First though a preliminary evaluation of the system in terms of presenting two distinct personae is presented prior to full evaluation of the agent with students.

7.2 SILA Preliminary Evaluation.

A preliminary evaluation was carried out with members of the Computer Based Learning Unit (CBLU) at the University of Leeds. Although this group are not ideally representative of the target group of users, this preliminary study did allow an early impression to be formed of people's perceptions of the SILA Affectations and ensured that a target group of students would be available for the full evaluation study to be described later. In other words, due to the lack of available students in the target group it was important not to use these at the pilot stage.

7.2.1. SILA Preliminary Evaluation Setting and Materials.

This evaluation took place as part of the CBLU "round" which existed to allow researchers to present their work to each other, whether they were postgraduate students, academics or other researchers from the University of Leeds or, indeed, occasionally from other Universities or research bodies. The software was demonstrated to 7 participants (postgraduate students and members of staff) in a University classroom. This classroom consisted of a number of networked computers around three of the walls with an electronic whiteboard attached to the wall at the front of the classroom. The SILA software was pre-installed on a machine at the front of the classroom which, in turn, was connected to the electronic whiteboard serving, in this case, as a large display. Speakers were also utilised to broadcast the agent dialogue and sound effects to all of the participants in the room. Thus, all of the participants were clearly able to witness the experimenter-agent interactions. The actual version of SILA used for this study is described in section 6.4.3 of Chapter 6. An 'Affect Scale' was also utilised, as described in the next section, an excerpt of which can be seen in Figure 7.3 (for the complete version see Appendix J).

7.2.2. SILA Preliminary Evaluation Method.

First, the Affect Scale was handed out which includes the instructions shown in Appendix J. These instructions were also read out to the participants. Furthermore, the purpose of the session was explained to the participants, notably that software will be demonstrated which is designed to help Advanced level students with the summary note-taking of Physics VLE materials. It is worth noting that the Physics VLE had been previously demonstrated in a similar session during its development. Hence, although not really required, the participants were familiar with the VLE. However, the focus of this session was the SILA software. With this in mind the demonstration proceeded with an outline of the functionality provided by SILA, explaining its menus, the agent interaction rationale, etc. all in a similar vein to the walkthrough presented in Chapter 8.

Thus, with some idea of the core functionality provided by SILA, the experimenter began a typical SILA session, summarising the "What is a Lens ?" section of the Physics VLE notes. At the end of the demonstration session participants were reminded to complete the Affect Scale if they had not already done so. On completion of the Affect Scales a Focus Group discussion was held. This Focus Group discussion took place in the same setting alluded to above in section 7.2.1. It simply consisted of asking the participants for their impressions of the software demonstration. The participants were quite forthcoming with observations, criticisms and recommendations related to their impressions of SILA formed during the evaluation so little encouragement was required. Subsequent to this discussion notes were made to record the key points for further deliberation.

7.2.3. SILA Preliminary Evaluation Results.

Table 7.1 shows how each of the 'Affectations' (along the x axis) relates to an overall score for all subject's impressions of each Affectation for both Peedy and Genie shown in Figure 7.1 and 7.2 respectively. The results shown are as a percentage of the total possible marks for each Affectation. They show a clear difference between Genie and Peedy with Genie appearing to be largely perceived as having negative affect characteristics and Peedy being perceived as having largely positive affect characteristics. The derivation of the Affectation types were previously discussed in Section 6.4.3.5 of Chapter 6.

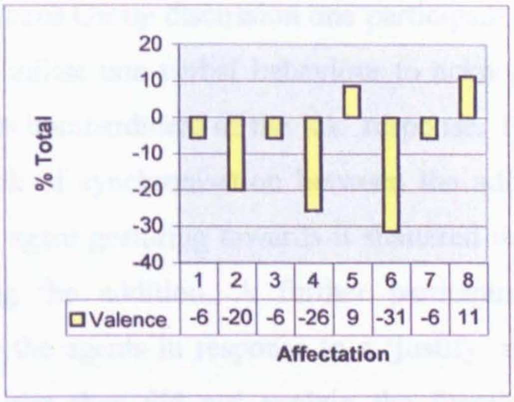


Figure 7.1 Genie Preliminary Affect Results.

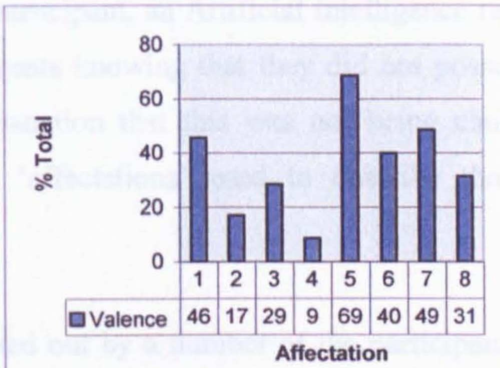


Figure 7.2 Peedy Preliminary Affect Results.

Affectation	negative affect	positive affect
1	*impatient*	*patient*
2	*derisive*	*encouraging*
3	*obstinate*	*amenable*
4	*macho*	*wimpy*
5	*scary*	*friendly*
6	*curt*	*laid-back*
7	*impolite*	*polite*
8	*contrived*	*genuine*

Table 7.1 SILA Preliminary Evaluation Affectation Adjectives.

impatient	-5	-4	-3	-2	-1	0	1	2	3	4	5	**patient**
derisive	-5	-4	-3	-2	-1	0	1	2	3	4	5	**encouraging**
obstinate	-5	-4	-3	-2	-1	0	1	2	3	4	5	**amenable**
macho	-5	-4	-3	-2	-1	0	1	2	3	4	5	**wimpy**
scary	-5	-4	-3	-2	-1	0	1	2	3	4	5	**friendly**
curt	-5	-4	-3	-2	-1	0	1	2	3	4	5	**laid-back**
impolite	-5	-4	-3	-2	-1	0	1	2	3	4	5	**polite**
contrived	-5	-4	-3	-2	-1	0	1	2	3	4	5	**genuine**

Figure 7.3 Affect Scale Excerpt.

In terms of the Focus Group discussion one participant pointed out that it would be more appropriate to utilise non-verbal behaviour to acknowledge a user command instead of the continuous bombardment of the 'ok' response. This same participant also pointed out that the lack of synchronisation between the addition of a phrase to the summary and the SILA agent gesturing towards it shattered the illusion that it was the agent that was making the addition. A further participant pointed out that the explanations offered by the agents in response to a 'justify' move from the user were inappropriate; in particular they did not explain the function of the phrase in the summary. Finally, one participant, an Artificial Intelligence researcher, was reluctant to attribute affect to the agents knowing that they did not possess any 'model' of affect. However, upon an explanation that this was not being claimed, this participant did concede that the term 'affectations' used to describe their behaviour was a fair description.

It was also pointed out by a number of the participants that it was not obvious what was being asked for on the final scale on the Affect Scale, that is 'Contrived-Genuine'. This scale was in fact a late addition to the Affect Scale and was an attempt to derive the degree of convincing behaviour the agents displayed in terms of how believable they were. However, a significant number of participants did not understand this objective.

7.2.4. SILA Preliminary Evaluation Conclusions.

Although, it cannot be concluded that these findings will generalise to the target group of student users the study did indicate at an early stage that Peedy was perceived with more positive Affect than Genie which was one of the stated design aims. In addition to completing this early version of the Affect Scale these participants provided valuable feedback in the Focus Group discussion leading to improvements in the software implementation such as more use of non-verbal feedback in preference to spoken 'ok' utterances, synchronisation of animation with environment actions and improved SILA explanations along the lines suggested. Additionally, the final scale on the Affect Scale was removed, as mentioned, as it was ambiguous and difficult for participants to assess. Instead, the degree of 'believability' will be assessed qualitatively as reported in the sections 7.5.5 and 7.5.6 particularly.

With these important insights gained from the Preliminary Evaluation and subsequent software refinements the main evaluation was ready to take place with a view to gaining an understanding of the detailed issues involved when SILA is used by the target students.

7.3 SILA Evaluation Methods.

As stated, this Evaluative study is intended to validate, evaluate and assess SILA in order, primarily, to ascertain if the relevant Research Aims have been met. To this end the following sections recount the experimental design and procedure and present an overview of data analysis techniques used and their rationale. These are followed by discussions of the analysis that took place then followed by the results which are further subdivided into eight groups. Finally the conclusions are presented then the study is discussed in more general terms.

7.3.1. SILA Evaluation Design.

As the number of students at CHALCS was very limited and previous studies had therefore involved both CHALCS students and others from local schools it was felt that the final evaluation should also represent these groups. Therefore, students were recruited for the study from:

1. Current CHALCS Adavanced Level Physics Students (S1, S2);
2. Current Notre Dame 6^{th} Form College Advanced Level Physics Students (S5, S6, S7, S8, S9, S10);
3. The 16 year-old GCSE Physics students involved in the Wizard of Oz studies (S3, S4).

To elaborate, there were 2 CHALCS students, 6 Notre Dame students and 2 'Wizard of Oz' students. Additionally, a PhD student piloted the evaluation study procedure. Abilities of the students varied across the board ranging from one student being "top-of-the-class" for Physics to another student experiencing "psychological problems" resulting in poor attendance and consequentially poor results. Both of these students attended Notre Dame, the observations are as reported by their teacher. Of the 10 students 3 of them were female (S5, S6, S9) with again 3 of the 10 belonged to ethnic minority groups (S1, S2, S3).

Artefacts gathered for subsequent analysis during this evaluation included:

1. Observational notes (made by the experimenter during the studies);
2. SILA log (showing time, participant e.g. student/Peedy/Genie and 'move' e.g. 'me next', 'condensed', 'ok' and 'phrase' – see Figure 8.2 in Chapter 8 for an example);
3. The final summary document (including colour-code information i.e. SILA's addition, SILA's condensed phrase, student's addition – e.g. see Appendix K);
4. The SILA Affect Scale (see Appendix J);
5. A transcript of a semi-structured interview designed to help students think about summarisation;
6. A transcript of a post-evaluation semi-structured interview designed to elicit students' impressions of using SILA (see Table 7.2);
7. Notes based on a post-evaluation video analysis.

Semi-Structured Interview Question	
1.	Can you name any emotional reactions you may have experienced towards Genie and/or Peedy?
2.	Was there any change in your mood before and after using Genie and/or Peedy ?
3.	What improvements would you make to Genie and/or Peedy's personality ?
4.	What differences did you notice in Genie and Peedy, which would you prefer and why ?
5.	Please name your least and most liked aspects of Genie and/or Peedy's characters.
6.	What other kind of characters would you like to see ?
7.	Did you see Genie and/or Peedy as an equal or less or more adept at summarisation than you ?
8.	What in your eyes is a good summary, what should it contain ?
9.	Were Genie and/or Peedy helpful in constructing the summary ?
10.	How would you improve the summary you've created with Genie and/or Peedy (e.g. do you have enough detail) ?
11.	Is your summary laid out well, (e.g. have you got the key points in the right order) ?
12.	Are there any improvements you would like to make to the environment ?
13.	Could you do exam questions on the basis of these notes ? Would you have enough material ?
14.	If you had more time would you cut these notes down ?

Table 7.2 Questions Used in the Post-Evaluation Semi-Structured Interview.

The methodology used for the evaluation of SILA is a derivative of the 'think aloud' protocol. The rationale for this approach is that as a user interacts with a system they are asked to verbalise their thoughts, for instance elaborating upon their motivations for a particular menu selection, etc. However, a problem with the think aloud method is that subjects sometimes find it difficult to elucidate their criteria for selecting an interface function. In addition, the very act of verbalising thoughts results in an alteration of the process being described or, more fundamentally "the process of observation can alter the way that people perform tasks and so provide a biased view" (Dix et al, 1997). On the other hand, the think aloud method is a simple technique to carry out. It can also be useful throughout the design process, for example, using think aloud techniques in conjunction with storyboards before a system is actually implemented since it provides richer, user-centred information than observation alone. However, a more flexible approach than 'think aloud' as traditionally conceived would allow the experimenter to also encourage users to critically evaluate or comment upon the prototype (in addition to simply 'doing the task' and commenting on this). Such a flexible approach also allows the User to clarify points of confusion concerning the use of the system or task instructions.

One possible solution to the problems with the traditional 'think aloud' method is an alternative approach sometimes termed 'cooperative evaluation'. In this method it is emphasised that the experimenter and subject are allowed to interact. This in effect means that the experimenter can ask for clarification of key points which perhaps the subject did not suitably elucidate or for their part, the subjects may ask for explanations of system functionality which are not obvious to them. Of course, it is still important that the experimenter does not lead subjects but merely facilitates the studies as much as possible. Thus, for instance, in the case of the SILA evaluation, although key functionality may be hinted at during the beginning of the session this kind of scaffolded support is removed as the evaluation proceeds and the User becomes more proficient. In this case the experimenter must have in mind guidelines as to what is reasonable intervention. For example, in the SILA evaluation 'reasonable intervention' would be offering explanations of the turn-taking mechanisms at a general, abstract level but unreasonable intervention would be advising the User as to precisely which dialogue moves they should be selecting.

During this 'cooperative evaluation' method a number of techniques can be used for the purposes of recording the protocol (Dix et al, 1997). These consist of :

1. Paper and pencil – cheap but limited to the analysts writing speed;
2. Audio recording – useful for 'think aloud' but may miss vital actions;
3. Video recording – the analyst can see what the User is doing but camera angles need to be carefully chosen;
4. Computer logging – cheap and unobtrusive but very low-level;
5. User notebooks – useful in longitudinal studies or for recording unusual events but they are interpretive and usually entries are infrequent.

Furthermore, some tools do now exist to help with the subsequent protocol analysis but a clear disadvantage of these methods is the sheer amount of data generated and the work required to extract and analysise information from them. For instance, the SILA evaluation results in video, audio, logs, and notes all of which require analysis and, prior to analysis, video and audio require transcription or mark-up and log files also require annotation to extract significant events from low-level command sequences. The later Discussion section will further reflect upon some of the issues arising from this type of evaluation. Firstly, the SILA Evaluation experimental procedure is described.

7.3.2. SILA Evaluation Procedure.

The notebook computer running SILA was set up on a desk with a chair placed directly in front of it for the student to sit on and a video camera to the right of this focused on the notebook's screen. The experimenter then sat next to the student (either on the right or left, whichever was the most convenient). In addition to the video camera an audio tape recorder was also placed between the student and experimenter to capture the verbal discussions.

The first step of the study involved the experimenter reading from a prepared sheet of instructions. This ensured that each student was given exactly the same instructions prior to commencing the study. These instructions first explain the purpose of the study, the steps involved in the evaluation, a common English description of the 'cooperative evaluation' technique and a declaration of subject anonymity. This is designed to make the students feel at ease by reassuring them about the purpose of the video recording and it's use, it usually just takes one minute.

Following the instructions the experimenter then demonstrated SILA including all of the relevant menu commands, the course notes window (including right mouse clicking in order to use cut and paste) and the summary window (including the text formatting options, the edit menu and the SILA key). In addition, interaction with Peedy and Genie was described, i.e. right mouse clicking over a character in order to select a menu command plus the fact that they are awaiting a reply when they raise an upper limb to their ear, etc. Five minutes were allowed for this stage.

After the system has been successfully demonstrated and the student has no further questions a discussion took place about summarisation. The purpose of this discussion is to highlight techniques, methods and motivations for summarising course notes, particularly relating to Physics Study Skills. To this end, questions were asked such as: "What makes a good summary?", "How do you decide what to keep and what not to keep?" and "What kinds of things are important in a Physics document for instance?". These questions are based on transcripts of discussions that took place as part of the original Summarisation Studies described in Chapter 4. Again five minutes were allowed for this phase.

Following the discussion the student was encouraged to start using the system by selecting 'instructions' from the Help menu. Again, they are reminded to 'think aloud' and ask questions as appropriate. This stage usually took about 20 minutes to complete. It is important that prior to the student commencing use of the system the experimenter opens the log thus enabling the automatic gathering of the student's interactions with the system (as described and illustrated in section 6.8.11 in Chapter 6). Equally as important is that the experimenter also remembers to close the log when the student announces that they have completed their summary. Also, the experimenter should remember to save the summary document for later analysis.

The next stage in the evaluation was to commence the post-evaluation semi structured interview as per the questions above (Table 7.2). Note that although these questions are important they should serve as guidelines only. If something of interest occurred during the previous hands-on session then these questions should be elaborated or, indeed, if they reveal interesting answers they should also be expanded. This is the inherent flexibility of the semi-structured interview. The interview lasted no longer than five minutes.

Finally, the student was thanked and rewarded and the equipment and tools were made ready for the next student.

7.3.3. SILA Evaluation Data Source Rationale.

Table 7.3 presents each of the seven data sources together with the rationale for their utilisation, an overview of the analysis technique used (note that the next section further details this analysis) then finally, a reference to the results reported in section 7.5. The intention then is that it is possible to match each result to a data source hence gaining an insight into a result's general rationale and how it was arrived at.

7.4 SILA Evaluation Analysis.

As mentioned in the introductory Rationale section above, through an interplay between the top-down, research aim driven method and the bottom-up, data driven method and in common with the general methodological philosophy described in section 1.4 of Chapter 1 the following categories can be derived for the purposes of analysis (the number in brackets indicates the data source used to derive results in these categories by referring to those listed in Table 7.3):

1. SILA Ease-of-use and Effectiveness (6 & 7);
2. Effects of the SILA Environment and Tools (1, 6 & 7);
3. Results of Summarisation with SILA (6 & 7);
4. General Summarisation Results (3, 5, 6 & 7);
5. Student Affective Response and Perception (1, 6 & 7);
6. SILA Affectations Results (2, 4, 6 & 7);
7. Effects of Carrying out the SILA Evaluation (1, 6, & 7);
8. Evidence of Student Summarisation Learning (3, 6, & 7);
9. Potential SILA Future Work Arising from the Evaluations (6 & 7).

DATA SOURCE	RATIONALE	ANALYSIS TECHNIQUE	RESULTS
1. Observational Notes	to record any general point-of-interest not easily covered by other data sources	group observations into more general classes	B1, B3, D2, D6, E1, E5, G3, G4
2. SILA Log	to record student-SILA interactions looking for effectiveness etc.	categorise types of exchanges to provide quantitative summary	F4
3. Summary Document	to record final summaries to ascertain quality, look for improvements and general techniques	use marking criteria to (quantitatively) assess summaries plus note any (qualitative) peculiarities, e.g. formatting, use of titles etc.	D5, D7, H2, H3
4. Affect Scale	to ascertain student's impression of assistant's affectations	total scores on likert-type scale, statistically anlalyse, plot graphs, looking for trends	F11, F12, F13
5. 'Summarisation' Interview	to assess and facilitate student's prior knowledge of summarisation techniques, strategies, etc.	group similar results together to produce 'categories' of responses	D8
6. Post-Evaluation Interview	to assess student's impressions of using SILA, e.g. moods, SILA effectiveness, future work, etc. (see Table 7.2)	group similar results together to produce 'categories' of responses	A2, A4, A5, A6, A7, A8, B2, B4, C2, C4, C5, C7, D1, D2, D3, D4, D5, D7, E1, E2, E3, E4, F1, F3, F5, F7, F8, F9, F10, G5, I1, I2, I3, I4, H1
7. Video Notes	to capture subtle nuances of SILA-student-experimenter interactions not captured with other data sources	group similar results together to produce 'categories' of responses	A1, A3, A5, B3, C1, C3, C6, D6, E5, F2, F6, F8, G1, G2, H4, I2

Table 7.3 SILA Evaluation Data Source Rationale.
(Key: Results Reference refers to Results in Section 7.5)

These categories then provide a framework by which to analyse the data utilising both qualitative (for example, student's answers to a post-test semi-structured interview) and quantitative (for example answers to the SILA Affect Scale) techniques. Each of these categories is visited in turn in the subsequent Results section.

First though the methods utilised for analysis of the 7 data artefacts are described. As the student is using the system the experimenter made notes on issues that arose during the course of the interactions. These can be general issues such as preferences for using either Peedy or Genie or more acute observations such as noting when a student is displaying affective behaviour, such as waving at the computer. In addition, the summarisation discussion and the post-test evaluation resulted in transcripts relating to the study albeit more focused than the observational notes. Furthermore, the video of the computer screen was analysed with a view to elaborating upon the observational notes or capturing the subtleties of student-system interactions as outlined in the computer log. Thus observational notes were produced based on the video analysis. These transcripts and notes were analysed in terms of the categories described above which emerged as a result of this analysis, the analyst looking for comments by the participants which are pertinent and interesting, i.e. particularly in terms of the SILA Evaluation Objectives of section 7.1. These comments then form 'codings' (Strauss & Corbin, 1998) which are subsequently grouped and triangulated with other data to arrive at conclusions.

In addition to the written data mentioned above the system also provides a log of interactions as previously discussed. The mechanisms by which this artefact was analysed consisted of categorising certain move combinations as being of one type then matching the log with occurrences of these exchange types. This resulted in 16 exchange types which in turn were placed into 4 categories of the exchange either being related to a positive (confirming) collaboration, or a negative (disafirming) collaboration, an enquiry dialogue or an SILA act exchange. For example, Table 7.4 shows a positive collaborative exchange whereas Table 7.5 shows a negative collaborative exchange, Table 7.6 shows an enquiry exchange and Table 7.7 shows a SILA act exchange. Appendix L further elaborates upon the exchange categories and includes the appropriate tally for each student.

Agent	Move
SILA	me next/start
(student	OK)
SILA	look

Table 7.4 Positive Collaboration Exchange Type Example: *agent_turn_ok.*

Agent	Move
SILA	me next/start
student	no
SILA	look

Table 7.5 Negative Collaboration Exchange Type Example: *agent_turn_ignore_student.*

Agent	Move
student	added
SILA	why
student	because

Table 7.6 Enquiry Exchange Type Example: *student_explain.*

Agent	Move
SILA	condensed

Table 7.7 SILA Act Exchange Type Example: *SILA_condensed.*

The summaries produced by the SILA-student collaboration were marked by the Physics 'expert' (that is the author of the original WebCT material) and a Study skills 'expert' (a University Academic). A 'Summary Feature Appraisal Scale' was utilised which was based on Tawalbeh's Scale of the same name (Tawalbeh, 1994, page 160) with modifications appropriate to the differences between Tawalbeh's and these studies. Notably, Tawalbeh was more concerned with 'abstracting' than 'summary note-taking'. The scale breaks the summary down into the four phases shown in Figure 7.4: selection, staging, condensation and organisation:

SELECTION

4 Summariser selects **all or nearly all** important information and includes **no or nearly no** unimportant information in text;

3 Summariser selects **most** important information and includes **little** unimportant information in text;

2 Summariser selects **some** important information and includes **some** unimportant information in text;

1 Summariser selects **little** important information and **a lot of** unimportant information in text.

STAGING

4 The Summariser **clearly** signals the relative importance of ideas in an appropriate sequence and with no distortion of the meaning of the text;

3 Summariser **for the most part** clearly signals the relative importance of ideas in an appropriate sequence and with no distortion of the meaning of the text;

2 Summariser **sometimes** signals the relative importance of ideas in an appropriate sequence and with little or no distortion of the meaning of the text;

1 Summariser for the most part **does not** maintain the relative importance of importance of ideas in an appropriate sequence and includes some material which significantly distorts the meaning of the text.

CONDENSATION

4 Summariser skilfully employs **all or nearly all** appropriate summarisation strategies to transform and condense sentences;

3 Summariser skilfully employs **some** appropriate summarisation strategies to transform and condense sentence;

2 Summariser employs **few** appropriate summarisation strategies to transform and condense sentences;

1 Summariser **does not** employ appropriate summarisation strategies to transform and condense sentences.

ORGANISATION

4 Summariser **skilfully** varies sentence construction to maintain the readability of the summary with **few or no** grammatical and mechanical errors;

3 Summariser **usually** varies sentence construction to maintain the readability of the summary with **few** grammatical and mechanical errors;

2 Summariser **sometimes** varies sentence construction to maintain the readability of the summary with **some** grammatical and mechanical errors;

1 Summariser **does not** vary sentence construction to maintain the readability of the summary with **many** grammatical and mechanical errors.

Figure 7.4 The Summary Appraisal Scale.

Thus, by using a scale for each of these phases of summarisation ranging from 4 marks for a good attempt to 1 mark for a poor effort in each phase it was possible to derive an overall mark for the quality of the summary produced by each student.

Additionally, the guidance notes shown in Figure 7.5 were provided elaborating on the four stages:

SELECTION – look for inclusion of information important to the meaning of the original text (e.g inclusion of high level rhetorical predicates) and exclusion of information not considered central to the text meaning;

STAGING – recognising and accepting the original (rhetorical) structure of the text and/or interpreting a novel and suitable structure thus maintaining a relevant relative structure;

CONDENSATION – employment of condensation summarisation strategies (e.g. replace, construct, incorporate, interpret, delete, etc. – see Tawalbeh's Model of Summarisation), number of words (amount of condensation);

ORGANISATION – use of cohesive ties and discourse markers, sentence expansion techniques, variation of sentence opener, number of sentences used (amount of cohesion), grammatical and mechanical errors.

Figure 7.5 Summary Appraisal Scale Guidance Notes.

Furthermore, a number of other documents were supplied including Tawalbeh's model of summarisation (see Appendix C), a rhetorical analysis of the original course notes (see Appendix H) and, of course, a hardcopy of the course notes themselves. The actual instructions given to the markers are shown in Appendix P. Additionally the differences between the different coloured texts were explained such that "red = student-entered phrase shown to SILA; blue = phrase SILA has added; magenta = phrase SILA has 'condensed'".

Finally, the adjective-based Affect Scale relied on a numerical scale which students selected ranging from –5 to +5 for each of the 7 Affecations (an excerpt was shown in Figure 7.3. Although, the final Affect Scale does not include the contrived-genuine scale). The individual Affectations across all students were summed to arrive at an overall impression of students' perceptions of each Affectation for Genie and/or Peedy as well as each individual student's overall impressions. That is, the *valence* of the student's perception of affect is ascertained by the overall total of the Affectations' score. This is in-keeping with the notion of *valence* as defined by researchers. For example, Picard (1997, p.6) defines valence as: "whether you feel positive or negative towards something". This concept is also utilised in the OCC appraisal model of emotion described in section 2.4.4 of Chapter 2. However, Picard states "For mood these descriptions are sufficient for most purposes. However, we know that there is more differentiation among emotions than simply valence and intensity. A negative input that contributes to anger may not contribute to sadness, and vice-versa" (Picard, 1997, p. 158). Additionally, Picard differentiates between physical and cognitive inputs to an emotional system but goes on to state that "...the majority of people do not make such distinctions. It is common to hear someone say 'I am feeling pretty good' or 'I am not feeling so good..." (Picard, 1997, p. 159). In this way, the Affectation score should provide an indication of the overall perception of 'mood' of the agents in the eyes of the student Evaluation participants. (The Affect Scale is included in Appendix J).

Note that the data sources, their rationale, analysis and results will be summarised in the next section which describes the results of the SILA Evaluation.

7.5 SILA Evaluation Results.

This Results section is subdivided into the following categories which emerged through an interplay with the data. The results are drawn from the 7 data artefacts previously highlighted. Each result has an associated reference (indicated in square brackets) referring to the results column of Table 7.3 and used subsequently in the conclusions – see Table 7.13.

7.5.1. SILA Ease-of-use and Effectiveness.

Figure 7.6 shows the results of the exchange category analysis for positive collaborative exchanges (see Appendix L for the full results table). As can be seen on these graphs, the majority of exchanges were of a positive nature with negative collaborations (i.e. Figure 7.7, disagreements) being in the minority. This is due to students, in the main, agreeing to collaborate with SILA to perform the summarisation task, i.e. letting SILA take turns and offer suggestions by adding something to the summary construction window. In addition SILA itself is designed to be more positively collaborative than negatively so this result, in terms of SILA interactions, is not unexpected.

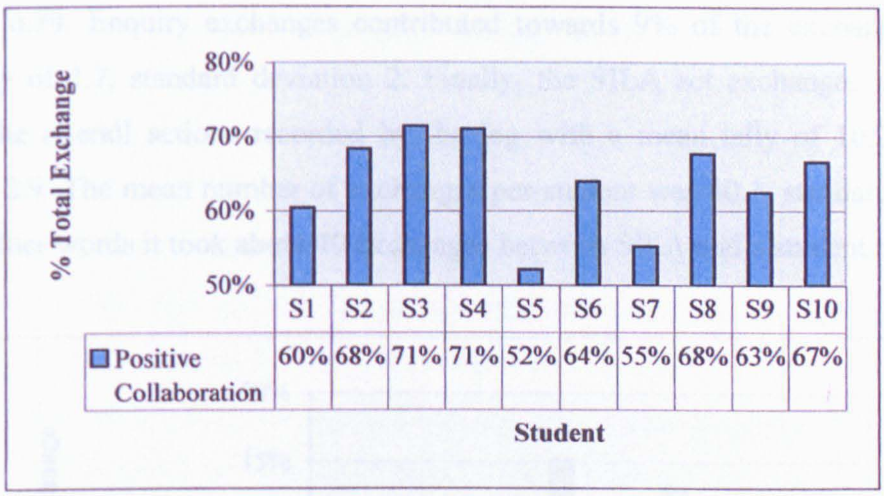


Figure 7.6 SILA-Student Positive Collaboration.

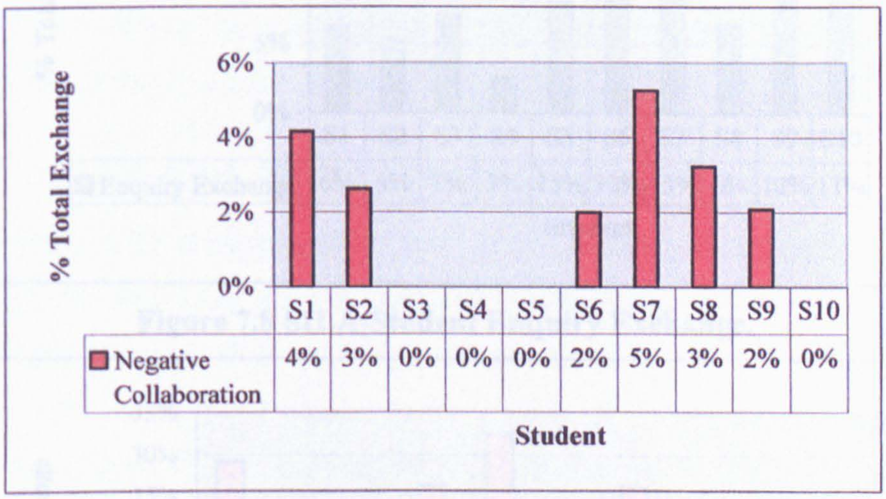


Figure 7.7 SILA-Student Negative Collaboration.

In addition Figure 7.8 shows that a moderate amount of time was spent in so-called 'enquiry' activities. These activities would involve generating explanations of summarisation behaviour. For example SILA may explain that a word was added because "this is the title !" or the student may offer an explanation for including a phrase in their summary with which SILA is unfamiliar. However, students did at times find writing explanations difficult [A6]. One student (S6) said "I really didn't know what to write down" in response to SILA's request for an explanation. Finally, Figure 7.9 shows the percentages of 'SILA act' exchanges that occurred in the logs. These are basically SILA 'condensing' or 'organising' a summary – as explained in Chapter 6.

The positive collaborations contributed 63% of the overall total exchanges with a mean tally per student of 25.4, standard deviation 4.12. The negative collaborations contributed 2% of the total exchanges with a mean tally per student of 0.8, standard deviation 0.79. Enquiry exchanges contributed towards 9% of the exchanges with a mean tally of 3.7, standard deviation 2. Finally, the SILA act exchanges contributed 26% of the overall actions recorded by the log with a mean tally of 10.2, standard deviation 2.9. The mean number of exchanges per student was 40.1, standard deviation 7.58. In other words it took about 40 exchanges between SILA and a student to write the summary.

