

# Automated Explanation of Research Informed Consent by Embodied Conversational Agents

Submitted in partial fulfillment of the requirements of  
Master's Thesis in Computer Science – CS 7990

by  
Rukmal Fernando

Advisor  
Prof. Timothy Bickmore

Submitted on 17<sup>th</sup> December 2009

College of Computer and Information Science,  
Northeastern University

© 2009, Rukmal Fernando

## **Thesis Committee**

### **Thesis Advisor:**

Prof. Timothy Bickmore, PhD  
College of Computer and Information Science,  
Northeastern University, Boston MA 02115 USA.

### **Readers:**

- Prof. Carole Hafner, PhD  
College of Computer and Information Science,  
Northeastern University, Boston MA 02115 USA.
- Dr. Michael Paasche-Orlow, MD, MPH  
Section of General Internal Medicine, Department of Medicine,  
Boston University School of Medicine, Boston MA 02118 USA.

# Contents

1.	Introduction.....	8
1.1.	Research Informed Consent .....	8
1.2.	Adult Health literacy .....	9
1.3.	Readability of Research Informed Consent Documents .....	10
1.4.	Improving the Informed Consent process .....	10
1.5.	Explanation of Research Informed Consent Documents by Embodied Conversational Agents .....	11
1.6.	Proposed Solution .....	12
1.7.	Overview of Thesis .....	13
2.	Related work .....	14
2.1.	Informed Consent Workflow Systems .....	14
2.2.	Intelligent Tutoring Systems .....	14
2.2.1.	Pedagogical Agents.....	15
2.2.2.	Representing teaching goals and achievements.....	16
2.3.	Supporting technologies for teaching systems .....	16
2.3.1.	Generating explanations.....	17
2.3.2.	Automatic summarization of content.....	17
2.4.	Explanation of Research Informed Consent Documents by Conversational Agents .....	18
2.5.	Tailoring information to recipient .....	19
3.	Analysis of Document explanation by humans .....	21
3.1.	Introduction .....	21
3.2.	The explanation of documents by experts.....	21
3.2.1.	Observations of experts explaining documents .....	21
3.2.2.	Analysis of information content.....	21
3.2.3.	Analysis of the order of information.....	22
3.2.4.	Analysis of information delivery .....	22
3.2.5.	Analysis of explanation length.....	23
3.2.6.	Analysis of readability of sentences .....	25
3.2.7.	Summary of findings.....	26
3.3.	Further observations of document explanation by experts.....	26
3.4.	Providing domain knowledge .....	26
3.4.1.	Topic relevance ranking.....	27
3.4.2.	Empirical testing of Topic relevance ranking algorithm .....	28
3.5.	Conversational Agents explaining Research Informed Consent Documents – a first attempt .....	31
4.	Computational Model for Research Informed Consent document explanation.....	32
4.1.	Background of Research Informed Consent Document explanation task .....	32
4.2.	Describing the Document Explanation task.....	32
4.3.	Implementation of the Document Explanation system .....	32
4.3.1.	Producing Informed Consent documents.....	32
4.3.2.	Informed Consent documents as input.....	33
4.3.3.	A non-scripted Conversational Agent.....	33

4.3.4.	Document explanation strategies .....	34
4.3.5.	Enabling traversal of documents.....	35
4.4.	Example conversation transcript .....	37
4.5.	Parallels with Intelligent Tutoring Systems .....	41
4.6.	Realizing the Document Explanation task .....	42
4.6.1.	Facilitate a natural method of interaction .....	42
4.6.2.	Displaying the document .....	42
4.6.3.	Pointing out the section that the current dialogue relates to .....	42
4.6.4.	Where possible, allow the user to control the level of detail .....	42
4.6.5.	Explain the document in a logical order .....	43
4.6.6.	Provide any background information that may help the participant.....	43
4.6.7.	Ask interactive questions and provide necessary corrections.....	43
5.	Evaluation of document explanation system .....	44
5.1.	Background .....	44
5.2.	Hypotheses .....	44
5.3.	Experimental design.....	45
5.3.1.	Methods.....	45
5.3.2.	Materials .....	45
5.3.3.	Research Design.....	46
5.3.4.	Explanation strategies .....	46
5.4.	Measures.....	47
5.4.1.	Primary outcome measure - Comprehension.....	47
5.4.2.	Secondary outcome measure – Satisfaction.....	47
5.4.3.	Participant’s Health Literacy .....	47
5.4.4.	Need for Cognition .....	48
5.4.5.	Participant Demographics.....	48
5.5.	Results .....	48
5.5.1.	Participants.....	48
5.5.2.	Scoring Comprehension tests.....	52
5.5.3.	Testing for order effects – Comprehension .....	54
5.5.4.	Hypotheses testing – Comprehension.....	54
5.5.5.	Hypotheses testing – Satisfaction .....	55
5.5.6.	Other tests .....	58
5.6.	Qualitative review .....	61
5.7.	Discussion .....	61
6.	Conclusions and future work .....	63
6.1.	Conclusions.....	63
6.2.	Future work .....	64
7.	Acknowledgements.....	65
8.	References.....	66

## List of Figures

Figure 1 Sample Research Informed Consent Document.....	8
Figure 2 Anderson et al.'s Intelligent Tutoring System for LISP programming.....	15
Figure 3 Bickmore et al. Comprehension scores across Mode by health literacy level ...	19
Figure 4 Scatterplot for participant health literacy and session duration.....	23
Figure 5 Scatterplot for REALM score vs ST22 word ratio .....	25
Figure 6 Comparison of document explanation strategies.....	35
Figure 7 Statechart for Document Explanation .....	36
Figure 8 Participants by gender .....	49
Figure 9 Age distribution .....	49
Figure 10 Participant ethnicity.....	50
Figure 11 Health Literacy level results .....	50
Figure 12 Last grade of school completed.....	51
Figure 13 Experience with Computers .....	51
Figure 14 Need for Cognition for participants.....	51
Figure 15 Prior participation in research studies .....	52
Figure 16 Participation in research studies in 2009 .....	52
Figure 17 Distribution of all Comprehension scores .....	53
Figure 18 Scatter-plot of comprehension scores by Researcher and Judges .....	53
Figure 19 Means for Comprehension scores by Mode of presentation .....	54
Figure 20 Comprehension scores and Need for Cognition across Modes .....	55
Figure 21 Satisfaction scores by Mode of session (approaching significance) .....	56
Figure 22 Effect of Need for Cognition on Satisfaction across Modes .....	56
Figure 23 Effect of Need for Cognition on self-reported perceived Pressure .....	57
Figure 24 Effect of Need for Cognition on self-reported Likelihood to Sign .....	57
Figure 25 Means of "How much information did you get?" .....	58
Figure 26 Incorrect vs. Correct responses to Agent's questions on first