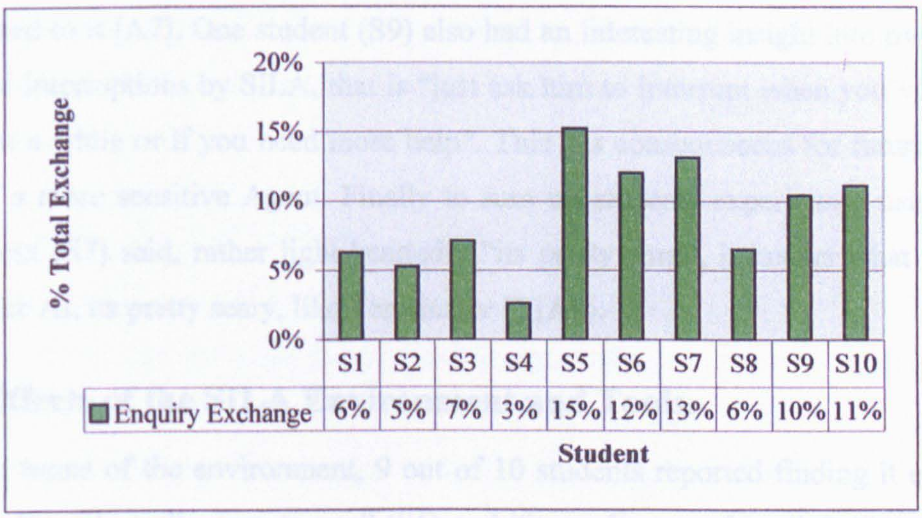


Figure 7.8 SILA-Student Enquiry Exchange.

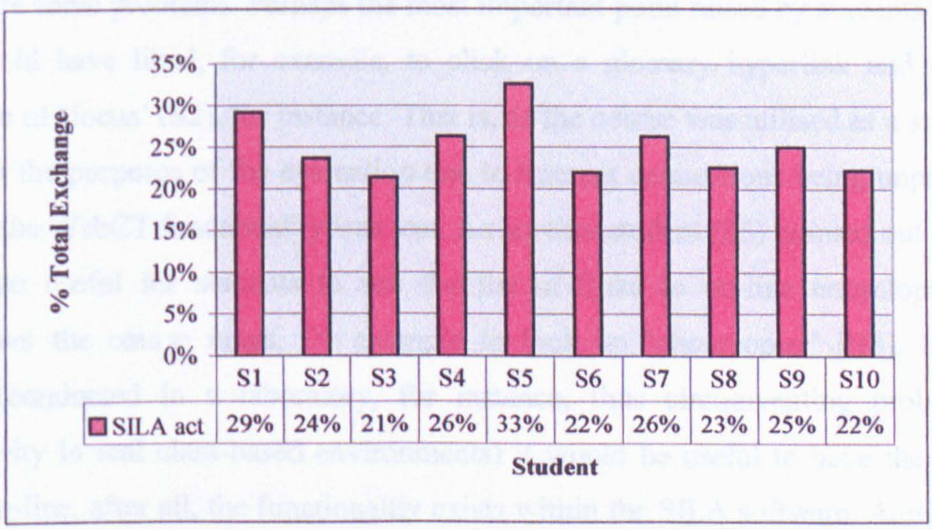


Figure 7.9 SILA Act Exchange.

All students quickly became accustomed to interacting with SILA albeit, in 7 out of 10 cases, after initial guidance by the experimenter [A1]. 9 out of 10 students also felt that the method of showing SILA text was easy to use although one student (S2) said it was "too confusing" [A2]. It did, however, emerge that SILA used the OK response in too many situations leading to ambiguity [A3]. It also emerged that the colour-coded text was sufficiently understood by all students although the purpose of the indentations (to show relative importance of phrases) was not so obvious [A4]. There was also evidence of 'edutainment' in the form of one student (S6) laughing and giggling and, in particular, choosing a character for a little light relief [A5]. As one student (S10) said "it stopped a program from just being a program, it was a bit of fun that was included with the software" although an interesting supposition was reported by the same student (S10), notably that students will continue using one character because they become accustomed to it [A7]. One student (S9) also had an interesting insight into overcoming unwanted interruptions by SILA, that is "just ask him to interrupt when you've stopped typing for a while or if you need more help". This has consequences for future work in terms of a more sensitive Agent. Finally to sum up students experiences using SILA, one student (S7) said, rather light-heartedly "its pretty smart, it knows what its doing almost like AI, its pretty scary, like Terminator !" [A8].

7.5.2. Effects of the SILA Environment and Tools.

In terms of the environment, 9 out of 10 students reported finding it eminently user-friendly, "its really easy to use" (S6) and "its really user friendly" (S4), although there were some problems. Perhaps the most important point raised by students was that they would have liked, for example, to click on a glossary hyperlink and obtain a definition of 'focus' (S2), for instance. That is, as the course was utilised as a static web page (for the purposes of the evaluation due to Internet connections being unpractical), some of the WebCT functionality was lost. As another student (S5) pointed out it would have been useful for students to use the list of links to on-line encyclopaedia to supplement the course notes, for example to look up 'photocopier' [B3]. In future studies (conducted in a laboratory, for instance, thus circumventing problems of connectivity in real class-based environments) it would be useful to have the WebCT course on-line, after all, the functionality exists within the SILA software. Additionally, one student (S1) commented that it would be useful if the environment could be tailored, saying "it would be good if you could customise the environment for the actual user" [B2].

Although care was taken in authoring suitable guidelines for the 'organisation' help material unfortunately, 9 out of 10 students did not use it [B1]. This could be, as one student (S4) reported, because the summary to be constructed was so small that he felt it did not require any organisation, "I think the passage there is like short enough for you to be able to know how to organise it. If it was a longer passage it would probably be more useful." Although as we shall see later in evaluating the quality of summaries there was scope for improvement here. However, one student (S1) did read about improving the grammar in a summary in the 'organisation' notes then subsequently amended their summary plus another student (S3) said that they might have used it if they were to spend time reflecting on the summary. Finally, 8 out of 10 students found the instructions useful and relevant although one student (S10) would have liked to have navigated back 'a page' and another (S5) thought they were too fast [B4].

7.5.3. Results of Summarisation with SILA.

Table 7.8 shows the results for each student for each of the markers. The scores are broken down into the score for each of the summarisation phases (selection, staging, condensation and organisation, as explained above in section 7.4) then totalled to give a result from each marker. The average mark of the two markers is then given (this is the overall mark as used in subsequent analysis) together with its equivalent percentage. These marks will be correlated with other data in the later 'conclusions' section. Note that 'marker 1' is the 'study skills' expert and 'marker 2' is the Physics expert.

From Table 7.8 it can be seen that the two markers were generally within 1 or 2 marks of agreement with each other on the overall mark for each summary's quality. We can see that the summary scores are generally high indicating that most students produced a good summary. Looking at the summaries (see example in Appendix K) it can be seen that one immediate effect of interacting with the agent is that none of the summaries are simply "cut and pasted" without condensation or use of own words (as was often the case in the study reported in Chapter 4). There is some tendency for the variability in summaries to be reduced – summaries made with the agent are more alike than summaries made by students on their own, evidencing the high acceptance of SILA's contributions to the final product. The summaries from these and the previous studies reported in Chapter 4 will be compared in section 7.5.8.

	S1	S2	S3	S4	S5	S6	S7	S8	S9
Marker1									
selection	2	2	3	3	2	3	3	4	2
staging	3	2	3	3	3	4	3	4	3
condensation	2	2	3	2	2	3	2	3	2
organisation	2	2	4	3	3	4	3	4	3
Total	**9**	**8**	**13**	**11**	**10**	**14**	**11**	**15**	**10**
Marker 2									
selection	3	2	3	3	3	4	3	4	3
staging	3	2	3	3	2	4	2	4	2
condensation	2	2	2	2	2	3	2	3	2
organisation	2	2	2	2	2	3	2	4	2
Total	**10**	**8**	**10**	**10**	**9**	**14**	**9**	**15**	**9**
mean score	**9.5**	**8**	**11.5**	**10.5**	**9.5**	**14**	**10**	**15**	**9.5**
% score	**59**	**50**	**72**	**66**	**59**	**88**	**63**	**94**	**59**

Table 7.8 Evaluation Summary Scores.

One student (S2) raised the interesting point of the purpose of the summary they were to produce, wondering if the aim of the software was to teach them how to write summaries or help with Physics notes [C1]. It is interesting that no one else raised this point, instead assuming that they were just creating their own notes for revision purposes. There is, indeed, a fine distinction between the software coaching generic summarisation skills or helping with creation of notes based on the Physics material. Of course the answer is that the system is providing both facilities. It is producing (it is hoped), through collaboration with the student, a *product* in the form of revision notes but it is also providing an opportunity for the appreciation of the *process,* i.e. summarisation, which should have an effect on student's other subject specialities.

The condensation strategy is a central technique of the software in terms of summarisation. Interestingly, only 1 student out of 10 (S2) had a problem understanding the keywords which this strategy utilises, saying "it's not quite obvious" although one (S10) does say "it makes it better and simpler" [C2]. Indeed, one student in fact stated "it turned out he was pretty good at condensing it" (S7). Two other students seemed positively impressed by the condensation strategy, saying of it "the summary is shorter now" (S7) and "put it in a better order where things are related to each other" (S10) [C5]. However, a main drawback of the condensation strategy was noticed by one student (S6) who pointed out that a phrase ("& onto the retina") had been appended to another unrelated phrase [C6]. This was due to the fact that the phrase to be appended had been matched with the new phrase on the basis of other key words but the one key keyword "the eye" was missing from the original thus resulting in the seemingly unrelated condensation. There was only one report of this kind of behaviour but nonetheless it is important since students may have a tendency to be uncritical of the agent's condensation suggestions (and given evidence from exchange analysis, may have generally been a little too ready to accept suggestions without explanation), and so any such errors may not be corrected by the student. One of the summary markers commented on this critical thinking. However, as discussed later, students were critical of SILA's strategies, particularly during the selection phase. In other words this propensity to accept SILAs condensations may reflect a lack of experience during this phase of summarisation. On the other hand, a SILA 'undo' facility may have made this student acceptance less prevalent. This is discussed further in the Conclusions.

In terms of summarisation strategies employed by the students 8 out of 10 would first read the notes then type their own summary notes in the summarisation window. However, two students (S3, S8) actually cut and pasted whole sentences from the Physics notes window into the summary window then would amend the sentences as necessary [C3]. It was also observed that only one student (S1) would amend SILA's contribution to the summary as they went along although two students (S3, S8) would delete an entire contribution from SILA, for example deleting "if we have no sight defect we can see clearly" because it was "common sense" (S8) [C4].

Most students also criticised SILA for including questions in the summary, one in particular representative comment being "it's a question isn't it, its not like a fact is it ?" (S7) [C7].

7.5.4. General Summarisation Results.

As to general summarisation issues arising from the evaluation, one of the most commonly-held beliefs was that summarisation could be enhanced by having a copy of the syllabus to hand of the course being studied [D1]. As one student (S2) said "at the end of the day you'd know you'd covered everything then". Another student raised the point of the syllabus being available on-line. An obvious conclusion concerning these points is to have a link within WebCT to the syllabus so that students can utilise it as part of the summarisation process.

6 out of 10 students, on reflection, considered at the end of the task that they had a good summary, saying for example "I've got enough of the gist of the material from that" (S7). However, one student (S5) said "it'd be helpful but I don't think I can depend on them" [D2]. In addition another student (S8) stated that they would make minor editorial changes, especially transformation of a long sentence into a shorter one. Furthermore, two students (S9, S6) stated that they would redraft the final summary [D4], one (S6) stating "I might go over it again cause I just do that...rewrite things" and in fact, did edit the final summary before being satisfied with it [D6].

In terms of techniques used in the summary creation two students (S6, S8) used the same technique of separating groups of related phrases by blank lines thus creating a 'chunking' effect in the summary [D5]. This was a completely untutored phenomenon so it was interesting to see two students use the same technique and may well suggest inclusion in future supports. 8 out of 10 students utilised short concise 'points' in their summary rather than grammatically correct prose, one student (S6) saying "one long sentence wouldn't help at all". This is in-keeping with strategies (such as the condensation techniques) employed by SILA. However, one student (S8) utilised a list and one long sentence and another (S9) utilised prose [D7]. Indeed, it became clear from the pre-evaluation discussion that each student in the Evaluation studies had a very good idea of what a summary should contain, the nature of the summarisation process itself and of techniques to be used in the summary. To elaborate, one student (S7) stated that a summary should be "short, to the point, very clear and self-explanatory"[D8]. Of course the techniques usually employed did vary between individuals, ranging from the use of diagrams, highlighting, lists, shorthand and copying out text and redrafting (S9) to making concise points (S8) (see Appendix K).

7.5.5. Student Affective Response and Perception.

Most students had a generally positively valenced (Picard, 1997) reaction to using SILA, Peedy in particular [E4]. For example one student reported "I'll stick with the parrot, he's doing a pretty good job" (S7) and adding "that's pretty good he's got a good point there". This same student actually waved 'goodbye' to Peedy as he exited the environment and said of Peedy "very helpful, almost like a kinship, you'd feel like he was really talking to you, not just saying what he was supposed to say" [E1]. This notion of empathic response in SILA is of course programmatically unfounded, in the strictest sense, as Peedy has no representation of the student's affective states or of its own. However, that a degree of empathy can be induced in the student indicates the power of the Affectations to create the illusion of an 'emotional life'. Other responses include one student repeatedly firmly pressing the mouse button suggesting annoyance (S3), speed-typing when suggesting feeling under pressure from SILA (S10), talking to Genie and/or Peedy even though the Evaluation agent did not possess speech recognition (S10), perception of sadness in Peedy, S9 saying he "looks a bit sad" on a denial of his request to go next and giggling (S9) or laughing (S5) at certain animation sequences. As one student (S6) said "the little parrot made me laugh" [E3].

These reported affective responses to SILA were perhaps most significantly accompanied by a relaxation in most of the students [E2]. One student (S1) stated "I used Peedy and I calmed down". In particular, the student cited in the previous paragraph (S7) who experienced the empathic phenomenon described their feelings before and after using SILA: "before writing the summary I was a bit on edge and nervous because I'm not really good at writing them...when I got the help from Peedy it helped a lot actually." However, not everyone had a positive reaction towards SILA.

One (quite able) student (S8) had quite a negative reaction to SILA [E5]. Overall he was not impressed by the advice that Peedy gave in constructing his summary but also there was almost a personality clash between the two. The student developed an 'attitude' in their interactions with SILA. One particular utterance made by the student to Peedy was "alright, don't get smart" then in relation to the Genie, "we won't have him 'cause he's irritating". In addition at one point he told Peedy to "shut up man". When later quizzed as to how he felt during these interactions the student stated that "I don't tend to get emotionally involved with computer characters". However, his behaviour suggested the opposite was true at least in this instance. Unfortunately, this particular student's perceptions of Peedy and Genie's personality by means of the Affect Scale were not collected but would probably have been negative which fits with the pattern of Affective Perception for Peedy later described in the section 7.6.

7.5.6. SILA Affectations Results.

This section is concerned with the Affectations of SILA. As one student (S8) pointed out both Peedy and Genie have a "personality in the sense that they have different characteristics" [F10]. So here it is pertinent to report student's impressions and reactions to these various 'characteristics', notably: speech, actions, language, etc. First though the following figures (Figure 7.10 and Figure 7.11) show the overall impression of the Affectations of both Genie and Peedy for the students that took part in the final Evaluation. Note that Table 7.9 provides the affectations for the indexes shown on the graph. These results can be compared to the preliminary results in Figure 7.1 and 7.2 with the exception that the contrived-genuine scale was not used in final Inventory (because it was deemed ambiguous by participants in the Preliminary Evaluation). Note that data are summed across all 9 students[1] for each affectation, the result being a percentage of total possible marks for each affectation. As can be seen Genie tends towards more negative/mediocre affectation experiences whereas Peedy is perceived more positively [F11]. Genie's dominant affectation is 'macho' whereas Peedy's is 'friendly' as would be expected [F12]. Interestingly, Peedy is seen as mildly 'impatient' (and similar in scale to Genie). Both are seen as 'encouraging', 'amenable', 'friendly' and 'polite' with Peedy scoring higher than Genie for these Affectations [F13].

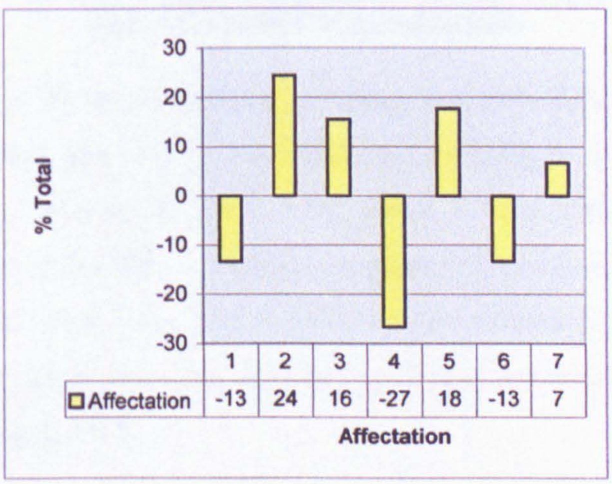


Figure 7.10 Final Affect Scale Results for Genie.
(see table 7.9 for the Affectation Indexes)

1 The experimenter was responsible for not obtaining an Affect Scale from one student (S8). This student has been omitted from these and the final results (to follow).

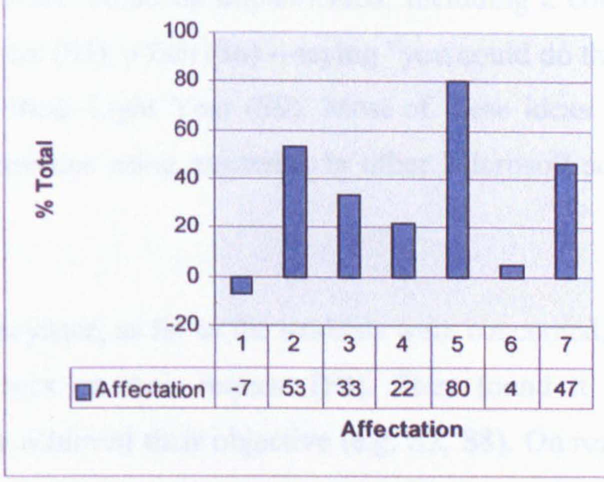


Figure 7.11 Final Affect Scale Results for Peedy.
(see table 7.9 for the Affectation Indexes)

Affectation Index	negative affect	positive affect
1	*impatient*	*patient*
2	*derisive*	*encouraging*
3	*obstinate*	*amenable*
4	*macho*	*wimpy*
5	*scary*	*friendly*
6	*curt*	*laid-back*
7	*impolite*	*polite*

Table 7.9 Affectation Adjectives.

Perhaps one of the most significant findings was that students did not like SILA interrupting them when they were either thinking or typing [F1]. Reports included "I didn't like when he interrupted" (S3), "the parrot is too impatient" (S6) and "its constantly butting in" (S8). This is further compounded, as observed with one student (S2), when Genie or Peedy may add a phrase to the summary document just as the student was about to add their own contribution resulting in confusion and wasted effort on the part of the student [F2].

In terms of the animations found in both Genie and Peedy one particular sequence where they explain their actions to the student tends to take slightly too long, particularly in Peedy's case [F3]. As one student (S7) pointed out "he takes a bit long about thinking about it". When asked if they found the characters suitable 8 out of 10 students agreed that they were appropriate although one student (S2) thought that they would appeal to a younger age range more, saying "I think for kids those two are excellent, specially the parrot cause it reminds me of one of the Disney characters in Aladdin, so…something they'll be able to identify with" [F5]. They suggested a number of other characters which could be implemented, including a computer (S1), Einstien (S2), a robot (S2), a cat (S3), a fish (S6) – saying "you could do the bubbles coming out of his mouth!" and Buzz Light Year (S9). Most of these ideas for further characters were based on experiences using assistants in other Microsoft software or, of course, experiences in films.

A further annoyance, as far as the students were concerned, occurred when SILA refused to end its turn on their request [F4]. They found it irritating at best but persevered until they achieved their objective (e.g. S2, S8). On reflection this particular design in the software may be too 'overbearing' for the students, possibly benefiting from a more negotiated ending rather than a point blank refusal. A further unfortunate misunderstanding was reported when one student (S3) construed Peedy's language as being slightly sarcastic and therefore critical, that is, when asked for an explanation saying "Wow ! Tell me why you did that !" [F6]. The student took this as a criticism. Again this raises questions concerning the age, ability or self-esteem of the students and their perception of the subtleties of language and prosody used.

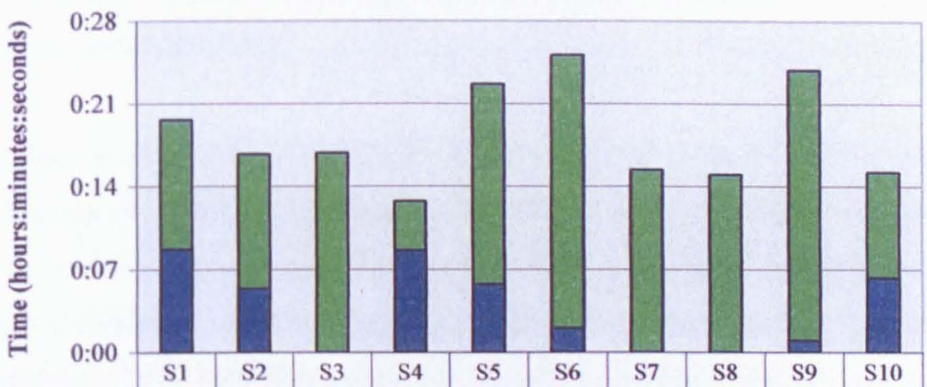
In terms of distinguishing between the two characters, Genie and Peedy, students cited language, voice, appearance and interruption time as the main factors which is unsurprising as these (as stated in Chapter 6, section 6.4.3.5.4) are the main variables to manipulate to create the Affectations [F7]. To illustrate, of the Genie's language one student (S5) who liked Peedy said "the Genie can be quite sarcastic", one student (S4) distinguished the two by voice "I think the voice is slightly different", another student (S9) justified her preference of Peedy by saying "maybe its 'cause I'm a girl and he's really cute" finally, one student (S8) preferred Peedy "because it doesn't interrupt as much and just lets you get on with it". In addition, one student (S5) may have found Genie too domineering, it actually added five phrases before she switched to Peedy then continued [F8]. Although this may have been because the student was allowing Genie to demonstrate the summarisation task. This will be revisited in section 7.5.8. She said of Peedy "he seems more user-friendly than Genie". Of course this kind of result was to be expected as the two characters were designed to be different in this way. However, it may have been the case that Genie's behaviour had the effect of putting off all potential users.

Finally, two students reported that the voices were unrealistic, one (S7) unfortunately joking that they sounded like Stephen Hawking ! This does, however, have serious consequences in terms of the perceived affect for, as one student (S10) pointed out "their voices were a bit cold because they were robotic" [F9].

7.5.7. Effects of Carrying out the SILA Evaluation.

Certain effects may be due to the experimenter's presence or intervention in the study together with problems with the SILA software and are recorded here for the sake of completeness and in the spirit of open reporting.

First, if the experimenter saw that a student was not working too well with a character then they would make it perfectly clear that the other characters existed and may warrant experimentation. Unfortunately, this sometimes resulted in the students changing characters immediately the suggestion was made [G1]. However, this occurred in four out of the ten individual cases, notably in the cases of S1, S4, S5 and S10. In the case of S1 the suggestion to change was made as the Genie seemed to be dominating the interactions. S4 changed from the Genie to Peedy too at a suggestion but said "I was about to change it" as did S5 who after a slight delay following a prompt stated of Peedy that she "just wanted to see what he was like". The intervention was more direct in the case of S10 who appeared to be getting stressed due to interactions with the Genie, (suggested by speed typing in order to keep up with this assistant). The student changed to Peedy immediately the suggestion was made. Aside from these four students three students (S3, S7 and S8) all started with Peedy and were content to stay with their initial choice. In fact S7 did attempt to end Peedy but Peedy refused so the student "decided to stick with him". The remaining three students (S2, S6, S9) did change assistants but completely of their own volition. In the ensuing discussions S6 revealed that "I just like the cartoon characters and I'd switch between them", laughing as she said this, i.e. the assistants hold some kind of entertainment value for her although she would like this to be varied. S9 also said of Peedy "I prefer this one, this one's funnier". Figure 7.12 provides an analysis of each student's interactions with each of the two assistants in terms of the total time taken, according to the log. It can be seen that with the exception of S4 students prefer to spend more time using Peedy to help with their summarisation task.



	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
Peedy	0:11:17	0:11:37	0:17:18	0:04:17	0:17:24	0:23:44	0:15:51	0:15:24	0:23:31	0:09:11
Genie	0:08:55	0:05:34	0:00:00	0:08:50	0:05:56	0:02:08	0:00:00	0:00:00	0:01:00	0:06:29

Student

Figure 7.12 An Analysis of the Proportion of Time Spent with Each Assistant.

Other interventions proved necessary in order to obtain results from the study. For instance, and as suggested in the 'cooperative evaluation' methodology the experimenter guided students through the first two or three interactions with SILA, making sure not to make the student's decision for them but merely facilitating their desired intentions in terms of SILA interactions e.g. negotiating the menus. After initial guidance students were generally able to use the system unaided. One student, however, (S10) needed prompting to interact with SILA and participate in creating their own summary as they were quite willing to allow SILA to construct the entire document [G4]. This is something which will be returned to in the final chapter when making recommendations for further work since it may be necessary to make sure some students do take turns rather than using the tool as an auto-summariser if they are to learn to take their own notes. One alternative might be to introduce 'sulking' behaviour on the part of SILA - refusing to play unless the student does their 'fair share'.

In addition the experimenter was responsible for not obtaining an Affect Scale from one student (S8). This student has been omitted from the final correlative conclusions. The log file for one student (S1) was not opened so the study was started again. This should not have had an effect on results [G3].

There were also some errors due to the software crashing. In one instance (S10) the software crashed right at the end of the session as 'save' was chosen by the student thus the summary was not saved and so they have also been omitted from the subsequent 'conclusions' [G2].

Finally, it became apparent that the study could have benefited if the system had been on-line for, as already mentioned, hyper-linked glossary entries could have been utilised as could links to external resources such as on-line encyclopaedia [G5]. However, the constraints of conducting the studies in the field at various locations made this possibility severely impractical, e.g. the necessity of using a real classroom without Internet connectivity.

7.5.8. Evidence of Student Summarisation Learning.

In terms of SILA tutoring students, one student in particular (S5) allowed SILA (Peedy) to demonstrate what was expected by taking the first few turns and continually asking for explanations then eventually taking back the initiative and continuing with the summarisation task [H1]. S1 also allowed SILA (Genie) to first demonstrate summarisation although this may have been less consensual as the Genie dominated interactions. In contrast though another student (S3) first completely ignored Peedy's requests to take a turn and when Peedy did take a turn would ignore it's contributions, i.e. carry on working on the sentence they were originally editing. In fact, Peedy's intervention was counter-productive as the student spent too much time moving the assistant to one side, saying "I didn't like when he interrupted", until they finally acquiesced thus making the collaborations more equal.

Figure 7.13 graphically compares the additions to the summaries made by SILA and the students as a percentage of each student's overall number of exchange categories. The overall mean number of additions SILA made to the summaries as a percentage of the total number of exchange categories per student was 21%, Standard Deviation 0.04 whilst the overall mean number of additions the students made again as a percentage of each student's exchange category total was 10%, Standard Deviation 0.04. It can be seen that S5 and S1 made far fewer contributions to the summaries as mentioned above. Furthermore, S10 and S3 made the most contributions themselves although the balance of additions between S3 and SILA is almost level (18% SILA, 14% S3). Unfortunately, as mentioned S10's summary was lost however S3 did create one of the better summaries, as judged by the markers (see Table 7.8). This would seem to indicate that this 'balance of power' may lead to a successful summarisation effort.

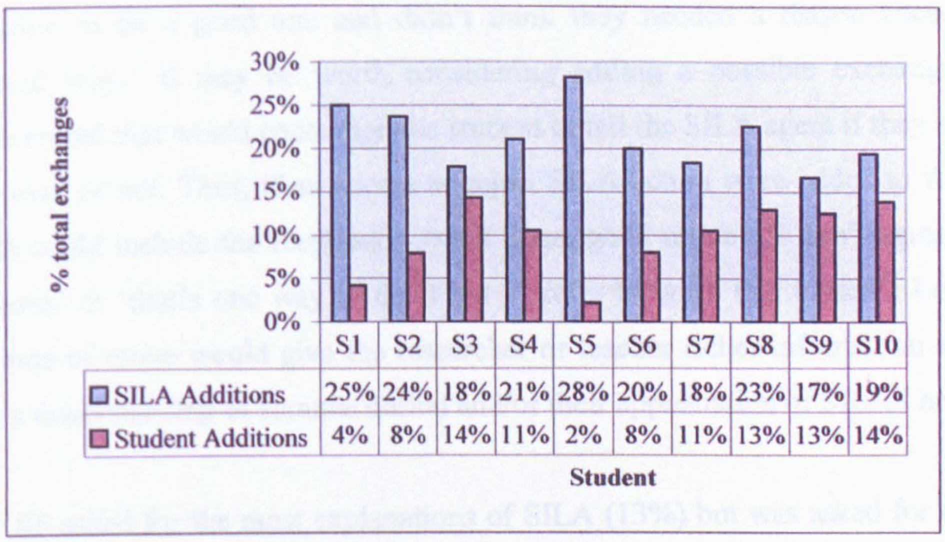


Figure 7.13 A Comparison of SILA-Student Summary Contributions.

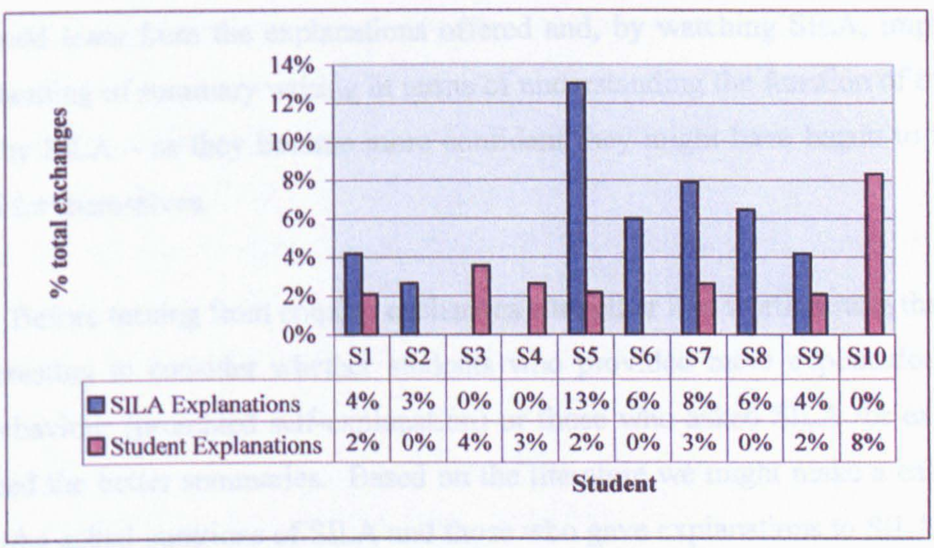


Figure 7.14 A Comparison of SILA-Student Explanations.

Furthermore Figure 7.14 compares explanations asked of SILA by the students and explanations offered by the students to SILA. Again the scores presented are shown as percentages of each student's overall number of exchange categories. This time the overall mean number of explanations SILA offered to the students as a percentage of the total number of exchange categories per student was 4%, Standard Deviation 0.04 whilst the overall mean number of explanations the students offered to SILA made as a percentage of each student's exchange category total was 2%, Standard Deviation 0.02. It can be seen that S10 offered the most explanations to SILA (8%) whilst asking for none in return. S3 also asked for no explanations of SILA. It is not simple to conclude what is occurring these instances. It may be that S3 accepted SILA's suggestions uncritically, lazily letting SILA auto-summarise or it may be that they judged the contribution to be a good one and didn't think they needed a reason because they understood why. It may be worth considering adding a possible exchange to the dialogue model that would encourage the student to tell the SILA agent if they approved of the move or not. Thus, if an accept or reject SILA action were added to the model then this could include the feedback – 'well done, good move', 'I don't agree that's a good move' or 'that's one way to do it but I prefer to write this myself'. Looking at these kinds of move would give the researcher or teacher richer information about the student's understanding of summarisation and/or their appreciation of SILAS help.

S5 asked for the most explanations of SILA (13%) but was asked for relatively few by SILA (2%). Again, it is not simple to conclude why but it may be that S5 felt they could learn from the explanations offered and, by watching SILA, improve their understanding of summary writing in terms of understanding the function of the phrases added by SILA – as they became more confident they might have begun to take more control for themselves.

Before turning from enquiry exchanges altogether it is worth noting that it would be interesting to consider whether students who provided more explanations of their own behaviour (prompted self-explanation) or those who asked SILA for explanations produced the better summaries. Based on the literature we might make a case both for those who asked questions of SILA and those who gave explanations to SILA to benefit more than those who do neither (Chi et al., 1989; Graesser & Person, 1994). It is not really possible to investigate this with the current set of data since it is not possible to control for student ability (prior summary note-taking skills) or be certain that each student had equal prompting from SILA for explanations but this is something which could be considered when designing future research using this or a future version of SILA. From Figure 7.13 it is noted that the two CHALCS students (S1 and S2) and S5 and S6 from Notre Dame have more SILA additions in their summary in relation to their own and also generally receive more SILA explanations than give their own explanations suggesting more passive and less active participation. Although these students all have some interaction with Genie they also have interaction with Peedy and those who only have interaction with Peedy (S3, S7, S8) do not seem on the face of it to have more self-explanations although the balance of their contributions to the summary is better split between SILA and student. It seems therefore that passivity is not simply a matter of being dominated by Genie and would bear further study. Further, when looking at summary performance S1, S2 and S5 (on the face of it rather passive interactions) do not perform particularly well. However, S6, also on the face of it rather passive performs very well. Those who interact only with Peedy (S3, S7, S8) and who therefore have more equal contributions to the summary generally seem to do rather better than S1, S2 and S5. However, S9 who has on the face of it rather a balanced interaction and spends most of their time with Peedy, does not perform well. However, to further make sense of these data it would be necessary to look at gains in performance pre and post SILA, to be able to control for prior ability and to have a larger sample of students.

Although it is not possible to compare the same student's performance with and without using SILA Figure 7.15 compares the summary scores of students from the summarisation studies described in Chapter 4 with the results reported as part of this evaluation. As can be seen in Table 7.10, exactly the same marking criteria and guidelines were used to mark the two studies by the same markers, i.e. the 'study skills' expert ('marker 1') and the Physics expert ('marker 2'). It can be seen from the graph that the range of summary scores is less in these evaluation studies 'with SILA' (50% to 88% 'with SILA', '28% to 81%' without SILA) although mean summary score is roughly equal, that is, 10.83 (68%) 'with SILA', Standard Deviation 2.29 and 10.65 (67%) 'without SILA', Standard Deviation 2.45 [H2].

Furthermore the two studies' summaries are qualitatively rather different with the Evaluation summaries ('with SILA') being less 'cut and pasted' although they are more alike due to the student acceptances of SILA's contributions [H3]. The appendices (Appendix F, G and K) contain examples of summaries produced during these two studies which illustrate these differences. Two typical comments from the same marker (marker 1, the 'study skills' expert) are provided below. The first is typical of those comments about the Summarisation Study summaries, reported in Chapter 4 and the second is typical of the marker's comments for the Evaluation summaries, reported in this Chapter.

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
Marker1										
selection	1	4	4	3	3	2	3	2	3	3
staging	1	3	4	4	4	4	3	3	3	4
condensation	1	4	4	3	3	2	2	2	3	2
organisation	2	3	4	1	4	3	3	3	4	2
Total	**5**	**14**	**16**	**11**	**14**	**11**	**11**	**10**	**13**	**11**
Marker2										
selection	1	2	2	2	3	3	3	3	4	4
staging	1	3	2	3	3	3	3	2	3	4
condensation	1	3	2	2	2	2	2	2	3	3
organisation	1	3	2	2	2	2	2	2	3	2
Total	**4**	**11**	**8**	**9**	**10**	**10**	**10**	**9**	**13**	**13**
Mean Score	**4.5**	**12.5**	**12**	**10**	**12**	**10.5**	**10.5**	**9.5**	**13**	**12**
% score	**28**	**78**	**75**	**63**	**75**	**66**	**66**	**59**	**81**	**75**

Table 7.10 Summarisation Studies Summary Scores (Without SILA)

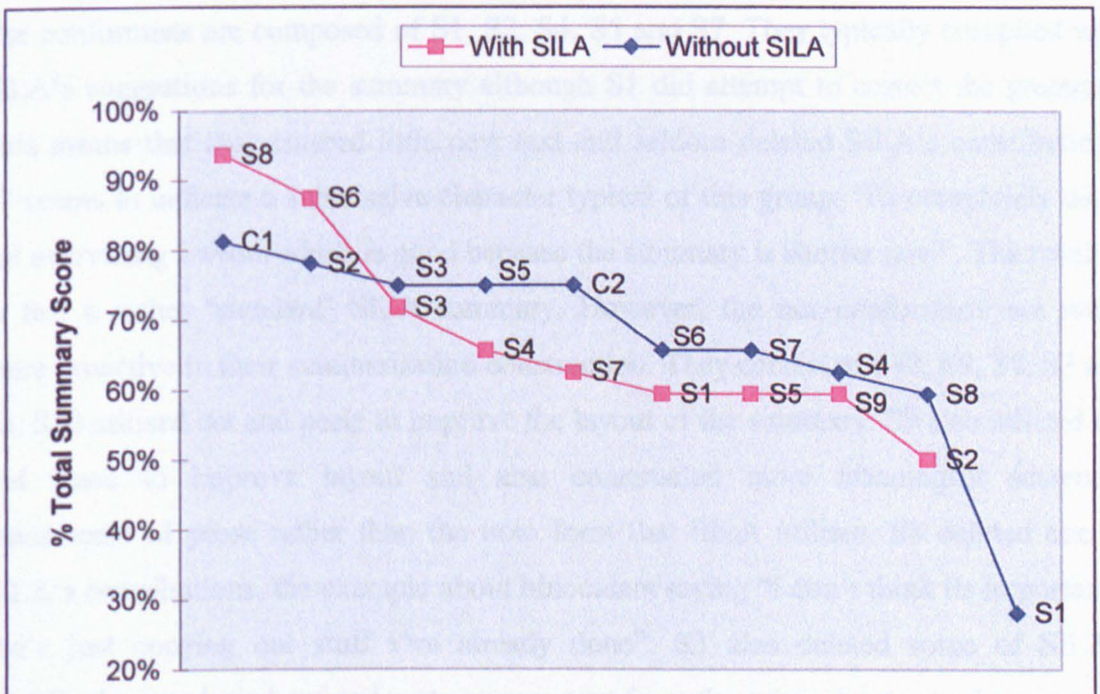


Figure 7.15 Comparison of Summary Scores With and Without SILA.

"This adopts a copying strategy which preserves too much and does not condense well though by preserving the authors words looks well presented." (S7, Summarisation Studies.)

"The summary is titled and begins with important information. The third sentence, on the "integration of image and brain" condensed by SILA, should probably have been moved further down to enable material about the eye to be drawn together in a single paragraph. The "binocular" example is included and the "sight defect" sentence which could have been deleted. The photocopier material is not included. Overall the summary is economical, well-condensed and generally readable." (S7, Evaluation Studies.)

The Evaluation summaries themselves can be typified as those produced by the '*conformists*' and those produced by the '*non-conformists*' [H4]. The conformists are more willing to accept SILA's suggestions for the summary whilst the non-conformists are more prone to question SILA's 'authority'. Interestingly these two groups each comprise of exactly half of the student population used during these Evaluative studies. The conformists are composed of S1, S2, S4, S5 and S7. They typically complied with SILA's suggestions for the summary although S1 did attempt to correct the grammar. This means that they entered little new text and seldom deleted SILA's contributions. S7 seems to indicate a submissive character typical of this group: "its completely taken out everything I wrote which is good because the summary is shorter now". The result is in fact a rather 'standard' SILA summary. However, the non-conformists are much more proactive in their summarisation construction. They consist of S10, S9, S8, S3 and S6. S10 utilised cut and paste to improve the layout of the summary. S9 also utilised cut and paste to improve layout and also constructed more meaningful sentences reminiscent of prose rather than the note form that SILA utilises. S8 deleted one of SILA's contributions, the example about binoculars saying "I don't think its important", "he's just copying out stuff I've already done". S3 also deleted some of SILA's contributions and used cut and paste to copy text from the original notes to the summary document then deleted superfluous words to arrive at a note form in a similar vein to SILA's condensation results. Finally, S6 elaborated on SILA's note form by adding text to the key words produced from the condensation and, as mentioned utilised 'chunking' to group associated phrases in the summary.

Overall, there is both a qualitative improvement in student's summary note-taking with SILA with student's displaying two distinct learning styles when using SILA which may be related to personality type. That is dominant students, the non-conformists mentioned, are less willing to accept SILA's suggestions whereas submissive students, the conformists, are less likely to question SILA. Obviously, without a personality assessment of the students, such as the Myers-Briggs scale (Myers & McCaulley, 1987), this possible cause is hypothetical, requiring further experimentation.

Note that the studies reported in Chapter 4 and those reported in this chapter do employ different cohorts due to them taking place months apart as a result of the intervening software development. However, the students are mainly drawn from consistent populations, i.e. CHALCS and Notre Dame with the Evaluation studies additionally utilising students from the WoZ studies as already explained. In this way consistency is maintained in keeping with the spirit of longitudinal studies, as mentioned in section 1.4 of Chapter 4 as a method of triangulation.

The next section though highlights one important conclusion based on results from the affect scales and the summarisation marking.

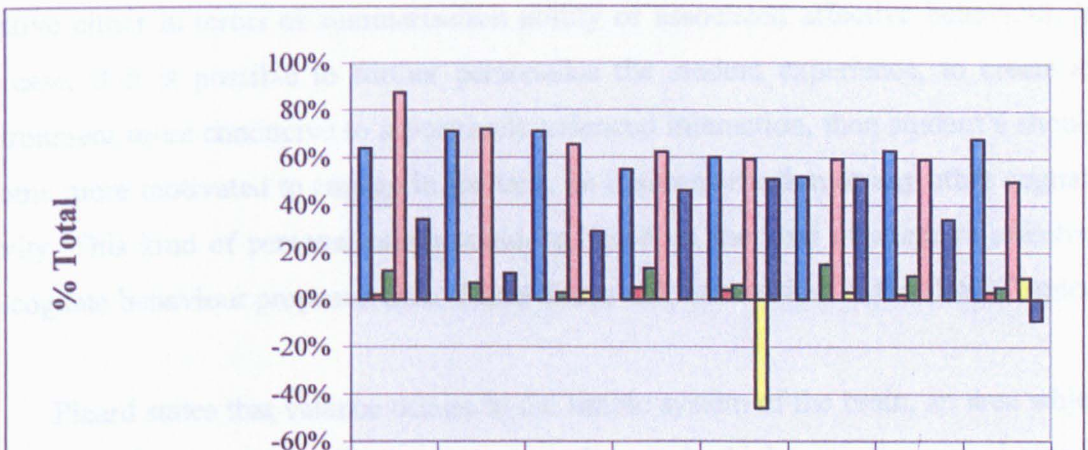
7.6 The Correlation of Affect and Performance.

The following figure (Figure 7.16) shows the student's performance statistics (positive collaboration, negative collaboration, enquiry exchange and summary score) and SILA's perceived overall Affect (resulting from the Affect Scale).

The most interesting result displayed on this graph is that the Affect of Peedy, as perceived by each student when they are ranked according to their summary score, follows a 'near normal' distribution (see Figure 7.17 for a clearer picture). The mode of the data is 34, the median is also 34 and the mean is 31 with a standard deviation of 21. To be a normal distribution Edwards (1999) states that the mode, median and mean should be at the same value which is not the case with Figure 7.16. Also the graph should be roughly "bell-shaped" which again is not quite the case with Figure 7.16 but it is not 'skewed' as the median equals the mode. Edwards goes on to state that for normal distributions "approximately 68% of all scores lie within 1 standard deviation either side of the mean, and approximately 95% within 2 standard deviations of the mean" (Edwards, 1999, p.138). In Figure 7.16 88% of the summary scores lie within 1 standard deviation whilst 100% lie within 2 standard deviations. The conclusion is that the curve is 'near normal'. Utilisation of a larger sample size would, it is hypothesised, further converge upon a true normal distribution although this of course would have to be confirmed by a future study.

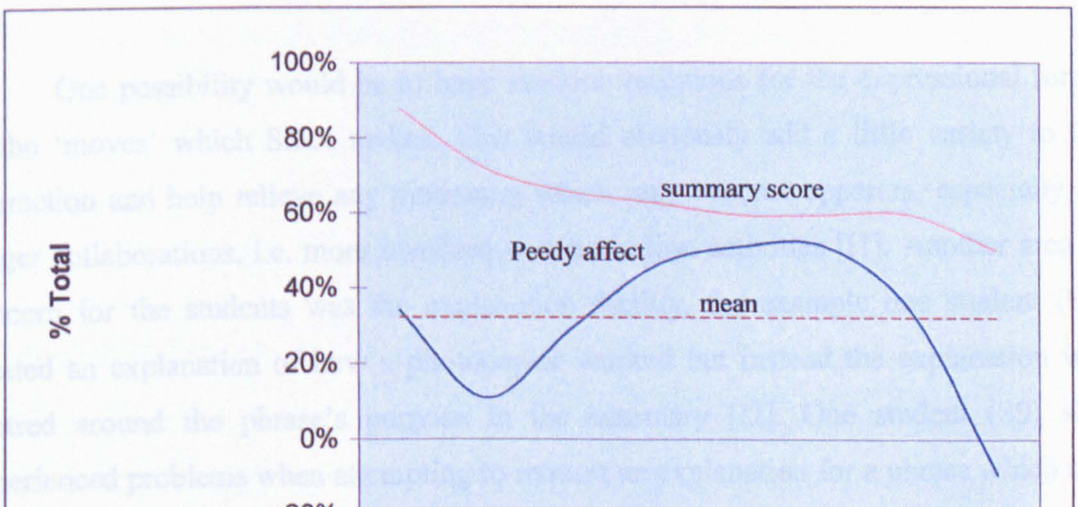
This result would suggest that students with 'average' summary performance tend to perceive Peedy as having more positive Affectations than those with higher or lower performance. Note that the highest summary score was obtained from the one student who had a strong negative reaction to SILA as mentioned above. It is unfortunate that the Affect Scale was not completed for this student as the anomaly present in the current correlative graph for the highest ranked student may have been reduced with a 'best fit' graph if this result was included. As all students positively collaborated with SILA to a moderately equal extent it is hypothesised that *students more adept at summarisation prior to using SILA and those who lack good summarisation skills prior to using SILA have a more negative valence towards SILA*. This could be due to the more able students feeling that they do not need to collaborate with SILA to write a good summary and thus interpreting SILA's attempts (that is those aspects of the Affectations) as interfering, i.e. negative or at least only having a moderate positive valence, whereas the less able students may feel that SILA has a threatening persona due to its knowledgeable stance on summarisation and thus experiencing its Affectations as negative or moderate valence.

As stated, it would be prudent to further test these concluding hypotheses with a larger sample size than above. If the hypotheses are proven then the result could have consequences for Artificial Study Companion design such as the possibility for Companions to adjust their Affect according to student prior skill or a number of Companions being made available to students along the range of perceived Affect. If, indeed, there is a link between summarisation ability and valenced reaction to SILA then it would be possible to adjust the affectations with, for instance, an assessment of student's prior summarisation skills. One obvious way to achieve this would be to ask students to complete a WebCT multiple choice quiz allowing SILA to access the result by means of the Application Programmers Interface then correspondingly adjust the Affectations. In this way it is hoped that the affectation score is perceived across the students in a more consistent way as intended by the design, for example resulting in the mean affect score.



	S6	S3	S4	S7	S1	S5	S9	S2
Positive Exchange	64%	71%	71%	55%	60%	52%	63%	68%
Negative Exchange	2%	0%	0%	5%	4%	0%	2%	3%
Enquiry Exchange	12%	7%	3%	13%	6%	15%	10%	5%
Summary Score	88%	72%	66%	63%	59%	59%	59%	50%
Genie Affect	0%	0%	3%	0%	-46%	3%	0%	9%
Peedy Affect	34%	11%	29%	46%	51%	51%	34%	-9%

Figure 7.16 Student Performance – SILA Affectation Correlation.



	S6	S3	S4	S7	S1	S5	S9	S2
Summary Score	88%	72%	66%	63%	59%	59%	59%	50%
Peedy Affect	34%	11%	29%	46%	51%	51%	34%	-9%

Figure 7.17 Correlation of Peedy Affect with Student Summary Score.

Note that simply improving the dialogue model, although perhaps resulting in a more individualistic experience by the students, will not ensure that the agents are more adaptive either in terms of summarisation ability or associated affective behaviour. In any case, if it is possible to further personalise the student experience, to create an environment more conducive to a positively valenced interaction, then student's should become more motivated to engage in the task, be it summarisation or any other cognate activity. This kind of personalisation could be based on the kind of adaptive affective and cognate behaviour proposed here. These issues will be revisited in the final Chapter.

Picard states that valence occurs in the limbic system of the brain, an area which attaches positive or negative feelings to perceptions and which is the "primary home of the emotions" (Picard, 1997). Such manipulation of valence as evidenced in appropriately designed Affective interfaces (including the Affectations utilised here) can only serve to improve the effectiveness of Human-Computer Interface Design.

7.7 Potential SILA Future Work Arising From the Evaluations.

A number of recommendations for future work were advised by the students themselves. In-keeping with the previous results section the reference in square brackets is referenced in Table 7.3 and the Conclusions.

One possibility would be to have random variations for the expressional forms of the 'moves' which SILA makes. This would obviously add a little variety to the interaction and help relieve any monotony which may become apparent, especially on longer collaborations, i.e. more involved summarisation activities [I1]. Another area of concern for the students was the explanation facility, for example one student (S5) wanted an explanation of how a photocopier worked but instead the explanation was centred around the phrase's purpose in the summary [I2]. One student (S9) also experienced problems when attempting to request an explanation for a phrase which had been added many turns before. In the current implementation SILA will only provide an explanation for the phrase which was most recently added and the explanation will be a justification for selection, condensation or organisation i.e. relate to the note-taking skill rather than to physics content. Both approaches to explanation could be implemented using a hierarchical dialogue model to enable students to step back through the hierarchy to highlight the phrase they require an explanation for. It would not involve too much effort to improve the facility along these lines but a dynamic understanding or parsing of the intent of the student's question based on an internal model of the current focus, previous foci from which it was derived (the recorded structure of the conversation) would require considerable AI modelling techniques and programming effort.

Two further ideas were suggested by one student. The first was suggested in order to make the behaviour of SILA more adaptive by, for example, changing the rate of interruption [I3]. The student (S9) said "just ask him to interrupt when you've stopped typing for a while or if you need more help". This would serve to overcome some of the problems mentioned earlier of SILA interrupting at inconvenient times and really is worth considering as an important future improvement to the system. Finally, the same student (S9) suggested that SILA could "underline key words" in the summary [I4]. This is, maybe, an obvious facility for the system to posses and would, in fact, require little programming effort although perhaps not all students would want this kind of help.

7.8 The Post-Evaluation Prototype.

The Post-Evaluation Prototype was construed as a stable, robust version of SILA which could be made generally available from a website[2]. This meant that a number of changes had to be made, most notably providing facilities for teachers to author the data required by SILA to construct a summary. This, in turn, meant removing the hardwired World Model in favour of a more flexible approach.

7.8.1. Requirements Analysis.

The main addition in the Post-Evaluation Prototype is the implementation of a 'teacher' menu which would allow a teacher to login using a password then access a special area of SILA. This area would now include the open and close log menu options as well as the ability (and associated instructions) to author the World Model. This editing would then be carried out in the usual 'summary window', saved then consulted on a future session. The storyboard in Figure 7.18 shows how this would work.

---

[2] currently www.synthoids.net (15-Aug-03)

7.8.2. Architectural Specification.

The architecture for this version of SILA is identical to the previous prototype version reported in Chapter 6, see Figure 6.11.

7.8.3. Detailed Specification.

The main functionality of ProSILA is identical to the Evaluation Prototype so further pseudo code is unnecessary. The only significant addition of a file to hold the summary data is described in the next section.

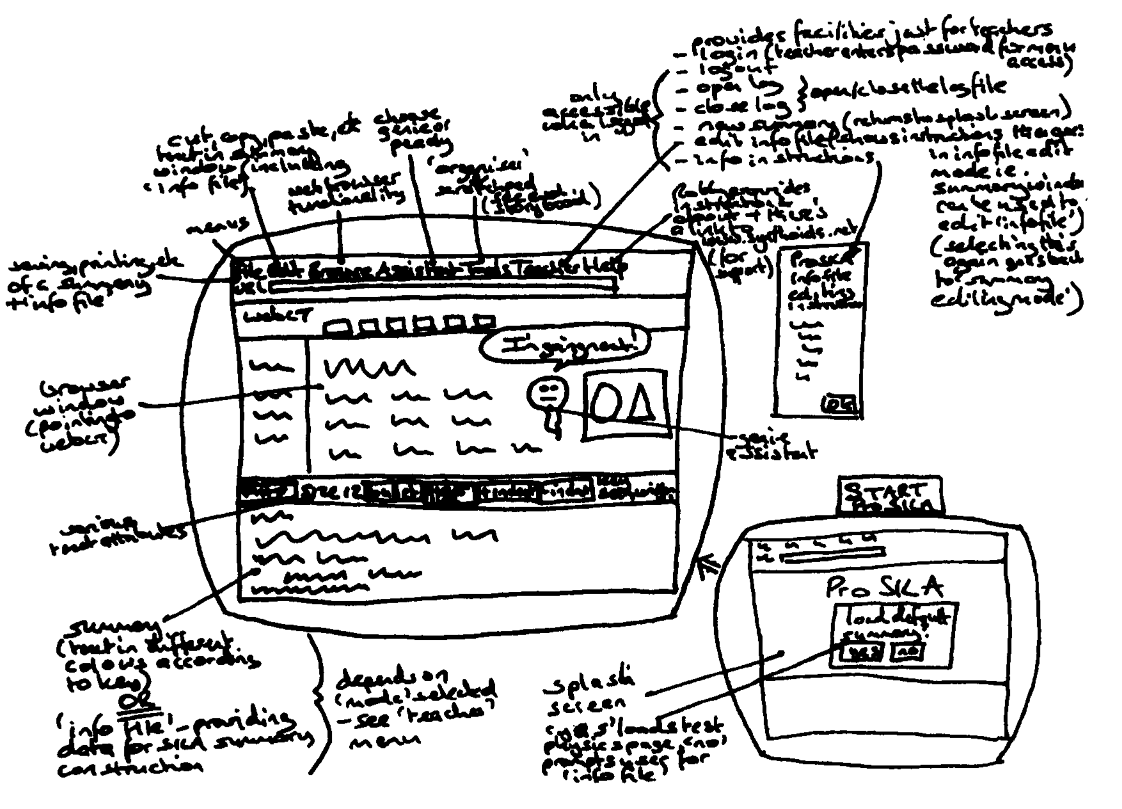


Figure 7.18. The Post-Evaluation Prototype Storyboard.

7.8.4. Implementation.

The implementation of the World Model or 'info file' was fairly straightforward, consisting of simply providing the teacher with a common file dialogue to open the file then reading each item sequentially into the previously static variables. Additionally, the 'info file' editing capability utilised the same editing facilities as used by students when writing summaries, the only difference being that it is saved with a different extension (.dat instead of .rtf). The format of the file is as follows:

0. text filename or URL (the source of the summary to be produced)
1. number of phrases
2. phrase summary
3. summary justification
4. number of keywords (n)
5. keyword #1
6. keyword #n
7. relative importance

(where 2 to 7 are repeated for each phrase)

Additionally, a 'splash screen' was added to give the user a chance to load an example default course notes page or point to an 'info file' and an option was added to the help menu to open a browser pointing to a web-based support page[3].

The overall result of the prototype development is, it is felt, a system which is robust enough to be used in a real, everyday classroom environment. Although, there is still room for improvement particularly in terms of the support provided to teachers to author the 'info file', the kinds of improvements mentioned by students in the previous section and those reported in the final chapter.

7.9 Conclusions of the SILA Evaluations.

Table 7.13 provides an overview of the conclusions based on the results previously presented in this section. As can be seen (where more than one is given) data sources are sometimes triangulated to arrive at particular conclusions, in line with the research methodology discussed in section 1.4 of Chapter 1. As was the intention with this chapter a number of metrics arose to measure the effectiveness of the SILA software. These metrics were based on the final Research Aim and delineated at the start of this chapter. They consisted of Validation, Evaluation and Assessment. In terms of Validation, as the system functionality is being compared to the WoZ experiment, the final WoZ recommendations are each validated in turn in the conclusions with the exception of the need for a 'processing' animation as this becomes irrelevant in the

[3] currently http://www.synthoids.net/support.htm (15-Aug-03)

current implementation (SILA does not take as long as an experimenter to think of a summary addition). In terms of Evaluation, upon reflection the two objectives of evaluating whether the affectations are motivating and believable and evaluating evidence of the differences between characters become subsumed by the WoZ metrics. Finally, in terms of Assessment a further metric is derived, notably whether SILA provides reasonable summary note-taking support.

Each evaluation metric is listed in Table 7.11 then referenced in Table 7.13 together with a reference to the relevant result summarised in Table 7.12, the data source it is derived from then finally the conclusion itself. The result is a list of 20 conclusions.

Overall students reported favourably on their experiences when summarising with SILA, for instance being impressed by the condensation mechanism (as highlighted above) although it was noted that the keywords must be carefully chosen so that the resultant condensation is coherent. However, some students reported that they would have liked more detail in their notes. Students also reported that SILA was entertaining and were generally enthusiastic about the software. They also reported a calming effect when using SILA, leading to less anxiety about the summarisation task. Furthermore, the environment as a whole was well received although the 'organisation' help was seldom utilised as students felt it was unnecessary for the short summary produced during this evaluation. It may be though that it would gain more importance if better integrated into the main dialogue model.

As for the characters themselves, it was clear that students perceived them to be distinct in terms of the two extremes of personalities, citing the subtle expressivities of language, voice, appearance and timing as the main distinguishing features. However, it was reported that the speech output could be more natural. There was also evidence of a preference for the positive Peedy character over the negative Genie character which was seen as too domineering with students often switching from Genie to Peedy and staying with Peedy as they became used to it. More research is needed into the learning and personality styles of students and their preference for such characters but the notion that students would be entertained or challenged by a fierce character was not well-supported here.

REF.	EVALUATION METRIC
EM1	adaptivity/provision of a number of different personalities
EM2	demonstrate praise, justification and patience
EM3	dialog model effectiveness including utility of 3 phases and 'why' move
EM4	value of including summarisation process mode and rhetorical analysis
EM5	ability to encourage student collaboration through motivating personality
EM6	supporting a range of text editing and appropriate assistance
EM7	including help and hints of summarisation
EM8	improvement of speech input/output
EM9	utility of dominant and submissive personalities
EM10	provision of reasonable summarisation note-taking support in a VLE

Table 7.11 SILA Evaluation Metrics.

REF.	RESULT SUMMARY
A1	S1, particularly S10, S4 OK with interaction, most require guidance
A3	OK response ambiguous (S4) e.g. in response to 'me next' & 'look'
A5	evidence of 'edutainment'
A6	students don't find self-explanations easy, e.g. S6
A7	students stay with current assistant because its easier/they get used to it
A8	students generally enthusiastic about SILA
B1	only S1 read 'organisation' then improved SILA's grammar
B2	Most happy with environment, S1 talks about making it adaptive
B3	A few (particularly S5) required more involved, detailed notes
B4	Most OK with instructions, S2 would like 'back' & S5: "a bit quick"
C2	Only S2 had a problem with condensation key words
C3	S3, S8 copy & paste sentences, most type, S1 edits SILA's contribution
C4	students delete unwanted notes: e.g. S3, "common sense" (S8)
C5	Condensation well received, e.g. S4, S7, S10
C6	condensation can result in adding keywords to unrelated text e.g. S6
D5	some students use 'chunking' to group similar material e.g. S6 & S8
D6	most students didn't edit Summary (S6 is an exception)
E1	S7 very positive + he actually waved !!!
E2	A few reported 'calming effect' of using software, e.g. S1, S6, S7
F4	a few struggled with ending Peedy although persevered e.g. S8, S2
F7	factors distinguishing Genie & Peedy: language,voice,appearance,timing
F8	Genie can be domineering, e.g. S5
F9	speech synthesis sounded like Stephen Hawking (S7)
F10	S8: "personality in the sense that they have different characteristics"
F11	Genie seen as negative/mediocre valence, Peedy as positive
F12	Genie's dominant affectation is 'macho', Peedy 'friendly'
F13	Both assistants seen as 'encouraging', 'amenable', friendly', 'polite'
H1	some students let SILA demonstrate summarisation (e.g. S5)
H2	'with SILA' summary score range less and qualitatively better
H3	'with SILA' summaries are more alike but less cut & pasted
H4	conformist and non-conformist learning styles are displayed
I2	explanation facility could explain Physics, e.g. S5, S9
I3	more adaptive behaviour beneficial? e.g. change rate of interruption

Table 7.12 SILA Results References (Abridged).

EVAL. REF.	RES. REF.	DATA REF.	CONCLUSION
EM 1	I3	6	1. Providing true adaptive behaviour may prove useful.
EM1	A7	6	2. Students stay with a personality because they get used to it.
EM1	F10	6	3. Characteristics validate personality design.
EM1	F12+ F13	4	4. Some similarities and some differences between assistants are evident.
EM1	F7	6	5. Subtle expressivities do distinguish assistants effectively.
EM2	I2	6	6. Interactional & propositional aspects of 'explanation facility' could be improved.
EM3	F4	2	7. 'Stubborn' ending is not a good idea.
EM3	A1+ A3+ A6	7 + 6	8. The dialogue model needs to be learned by students and could still be improved.
EM4	C2+ C5+ C6	6 + 7	9. Generally condensation is well thought-of although keywords need to be carefully chosen.
EM5	A5+ A8+ E1	6	10. Students are generally entertained and enthusiastic about SILA.
EM5	E2	6	11. SILA calmed students so they were less anxious about summarising.
EM6	C3 + D5 + H4	7 + 6	12. A range of summary note-taking techniques were utilised grouping into *conformist* and *non-conformist* learning styles.
EM6	C6 + C4	7 + 6	13. Assistance was not always appropriate for some individuals.
EM7	B1 + D6	1 + 7	14. Help and hints are included in 'organisation' but few students used it as they felt their summary was sufficient.
EM8	F9	6	15. Speech output could be less 'robotic'.
EM9	F8 + F11	6 + 4	16. Genie may be too domineering and is perceived more negatively than Peedy.
EM10	H1	6	17. Some students allow SILA to demonstrate summarisation first.
EM10	H2 + H3	3	18. Summaries with SILA are more alike but generally of a slightly better quality (and simple copying and pasting is eliminated).
EM10	B2 + B4	6	19. Most students are happy with the SILA environment.
EM10	B3	7	20. Some students would like more detailed notes.

Table 7.13 SILA Evaluation Conclusions.
(Key: Evaluation Metric Reference refers to Table 7.11,
Results Reference refers to results in section 7.5 and summarised in Table 7.12,
Data Source Reference refers to Table 7.3)

There was evidence of students allowing SILA to take turns and explain moves which indicated that there was scope for the student to learn from the agent. However, there was room for improvement in terms of the interactional and propositional aspects of the explanation facility. Additionally, the dialogue model had an associated 'overhead' in terms of students having to become accustomed to using it. Also, the tendency of SILA to sometimes disallow a student 'end' move was not well received by students. With a mean score of 68% for student's summarisation with SILA compared to 67% without SILA the quantitative assessment seems to indicate no increase in overall summarisation performance (see the discussion in section 7.10 for possible mitigating factors) although the associated qualitative improvement shows that the summaries produced were of a generally higher standard although they were less individualistic. Extreme forms of copying and pasting entire notes (no condensation) were eliminated with SILA. There was some tendency from some students to passively allow SILA to complete too much of the summary. In general we might have wished that students were a little more questioning and a little less accepting of SILA's suggestions particularly at the condensation level. As pointed out this may be due to their personality type and learning style or it may be due to the general "hassle" of rejecting or undoing the changes made. Adding a loop to the condensation interaction where the student could either accept or reject the change (as for example in collaborative authoring using Micorsoft Word's track-changes facility) might reduce both the sense of Genie being too dominant and result in students taking more control over the condensation phase. This requires further investigation.

An important result was that students utilised a number of summary note-taking techniques, again depending on their own learning style for instance. So that SILA supports a number of approaches to note-taking the idea of providing different personalities could be further developed in terms of creating adaptive assistants (also suggested by students) which adjust in terms of behavioural as well as the personality dimensions. In this way it is hoped that students will find that the assistance provided by SILA is more appropriate.

7.10 Discussion of the SILA Evaluations.

The 'cooperative evaluation' technique described above provided a humanistic observational method for eliciting a student's decision-making processes as they interacted with the system. This method also allows the student to feel less like the object of study and more a collaborator in evaluating the system. Although, it is likely that students still felt some degree of tension – initial concern that they were being 'tested' as evidenced by the reports of nervousness at the beginning – during the course of the interaction students generally 'calmed down' and felt at ease. This in part is likely to be through a growing understanding that the situation was one in which the software and not the student was the real 'subject'. Additionally, the experimenter could demonstrate an example interaction with SILA so that the student quickly gained confidence in using the interaction mechanisms and was able to use the system unaided. That is, instructions given to each student at the beginning of the study were elaborated by concrete examples. Although, as reported it was important that the experimenter did not lead the entire session.

The exchange analysis used to analyse the SILA logs does not capture the individual subtle nuances of the interactions and instead gives an overall approximate impression of the exchanges. That is, localised pockets of continued disagreement or agreement fail to show up in the analysis as they are summed together for the overall impression. Such local clustering could be indicative, for example, of an intolerance that it would be wise to take heed of. However, the picture presented above does give an easily accessible account of the overall impressions of exchanges that could be followed up with a more detailed analysis in particularly interesting cases.

As well as potential problems due to clustering there were other problems with coding the exchanges. Notably, sometimes a student may miss a 'no' response resulting in a 'null' command being inserted into the log. Obviously the student's intention was to respond 'no' to the system but due to mis-timing their intention was, in fact, misrepresented. Again, it is these kind of subtleties that the above analysis fails to capture. In addition a small number of particular episodes did not easily fit into the derived categories thus resulting, again, in inaccurate portrayals of the interactions. For example, if a student selects a 'no' command after SILA adds a phrase to the summary then this will not be coded by the scheme. This can be justified by way of SILA actually not recognising this command in this context. However, the student's intent was perfectly valid.

A further example would be when a student replies 'because' to a question from SILA to explain a phrase addition but instead of giving a reason they actually enter 'I don't know'. SILA would not recognise that this response from the student was not an explanation and reflects the wider problem of checking student's understanding by evaluating the quality of their explanations. This is clearly beyond the scope of the current work requiring a knowledge rich approach and natural language understanding. Unfortunately, this also means that in classifying data to arrive at an overall picture of interactions we cannot tell whether the explanations offered by students were meaningful without further analysis. However we can say that students were offering something when asked for an explanation and such prompts have been found in other contexts to improve performance.

Returning to the overall aims of the evaluation set out at the beginning of this chapter SILA appears to give a reasonable approximation to the behaviour of the wizard in the WoZ study and certainly addresses all of the recommendations of that study by presenting affectations in the form of two characters which were perceived to be distinctly different in the directions expected. In the case of Peedy at least the 'persona' was motivating and believable and provided reasonable support for student's summary note-taking resulting in summaries which were judged by tutors to be of a good quality. As demonstrated, the summaries produced with SILA showed improvement over previous attempts produced without the help of SILA.

One possible factor affecting the seemingly similar quantitative results of the marks for the Evaluation summaries and the marks for the Summarisation Study was the apparent inadequacies of the marking scheme. It transpired that the scheme did not penalise students for simply cutting and pasting (most of) the original text. That is, a student who selected all of the original material would gain marks for selecting all of the "important material" but would not necessarily be penalised for selecting all of the "unimportant material". Additionally, by copying the whole text the student could achieve a high 'staging' mark as there is no distortion of the meaning, it is in the same order and usually the "relative importance" is signalled. Both markers commented on these faults in the marking scheme. In retrospect the scheme could have taken account of this copying strategy, penalising students appropriately. However, this action was not anticipated and certainly Tawalbeh's scheme, upon which this scheme was based, did not take this into account. The outcome is that the quantitative results should be used only as a guide and it is suggested that the overall 'true' value of the mean result for summarisation without SILA may be slightly lower indicating that SILA also showed a quantitative improvement in summarisation performance. Obviously future studies would need to revisit the marking scheme.

It is, however, clear that students who use SILA do not end up with a "cut and paste" intact version of the original notes as the end product. As teachers in schools increasingly face marking work students synthesise from web-based sources by cutting and pasting, with all the inherent difficulties in terms of judging the extent to which students are offering their own work (and the related problem of plagiarism faced in Higher Education) any tool which can encourage students to learn to think more critically about the material they select and how they condense and organise it to meet the needs of a particular task will be welcomed. Moreover, students engaged in collaboratively summarising with SILA seemed to enjoy the experience more than the summary task without SILA, making it more likely that they would complete the task and produce a summary as an outcome when using such a tool.

Therefore, overall, the SILA software has been shown to meet the goal of producing a limited proof-of-concept solution to the problem of providing assistance to students during their VLE experiences although, of course, there is still room for improvement. These potential improvements will partly be the subject of the final chapter which also discusses the overall aims of this research project, evaluates whether or not these have been met and makes recommendations for further work. First though the SILA system is presented in terms of a 'walkthrough'.

Chapter 8: SILA Walkthrough.

8.1 Introduction.

The following walkthrough is intended to show how SILA will perform in a typical session. The first section explains the SILA facilities then the subsequent two sections provide two typical session scenarios intended to illustrate the software, one showing a student's interactions and the other showing those of a teacher.

8.2 SILA Facilities.

The explanation of facilities are broken down into instructions for using the two Windows, i.e. the web browser and the document editor then instructions for using the various menu commands. Finally, interaction with the Agents themselves is explained. Note that basic Windows familiarity is assumed. Additionally, the startup dialogue is explained in 8.2.8 Teacher Menu, 'Test New Info File' option.

8.2.1. Web Browser.

This window provides full World Wide Web browser functionality based on MS Internet Explorer technology. That is, any hyperlinks contained within a document displayed in this window can be clicked upon to open up the page they point to and, furthermore, any passwords required, such as those requested by WebCT, will be handled as usual. Obviously, these do not work unless an Internet connection is running on the local computer. In addition by right mouse clicking over the document the student can use the normal editing facilities of cut, copy, paste, etc. Finally, the student can enter any valid URL into the box above the window to force the browser to load any web page. (N.B. see the later section 8.2.5, 'Browse Menu' for an explanation of the remaining functionality.)

8.2.2. Document Editor.

This window contains the jointly constructed summary document. There are a variety of formatting functions available including selecting the font size from the drop-down menu, changing the size of text by typing in a *font size* (this will either affect subsequent text or currently selected text), using *bullet points* (click on the button to add bullets then again to remove them as indicated by the text on the button), use of *underlining* (again either on subsequent text or selected text) and finally, it is possible to increase *indentation* using the '+' button or decrease using the '-' button (note that this will only work on one line at a time). Note that indentation shows the relative importance of phrases, the more the indentation, the less the importance. SILA currently automatically changes any indentation to reflect its opinion as to a phrases' importance. The Key above the window and to the right indicates that red text has been 'seen' by SILA, blue was written by SILA and magenta text indicates that phrases have been condensed. Right mouse clicking over the window results in the display of a pop-up editing menu consisting of cut, copy and paste.

8.2.3. File Menu.

The File Menu contains the following options which provide all file input/output facilities:

- **new summary** – clears the document window;
- **open summary** – loads an existing document, the student is presented with a common Windows file dialogue;
- **save summary** – saves the current document, again the student is presented with a common Windows file dialogue;
- **print summary** – prints the entire document or just the selected text, the student is presented with a common Windows print dialogue;
- **exit SILA** – quits the SILA environment.

8.2.4. Edit Menu.

The Edit Menu simply provides cut, copy and paste facilities for the summary document window *only*. Additionally, there is a 'clear' option which clears the summary window.

8.2.5. Browse Menu.

The Browse Menu provides additional functionality for the Web Browser window. The individual functions are explained below but should be mostly familiar to browser users:

- **open file** – opens a common Windows file dialogue to load a file into the browser window;

- **stop** – halts the current loading of a web page in the browser;
- **refresh** – reloads the current page (pointed to by the URL) into the browser;
- **back** – navigates backwards to display the previous page in the history in the browser;
- **forward** – navigates forwards to display the more recent page in the history in the browser;
- **home** – displays the current homepage in the browser (currently set to www.timbarker.org);
- **search** – displays a (MSN) search tool in the browser.

8.2.6. Assistant Menu.

The assistant menu simply allows the student to choose one of the two Companions, Peedy or Genie. Once selected the appropriate menu item will have a tick placed next to it and the other Companion will be deselected (appear greyed-out).

8.2.7. Tools Menu.

The Tools Menu contains additional facilities provided for the student as detailed below:

- **sketchpad** – this loads the 'jotter' facility onto the screen which allows a student to make miscellaneous notes separate from the main summary document. The edit menu can be chosen or right mouse click selected to provide *cut*, *copy*, *paste* and *clear* editing facilities. Clicking 'OK' at the bottom right of the sketchpad will exit and return to the main SILA window.
- **organisation** – this provides "help with organising the summary", for example, providing help on what to do when "my summary is too long" or "how do I show something is important in my summary ?". It is a HTML document so that links (underlined text) can be selected to jump to other parts of the document. In addition one link in particular, '*discourse markers*', leads to a separate document containing a table of words which can be used, for example, in "collections and lists" and "similarities and differences". The usual editing facilities can be selected by right mouse clicking, for instance

enabling a word used for lists to be directly pasted into the summary document. Finally, selecting OK at the bottom right of the window will exit it.

8.2.8. Teacher Menu.

This menu provides administrative features which a teacher/experimenter may wish to keep hidden from students.

- **Login** - A teacher/experimenter must first log in to access the functions provided by this menu. Currently the password is "Eggplant" (note the use of upper case E). Select the 'Login' option and enter the password in the pop-up dialogue. Once logged in the following options appear.
- **Logout** – Logout from the Teacher Menu and thus hide all options apart from 'Login'.
- **Open Log** – will first present a dialogue requesting the student's name then a common Windows file dialogue to assign a filename to the log file. It will be greyed-out and ticked if a log is already open (the nature of the log file is explained later in section 8.2.11).
- **Close Log** – this will close the log file but will be greyed-out if a log is not open.
- **New Summary** – will first prompt you to load the default summary. If you answer 'yes' then a default sample Physics page will load into the web browser window. If you answer 'no' then you will be presented with a common Windows file dialogue to point to an 'Info File' which contains all of the information SILA requires to co-create a new summary (more details below).
- **Edit Info File** – the teacher/experimenter can author new information SILA requires for a new summary co-creation scenario. First instructions pop-up outlining the format of the 'Info File', as follows, for each line in the Info File (note that all text must be in quotes and that 3 to 7 are repeated for each phrase):

1. the filename or the URL of a document to be summarised;
2. the number of phrases in the summary to be created;
3. a phrase's summary;
4. a justification for summarising that phrase;
5. the number of keywords (to be used to recognise a student's added phrase);
6. each keyword *on a separate line*;
7. the importance of that phrase relative to other phrases (with 1 being the most important).

When this option is selected a tick will appear next to it indicating that the Document Editor can be used to edit the 'Info File'. The usual File Menu can thus be used to load, save, print, etc. the 'Info File' in the normal way. Although note that the file extension is different. The Edit Menu may also be used in the normal way. Finally, the 'Info File' must be saved then by selecting Edit Info File, again the tick will disappear and the Document Editor can once again be used to co-create a summary document. (By way of example, there is a sample infofile, "silainfo.dat", in the SILA install directory).

- **Info Instructions** – will again display the instructions for editing the 'Info File'.

8.2.9. Help Menu.

The Help Menu provides brief instructions on the basic use of the system and background information such as the version and authors, as follows:

- **instructions** – Robby the Robot appears (another MS Agent character) and explains the assistant menu, the differences between Agents personalities, how to interact with them (including showing them single phrases), indentation, speech input and the file and browser menu functionality.

- **about** – this displays the version number of the current prototype, a note that this forms part of a PhD research project and those who were involved in the project.
- **SILA web** – will open up a browser window pointing to the support page at http://www.synthoids.net/support.htm where up-to-date support is provided for SILA.

8.2.10. Interacting with Genie/Peedy.

This section details the reasoning behind the SILA interaction style and how this is achieved through the use of menu commands and speech.

8.2.10.1. Rationale.

It is important to stress that the Companions are there to be collaborated with, that is, there is an expectation that students will interact with the Agents and either not let the Companion do all of the work or not ignore it completely. The idea is to share the task of constructing the summary.

8.2.10.2. Menus.

To interact with a Companion it is necessary to right mouse click over it, i.e. place the cursor on the character then click (although see 'speech option' later in section 8.2.10.4). The time to do this is when they have either their hand or wing, as appropriate, to their right ear. A menu will then appear thus allowing an option to be selected by dragging the mouse over it and left mouse clicking to select. The menu is context-dependant which means that it will change depending on the nature of the interaction required. The commands below the separator line are the ones to control the dialogue interaction. The commands above the line will open or close the voice commands window (see 8.2.10.4 Speech Option later) or hide the Companion. However, it is preferable to use the 'end of session' command to hide the Agent as the 'hide' command could result in unforeseen outcomes. 'Advanced Character Options' control such MS Agent aspects as which key represents the commencement of speech input, etc.

8.2.10.3. Turn-Taking.

Figure 8.1 shows the turn-taking process in action. Genie has his hand to his ear so the student has right mouse clicked over him to get the menu of commands then navigate the menu to the "Why did you do that ?" option which will subsequently be left mouse clicked with the net effect of SILA offering an explanation for the recent addition of a phrase in the summary.

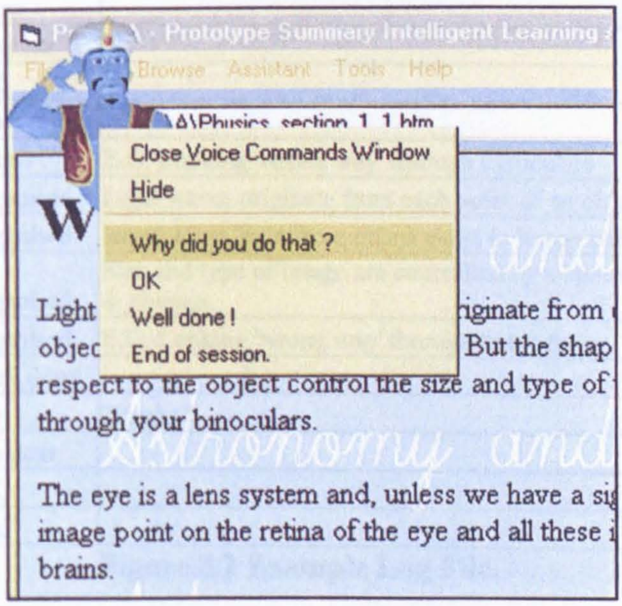


Figure 8.1 Interacting with SILA.

8.2.10.4. Speech Option.

SILA can recognise speech input for each character. The student should press the 'scroll lock' key (unless this has been changed in the 'Advanced Character Options') when they want either Genie or Peedy to listen then speak clearly into the microphone. The microphone requires careful calibration during a test run of the system prior to allowing a student to use it. The level of the microphone needs to be carefully adjusted using the Windows 'Volume Control' to obtain optimal performance of the speech recognition system. As mentioned, a window can be displayed which shows the voice commands that SILA recognises at any instance.

8.2.11. Logs.

Note that the logs obtained from selection in the Teacher Menu contain the time of the interaction, the command issued (e.g. 'condense', 'explain' etc.) and the Agent (human or artificial) issuing it and any supplementary information such as the phrase added. It is intended that these be used during evaluation of the software but tutors may find it useful to have a record of the student-SILA interactions which could be used for reflective purposes in class, etc. The log is a simple comma-delimited text file which can be imported into standard packages like MS Excel for later analysis. An example is shown in Figure 8.2.

SILA Beta 1.0 log of tim on April 11th, 2003 at 20:36:34			
**			
20:36:48	Peedy	show	
20:36:48	Peedy	me next	
20:37:06	tim	OK	
20:37:06	Peedy	look	
20:37:13	Peedy	added	E.G. Looking 'wrong way' through binoculars
20:37:13	Peedy	organised	Light waves originate from each point of an object
20:37:13	Peedy	organised	lenses focus light from object point to image point
20:37:13	Peedy	organised	Size and type of image are controlled by a lenses shape, material & placing
20:37:14	Peedy	organised	E.G. Looking 'wrong way' through binoculars
20:37:26	tim	Well done	
20:37:26	Peedy	say	Thanks !
20:37:51	tim	Me next	
20:37:51	Peedy	OK	
20:37:59	tim	End	

Figure 8.2 Example Log File.

8.3 SILA Student Scenario.

This section describes the main student features to be found in the SILA software. It is not a complete description of all the minutiae of features, rather it portrays the essential elements in order to give a taste of what it is like to use the software as a student. The intention is to provide a walkthrough of the software as if using it in a typical summarisation session. Note that this walkthrough is when using SILA[1] with the WebCT Physics course although it should also give a flavour of its use with any material.

1. Starting. To start the software the student must click on the icon on the desktop identified as 'SILA' plus a version number. This should result in the main SILA

[1] The screenshots show development version 4.0, a previous version to Beta 1.0. All information is still correct.

windows displaying (select 'no' to the default summary prompt in SILA Beta 1.0). It is then necessary to enter the URL of the CHALCS WebCT course into the box provided, that is enter: http://chalcs.org.uk:8900/ Assuming there is a live Internet connection the 'myWebCT' window will be displayed in the top window. If the student does not have a myWebCT global i.d. then 'create myWebCT' will have to be chosen and the procedure carried out as subsequently described on screen. If the student has a myWebCT account then selecting 'log on to myWebCT' will result in the 'Enter Network Password' dialogue being displayed where, obviously, the Username and Password must be entered. Figure 8.3 shows this dialogue with the myWebCT window in the background.

2. Homepage. Having gained entry to myWebCT the student must navigate to the Physics course. This is achieved by selecting "CHALCS Physics" in the myWebCT page. The result should be the 'Astronomy and Optics' homepage shown in Figure 8.4. This page shows the tools available within WebCT as well as the link to 'Course Content'. Selecting this link results in further links to the various content available. Selecting 'Lenses' leads to the next step.

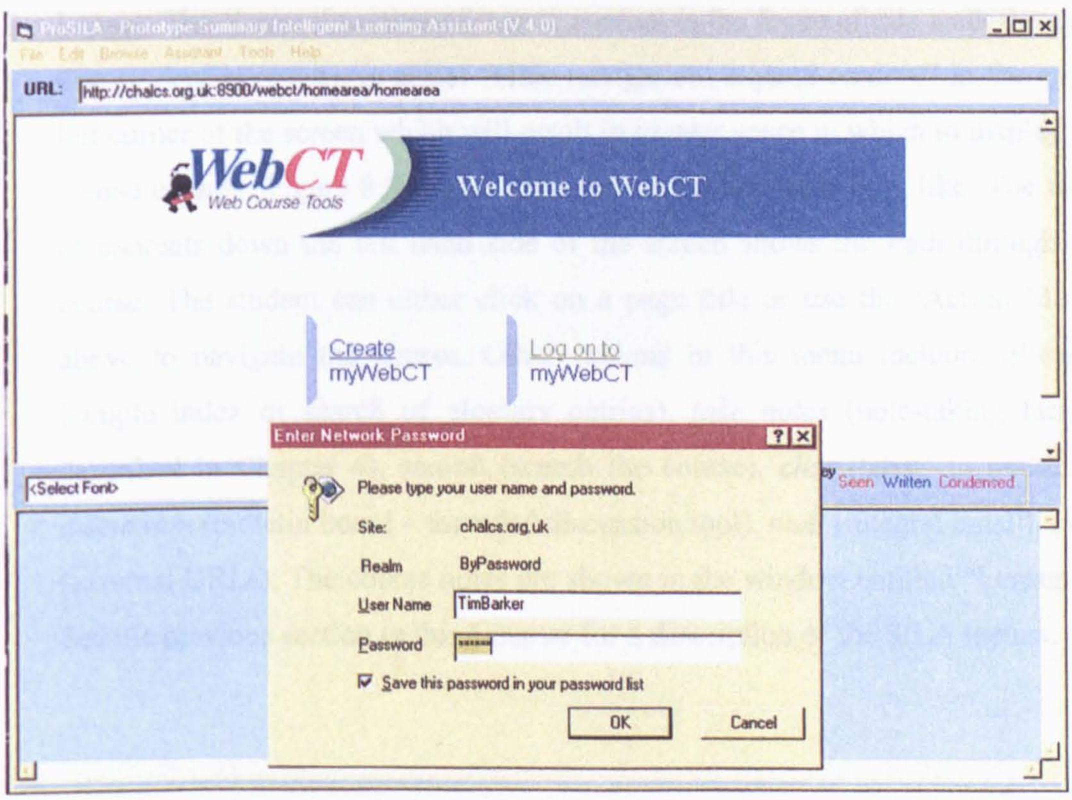


Figure 8.3 Starting WebCT.

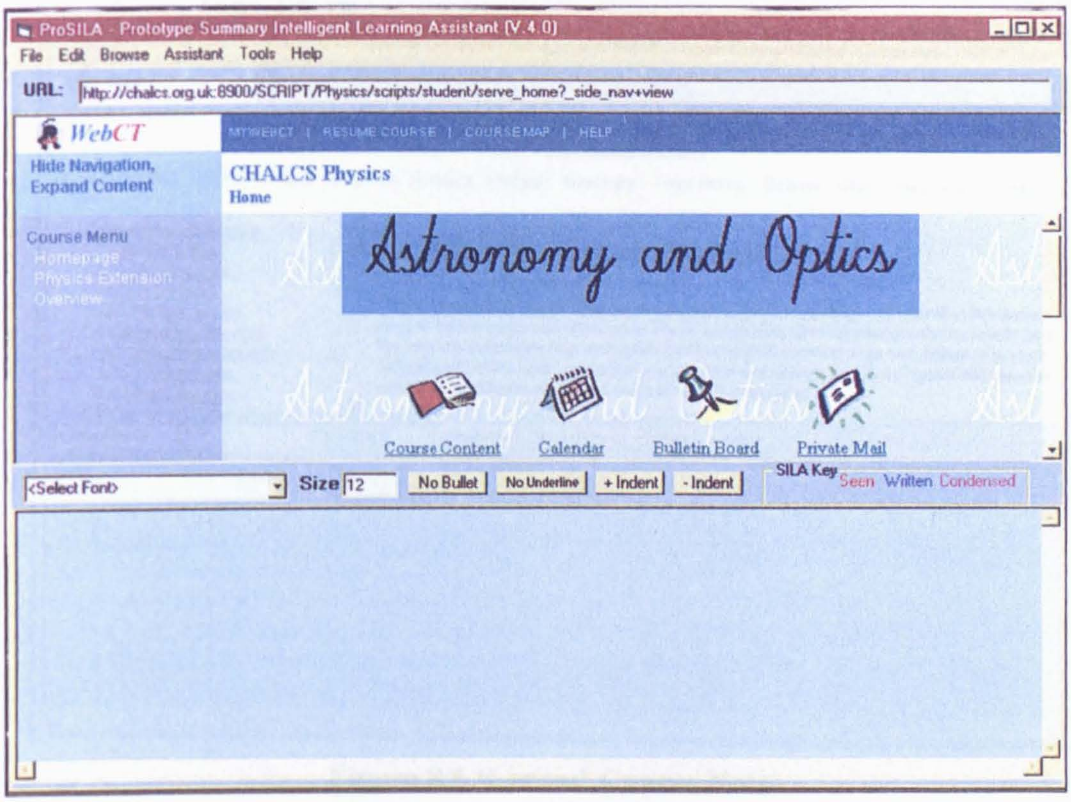


Figure 8.4 Astronomy And Optics Homepage.

3. Lenses. The 'Lenses' section of course content is the focus of this walk through. The student may want to select "Hide navigation, expand content" in the upper left corner of the screen which will result in greater space in which to display the course content. Figure 8.5 shows what the screen will now look like. The table of contents down the left hand side of the screen shows the Path through the course. The student can either click on a page title or use the 'Action Menu' above to navigate the course. Other options in this menu include: *glossary* (simple index or search of glossary entries), *take notes* (note-taking facility described in Chapter 4), *search* (search the course), *chat* ('chat' in realtime), *discussion* (bulletin board – threaded discussion tool), *mail* (integral email), *links* (external URLs). The course notes are shown in the window entitled "Lenses…" See the previous section in this Chapter for a description of the SILA menus.

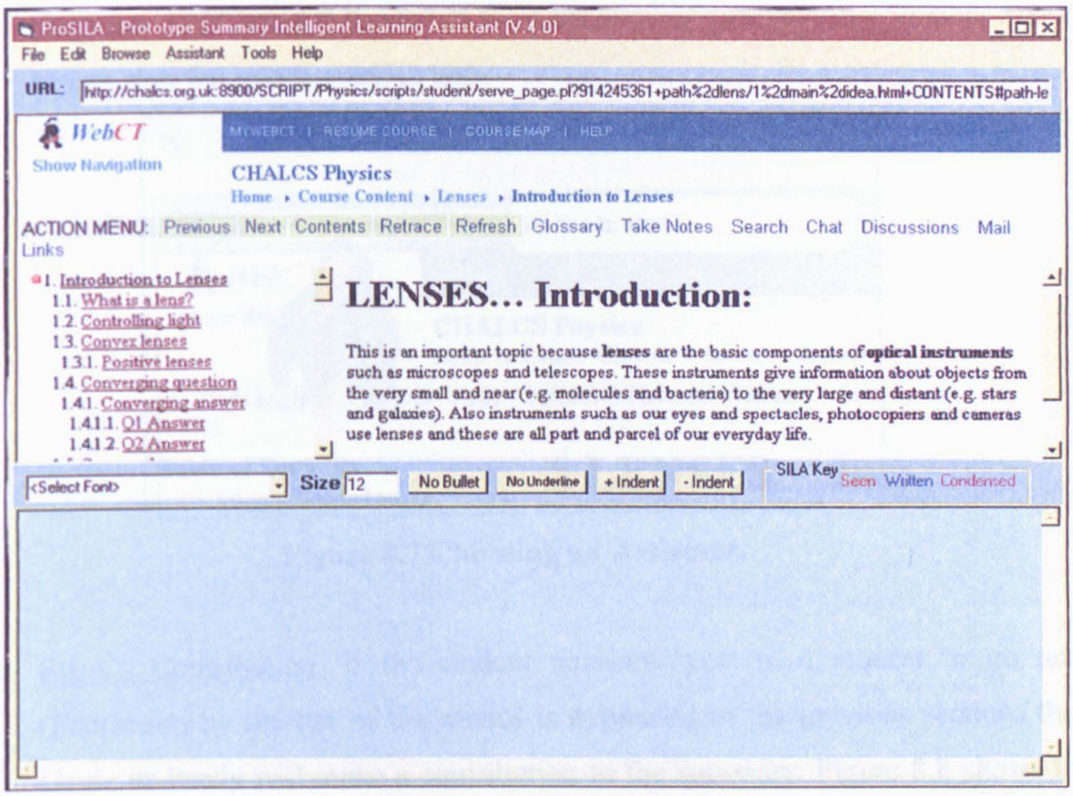


Figure 8.5 'Lenses' Course Notes.

4. Instructions. By selecting 'Instructions' in the SILA 'Help' menu Robby the Robot will appear and guide the student through the basics of using the system including the differences between Peedy the Parrot and The Genie and how to interact with them. Robby is shown in Figure 8.6.

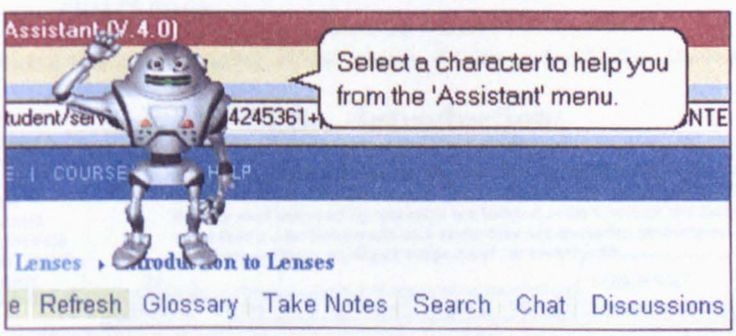


Figure 8.6 Robby's Instructions.

5. The Assistants. The first task for the User is to choose which assistant they would like to use by selecting the appropriate character from the 'assistant' menu. As is identified throughout this thesis, Genie can be thought of as a 'strong', 'dominant' character whereas Peedy is a 'friendly', 'more amenable'

character. In Figure 8.7 Peedy has been chosen and, after an introduction, asks if he can start the summarisation task.

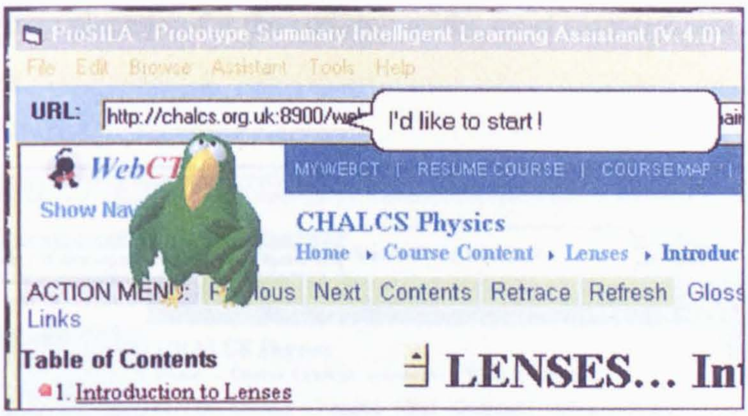


Figure 8.7 Choosing an Assistant.

6. SILA's Contribution. If the student answers 'yes' to a request to go next (interaction by the use of the menus is explained in the previous section) then Genie or Peedy will make a contribution to the summary. Figure 8.8 shows the Genie gesturing towards a phrase that he has added in the summary document window. The greater the indentation of a phrase, the less it's importance is to the summary. All of SILA's contributions will appear in blue text.

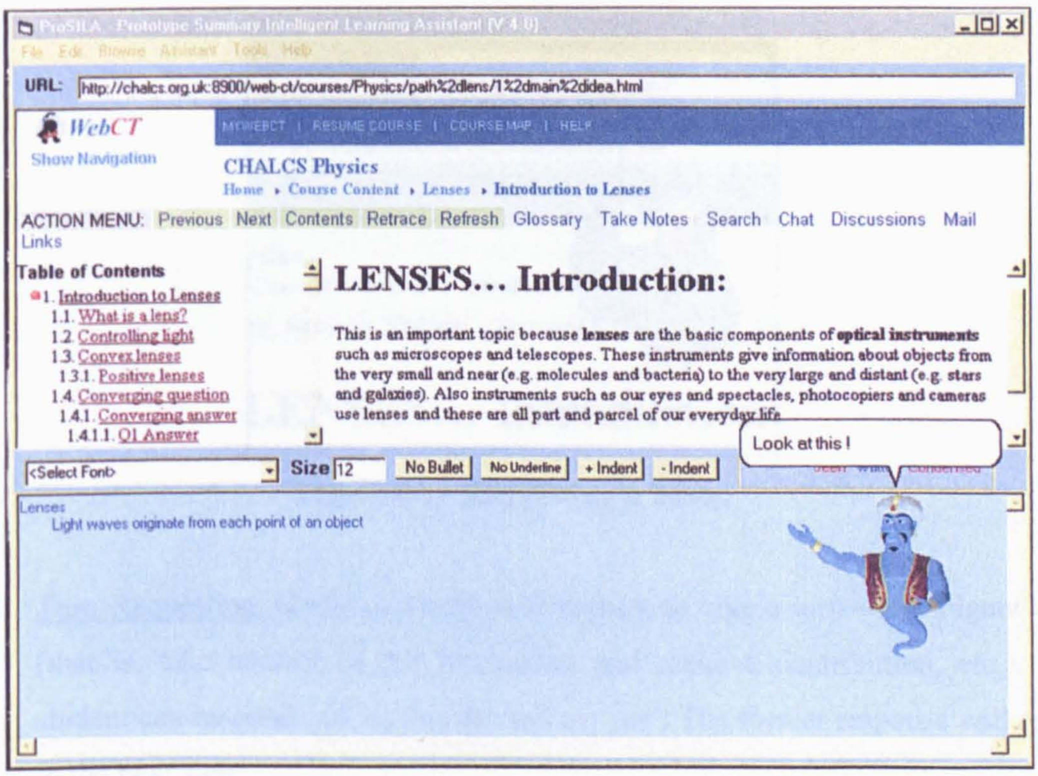


Figure 8.8 SILAs Contribution.

7. Justification. The student can ask SILA to justify the addition of a phrase by selecting 'why' in response to the Assistant. Figure 8.9 shows Genie beginning to offer an explanation for the addition of the most recent phrase in the summary document.

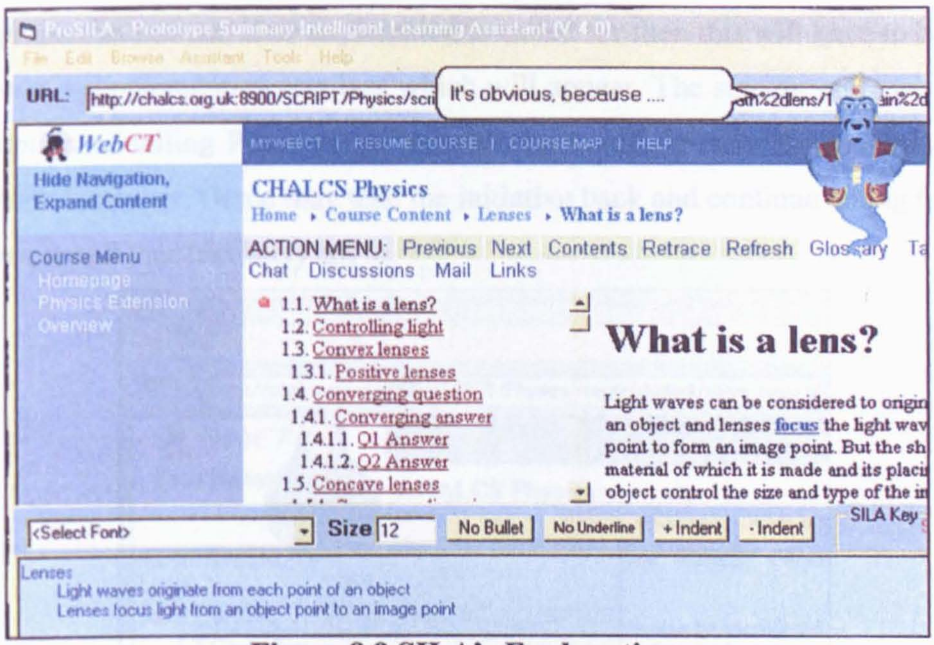


Figure 8.9 SILA's Explanation.

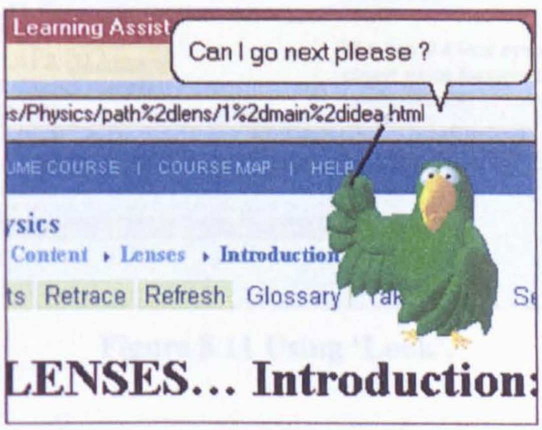


Figure 8.10 Requesting a Turn.

8. Turn Requesting. Genie or Peedy will request to take a turn – see Figure 8.10 (that is, take control of the interaction and make a contribution, etc.). The student can respond 'no' to this as well as 'yes'. The former response will result in the next step.
9. Showing SILA a Contribution. If the student does respond 'no' to Genie or Peedy's request for a turn then they may allow the User to go next. If so, then

the student should highlight their addition to the summary then 'show' it to the Assistant by selecting 'look', see Figure 8.11. The Assistants will subsequently analyse the contribution and it will change from black to red text if it is recognised. However, the characters may ask for an explanation, for instance if it is considered unimportant. Alternatively, if it is a good contribution they may praise the student. If an explanation is called for then this will have to be entered into a supplementary text box which will appear. The student will also have the option of telling Peedy to go next or requesting the turn themselves after one turn. However, Genie may take the initiative back and continue taking turns until requested (successfully) not to.

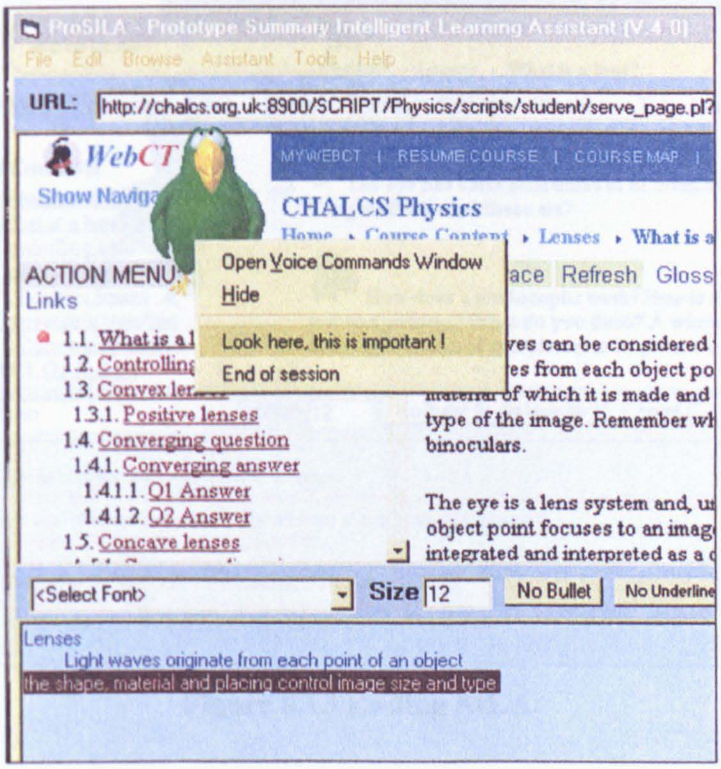


Figure 8.11 Using 'Look'.

10. Condensation. On occasions when SILA can condense one or more phrases it will do so on the turn of the Assistants. The results of the condensation will be temporarily highlighted as the process is executed. The resultant condensed phrases will appear in magenta in the summary document window. The original phrases will be deleted. Figure 8.12 shows this process as Peedy is condensing "lenses focus light from an object point to an image point" and combining it with "object point focuses to image point and onto the retina".

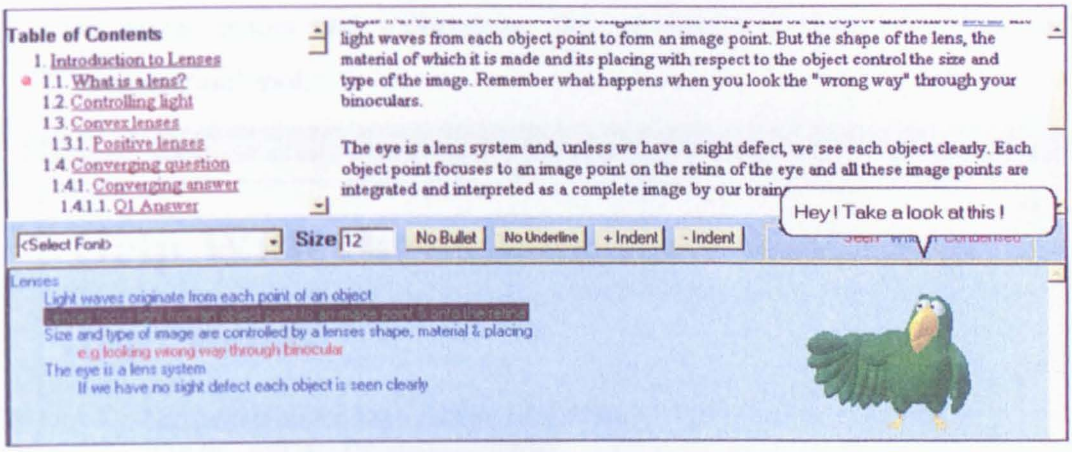


Figure 8.12 Condensation.

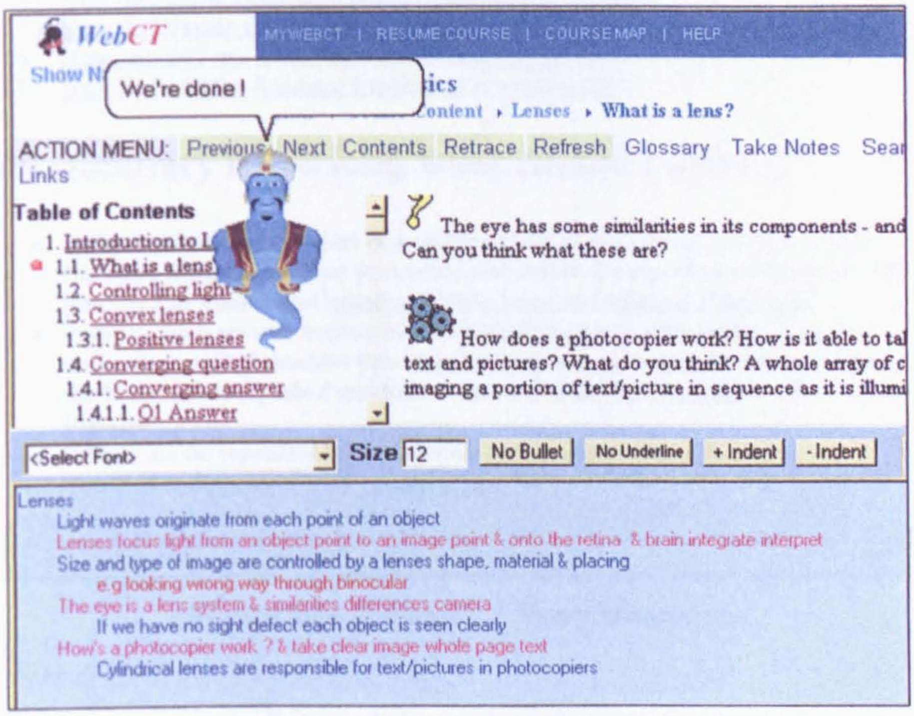


Figure 8.13 Ending SILA.

11. Ending When SILA contributes the last clause it will end. The Assistants will also end if they are told to 'end' at any time although the Genie may be 'stubborn'. Figure 8.13 shows the Genie applauding at the end of the summary shown in the bottom window. The assistant will then disappear and it will be the responsibility of the student to 'organise' the summary.
12. Organising. This final stage employs one of the 'tools' available to the student in the menus within SILA. On selection of the 'organise' tool a hypertext 'help file' will be displayed offering advice on, for example, deleting redundant information or the use of bullet points, etc. The student would then need to close the help file and edit the summary document accordingly using cut, paste, copy,

the bullet points and indentation, etc. as required. Figure 8.14 shows the 'organisation' tool.

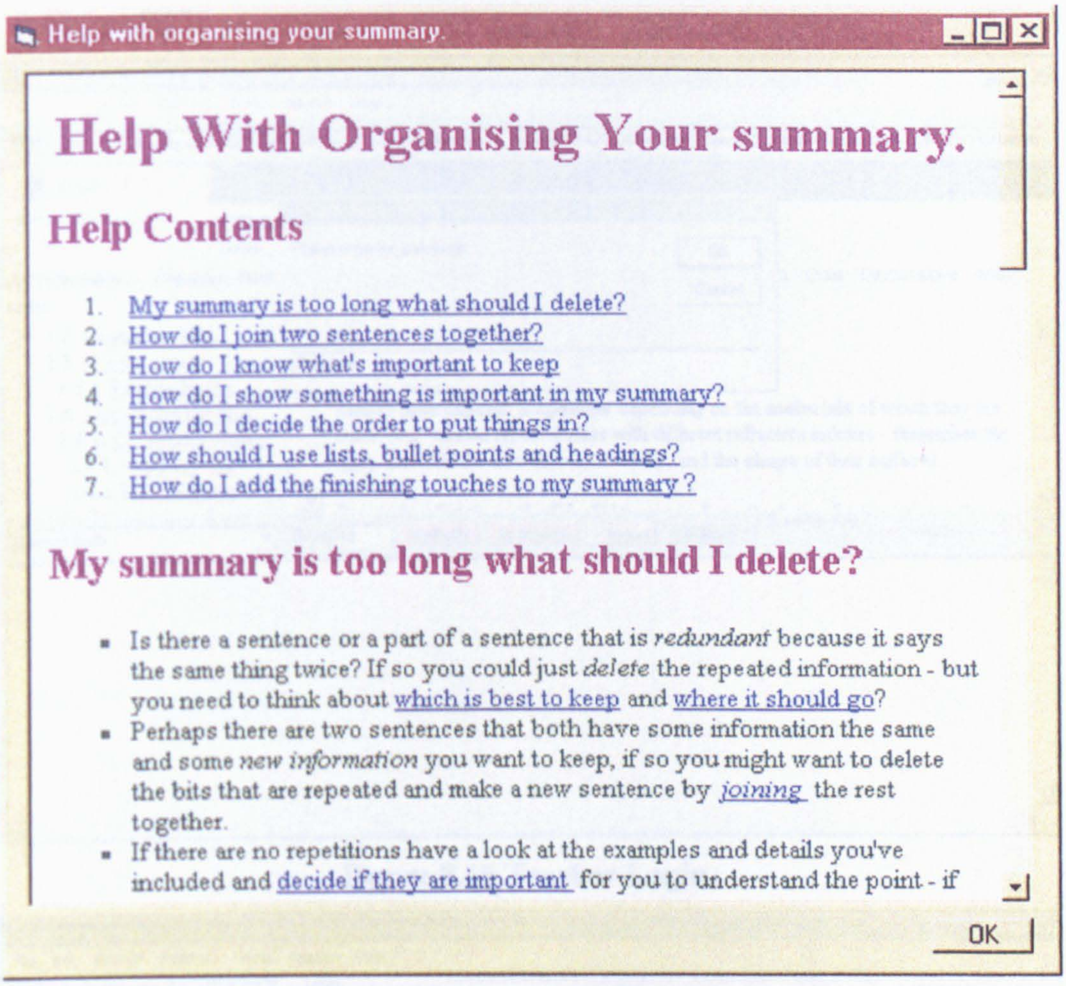


Figure 8.14 Organising Your Summary.

It is then the responsibility of the student to save the summary (using the 'file' menu) as rich text format which can be subsequently imported into Microsoft Word, for instance.

8.4 SILA Teacher Scenario.

As well as the typical student scenario presented in the previous section it is worthwhile providing an overview of a typical teacher interaction with SILA as there are a number of features specifically designed for the classroom instructor rather than the general student. Note it is assumed that steps 1 to 3 in section 8.3 have been followed.

1. Login. The first step for the teacher involves selecting 'login' in the 'teacher' menu then entering the password (currently set to 'Eggplant'). This is shown in Figure 8.15.

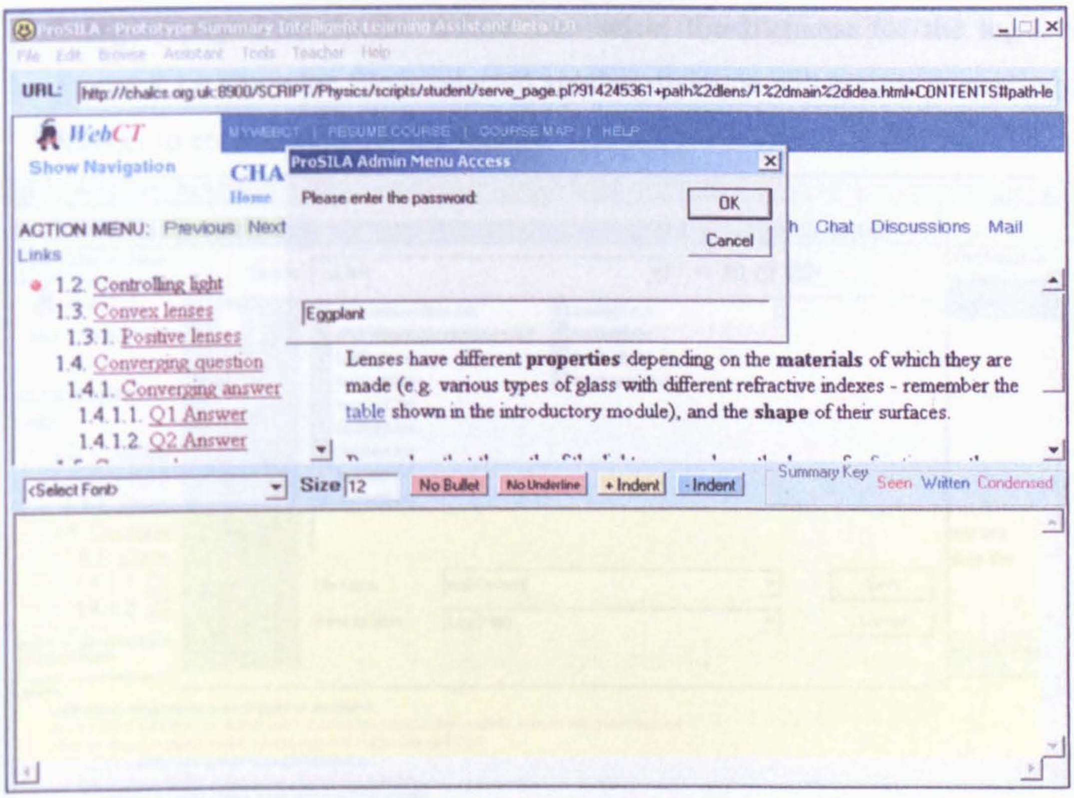


Figure 8.15 Teacher Login.

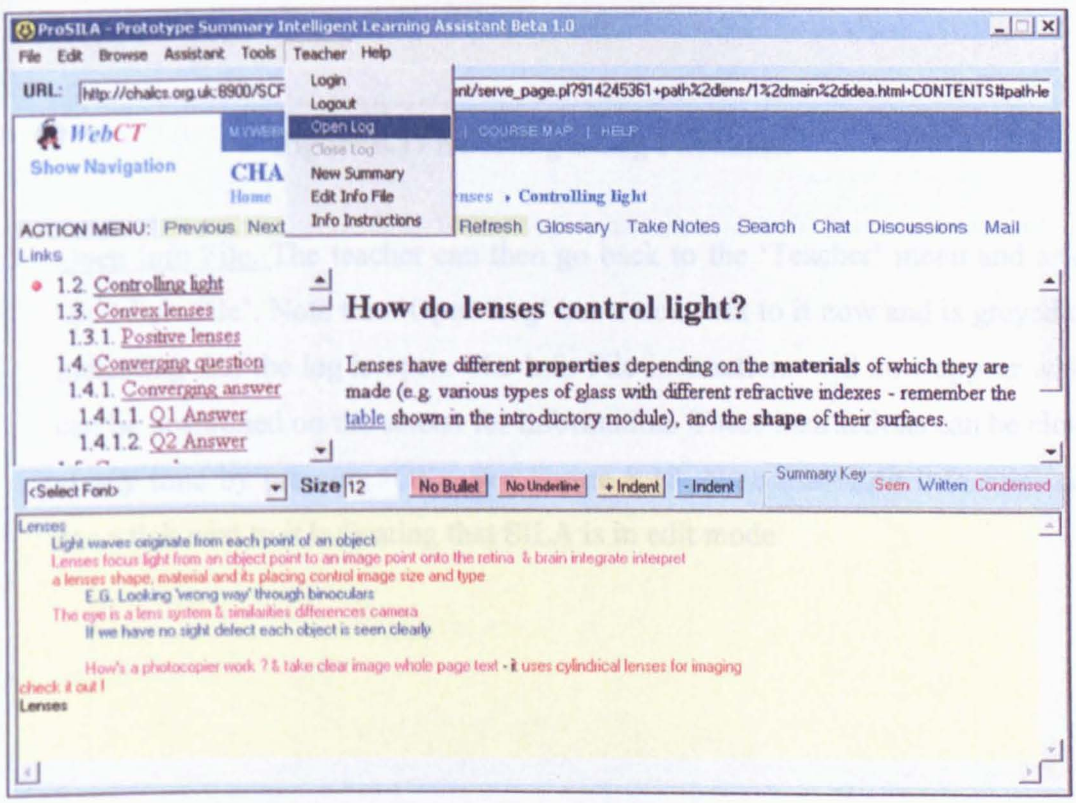


Figure 8.16 Opening a Log.

2. Open Log. Next the teacher will be able to select 'Open Log' as the full Teacher menu will be displayed – see Figure 8.16. This results in the Open Log dialogue being displayed where the teacher can select the filename for the log – see Figure 8.17. Once the file is selected another dialogue will appear prompting the teacher to enter a name for the student.

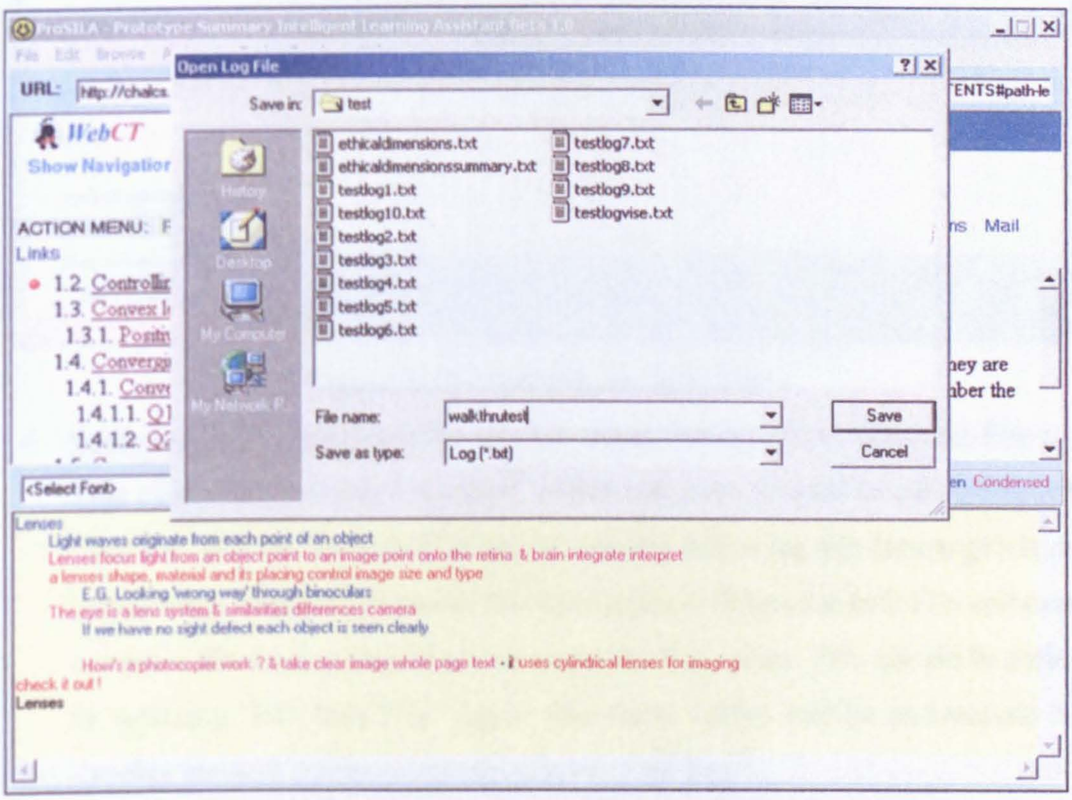


Figure 8.17 Selecting a Log Filename.

3. Open Info File. The teacher can then go back to the 'Teacher' menu and select 'Edit Info File'. Note that 'Open Log' has a tick next to it now and is greyed out indicating that the log is open. The Info File instructions will now appear which can be positioned on the screen for information. These instructions can be closed at any time by pressing 'OK'. See Figure 8.18. Note that 'Edit Info File' now has a tick next to it indicating that SILA is in edit mode.

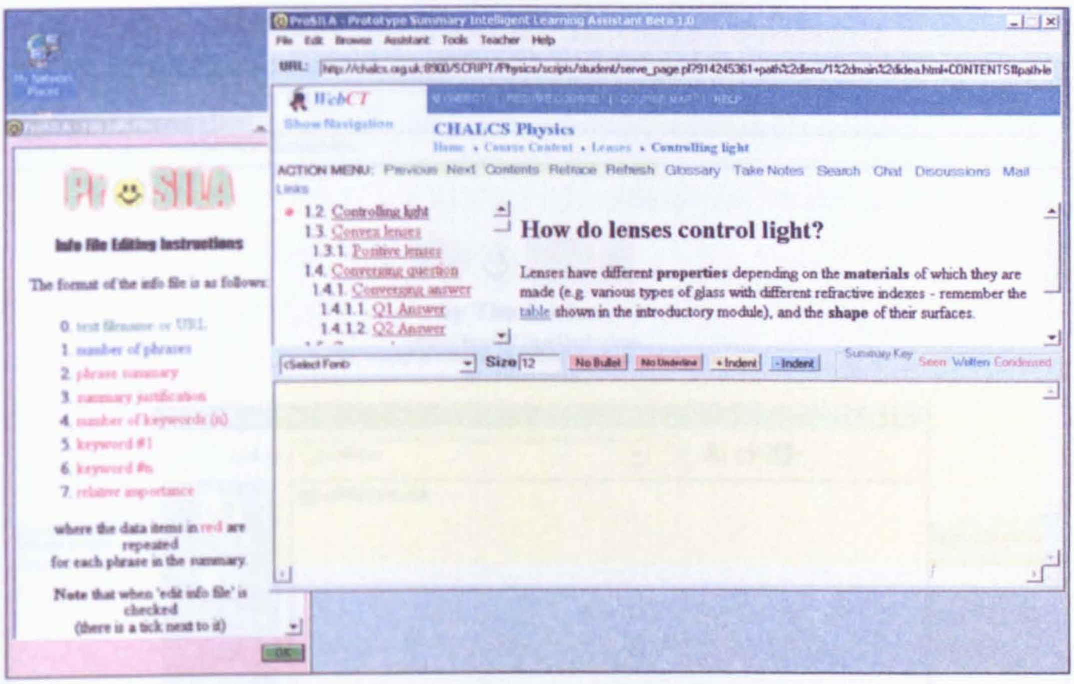


Figure 8.18 Opening an 'Info File'.

4. Enter Info File Data. Next the teacher should select 'New' from the File menu. This clears the 'summary window' which can now be used to edit the Info File. Thus the data for the Info File can be entered following the format given in the instructions. Figure 8.19 shows this taking place. When the Info File authoring is complete the teacher should save it using the File menu. This should be followed by selecting 'Edit Info File' again (the menu option will be unchecked) in the Teacher menu to return to normal summary editing.

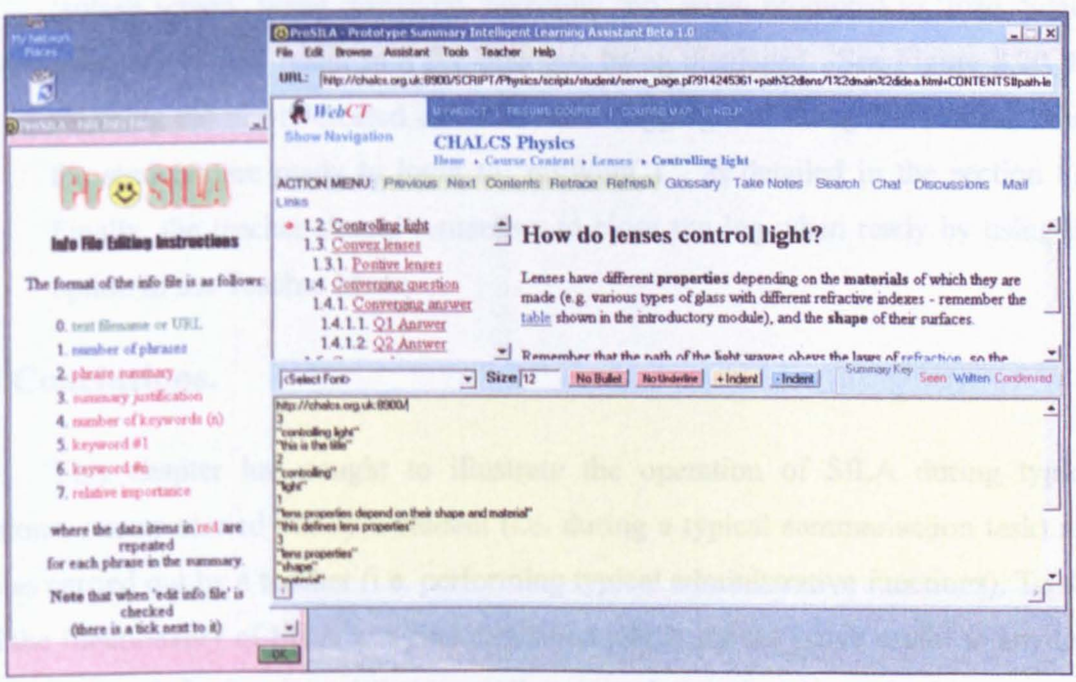


Figure 8.19 Entering 'Info File' Data.

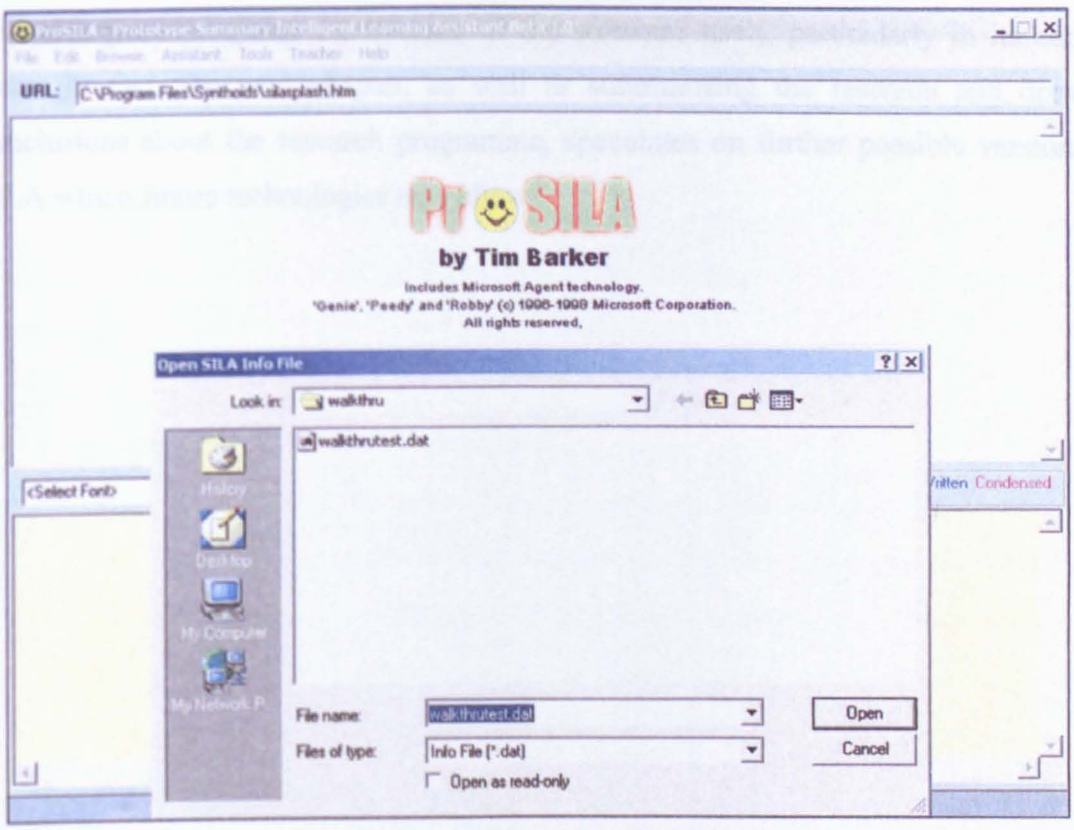


Figure 8.20 Loading a newly created 'Info File'.

5. Test New Info File. The teacher can then test their newly authored Info File by selecting 'New Summary' in the Teacher menu which will result in the initial 'splash screen' being displayed. Selecting 'no' when prompted to 'load default summary ?' will result in a file dialogue being displayed. See Figure 8.20. By selecting the newly-created Info File then 'logging out' using the Teacher menu the students are ready to login to 'myWebCT', as detailed in the section 8.3. Finally, the teacher should remember to close the log when ready by using the option in the Teacher Menu.

8.5 Conclusions.

This chapter has sought to illustrate the operation of SILA during typical sessions, one as carried out by a student (i.e. during a typical summarisation task) and one as carried out by a teacher (i.e. performing typical administrative functions). To this end the functionality of SILA was first described which should prove useful to any user of the system. It is thus hoped that as well as completing the picture in the context of this thesis by illustrating the software developed during the research this walkthrough will also persist beyond the lifetime of the software itself, particularly in its current form. In fact, the next chapter, as well as summarising the research and drawing conclusions about the research programme, speculates on further possible versions of SILA which future technologies may allow.

Chapter 9: Summary.

9.1 Introduction.

This final Chapter summarises the research presented in this thesis. This will involve a final evaluation of the outcomes of the research in relation to research goals including:

- A summary of the work completed;
- An examination of the 'contributions' of this research to knowledge in the field;
- An exploration of further work to extend and develop this research further;
- A discussion of the relevance of this work in relation to other more recent developments in the field outside the scope of this thesis but which may be complementary to future development of the field.
- Final conclusions.

9.2 Summary.

The Aims of this research were stated in Chapter 1. This section of this final Chapter will now revisit each of these five aims with a view to summarising the outcomes of the research project as a whole.

1. To design a pedagogical model to support both students and tutors working within a Virtual Learning Environment (VLE) in line with Chan's Social Learning System model - An argument was established in Chapter 2 to support the introduction of Virtual Learning Environments in schools such as CHALCS where there are few students studying specialised subjects so that, for example, the possibility of finding a 'peer' at a distance is facilitated (akin to Social Learning Systems). 'Collaborative Learning' was reviewed during the early days of this research (as reported in Chapter 2), especially Computer Supported Collaborative Learning. The arguments were based upon a historical account of the origins of (Collaborative) Learning from Educational and Psychological perspectives. Examination of this literature, together with observations at CHALCS, led to the development of a Pedagogical Framework.

2. To evaluate existing VLE design tools and to implement the pedagogical model as far as possible with a particular selected VLE instantiated with materials and tasks to support an area of the curriculum appropriate to CHALCS - To evaluate the effectiveness of the Pedagogical Framework the exemplar Physics course was developed in the Virtual Learning Environment WebCT (as reported in Chapter 3) which was chosen based on a review of VLEs available at the time. This led to a number of conclusions for the further development of the technological infrastructure, as follows.

3. To evaluate the potential of this instantiated VLE for supporting the development of key skills within this area of the curriculum - The study reported in Chapter 4 was carried out in order to ascertain the nature of further support required, if any, and the effectiveness of computer-based support for the development of key skills. In addition key skills were examined in the context of writing skills and summarisation for the purposes of this research. The main outcome of this stage was the development of a process model of summarisation based upon the empirical studies. Additionally, automatic summarisation by computer was examined with a view to informing the possible design of suitable software.

4. To consider key skills in this curriculum area which might benefit from the addition of Agent-based support to the VLE and, focusing on one such skill, develop an exemplar Learning Companion to support it - the conclusions drawn (see Chapter 4) were that WebCT provided inadequate support for the summarisation process as enacted by the CHALCS A Level Physics students and thus the fundamental primary requirements of the pedagogical model, that of Acquisition, were not sufficiently being met. Therefore a Learning Companion was designed and developed (utilising a User-Centred methodology, see Chapters 5 and 6) to aid students when summarising their course notes.

5. *To trial the Learning Companion to suggest Agent design characteristics which may prove motivating and effective in assisting students to learn from the VLE material* - the prototype Summary Intelligent Learning Assistant (SILA) was validated with sample students from CHALCS and a local school (see Chapter 7).

In addition, the notion of 'Affectations' was proposed which sought to further engage students in the summarisation tasks thus resulting in greater learning gains. The result was a validation of certain implemented agent design characteristics as well as recommendations for future characteristics.

9.3 Research Outcomes and Contributions.

The main hypothesis at the centre of this research and behind the development of the prototype SILA system was:

> *"The introduction of an Animated Pedagogical Agent displaying appropriate 'Affectations' into a VLE can help students summarise web-based course notes and may result in a motivating, engaging and rewarding experience."*

The aim of the research was to realise Chan's (1996) vision of Virtual Learning Companions at CHALCS where it was hoped disaffected students would be motivated by the use of affective agents in VLEs employing *deliverable technologies* to help with note-taking - a key study skill necessary for initial 'acquisition' of web-based notes. By encouraging an active learning approach to note-taking through collaboration with the agent, it was hoped that students would sustain attention sufficiently to produce a joint product that they could use for revision purposes and, in the process of constructing this joint product, they would think more reflectively about the processes of selecting and condensing materials to form a 'good' summary.

The following list of seven points expand upon this aim and suggest the main contributions of the work.

1. *Synthesising and partially evaluating an appropriate 'Pedagogical Framework' for the successful execution of a web-based Virtual Learning Environment.* The pedagogical framework reported in Chapter 2 is the result of active research alongside tutors in an authentic learning environment and as such it represents a response to both student and tutor requirements. Important outcomes of this work include observations related to the requirements of a "my notes" facility.

Results showed that the current WebCT version of this facility does not have the necessary functionality to meet student needs and that, moreover, students require additional interaction and stimulation (if they are to take adequate notes) that would be costly for a human tutor to provide.

2. *Implementing a Summarisation CBL system using the Wizard of Oz Prototyping Methodology in Animated Pedagogical Agent design.* The Wizard of Oz prototyping method is gaining popularity and the technique has been used in the design of intelligent help systems (Maulsby, 1993 and Breuker, 1990) and more recently with Primary school children at the University of Edinburgh – Robertson, 2001). The application of the technique to the design of affectations that motivate interaction has to the author's knowledge not been reported elsewhere. This contribution is reported in Chapter 5 and has been shown to be an effective technique for User-Centred system design.

3. *Providing 'Affectations' of an Animated Pedagogical Agent (particularly as a first step to "having" emotions).* 'Affectations' are a central concept of the Affective element of the Study Companions. Although there is no central model of either the Agent's emotional state or that of the student, it is hypothesised and, to a degree, justified by the results of the evaluation in Chapter 7, that these Affectations portray convincing behaviours in the sense of being both believable and engaging whilst being implemented using robust "off the shelf" technologies.

4. *Integrating research on (Automatic) Summarisation with collaborating, personified Agents (or 'Learning Companions').* The research on summarisation, particularly automatic summarisation, together with elements of Tawalbeh's model of summarisation which made use of the rhetorical structural properties of text was reviewed in Chapter 2 and informed the implementation of the system described in chapter 6 and 7 of the thesis. In particular, this previous research on summarisation and the results of the empirical work on summarisation (as reported in Chapter 4) led to a process model of summarisation which when combined with knowledge of text structure was used to inform the design of a Learning Companion capable of collaborating with students to construct a

summary. The agent was capable of suggesting moves at appropriate stages in summary construction based on rules governing strategic content selection and condensation (these rules are described in Chapter 6).

5. *Investigating Usability of diametrically opposed characterisations of Animated Pedagogical Agents as realised in multiple modalities.* This claim is related to the two distinct 'personality' types of the two Agents. Using the opposite adjectives (as explained in Chapter 6) to typify the Agents, as realised in their design utilising the five subtle expressivities, one was expected to be perceived as 'dominant' and the other as 'passive'. Results of the evaluation showed that the design was effective in creating the impression of two clearly distinct 'personae' based on ratings on opposite adjective scales. Subsequent analysis allowed the usability of each Agent persona to be evaluated and it was possible to show a distinct preference for one of the 'persona' by this means.

6. *Investigating Summarisation processes and strategies in post-compulsory Physics Education.* As already mentioned, one outcome of the study of summarisation in Chapter 4 was a process model of summarisation and the demonstration of individual variance in approach. Findings reported in Chapter 4 were further supported by the evaluation study in chapter 7. The summarisation study, in addition to showing that some students were more serial in their processing of text whilst others were more holistic (exploiting hyperlinks), also showed that some students used a strategic approach based on the knowledge of discourse structure – skipping to the end to read the summary first. These findings are interesting in their own right and also serve to suggest the extent to which future agents may be improved by being able to flexibly adapt to the style of the individual.

7. *Making recommendations for future Animated Pedagogical Agent (summarisation) systems based upon empirical evaluation.* Chapter 7 and the remainder of this chapter state a number of directions for future research in the area of Personified Pedagogical Agents, particularly those related to the support of students' study skills.

9.4 Future Work.

This section on future work has been divided into work which could be carried out on the Learning Companion and that which could be carried out on the implementation of further domains. Additionally, it is expected that the studies reported in the previous chapters (Chapter 3: The Development of a Web-Based Physics Course, Chapter 4: The Summarisation Studies, Chapter 5: User-Centred Design of an Affective Learning Companion: Utilising The Wizard of Oz Study and Chapter 7: Validation of the Learning Companion) could benefit from being repeated with, for example, a larger sample size than found possible within the context of this PhD research due to the pressure on the schools' curriculum and the small and unstable population of A level Physics students at CHALCS.

9.4.1. Further Learning Companion Development.

Further work could be carried out on the development of the final prototype SILA relating to improvements in the environment, improvements to the Agents and system-level improvements, as follows:

9.4.1.1. Improvements in the Environment.

1. **update the organisation document** – the HMTL help document designed to give additional tutoring on summarisation techniques is by no means complete or exhaustive and additional content support could be added. More ambitiously, the organisation of a summary could be better integrated into the main dialogue model by having the agent refer to it, for instance at the end of the session. As it stands the organisation content is static rather than dynamically created or generated in response to a particular move on the part of the student. For example, another means of integrating it further would be to call up chunks of the text in response to a student request for an explanation from SILA.
2. **implement lists** – this would allow lists of points to be included in a summary and subsequently understood by SILA (i.e., both originated by the student and SILA), an essential summarisation technique. Presently SILA recognises just one phrase at a time.
3. **implement links to WebCT** – if a student could embed hyperlinks within their summary document to WebCT this would also tighten the coupling between the

two systems. So, for example, if the student finds at revision that they have been over selective in constructing their summary they can follow up more detailed explanations, definitions and/or examples.

4. **implement 'undo'** – this would allow a student to undo a condensation, for instance – this might help students reject some of SILAs condensations more easily rather than just "go with them" anyway because they are hard to reverse. This could be handled through adaptation of the dialogue model to enable a stage in which the student actively accepts or rejects the suggested change and (optionally) provides SILA (and therefore the tutor looking at the record of the interaction) with feedback as to why the change is liked or disliked. This could also prove instructive in suggesting the maturity of the student's own condensation skills.

9.4.1.2. Improvements to the Agents.

5. **implement 'you finish'** – currently this move is not implemented, it would allow a student to direct SILA to complete the summary. [NB. this was not improved as it may not be desirable since some students may use this facility as a lazy way to auto-summarise – an alternative would be to give the tutor the option of disabling this function to prevent students from using SILA as an auto-summariser when the aim was for students to actively rather than vicariously learn how to summarise for themselves].
6. **implement multiple predicate recognition** – currently SILA only recognises single phrases shown to it by the student therefore if the student has constructed a number of phrases in their summary they must show each one to SILA in turn, i.e. by highlighting each phrase and selecting 'look'.
7. **implement 'moods'** – two kinds of 'moods' were envisaged for SILA in an attempt to further personify the system. The first, 'temper' or 'sulk', would involve the Companions quitting if disallowed a number of turns in sequence. The second 'confidence' would ensure that they continue with their turns if allowed a number of turns sequentially. [Whether or not these prove helpful is an open question which would need to be investigated].
8. **implement student selected summary detail** – simply by including phrases of different ranges of hierarchical level depending on student preferences different levels of detail of summaries would be produced.

9. **add varied moves and new characters** – vary choice of moves and add new characters as assistants resulting in less repetitive and potentially demotivating extended interactions.

9.4.1.3. System-Level Improvements.

10. **develop a mark-up tool** – it is envisaged that a future version of SILA will include a tool to help tutors author the kind of information contained in the current 'info file' i.e. a phrase summary, justification, keywords and hierarchical place. This tool could use statistical analysis, together with knowledge of discourse markers, to automatically highlight keywords, for instance, which could be accepted or rejected by tutors. In this way the process of mark-up would be much streamlined.
11. **dialogue model editor** – a tool to edit the dialogue model, e.g. graphically, could be implemented enabling a non-programmer to more easily alter the core behaviour of SILA. For example, tailoring it to specific scenarios or varied student groups.
12. **adapation** – adjusting the affective and summarisation ability of the agents and hence personalising interactions would further work towards the goal of student motivation, engagement and reward.

To conclude, the present system although leaving room for improvement nevertheless suggests the potential utility of the approach, forming a solid foundation for future development along the lines alluded to here.

9.4.2. Further Domains.

As reported in Chapter 3, the rationale for the development of an Advanced Level Physics course was that it was a subject with which students most wanted help at CHALCS. However, the development of a note-taking assistant has the potential to help students in a wide range of subjects including the Humanities and Social Sciences as well as subjects such as Biology and Chemistry. One of the original attractions of the note-taking context was that the skill is not one which is highly domain specific and potentially the approach is extendable to other domains.

Additionally, there is a large number of younger students at CHALCS who study a range of subjects. As study skills (or 'key skills') are seen as important at CHALCS it may be beneficial to explicitly represent these as a further WebCT course to be used by these younger pupils and their teachers both in face-to-face and distance learning modes. In particular there is a long established 'Literacy' project at CHALCS with a dedicated and well-motivated tutor who may also benefit from on-line learning materials. Of course, before any further development takes place a feasibility study would need to be conducted.

9.5 Discussion.

As mentioned, this section seeks to seed discussion of the more esoteric aspects of the research where perhaps no hard answers are yet possible. From discussion of concerns regarding Pedagogical issues to those based upon experiences of cultural impact on the project and future developments in technology these issues are elaborated and discussed. Finally, the question of appropriate methodology for this research is discussed before the thesis is rounded off with a summary of the work.

9.5.1. Pedagogical Concerns.

A pedagogical concern for this research was the initial scepticism inherent in the CHALCS tutors towards the development of their own WebCT course notes. As part of the remit of this research was to deliver learning using ICT solutions it also became clear that some exemplar course notes would have to be developed by the teams mentioned in Chapter 3. This did, however, have the result of causing apprehension amongst the CHALCS tutor and an unwillingness to accept responsibility first for their use within the normal lessons and second for their subsequent updating to bring them more into line with their own thoughts on the subject. Eventually, the tutor was persuaded to utilise some of the material contained within the WebCT course and consequently developed their own material to supplement that developed by the Physics expert associated with this research.

These kind of issues have consequences for the development of content by authors who are not the primary face-to-face tutors involved in the delivery of the material. As content is authored by outside parties to the physical school system (the model currently adopted by WebCT) tutors based in that school may have problems accepting the material as their own leading to alienation and a decrease in motivation amongst staff with obvious consequences for the education of the students. On the other hand, it is not an easy task to train a tutor to enter their own notes within packages such as WebCT as it presents them with an additional burden when already over-burdened. The solution, perhaps, is to establish teams (perhaps at the level of Local Education Authority) which involve the tutors at ground level in a tight consultation process, something akin to the user-centred design philosophy espoused in this research. It is then hoped that this will alleviate the feelings of lack of ownership of the developed material and subsequently encourage tutors to take a more active role in the development of ICT delivered courses.

Clearly in this context, the addition of intelligent agent-based support presents similar problems. The extra burden of providing suitable mark-up to course notes to enable the assistant to perform its help-role could be prohibitive. The requirement to provide tools that would render this process as automatic as possible would be essential. In addition the implementation of such agents would need to be sensitive. In this research the decision was taken to implement the note-taking guide, in part, for tasks normally considered to be tasks undertaken by students on their own and in part because this would preserve the human tutor in the primary role of giving assistance within his specialist domain such as help with problem-solving.

9.5.2. Cultural Concerns.

There is undoubtedly an issue of appropriate cultures with which to introduce ICT-based models of learning. The culture at CHALCS is geared towards the integration of ICT into the curriculum as it is central to the Aims (as stated in Chapter 1). This means that teaching practice already exists which accommodates ICT into day-to-day lessons. Obviously, this means that CHALCS had an initial head start over the feeder schools, particularly in regard to the provision of hardware, software and internet access at the beginning of this project. It would be fair to say that the positions in this regard have changed dramatically as a result of government initiatives to equip schools, provide technical support and training. However, at the time the research was carried out schools were not ready to respond. It can only be hypothesised that this was due to them not seeing such work as a priority in an already stretched timetable. Although this might seem to cast doubt on the general utility of the work for mainstream schools it is likely that as ICT literacy has become so much more central to the curriculum within the state Educational system during the lifetime of this project (as evidenced by the NGfL initiative, for instance) that were schools to be asked to trial a VLE today considerably more interest would be generated. It is hoped that the studies which took place at the CHALCS feeder school and also with the younger students from other outlying schools (see Chapters 4, 5 and 7) will help lend credibility to the notion that VLEs have a useful role to play in teaching and learning

9.5.3. Future Technologies.

A limiting fact with the design of SILA is the amount of 'understanding' involved in recommending, for instance, phrases to include within a summary. As part of the initial design investigations schemata were constructed which represented the relationships between key concepts, properties or instruments, etc. These schemata exposed the underlying relationships inherent in the *semantic web* of the course notes. Berners-Lee and colleagues at the World Wide Web Consortium have proposed a revolutionary vision for the next generation of web-enabled information which they are calling the Semantic Web.

> "The Semantic Web is the web of connections between different forms of data that allow a machine to do something it wasn't able to do directly."
>
> (Berners-Lee & Fischetti, 2000)

The consortium has proposed a Standard, Resource Description Framework (RDF), for representing this kind of meta-data which will bring to light the relationships between data such as those found in many of the relational databases driving the statically-rendered HTML pages of the present day. RDF works by encoding meaning in triples of things, properties and values which can be expressed in XML tags. The hope, then, is that automatic reasoners or agents will begin sifting through the web of meta-data using appropriately defined Ontologies. It is thus speculated that in the not-to-distant future SILA may be able to reason on and 'understand' the course notes which will result in a much more 'knowledge-rich' based approach to summarisation than is currently practically possible and which would overtake the need to develop author mark-up tools along the lines described here.

A further possibility which lies in the nearer future is the advent of affectively aware computers. Such computers will use advanced affect recognition techniques, such as voice and facial analysis (Picard, 1997) to update their model of the affective state of their User. In addition, further, more intrusive, techniques such as psychophysiological measurements (like heart and breathing rate) may allow a more complete picture to be built of the affective state of an interlocutor. Over time an appropriately designed computer-based system could build up a picture of a User, looking for patterns of moods for instance, so that it could predictively adjust it's affective qualities which it knew from experience will result in a productive interaction. Indeed some work in this area has already taken place (Kapoor et al., 2001).

The relevance to this work would be for SILA to adopt certain personality types based not just on self-assessment of either affective predisposition or level of attainment (as has been proposed) but also on its own assessment of the student's state during prolonged and successive interactions. Of course, there are further research questions to be answered here which will have a bearing on such developments, such as the correlation between *perceptions* of affective states and the *experience* of those states. These will require further empirical and theoretical investigation. Indeed Barker (2003) proposes such further investigation to test the *adaptation-performance hypothesis*: "adapting agent affect to affect student affect based on student prior performance will affect student subsequent performance". That is, by basing the adaptivity of a Learning Companion's interactional or 'social skills' aspects on a student's prior performance at a domain task, e.g. summarisation, the student will better respond affectively to the Companion resulting in an improvement in performance. This hypothesis is based on the results expressed herein as well as a 'contagion assumption' which states that student perception of agent affect leads to a corresponding student affective response. These experiments would ideally utilise typical control groups as well as performance pre and post tests together with student self-reporting and psychophysiological measurements. The wider vision, in line with other researchers, e.g (Johnson et al., 2003; Prendinger, 2003), is to enable 'Social Computing' through provision of such socially responsive, personalised educational agents.

Furthermore these varied personality types will require new technologies to express them. By way of illustration Thomas and Johnston provide four major rules used at Disney for expressing dialogue. Of interest here is Rule 4:

> "Dialogue must be written so the actor doing the voice can contribute something. Without changing the meaning of the line, it is usually possible to give it more life by rephrasing or adding a touch that gives the actor a better chance. Often a chuckle, a sigh, a stutter, swallow, or gulp will reveal more of the personality than the words themselves. Other times, the actor may have an intimate knowledge of folk phrases from certain regions that will help build a more interesting character."
>
> (Thomas & Johnston, 1981)

In other words, there is a range of improvisational devices or indeed well-rehearsed techniques that an actor can bring to bear on dialogue to 'bring it to life'. Such embellishments inevitably lead to the portrayal of a richer and subsequently more convincing character. However, as Lewis Johson points out (Johnson, 2003) speech synthesizers are not designed to produce a "chuckle, a sigh, a stutter, swallow or gulp". Hence this technology and most likely others require further humanising if the goal of effective character portrayal in personalised, social agents is to be achieved.

9.5.4. Methodological Concerns.

Out of necessity a range of research techniques have been used from the use of the semi-structured interview of students in the early summarisation studies to the use of the Wizard of Oz studies in the design of the SILA software. It can be seen that by examining the array of results necessitated in such an elaborate enquiry that no one technique, or indeed, methodology is entirely suitable for the task at hand. Instead, both qualitative and quantitative data gathering and analysis lived side-by-side in an attempt to form a more holistic picture of the object of study.

However, what was clear was that the classical notion of 'scientific enquiry' with its associated control groups and dependant and independent variables were entirely inappropriate for the kind of case-study research that was being undertaken. That is, it would have been impossible, for instance, to randomly sample a group and assign them to separate studies (a) it would not have been allowed by CHALCS who needed to retain control of group composition and (b) the groups being studied were relatively small. This small sample size necessitated or enabled more in-depth study of the phenomenon being studied and hence lent itself more to a qualitatively-based analysis. On the other hand analysis of the data provided by WebCT such as statistics for page accesses or the results from the Affect Scale lend themselves to more quantitative techniques. When both quantitative data and qualitative data are taken together and triangulated the result is a rich description of the object of study revealing more of the processes which students are engaged in during learning and this can help generate research hypotheses which may be suitable for further investigation, such as the adaptation-performance hypothesis proposed above.

9.6 Final Conclusions.

This research has taken place both at Chapeltown and Harehills Assisted Learning Computer School and a CHALCS 'feeder' school, Notre Dame Sixth Form College. It has sought to introduce an ICT-based solution to the problem of few tutors and peers and consequently improve access through the development of an exemplar course in Advanced Level Physics utilising WebCT.

Through careful evaluation of the WebCT Physics course at CHALCS and other schools it was felt that the support provided which enabled students to create their own notes was lacking. It was noted that it did nothing to aid the process of summary note-taking as well as failing to provide basic editing facilities. Based on the literature, summarisation was seen as an important element of active learning, facilitating engagement with the material and deeper understanding, an important activity in the initial 'acquisition' phase of the prescribed pedagogical model. It has also been highlighted by Government initiatives as a key skill for students to acquire.

A solution to the problem of inadequate summarisation support was proposed in the form of an Agent-based artificial Study Companion realised as deliverable computer technologies and operating in the same WebCT environment as the student. To this end a user-centred design of such a companion was undertaken. It was further hypothesised that to enhance students' experiences of interacting with such an Agent it should utilise appropriate 'Affectations'. These Affectations can be thought of in terms of behaviour-based Affective qualities of the Agent which, it was demonstrated, can help motivate students and lead to engagement with subsequent learning gains.

The Companion was implemented in a prototype system and evaluated with students mostly from CHALCS and Notre Dame. It was found that students, in the main, preferred to interact with the character which they perceived showed the most positive valence (Picard, 1997) in terms of its affective qualities. In addition, it was generally found that students who performed the best and the worst on the summarisation task (as evaluated by independent experts) perceived the most negative valence of the characters. In other words, the 'average' performing students perceived the greatest positive affect in the Companions. It was hypothesised that this was due to the more able students finding the character 'annoying' and 'interfering' whilst the less able students found the character 'smart' and 'condescending'. These hypotheses require further exploration since the prior ability or summarisation skill of the students was not available for all students and the numbers of students involved was low. However, the suggestion is that future agents should adapt affectively and in terms of ability based upon an evaluation, at least, of student ability resulting in an increase in student performance.

Finally, further work was proposed in terms of the Study Companion development such as that possible due to future technologies. Further domains for the application of the SILA system were also discussed beyond the Physics course described here. It is hoped that a research programme will be established building on the work reported here which will continue to take the research forward in the directions outlined in this Chapter and elsewhere, i.e. towards a more socially-oriented style of Human-Computer Interaction deemed particularly relevant to Computer Based Learning.

It was stated at the beginning of this thesis that for a modern society to function a host of complex knowledge sources must be available to its citizens. As more and more constraints are placed upon these global citizens, in terms of time and distance limitations, Virtual Learning Environments (VLEs) may help to bridge the knowledge divide, particularly between poorer and richer nations, by widening provision through flexible and life-long learning. However, VLEs by themselves are certainly no panacea to the problems of these new styles of learning. VLEs need to be utilised by educators in appropriate ways and furthermore require further technological development if they are to solve some of the emerging problems. However, the utilisation of the kind of techniques presented in this thesis used in conjunction with future developments such as those discussed in this chapter could move us closer to the goal of an egalitarian educational provision with resultant implications for the currently prevalent poverty to be witnessed around the world.

Bibliography.

Aaronson, E., Blaney, N., Stephan, C., Sikes, J., & Snapp, M. (Eds.). (1978). *The Jigsaw Classroom*. Beverley Hills, CA: Sage Publications.

Abou-Jaoude, S., & Frasson, C. (1999). *Integrating a Believable Layer into Traditional ITS.* Paper presented at the AIED99, Le Man, France.

Aimeur, E., & Frasson, C. (1996). Analyzing A New Learning Strategy According to Different Knowledge Levels. *Computers in Education, 27*(2), 115-127.

Andre, E., & Rist, T. (2001). Controlling the Behaviour of Animated Presentation Agents in the Interface. *AI Magazine*(Winter), 53-66.

Andre, E., Rist, T., van Mulken, S., Klesen, M., & Baldes, S. (2000). The Automated Design of Believable Dialgogues for Animated Presentation Teams. In J. Cassell & J. Sullivan & S. Prevost & E. Churchill (Eds.), *Embodied Conversational Agents* (pp. 220-255). Cambridge, MA: MIT Press.

Andrews, R. (1995). *Teaching and Learning Argument*. London: Cassell.

Armbruster, B., Anders, T., & Osterag, J. (1987). Does Text Structure/Summarisation Instruction Facilitate Learning From Expository Text ? *Read Research Quarterly, 22*(3), 331-346.

Ausubel, D. (1985). Learning as Constructing Meaning. In N. Entwistle (Ed.), *New Directions in Educational Psychology 1: Learning and Teaching* (pp. 333). Lewes, E.Sussex and Philadelphia, P.A.: The Falmer Press.

Ball, G., Ling, D., Kurlander, D., Miller, J., Pugh, D., Skelly, T., Stankosky, A., Thiel, D., Van Dantzich, M., & Wax, T. (1997). Lifelike Computer Characters: The Persona Project at Microsoft Research. In J. M. Bradshaw (Ed.), *Software Agents* (pp. 480). Cambridge, Massachusetts. London, England.: AAAI Press/The MIT Press.

Barker, T. (1999a). *Community Based Virtual Learning : A WebCT Physics Course.* Paper presented at the WebCT99 : From Innovation to Implementation, Vancouver, Canada.

Barker, T. (1999b). *Integrating a Note-Taking Learning Companion Within the WebCT Virtual Learning Environment.* Paper presented at the Workshop on Animated and Pedagogical Agents, Le Mans, France.

Barker, T. (2002, 4-5th April). *How Are You ? Synthetic Personalities for Edutainment.* Paper presented at the Society for the study of Artificial Intelligence and Simulation of Behaviour (AISB02): Workshop on Animating Expressive Characters for Social Interactions, Imperial College, London, UK.

Barker, T. (2003, 15th July). *The Illusion of Life Revisited.* Paper presented at the Autonomous Agents and Multiagent Systems, Workshop 10: Embodied Conversational Agents as Individuals, Melbourne, Australia.

Barker, T., & Pilkington, R. (2000). *An Investigation and Design of Networked Learning in Inner-City Leeds.* Paper presented at the Networked Learning 2000: Innovative Approaches to Lifelong Learning and Higher Education Through the Internet, Lancaster.

Barker, T., & Pilkington, R. M. (2001, 21st.-24th. March). *Simulated Affectations of an Animated Pedagogical Agent.* Paper presented at the The Society for the study of Artificial Intelligence and Simulation of Behaviour (AISB01): Symposium on emotion, cognition and affective computing, University of York, U.K.

Barret-Lennard. (1962). Dimensions of therapist response as causal factors in therapeutic change. *Psychological Monographs, 76*(43).

Bates, J. (1994). The Role of Emotion in Believable Agents. *Communications of the ACM*(July).

Bates, J., Loyall, A. B., & Reilly, W. S. (1992, July 29-31). *An Architecture for Action, Emotion and Social Behaviour.* Paper presented at the European Workshop on Modelling Agents in a Multi-Agent World, MAAMAW'92, Italy.

BECTA. (2001). *Information Sheet on the National Grid for Learning* [pdf]. British Educational Communications and Technology Agency. Retrieved 13th June, 2002, from the World Wide Web: http://www.becta.org.uk/technology/infosheets/html/ngfl.html

Bell, P. (1997). *Using Argument Representations to Make Thinking Visible for Individuals and Groups.* Paper presented at the CSCL97 : The Second International Conference on Computer Support for Collaborative Learning, Toronto.

Berieter, C., & Scardamalia, M. (1987). *The Psychology of Written Composition.* Hillsdale, NJ: Lawrence Erlbaum and Associates.

Berners-Lee, T., & Fischetti, M. (2000). *Weaving the Web: the past, present and future of the World Wide Web by its Inventor*. London: Texre.

Bielaczyc, K., Pirolli, P., & Brown, A. L. (1995). Training in self-explanation and self-regulation strategies: Investigating the effects of knowledge acquisition activities on problem solving. *Cognition and Instruction, 13*(2), 221-252.

Bigge, M. L., & Shermins, S. S. (1999). *Learning Theories for Teachers*: Addison Wesley Longman.

Bostock, S. (2000). *A Review of Virtual Learning Environments* [Web Page]. Staff and Educational Development Association. Retrieved 22 September, 2003, from the World Wide Web: http://www.seda.ac.uk/ed_devs/vol1/Virtlearn.htm

Breuker, J. (Ed.). (1990). *EUROHELP: Developing Intelligent Help Systems.* Kopenhagen, Amsterdam, Manchester, Leeds.

Britton, B., & Glynn, S. (1985). Cognitive Demands of Processing Expository Text: A Cognitive Workbench Model. In B. Britton & J. Black (Eds.), *Understanding Expository Text* (pp. 227-247). New Jersey: Lawrence Erlbaum Associates.

Brna, P., Cooper, B., & Razmerita, L. (2001, July 2001). *Marching to the Wrong Distant Drum: Pedagogic Agents, Emotion and Student Modelling.* Paper presented at the Workshop on Attitude, Personality and Emotions in User-Adapted Interaction in conjunction with User Modelling 2001, Southofen, Germany.

Brooks, R. A. (1991). Intelligence Without Representation. *Artificial Intelligence*(47), 139-159.

Bruner, J. S. (1966). *Toward a Theory of Instruction.* New York: Norton.

Bruner, J. S. (1975). Beyond the Information Given. In N. Entwistle & D. Hounsell (Eds.), *How Students Learn* (pp. 105-116). Lancaster: Institute for Research and Development in Post-Compulsory Education.

Burnett, R. E. (1993). Decision-Making during the collaborative planning of coauthors. In A. Penrose & B. Sitko (Eds.), *Hearing Ourselves Think: Cognitive Research in the College Writing Classrooom* (pp. 125-146). Oxford: Oxford University Press.

Burton, M., Pilkington, R., & Brna, P. (2000). Modelling Learning to Collaborate through Problem-Solving: An Agent-Based Approach. University of Leeds.

Carrell, P. L. (1985). Facilitating ESL Reading by Teaching Text Structure. *TESOL Quarterly, 19*(4), 727-752.

Cassell, J. (2000). Nudge Nudge Wink Wink: Elements of Face-to-Face Conversation for Embodied Conversational Agents. In J. Cassell & J. Sullivan & S. Prevost & E. Churchill (Eds.), *Embodied Conversational Agents* (pp. 1-27). Cambridge, MA: MIT Press.

Cassell, J. (2001). Embodied Conversational Agents: representation and intelligence in user interfaces. *AI Magazine*(Winter), 67-83.

Cassell, J., Bickmore, H., Vilhjalmsson, H., & Yan, H. (2001). More Than Just a Pretty Face: Affordances of Embodiment. *Knowledge Based Systems, 14*, 55-64.

Cassell, J., Bickmore, T., Campbell, L., Vilhjalmsson, H., & Yan, H. (2000). Human Conversation as a System Framework: Designing Embodied Conversational Agents. In J. S. Cassell, J. Prevost, S. Churchill, E. (Ed.), *Embodied Conversational Agents* (pp. 29-63). Cambridge, MA: MIT Press.

Chan, T.-W. (1996). Learning Companion Systems, Social Learning Systems , and the Global Social Learning Club. *Artificial Intelligence in Education*, 7(2), 125-159.

Chan, T.-W., & Baskin, A. B. (1988, June). *"Studying With the Prince" : The Computer as a Learning Companion*. Paper presented at the Intelligent Tutoring Systems 88, Montreal, Canada.

Chan, T.-W., & Baskin, A. B. (1990). Learning Companion Systems. In C. Frasson & G. Gauthier (Eds.), *Intelligent Tutoring Systems* (pp. 6-33). New Jersey: Ablex.

CHEST. (2002). *VLE Comparison Grid*. Retrieved 9th October, 2003, from the World Wide Web: http://www.chest.ac.uk/datasets/vle/

Chi, M., & Bassock, M. (1989). Learning from examples. In L. B. Resnick (Ed.), *Knowing Learning and Instruction: Essays in Honour of Robert Glaser*. Hillsdale, NJ.: Lawrence Erlbaum.

Chi, M. T. H., Lewis, M. W., Reimann, P., & Glasser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145-182.

Chou, C.-Y., Chan, T.-W., & Lin, C.-J. (2003). Redefining the learning companion: the past, present and future of educational agents. *Computers and Education, 40*, 255-269.

Chou, C.-Y., C-J., L., & Chan, T.-W. (1999). *User Modelling in Simulating Learning Companions*. Paper presented at the Artificial Intelligence in Education.

Cohen, L., & Manion, L. (1994). *Research Methods in Education* (4th. ed.). London: Routledge.

Cole, M., & Wertsch, J. V. (1996). Beyond the Individual-Social Antimony in Discussions of Piaget and Vygotsky. *Human Development, 34*(5), 250-256.

Collins, A. (1977). Processes in Acquiring Knowledge, *Schooling and the Acquisition of Knowledge* (pp. 339-363). Hillsdale, NJ.: Lawrence Erlbaum Associates.

Columbetti, G. (2001, 21st-24th March). *The somato-cognitive system.* Paper presented at the AISB'01: Symposium on Emotion, Cognition and Affective Computing, University of York, U.K.

Conklin, J., & Burgess Yakemovic, K. C. (1991). A process-oriented approach to design rationale. *Human-Computer Interaction, 6*(3 & 4), 357-391.

Cooper, B., & Brna, P. (2002). Designing for Interaction: creating and evaluating an empathic ambience in computer integrated learning environments. In K. Dautenhahn (Ed.), *Socially Intelligent Agents: creating relationships with computers and robots*: Kluwer.

Crook, C. (1994). *Computers and the Collaborative Experience of Learning.*

Crook, C. (1998). Children As Computer Users: The Case of Collaborative Learning. *Computers in Education, 30*(3/4), 237-247.

Damasio, A. R. (1994). *Descartes' Error.* New York: G.P. Punam's Sons.

Denzin, N. K. (1970). *The Research Act in Sociology: A Theoretical Introduction to Sociological Methods.* London: The Butterworth Group.

DfEE. (1999). *Key Skills*: Department for Education and Employment.

Dillenbourg, P. (1999a). *Collaborative Learning: cognitive and computational approaches.* Amsterdam: Pergamon.

Dillenbourg, P. (1999b). Introduction: What do you mean by "Collaborative Learning" ? In P. Dillenbourg (Ed.), *Collaborative Learning: cognitive and computational approaches* (pp. 1-19). Oxford: Elsevier Science.

Dillenbourg, P. (2000). *Virtual Learning Environments.* Paper presented at the EUN Conference 2000: Learning in the new millenium: building new education strategies for schools.

Dillenbourg, P., & Self, J. A. (1992, June 10-12). *People Power : A Human-Computer Collaborative Learning System.* Paper presented at the International Conference on Intelligent Tutoring Systems, Montreal, Canada.

Dix, A., Finlay, J., Abowd, G., & Beale, R. (1997). *Human-Computer Interaction* (Second ed.): Pearson Education.

Dowd, C. A., & Sinatra, R. (1990). Computer Programs and the Learning of Text Structure. *Journal of Reading, 34*(2), 104-112.

du Boulay, B., Luckin, R., & del Soldata, T. (1999). *The Plausability Problem: Human Teaching Tactics in the 'Hands' of a Machine*. Paper presented at the International Conference on Artificial Intelligence in Education, Le Mans, France.

Elliot, C. (1992). *The Affective Reasoner: a process model of emotions in a multi-agent system*. Unpublished PhD.

Elliot, C., & Brzezinski, J. (1998). Autonomous Agents as Synthetic Characters. *AI Magazine, 19*(2), 13-30.

Elliot, C., Rickel, J., & Lester, J. (1999). Lifelike pedagogical agents and affective computing: An exploratory synthesis. In M. Woolridge & M. Eleso (Eds.), *Artificial Intelligence Today* (Vol. 1600, pp. 195-212): Springer-Verlag.

Gagne, R. M. (1975). Domains of Learning. In N. J. Entwistle & D. J. Hunsell (Eds.), *How Students Learn* (pp. 15-25). Lancaster: Institute for Research and Development in Post-Compulsory Education, University of Lancaster.

Garner, R., & McCaleb, J. L. (1985). Effects of Text Manipulations on Quality of Written Summaries. *Contemporary Educational Psychology, 10*, 139-149.

Goodman, B., Soller, A., Linton, F., & Gaimari, R. (1998). Encouraging Student Reflection and Articulation using a Learning Companion. *International Journal of Artificial Intelligence in Education, 9*, 237-255.

Gore, K. (1997). Cogito Auto Sum. Slate: MSN.

Graesser, A. C., & Person, N. K. (1994). Question Asking During Tutoring. *American Educational Research Journal, 31*(1), 104-137.

Hahn, U., & Mani, I. (2000). The Challenges of Automatic Summarization. *Computer*(November), 29-36.

Hare, V. C., & Borchardt, K. M. (1984). Direct Instruction of Summarisation Skills. *Research Quarterly, 20*(1), 62-78.

Hartley, J. R., Byard, M., & Mallen, C. (1991). *Qualitative Modelling and Conceptual Change in Science Students*. Paper presented at the The International Conference on the Learning Sciences, Charlottesville, Virginia.

Hawkridge, D. (2000). Collaborative Learning: Cognitive and Computational Approaches. *Computers and Education, 35*, 81-94.

Hayes-Roth, B., van Gent, R., & Huber, D. (1997). Acting in Character. In R. Trappl & P. Petta (Eds.), *Creating Personalities for Synthetic Actors* (pp. 92-112). Berlin: Springer-Verlag.

Hietala, P., & Niemirepo, T. (1998). The Competence of Learning Companions. *Artificial Intelligence in Education, 9*.

Hudlicka, E., & Fellous, J.-M. (1996). Review of Computational Models of Emotion: Psychometrix.

Hume, G., Michael, J., Rovick, A., & Evens, M. (1996). Hinting as a Tactic in One-on-One Tutoring. *The Journal of the Learning Citizens, 5*(1), 23-47.

Isbister, K., & Nass, C. (2000). Consistency of personality in interactive characters: verbal cues, non-verbal cues, and user characteristics. *International Journal of Human-Computer Studies*(53), 251-267.

John, O. P. (1990). The "Big Five" Factor Taxonomy: dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of Personality Theory and Research* (pp. 66-100). New York: NY: Guilford Press.

Johns, A. M. (1985). Summary Protocols of "Underprepared" and "Adept" Univeristy Students: Replications and Distortions of the Original. *Language Learning, 35*, 495-512.

Johns, A. M. (1988). Reading for Summarising: An Approach to Text Orientation and Processing. *Reading in a Foreign Language, 4*(2), 79-90.

Johnson, K. (1981). *Communicate in Writing : A functional approach to writing through reading comprehension*. London: Longman.

Johnson, W. L. (2001). Pedagogical Agent Research at CARTE. *AI Magazine*(Winter), 85-94.

Johnson, W. L. (2003). Performance. In T. Barker (Ed.). Stafford: Unpublished.

Johnson, W. L., Kole, S., Shaw, E., & Pain, H. (2003). *Socially Intelligent Learner-Agent Interaction in Tactics*. Paper presented at the International Conference of Artificial Intelligence in Education, Sydney, Australia.

Johnson, W. L., & Rickel, J. W. (2000). Animated Pedagogical Agents: Face-to-Face Interaction in Interactive Learning Environments. *International Journal of Artificial Intelligence in Education, 11*, 47-78.

Johnson-Laird, P. N. (1983). *Mental Models: Toward a Cognitive Science of Language, Inference and Consciousness*. Cambridge: Cambridge University Press.

Johnson-Laird, P. N. (1988). *The Computer and the Mind : An introduction to Cognitive Science* (2nd. ed.). London: Fontana Press.

Jordan, R. R. (1997). *English for Academic Puposes*. Cambridge: Cambridge University Press.

Kanselaar, G., Erkens, G., Jaspers, J., & Schijf, H. (2001). Essay Review: Computer supported collaborative learning. *Teaching and Education, 17*, 123-129.

Kapoor, A., Mota, S., & Picard, R. W. (2001, 2-4 November). *Towards a Learning Companion that Recognizes Affect.* Paper presented at the Emotional and Intelligent II: The Tangled Knot of Social Cognition, North Falmouth, MA.

Katz, S. (1997). *Peer and student-mentor interaction in a computer-based training environment for electronic fault diagnosis* (Technical Report). Pittsgurgh, PA: Learning Research and Development Center.

Kemp, R. H. (1997). *Using the Wizard of Oz Technique to Prototype a Scenario-Based Simulation Tutor.* Paper presented at the International Conference on Artificial Intelligence in Education, Kobe, Japan.

Kiewra, K. A., Benton, S. L., Kin-Sung, I., Risch, N., & Christenson, M. (1995). Effects of Note-Taking Format and Study Technique on Recall and Relational Performance. *Contemporary Educational Psychology, 20*(2), 172-187.

Kintsch, W., & Van Dijk, T. A. (1978). Toward a Model of Text Comprehension and Deduction. *Psychological Review, 85*(5), 363-394.

Kolb, D. A. (1984). *Experiential Learning: experience as the source of learning and development*: Prentice Hall.

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. Chicago: Chicago U.P.

Kupiec, J., Pederson, J., & Chen, F. (1999). A Trainable Document Summarizer. In M. Inderjeet & M. T. Maybury (Eds.), *Advances in Automatic Text Summarization.* Cambridge, MA: MIT Press.

Langton, C. G. (Ed.). (1997). *Artificial Life: an overview* (Paperback ed.). Cambridge, MA: MIT Press.

Le Doux, J. (1996). *The Emotional Brain*. New York: Simon and Schuster.

Lester, J. C., Converse, S. A., Stone, B. A., Kahler, S. E., & Barlow, S. T. (1997). *Animated Pedagogical Agents and Problem-Solving Effectiveness: A Large-Scale Empirical Evaluation.* Paper presented at AIED 97, Kobe, Japan.

Lester, J. C., Stone, B. A., & Stelling, G. D. (1999). Lifelike Pedagogical Agents for Mixed-Initiative Problem Solving in Constructivist Learning Environments, *User Modelling and User-Adapted Interaction*.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Inquiry*: Sage.

Luhn, H. P. (1999). The Automatic Creation of Literature Abstracts. In I. Mani & M. T. Maybury (Eds.), *Advances in Automatic Text Summarization* (pp. 15-21). Boston: Kluwer Academic.

Maes, P. (1997). Agents that reduce work and information overload. In J. M. Bradshaw (Ed.), *Software Agents* (pp. 145-164). Menlo Park, California: AAAI Press/MIT Press.

Mann, W., & Thompson, S. (1988). Rhetorical Structure Theory: Toward a Functional Theory of Text Organization. *Text, 8*(3), 243-281.

Mann, W. C., Matthiesesson, C., & Thompson, S. (1989). *Rhetorical Strucuture Theory and Text Analysis* (ISI/RR-89-242): Information Sciences Institute, CA.

Mannes, S., & St.George, M. (1996). Effects of prior knowledge on text comprehension: a simple modelling approach. In B. Britton & A. C. Graesser (Eds.), *Models of Understanding Text*. New Jersey: Lawrence Erlbaum Associates.

Marcu, D. (1999). Discourse Trees Are Good Indicators of Importance in Text. In I. Mani & T. Maybury (Eds.), *Advances in Automatic Text Summarization* (pp. 123-136). Boston: Kluwer Academic Publishers.

Marcu, D. (2000). *The Theory and Practice of Discourse Parsing and Summarization*. Cambridge, MA: MIT Press.

Maulsby, D., Greenberg, S., & Mander, R. (1993). *Prototyping an Intelligent Agent through Wizard of Oz*. Paper presented at INTERCHI93.

McCrae, R. R., & Costa, P. T. J. (1989). The structure of interpersonal traits: Wiggins's circumplex and the five-factor model. *Journal of Personality and Social Psychology, 56*, 586-595.

McCrae, R. R., & John, O. P. (1992). An introduction to the five factor model and its applications. *Journal of Personality, 60*, 175-215.

McKeown, K. R. (1985). *Text Generation*. Cambridge: Cambridge University Press.

Mercer, N., & Wegerif, R. (1999). Children's Talk and the Development of Reasoning in the Classroom. *British Educational Research Journal, 25*(1), 95-111.

Meyer, B. J. (1975). *The Organisation of Prose and its Effects on Memory*. Amsterdam: North Holland.

Meyer, B. J. F. (1984). Text Dimensions and Cognitive Processing. In M. Stein & T. Trabasso (Eds.), *Learning and Comprehension of Text*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Microsoft. (2003). *MS Agent Homepage* [web page]. Retrieved 26th September, 2003, from the World Wide Web: http://www.microsoft.com/msagent/default.asp

Minsky, M. (1987). *The Society of Mind*. Boston, MA: MIT Press.

Mohammed, Z. B. (1996). *An Evaluation of an Extra-School Science Programme in a Multi-Ethnic Community: A Computer Based Case Study*. University of Leeds, Leeds.

Moore, D. (1993). *Dialogue Game Theory for Intelligent Tutoring Systems*. Unpublished PhD, Leeds Metropolitan University, Leeds.

Myers, I. B., & McCaulley, M. H. (1987). *Manual: a guide to the development and use of the Myers-Briggs Type Indicator*. Palo Alto, CA: Consulting Psychologists Press.

Nass, C., Isbister, K., & Lee, E.-J. (2000). Truth is Beauty: research in embodied conversational agents. In J. Cassell & J. Sullivan & S. Prevost & E. Churchill (Eds.), *Embodied Conversational Agents* (pp. 374-402). Cambridge, MA: MIT Press.

Newell, A., & Simon, H. A. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice Hall.

Newell, A., & Simon, H. A. (1976). Computer Science as Empirical Enquiry: Symbols and Search. *Communications of the ACM, 19*(3), 113-126.

OFSTED. (1999). *Raising the attainment of minority ethnic pupils : schoool and LEA responses*. London: Office of Her Majesty's Chief Inspector of Schools.

Ortony, A., Clore, G. L., & Collins, A. (1990). *The Cognitive Structure of Emotions* (paperback ed.). Cambridge: University of Cambridge.

Paiva, A., Machado, I., & Maartinho, C. (1999, July, 1999). *Enriching Pedagogical Agents with Emotional Behaviour: The Case of Vincent*. Paper presented at the AIED-99 Workshop on Animated and Personified Pedagogical Agents, Le Mans, France.

Piaget, J. (1970). *Science of Education and the Psychology of the Child*. New York: Orion Press.

Picard, R. W. (1997). *Affective Computing*: MIT Press.

Pilkington, R. (1992). *Intelligent Help: Communicating with Knowledge Based Systems*. London: Paul Chapman.

Pilkington, R., & Kuminec, P. A. (2002). Using a role-play activity with synchronous CMC to encourage critical reflection on peer debate. In M. Monteith (Ed.), *ICT for Curriculum Enhancement* (pp. 83-99). Bristol: Intellect.

Pilkington, R., & Parker-Jones, C. (1996). Interacting with Computer-Based Simulation: The Role of Dialogue. *Computers in Education, 27*(1), 1-14.

Pilkington, R. M. (1988). *Knowledge-Based Systems in Topic Learning*. University of Leeds, Leeds.

Pilkington, R. M. (1999). *Analysing Educational Discourse: The DISCOUNT Scheme* (Technical Report 99/2). Leeds: Computer Based Learning Unit, University of Leeds.

Pilkington, R. M., & Walker, S. A. (in press). Facilitating debate in networked learning: reflecting on online synchronous discussion in higher education. *Instructional Science, 31*(1-2), 41-63.

Ploetzner, R., & Kneser, C. (1998). *Steps towards the acquisition of expertise: shifting the focus from quantitative to qualitative problem representation during collaborative problem solving*. Paper presented at the 20th. Annual Meeting of the Cognitive Science Society.

Preece, J., Rogers, Y., & Sharp, H. (2002). *Interaction Design: beyond human-computer interaction*. New York: John WIley and Sons.

Prendinger, H. (2003). *Persona Effect Revisited: using bio-signals to measure and reflect the impact of character-based interfaces*. Paper presented at the Fourth International Conference on Intelligent Virtual Agents (IVA-03), Kloster Isree, Germany.

QCA. (2002). *The Key Skills Qualifications, Specifications and Guidance: Communication, Application of Number and Information Technology* (QCA/02/896): Qualifications and Curriculum Authority.

Ravenscroft, A. (1997). *Learning as Knowledge Refinement : A computer based approach*. University of Leeds, Leeds.

Ravenscroft, A., & Hartley, R. (1998). *Evaluation of Chapeltown and Harehills Assisted Learning Computer School (CHALCS)*. Leeds: University of Leeds.

Reeves, B., & Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places* (Paperback ed.). Cambridge, UK: Cambridge University Press.

Reilly, W. S. N. (1996). *Believable Social and Emotional Agents.* Unpublished Ph.D. Thesis, Carnegie Mellon University, Pittsburgh, P.A.

Rickel, J., & Johnson, W. L. (1999). Animated Agents for Procedural Training in Virtual Reality: Perception, Cognition and Motor Control. *Applied Artificial Intelligence*(13), 343-382.

Rist, T., & Schmitt, M. (2002). *Avatar Arena: An Attempt to Apply Socio-Psychological Concepts of Cognitive Consistency in Avatar-Avatar Negotiation Scenarios.* Paper presented at the Artificial Intelligence and Simulated Behaviour Convention Workshop on Animating Expressive Characters for Social Interactions, Imperial College, London.

Robertson, J., Good, J., & Pain, H. (1998). BetterBlether: The Design and Evaluation of a discussion tool for education. *International Journal of Artificial Intelligence in Education, 9*, 219-236.

Rogers, C. R. (1975). Empathic: An Unappreciated Way of Being. *The Counselling Psychologist, 5*(2), 2-9.

Roschelle, J. (1992). Learning by collaborating: convergent conceptual change. *Journal of the Learning Sciences*(2), 235-276.

Rosenshine, B., & Meister, C. (1994). Reciprocal Teaching : A Review of the Research. *Review of Educational Research, 64*(4), 479-530.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel Distributed Processing.* Cambridge, MA: MIT Press.

Russell, S., & Norvig, P. (2003). *Artificial Intelligence: a modern approach* (Second Edition ed.). New Jersey: Pearson Education, Inc.

Salmon, G. (2000). *E-Moderating*. London: Kogan Page Ltd.

Schwartz, D. L. (1995). The Emergence of Abstract Representations in Dyad Problem Solving. *The Journal of the Learning Sciences, 4*(3), 321-354.

Shaw, E., Ganesham, R., Johnson, W. L., & Millar. (1999). *Building a Case for Agent-Assisted Learning as a Catalyst for Curriculum Reform in Medical Education.* Paper presented at the AIED99, Le Mans, France.

Sheremetov, L., & Arenas, A. G. (2002). EVA: an interactive Web-based collaborative learning environment. *Computers and Education, 39*(2), 161-182.

Simon, H. (1979). Motivational and Emotional Controls of Cognition. In H. A. Simon (Ed.), *Models of Thought*. New Haven & London: Yale University Press.

Skinner, B. F. (1953). *Science and Human Behaviour*. New York: Macmillan.

Sloman, A. (1999). *How Many Separately Evolved Emotional Beasties Live Within Us ?* Paper presented at the Workshop on Humans and Artifacts, Vienna.

Slotte, V., & Lonka, K. (1999). Review and Process of Spontaneous Note-Taking on Text Comprehension. *Contemporary Educational Psychology, 24*(1), 1-20.

Soanes, C. (Ed.). (2001). *The New Pocket Oxford Dictionary*. Oxford: Oxford University Press.

Solomon, R. C. (1993). The Philosophy of Emotions. In M. Lewis & L.M. Haviland (Eds.), *Handbook of Emotions*. New York & London: Guilford.

Stevens, A., Collins, A., & Goldin, S. E. (1982). Misconceptions in Students Understanding. In D. Sleeman & J. S. Brown (Eds.), *Intelligent Tutoring Systems*: Academic Press.

Strauss, A., & Corbin, J. (1998). Grounded Theory Methodology. In N. K. Denzin & Y. S. Lincoln (Eds.), *Strategies of Qualitative Enquiry* (Vol. 2, pp. 158-183). London: SAGE Publications.

Suthers, D., & Jones, D. (1997). An Architecture for Intelligent Collaborative Educational Systems. *Artificial Intelligence in Education*, 55-62.

Tawalbeh, M. (1994). *Computer Based Methods for Improving Summarisation Skills in ESL Contexts*. Unpublished PhD., University of Leeds, Leeds.

Thomas, F., & Johnston, O. (1981). *Disney Animation: The Illusion of Life*. New York: Abbeville Press.

Trafton, J. G., & Trickett, S. B. (2001). Note-Taking for Self-Explanation and Problem Solving. *Human-Computer Interaction, 16*, 1-38.

Tuckman, B. W. (1965). Stages of small group development revisited. *Group and Organizational Studies, 2*, 419-427.

Uresti, J. A. R. (2000, June, 2000). *Should I Teach My Computer Peer ? Some Issues in Teaching Learning Companion*. Paper presented at the International Conference on Intelligent Tutoring Systems, Montreal, Canada.

Vygotsky, L. (1986). *Thought and Language* (A. Kozulin, Trans. Paperback ed.). Cambridge, MA: The MIT Press.

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, Massachusetts: Harvard University Press.

Walker, S. A. (2003). *The contribution of computer-mediated communication in developing argument skills and writing-related self-esteem.*, The University of Leeds, Leeds.

Walker, S. A., & Pilkington, R. M. (2001). *Facilitating computer-mediated discussion classes: a study of teacher intervention strategies.* Paper presented at the International Conference on Adavanced Learning Technologies 2001, Madison, U.S.A.

Wan, D., & Johnson, P. (1994, October 22-26). *Computer Supported Collaborative Learning Using CLARE : the approach and Experimental Findings.* Paper presented at the ACM conference on Computer Supported Cooperative Work, Chapel Hill, N. Carolina, USA.

Wang, W.-C., & Chan, T.-W. (2000). CAROL5: An Agent-Oriented Programming Language for Developing Social Learning Systems. *International Jourrnal of Artificial Intelligence in Education, 11.*

Wason, P. C., & Johnson-Laird, P. N. (1973). *Psychology of reasoning: Structure and content.* Cambridge, Mass.: Harvard University Press.

WebCT. (2003). *WebCT Standards* [Web page]. WebCT. Retrieved 19th August, 2003, from the World Wide Web: http://www.webct.com/standards

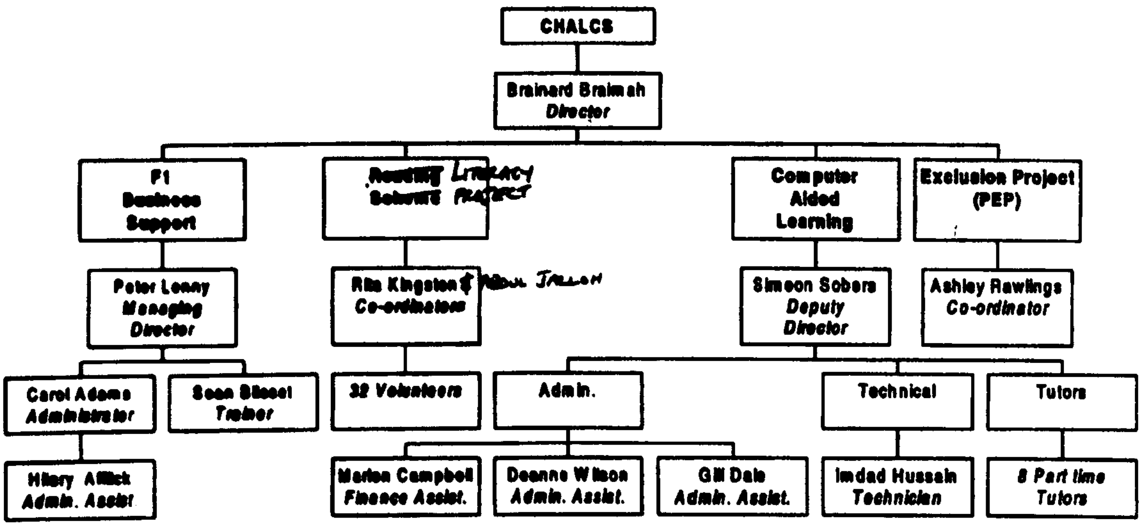
White, B. Y., & Shimoda, T. A. (1999). Enabling students to construct theories of collaborative inquiry and reflective learning: computer support for metacognitive development. *International Journal of Artificial Intelligence in Education, 10,* 151-182.

Wiggins, J. S. (1979). A Psychological Taxonomy of Trait-Descriptive Terms: The Interpersonal Domain. *Journal of Personality and Social Psychology, 37*(3), 395-412.

Wiggins, J. S. (1982). Circumplex models of interpersonal behaviour in clinical psychology. In P. C. Kendall & J. N. Butcher (Eds.), *Handbook of research methods in clinical psychology* (pp. 183-221). New York: Wiley.

Wooffitt, R., Fraser, N. M., Gilbert, N., & McGlashan, S. (1997). *Humans, Computers and Wizards : analysing human (simulated) computer interaction.* London & New York: Routledge.

Appendix A: CHALCS Organisational Structure.



Appendix B: Leeds Learning Network (www.leedslearning.net).

14. The Leeds Learning Network would be a private linking of school and authority computers. It would be connected to the Internet but in such a way as to protect all of the computers and users within it. The risks and threats of Internet use, (offensive content and breaches of network security), would be managed at one site instead of having to be controlled at over three hundred. The creation of this Virtual Private Network (VPN) and Intranet offers Leeds many strategic opportunities for the development of services to its schools. There is also the potential to provide additional functions and to serve other users within the wider community. The levels of physical security can be high – sufficient to satisfy the managers of Leeds corporate systems that certain authorised users of the VPN could have access to systems within the authority. The figure below schematically represents the proposed network.

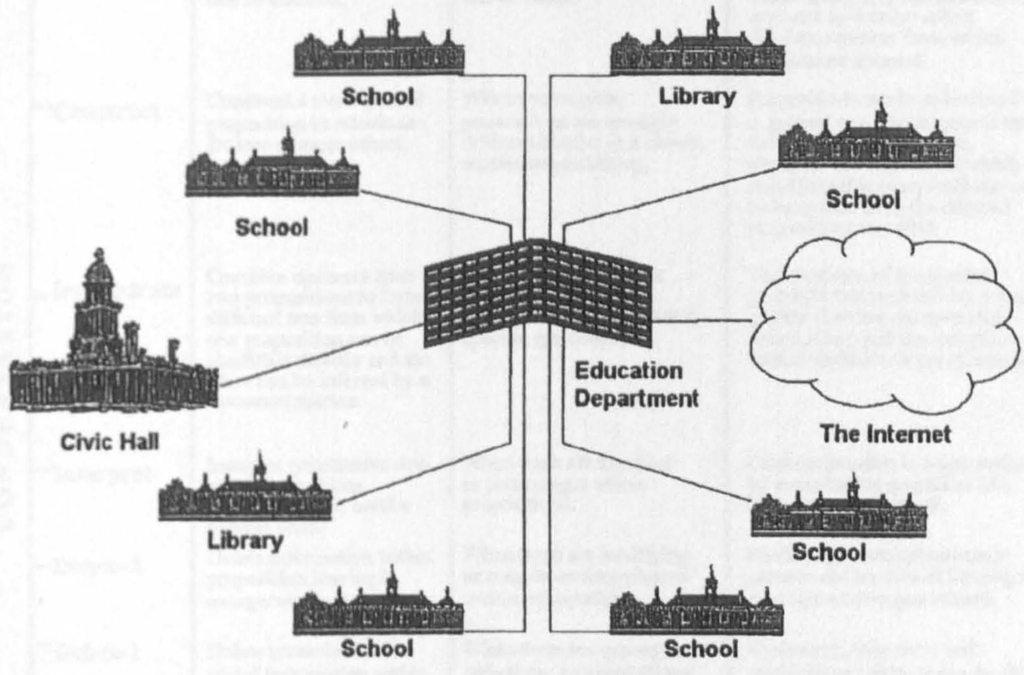


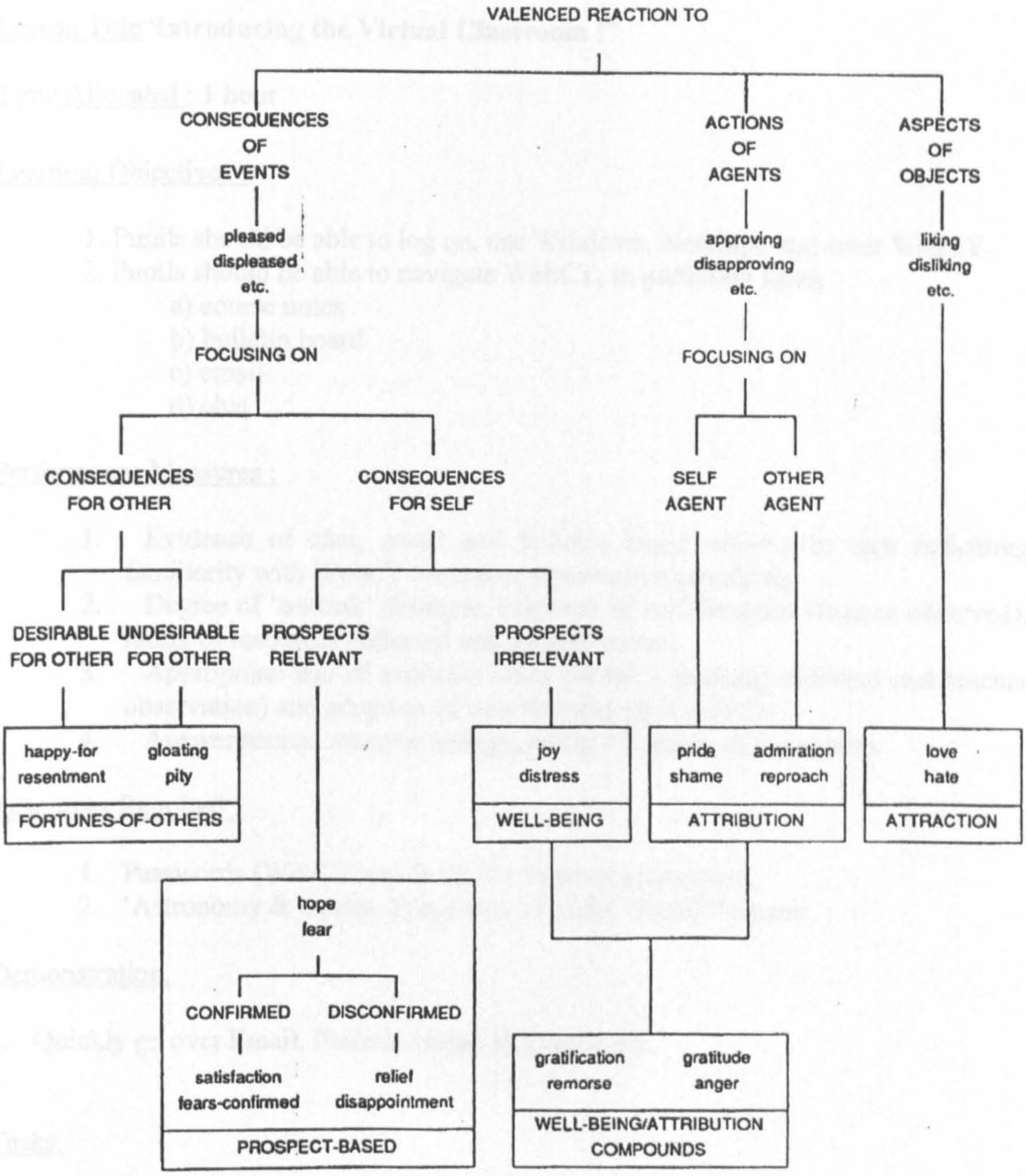
Figure 2 Schematic representation of the proposed VPN or Leeds Learning Network

15. Private content available on a Leeds Intranet can offer the wide learning community of the city a very large and valuable collection of resources to aid its work. Moreover, the concept promotes greater openness and employees at every level can have access to better information to help them in their responsibilities. Materials are not confined to the ownership of the person who received the printed copy or become inaccessible because the document is lost or out of print. A Leeds Intranet might ultimately become the key point of access to most of the Authority's knowledge, information giving and policy initiatives thus paralleling the approach being taken by government departments. The use of ICT can promote greater individual efficiency and effectiveness. It also offers a route towards more open management and individual involvement.

Appendix C: Tawalbeh's Summarisation Model (Tawalbeh, 1994).

PROCESS	STRATEGY	DESCRIPTION	WHEN IS THE STRATEGY APPLIED?	WHAT CAN BE DONE WHEN IT IS APPLIED?
SELECTION	Select-1	Select propositions which are relatively higher in the hierarchical structure.	When propositions contain a strong signal, giving them supremacy over others which they govern.	All propositions which pass this criterion at the highest level are selected.
	Select-2	Select propositions which are important to the rhetorical structure of text.	When a certain structure is inferred or signalled by an appropriate discourse marker.	All propositions with markers passing the criterion for that particular text type are selected.
CONDENSATION	Replace	Replace a proposition with a contrastive lexical item or discourse marker by which one or more propositions can be inferred.	When either there are (i) contrastive propositions or (ii) cause-effect situations on which the inference can be based.	(i) Contrastive propositions are replaced by a contrastive discourse marker or by negating a verb from which they can be inferred (ii) cause-effect propositions can be replaced by a cause-effect discourse marker from which they can be inferred.
	Construct	Construct a more general proposition to substitute for two or more others.	When two or more propositions are strongly different/similar in a certain attribute/aspect/thing.	Propositions can be substituted by a general one which ignores the details and only gives the attribute. The key words which stand for difference/similarity can be borrowed from the original propositions or added.
	Incorporate	Combine elements from two propositions to form a different one from which one proposition can be identified directly and the other can be inferred by a discourse marker.	When two propositions exist; one is general (collective) and the other is specific (particular).	The predicate of the general proposition is replaced by a discourse marker (leaving out redundant information) and the specific is either copied verbatim or paraphrased.
	Interpret	Interpret quantitative data within propsitions qualitatively (to build a general one).	When there are numbers or percentages within propositions.	Quantitative data is being replaced by a qualitative quantifier like high, large, low, small.
	Delete-2	Delete information within proposition leaving it overgeneralised.	When there are modifying or complementary phrases within a proposition.	Modifying or complementary phrases can be deleted leaving the proposition overgeneralised.
	Delete-1	Delete redundant and trivial information within a proposition.	When there are synonyms, adjectives, or prepositional phrases within a proposition.	Synonyms, adjectives and prepositional phrases can be deleted without affecting the sense of the proposition.
	Paraphrase	Reproduce a proposition using other words.	When it is not possible to apply any of the above strategies.	The proposition can be reproduced using other words. Paraphrasing might help to reduce the number of words.
	Copy	Copy a proposition verbatim.	When it is not possible to apply any strategy (last resort).	The whole proposition can be reproduced by copying it verbatim.
ORGANISATION	Tie And Link	If propositions have not been joined together as a result of the incorporation strategy, then tie and link them using cohesive ties and discourse markers appropriately.	When propositions share the same theme and grammatical structure.	Propositions (and discourse markers used at the condensation stage) can be coordinated and/or subordinated to vary the structure and length of sentences. Appropriate cohesive ties and discourse markers have to be used to do this.
	Vary Opener	Use different sentence openers such as conjunction, modifying or complementary phrase to vary the conventional subject-verb sequence.	When sentences have been linked and tied already.	If possible, discourse markers or any other words apart from nouns or pronouns are moved to be at the beginning of sentences taking into consideration the mechanical and grammatical constraints.

Appendix D: OCC Emotion Model (Ortony, Clore, Collins, 1990).



Appendix E: CHALCS Lesson Plan.

<u>Lesson Title</u>**"Introducing the Virtual Classroom !"**

<u>Time Allocated :</u> 1 hour

<u>Learning Objectives :</u>

1. Pupils should be able to log on, use Windows, Netscape and enter WebCT.
2. Pupils should be able to navigate WebCT, in particular using :
   a) course notes
   b) bulletin board
   c) email
   d) chat

<u>Performance Measures :</u>

1. Evidence of chat, email and bulletin board activity in logs indicating familiarity with WebCT tools plus observation (teachers).
2. Degree of 'on-task' dialogue, evidence of collaboration (logs or observed), rating of resources collected and answer scores.
3. Appropriate use of available tools (WebCT tracking statistics and teacher observation) and adoption of new learning style (see 2).
4. Answer scores, resource ratings, rating of quality of discussion.

<u>Resources Required :</u>

1. Passwords (WebCT and WinNT), Internet connection.
2. 'Astronomy & Optics- Properties of Light' WebCT course.

<u>Demonstration.</u>

1. Quickly go over Email, Bulletin Board & Chat again.

<u>Tasks.</u>

1. Use the **email** to take a term from this topic of optics and light and prepare a brief definition of the term for the glossary. Send it by private mail to 'Instructor (Physics)'.
2. Spend some time (not too much!) on the encyclopaedia looking under some of the concepts of this topic on light and optics e.g. refraction, speed of light or chromatic aberration. Use the **bulletin board** <u>as a group</u> to <u>share</u> and make any brief additional notes but comment on the probable value of these sources and the best ways of using them in your learning.
3. Use the **chat** to discuss answers to the question : "If we can only see visible light, how are we able to see and use images from infra-red and X-ray photographs?"

Appendix F: CHALCS Summarisation Study Example Summary (S1).

LENSES[PC1]… Introduction:

Lenses are the basic components of **optical instruments** such as microscopes and telescopes. These instruments give information about objects from the very small and near (e.g. molecules and bacteria) to the very large and distant (e.g. stars and galaxies). Also instruments such as our eyes and spectacles, photocopiers and cameras use lenses and these are all part and parcel of our everyday life. [PC2]

What is a lens?

Light waves can be considered to originate from each point of an object and lenses **focus** the light waves from each object point to form an image point. But the shape of the lens, the material of which it is made and its placing with respect to the object control the size and type of the image. Remember what happens when you look the "wrong way" through your binoculars.
The eye is a lens system and, unless we have a sight defect, we see each object clearly. Each object point focuses to an image point on the retina of the eye and all these image points are integrated and interpreted as a complete image by our brains. [PC3]

How do lenses control light?

Lenses have different **properties** depending on the **materials** of which they are made (e.g. various types of glass with different refractive indexes and the **shape** of their surfaces. [PC4]
Remember that the path of the light waves obeys the laws of refraction, so *the greater the angle of incidence the greater the angle of refraction*, as summarised in Snell's law.

In a convex lens light waves come from a point on the object (shown as rays giving the direction of the waves). Note that the outer rays have a larger angle of incidence and hence a greater angle of refraction, so this curvature of the lens brings rays from each object point to an image point.

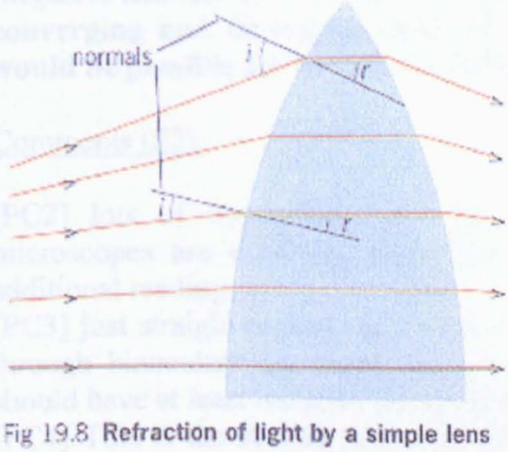
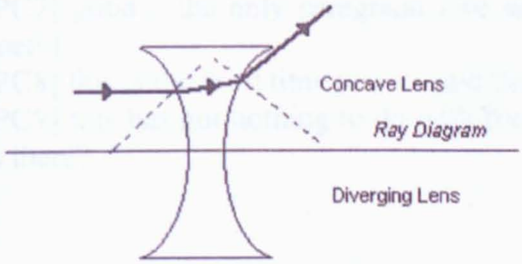


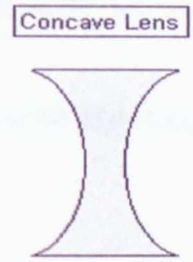
Fig 19.8 **Refraction of light by a simple lens**

[PC5]

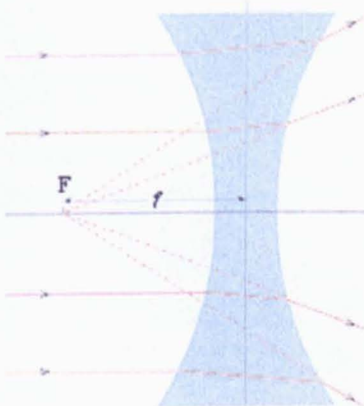
As a converging lens is a positive lens the image will be focused at the principle focus to the right of the lens. It would form a real image.

The surface of the concave lens are parts of a sphere. A concave lens produces a virtual image, it diverges parallel rays entering it .[PC6]





At the first face the light enters a denser medium and will bend towards the normal. It will diverge. And at the second face it will diverge more because the light ray enters a less dense medium (air) and is refracted away from the normal as the light waves pick-up speed. So the image will appear to come from behind the lens i.e. the same side as the object (project the final ray back to where it crosses the axis) and so is a **virtual** image.[PC7]



F is a virtual principal focus

Fig 19.11 A diverging lens has a virtual principal focus and a negative focal length

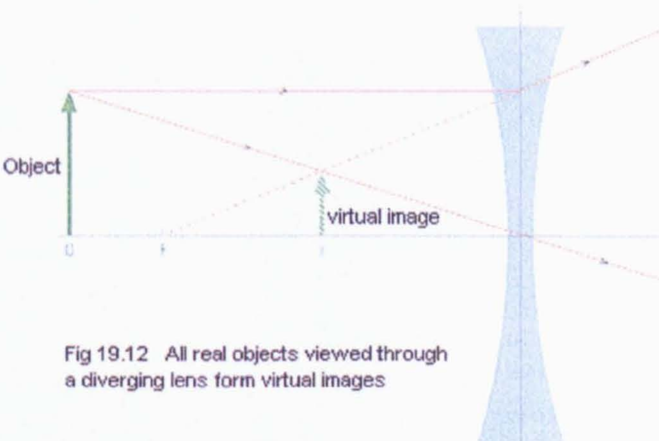


Fig 19.12 All real objects viewed through a diverging lens form virtual images

We classify lenses into converging lenses (*positive lenses*) and diverging lenses (*negative lenses*). By varying the curvature of the faces of the lens and its material, converging and diverging lenses of different powers can be constructed. [PC8] **It would be possible for us to make lenses for all electromagnetic waves.**[PC9]

Comments (S2).

[PC2] lots of unnecessary, obvious information, you know what telescopes and microscopes are used for, also "part and parcel", not really saying anything, just additional reading.

[PC3] just straight copied out, no own notes made, *what does happen when you look through binoculars the wrong way? If you were going to include the question you should have at least included the answer, otherwise it serves no purpose including it.

[PC4] This is the second time that this same sentence has been included, if you look carefully it is already in paragraph two!!!!!

[PC5] no explanation of what all the lines on the diagram are supposed to be showing, the only useful part and you've left it out.

[PC6] You've just started yacking on about concave lenses, you haven't told me what they are, come to think of it you haven't even told me what convex lenses are yet… this sentence has just been stuck in the middle of nowhere.
[PC7] good….the only paragraph I've seen so far where all the information has been useful.
[PC8] this is the third time you've said this
[PC9] this has got nothing to do with the rest of this paragraph, can you explain why it is there?

Appendix G:Notre Dame Summarisation Study Summaries (S3,S4,S7).

SUMMARY
Lenses Introduction

1. Lenses are the basic components of optical instruments
2. Lenses are part and parcel of everyday life
3. Lenses focus light rays from an object to form an image
4. Shape, material and position of the lens control the properties of the image
5. The eye is a lens system. The image is focused on the retina and interpreted by the brain
6. Lenses have different properties depending on their material and shape
7. The greater the angle of incidence, the greater the angle of refraction, this is Snell's Law
8. With convex lenses, outer rays form a large angle of incidence and therefore a large angle of refraction
9. Curvature of the lens converges the rays to an image point with the convex lens

---

This is my summary.

A light wave can be said to originate from each point of an object and lenses focus the light to form an image point. The eye is a lens system and, unless we have a sight defect, we see each object clearly. Each object point focuses to an image point on the retina of the eye and all these image points are integrated and interpreted as a complete image by our brains.

The path of light wave obeys the laws of refraction, so the greater the angle of incidence the greater the angle of refraction.

Suppose we have a **convex lens** shaped like this:

... and light waves come from a point on the object (shown as rays giving the direction of the waves). Note that the outer rays have a larger angle of incidence and hence a greater angle of refraction, so this curvature of the lens brings rays from each object point to an image point.

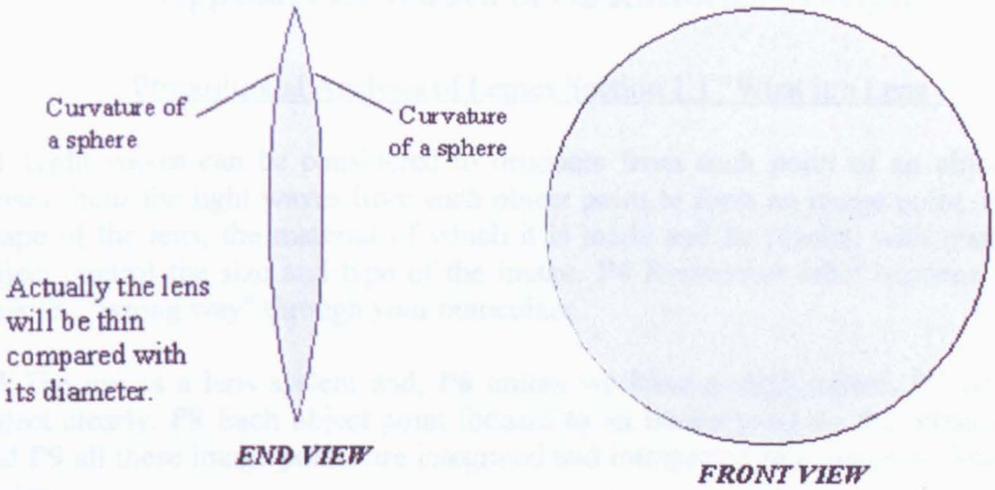
LENSES... Introduction:

This is an important topic because lenses are the basic components of optical instruments such as microscopes and telescopes. These instruments give information about objects from the very small and near (e.g. molecules and bacteria) to the very large and distant (e.g. stars and galaxies). Also instruments such as our eyes and spectacles, photocopiers and cameras use lenses and these are all part and parcel of our everyday life.
But what are the properties of lenses? How do they work and how are they used to our advantage?

What is a lens?

Light waves can be considered to originate from each point of an object and lenses focus the light waves from each object point to form an image point. But the shape of the lens, the material of which it is made and its placing with respect to the object control the size and type of the image. Remember what happens when you look the "wrong way" through your binoculars.
The eye is a lens system and, unless we have a sight defect, we see each object clearly. Each object point focuses to an image point on the retina of the eye and all these image points are integrated and interpreted as a complete image by our brains.

Suppose we have a **convex lens** shaped like this:



... and light waves come from a point on the object (shown as rays giving the direction of the waves). Note that the outer rays have a larger angle of incidence and hence a greater angle of refraction, so this curvature of the lens brings rays from each object point to an image point.

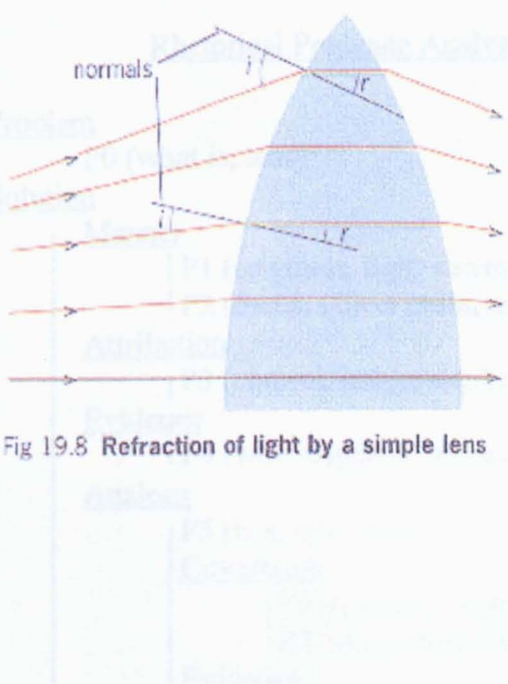


Fig 19.8 **Refraction of light by a simple lens**

Appendix H: Wizard of Oz Rhetorical Analysis.

<u>Propositional Analysis of Lenses Section 1.1 "What is a Lens ?"</u>

P1 Light waves can be considered to originate from each point of an object and **P2** lenses focus the light waves from each object point to form an image point. But **P3** the shape of the lens, the material of which it is made and its placing with respect to the object control the size and type of the image. **P4** Remember what happens when you look the "wrong way" through your binoculars.

P5 The eye is a lens system and, **P6** unless we have a sight defect, **P7** we see each object clearly. **P8** Each object point focuses to an image point on the retina of the eye and **P9** all these image points are integrated and interpreted as a complete image by our brains.

P10 The eye has some similarities in its components - and some differences - to a camera. Can you think what these are?

P11 How does a photocopier work? **P12** How is it able to take a clear image of a whole page of text and pictures? What do you think? **P13** A whole array of cylindrical lenses are responsible for imaging a portion of text/picture in sequence as it is illuminated. Check it out!

<u>Rhetorical Predicate Analysis of above (based on Meyer, 1975)</u>

<u>Problem</u>
- P0 (what is, lens)

<u>Solution</u>
- <u>Manner</u>
  - P1 (originate, light waves, object point)
  - P2 (focus, object point, image point)
- <u>Attribution</u>
  - P3 (control, image size/type, shape, material, placing)
- <u>Evidence</u>
  - P4 (what happens, binoculars, "wrong way")
- <u>Analogy</u>
  - P5 (is a, eye, lens)
  - <u>Covariance</u>
    - P6 (not have, sight defect)
    - P7 (see, object, clearly)
  - <u>Evidence</u>
    - P8 (focus on retina, object point, image point)
    - P9 (interpreted, by brain, image points)
- <u>Adversative</u>
  - P10 (differences, eye, camera)
- <u>Problem</u>
  - P11 (how work, photocopier)
  - <u>Specific</u>
    - P12 (take, clear image, whole page of text)

Solution

| P13 (responsible, lenses array, imaging)

Notes on Rhetorical Predicate Analysis 01/06/00

1. The thirteen propositions were decided by looking for verb, subject pairs, etc. ad-hoc.
2. RP suggested that most instructional texts are problem-solution types as the section starts with P0 "what is a lens ?" this was taken as the *problem* with the ensuing section paragraphs forming the *solution*
3. Meyer's (1975) Rhetorical Predicates were used as quoted in Tawalbeh's (1994) thesis p.59
4. P1 & P2 describe how an event complex is performed ie. *manner*
5. P3 described *attributes* of lens & their role
6. P4 *evidence* supports manner
7. An *analogy* P5 is made between eye & lens
8. If P6 then P7 ie. *covariance* (condition-consequence)
9. P8 & P9 support assertion of P5 ie. *evidence*
10. P10 highlights differences between eye & camera ie. *adversative*
11. P11 sets out another question ie. *problem* -solution plus P12 presents more specific information on the questions
12. P13 present the *solution* to the P11 problem

Appendix I: Sample Wizard of Oz Summary (S3).

- Lenses focus the light waves.
- The shape, material and the size of a lens controls the image.
- The eye is a lens.
- An object point focuses to an image point in the retina and all this points are interpreted by the brain. The eye has some differences and similarities to a camera, e.g. they both use lenses, the eye uses the brain to interpret the image.
- A whole array of cylindrical lenses are responsible for imaging a portion of text/picture in a photocopier.

Appendix J: Affect Scale.

Either during or after the session please circle a number on the continuum indicating your impressions of the characters' personalities. If you feel that neither extreme is appropriate then please circle a zero. Thanks !

Genie

impatient	-5	-4	-3	-2	-1	0	1	2	3	4	5	**patient**
derisive	-5	-4	-3	-2	-1	0	1	2	3	4	5	**encouraging**
obstinate	-5	-4	-3	-2	-1	0	1	2	3	4	5	**amenable**
macho	-5	-4	-3	-2	-1	0	1	2	3	4	5	**wimpy**
scary	-5	-4	-3	-2	-1	0	1	2	3	4	5	**friendly**
curt	-5	-4	-3	-2	-1	0	1	2	3	4	5	**laid-back**
impolite	-5	-4	-3	-2	-1	0	1	2	3	4	5	**polite**

Peedy

impatient	-5	-4	-3	-2	-1	0	1	2	3	4	5	**patient**
derisive	-5	-4	-3	-2	-1	0	1	2	3	4	5	**encouraging**
obstinate	-5	-4	-3	-2	-1	0	1	2	3	4	5	**amenable**
macho	-5	-4	-3	-2	-1	0	1	2	3	4	5	**wimpy**
scary	-5	-4	-3	-2	-1	0	1	2	3	4	5	**friendly**
curt	-5	-4	-3	-2	-1	0	1	2	3	4	5	**laid-back**
impolite	-5	-4	-3	-2	-1	0	1	2	3	4	5	**polite**

Please use the space below to make any further comments regarding the characters' personalities or anything in general.

Appendix K: Example Evaluation Summary (S1).

Lenses

Light waves originate from each point of an object

eg. binoculars (gets smaller when you look the other way)

Lenses focus light from an object point to an image point & onto the retina & brain integrate interpret

Size and type of image are controlled by a lenses shape, material & placing

The eye is a lens system & similarities differences camera

If we have no sight defect each object is seen clearly

How's does a photocopier work?

Take clear image of whole page text then prints the information.

Cylindrical lenses are responsible for text/pictures in photocopiers (The light which appears when you photocopy)

Appendix L: Log Exchange Analysis.

Category	Agent	Move	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	Total	Mean	S.D.
1. agent_turn_ok+	SILA	me next/start													
	(student	OK)													
	SILA	look													
			9	7	6	7	12	8	4	5	6	7	71	7.1	2.23
2. agent_turn_ ignore_student-	SILA	me next/start													
	student	no													
	SILA	look													
			2	0	0	0	0	0	0	0	0	0	2	0.2	0.63
3. agent_denied+	SILA	me next/start													
	student	no													
	SILA	OK													
	(student	look)													
			3	6	4	6	1	10	3	4	6	5	48	4.8	2.44
4. student_turn_ok+	student	me next													
	SILA	OK													
	(student	look)													
			2	1	1	0	2	2	3	2	3	3	19	1.9	0.99
5. student_pass_turn+	student	you next													
	SILA	look/me finish													
			2	0	1	1	0	2	4	0	4	0	14	1.4	1.58
6. student_end_denial-	student	end													
	(SILA	no)													
			0	1	0	0	0	0	1	1	1	0	4	0.4	0.52
7. student_explain*	student	added													
	SILA	why													
	student	because													
			1	0	1	1	1	0	1	0	1	3	9	0.9	0.88
8. SILA_explain*	SILA	added													
	student	why													
	SILA	because													
			2	1	0	0	6	3	3	2	2	0	19	1.9	1.85
9. agent_deny_student-	student	me next/end													
	SILA	no													
			0	0	0	0	0	1	1	0	0	0	2	0.2	0.42
10. student_explain_ denied*	student	added													
	SILA	why													
	(student	no)													
			0	1	1	0	0	3	1	0	2	1	9	0.9	0.99
11. student_no_explain+	student	added													
	SILA	ok													
			1	2	2	3	0	1	2	4	3	1	19	1.9	1.20
12.SILA_no_explain+	SILA	added													
	(student	OK/well done)													
			10	8	5	8	7	7	4	5	6	7	67	6.7	1.77

Category	Agent	Move	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	Total	Mean	S.D.
13. SILA_end_ok+	student	end													
	SILA	hide													
			2	2	1	2	1	2	1	1	1	1	14	1.4	0.52
14. SILA_organise^	SILA	organised													
			11	7	5	7	12	8	7	6	9	6	78	7.8	2.25
15. SILA_condensed^	SILA	condensed													
			3	2	1	3	3	3	3	1	3	2	24	2.4	0.84
16. SILA_finish+	SILA	me finish													
	student	ok													
			0	0	0	0	1	0	0	0	1	0	2	0.2	0.42
Total positive collaboration+			29	26	20	27	24	32	21	21	30	24	254	25.4	4.12
			60%	68%	71%	71%	52%	64%	55%	68%	63%	67%	63%		
Total negative collaboration-			2	1	0	0	0	1	2	1	1	0	8	0.8	0.79
			4%	3%	0%	0%	0%	2%	5%	3%	2%	0%	2%		
Total enquiry*			3	2	2	1	7	6	5	2	5	4	37	3.7	2.00
			6%	5%	7%	3%	15%	12%	13%	6%	10%	11%	9%		
Total SILA act^			14	9	6	10	15	11	10	7	12	8	102	10.2	2.90
			29%	24%	21%	26%	33%	22%	26%	23%	25%	22%	25%		
Overall Total Exchanges			48	38	28	38	46	50	38	31	48	36	401	40.1	7.58
			12%	9%	7%	9%	11%	12%	9%	8%	12%	9%			
SILA added			12	9	5	8	13	10	7	7	8	7	86	8.6	2.46
student added			2	3	4	4	1	4	4	4	6	5	37	3.7	1.42
total added			14	12	9	12	14	14	11	11	14	12	123		
SILA added % additions per student			86%	75%	56%	67%	93%	71%	64%	64%	57%	58%	70%		
student added % additions per student			14%	25%	44%	33%	7%	29%	36%	36%	43%	42%	30%		
SILA added % overall total per student			25%	24%	18%	21%	28%	20%	18%	23%	17%	19%	21%	**21%**	**0.04**
student added % overall total per student			4%	8%	14%	11%	2%	8%	11%	13%	13%	14%	9%	**10%**	**0.04**
SILA explained % overall total per student			4%	3%	0%	0%	13%	6%	8%	6%	4%	0%	5%	**4%**	**0.04**
student explained % overall total per student			2%	0%	4%	3%	2%	0%	3%	0%	2%	8%	2%	**2%**	**0.02**

Appendix M: Evaluation Affect Scale Results.

	S1	S2	S3	S4	S5	S6	S7	S9	S10	affect	affect%
Genie											
patience	-4	1	0	-2	-1	0	0	0	0	-6	-13
attitude	-1	3	0	3	2	0	0	0	4	11	24
likeability	3	0	0	0	0	0	0	0	4	7	16
demeanour	-2	-5	0	-3	-2	0	0	0	0	-12	-27
friendliness	-3	3	0	0	3	0	0	0	5	8	18
relaxation	-4	-1	0	0	-3	0	0	0	2	-6	-13
politeness	-5	2	0	3	-2	0	0	0	5	3	7
overall impression	-16	3	0	1	-3	0	0	0	20		
dominant/passive	-46%	9%	0%	3%	-9%	0%	0%	0%	57%		
Peedy											
patience	5	-3	-3	-1	2	-2	0	-1	0	-3	-7
attitude	2	1	2	4	2	2	4	3	4	24	53
likeability	4	0	3	0	1	2	2	-1	4	15	33
demeanour	0	4	0	0	0	4	0	2	0	10	22
friendliness	3	3	3	4	5	4	5	4	5	36	80
relaxation	0	-3	0	0	4	-1	0	2	0	2	4
politeness	4	-5	-1	3	4	3	5	3	5	21	47
overall impression	18	-3	4	10	18	12	16	12	18		
dominant/passive	51%	-9%	11%	29%	51%	34%	46%	34%	51%		

Appendix N: WebCT API (www.webct.com).

SECTION ONE: OVERVIEW OF THE API

OVERVIEW

What is the WebCT API?

The WebCT Application Program Interface (API) is a tool that provides you with an alternative method of maintaining the global user and student databases. Most users have found this method to be faster, and the API allows them to "automate" the process.

Using the API, you can modify or query both the WebCT global user and student databases directly, without having to use the WebCT administrator or designer web interface. You could, for example, use the API to synchronize WebCT student information with that in your Student Information System (SIS). Furthermore, you could configure the synchronization process so that any SIS changes are immediately reflected in WebCT. Instructors can benefit from this through a reduction in the time that they spend on course administration.

How do you use the API?

You can use the API in two ways:

1. through a command line interface
2. through a Web forms-based interface (URL)

If you choose to use the URL interface, you can only perform operations on single records, and not multiples. For example, through the command line interface you can add multiple student records from a text file (using `fileadd`); you cannot do this with the URL.

Two executable programs provide the API functionality: `webctdb` (used for the command line interface) and `serve_webctdb` (used for the URL interface). These programs allow you perform the following actions on the WebCT global user or student database:

Command Line (webctdb)	URL (serve_webctdb)
add a group of users or students from a file	add a single user or student
add a single user or student	delete a single user or student
delete a single user or student	find and retrieve the record of a single user or student
find and retrieve the record of a single user or student	modify the record of a single user or student
modify a group of user or student records from a file	
modify the record of a single user or student	

Appendix O: Tawalbeh Mark-Up (Tawalbeh, 1994).

<text type = contrast collec = one, few, many> The Polar Regions</text>

<prop pn = 1 L = 1 rp = DIFF>The <topic T = 1>northern polar region </topic> and <topic T = 2> southern polar region </topic>are different in <collec type = many> many ways </collec> </prop>.
<prop pn = 2 L = 1 rp = DIFF>The most important difference concerns the <header h = 1 significance = most imp>distribution of land and water</header></prop>.
<prop pn = 3 L = 2 rp = ATTRIB/ADV LINK = T1,h1>The <syn item = T1> northern Arctic regions</syn> are <value V = 3/h1>ice covered-sea</value></prop>,
<prop pn = 4 L = 2 rp = ATTRIB/ADV LINK = T1,h1>almost completely <value V = 4/h1> surrounded by land</value></prop>.
<prop pn = 5 L = 2 rp = ATTRIB LINK = T1,h1> The pole itself is in deep water </prop>.
<prop pn = 6 L = 2 rp = ATTRIB LINK = T2,h1>In the south, <syn item = T2> Antarctica </syn> </prop>
<prop pn = 7 L = 2 rp = ATTRIB/ADV LINK = T2,h1>Antarctica is a huge <value V = 7/h1>continent </value> </prop>
which is
<prop pn = 8 L = 2 rp = ATTRIB/ADV LINK = T2,h1>surrounded by a great <value V = 8/h1>ocean</value></prop>.
<prop pn = 9 L = 1 rp = COV>Because of this basic difference </prop>
<prop pn = 10 L = 1 rp = COV> other differences occur</prop> <arrow = h1:h2, arrow = h1:h3 arrow = h1:h4) </arrow>
<prop pn = 11 L = 2 rp = ATTRIB/ADV LINK = T1,h2>The <syn item = T1> arctic</syn> has a <value V = 11/h2>varied</value><header h = 2>climate</header> </prop>,
while the
<prop pn = 12 L = 2 rp = ATTRIB/ADV LINK = T2,h2>Antarctic <value V = 12/h2>climate varies little</value></prop>;
<prop pn = 13 L = 2 rp = ATTRIB/ADV LINK = T1,h3>the Arctic has <value V = 13/h3>much <header h = 3>plant life</header></value></header> </prop>
but
<prop pn = 14 L = 2 rp = ATTRIB/ADV LINK = T2,h3>the Antarctic is <value V = 14/h3>an empty desert</value></prop>.
And whereas the
<prop pn = 15 L = 2 rp = ATTRIB/ADV LINK = T1,h4> Arctic has been <header h = 4> <value V = 15/h4>exploited economically</header> for centuries</value> </prop>,
<prop pn = 16 L = 2 rp = ATTRIB/ADV LINK = T2,h4> trade has never really touched Antarctica </prop>.

Alternatively the noun phrase of the first sentence can be marked like this:
The <T1 part = 1>northern</T1> and <T2>southern polar regions</T2> <T1 part = 2> polar region <T1>.

or synonyms can be used instead.
<syn item = T1> northern Arctic region </T1>

Appendix P: Instructions for Summary Markers.

Could you please mark the enclosed summaries which were written by students at CHALCS and Notre Dame ? Your marking should be based on Tawalbeh's Marking Guide (enclosed) which requires that you are familiar with his Model of Summarisation (also enclosed) and the rhetorical structure analysis of the original text (also enclosed). Could you please break down your marking in terms of the four stages, ie. **content**, **staging**, **condensation** and **cohesion and style**[1]. You may find the following table useful:

Name	content	staging	condensation	cohesion and style	Comments

Thanks for your help ! If you have any questions please ask !

[1] "cohesion and style" is synonomous with and was later replaced by "organisation"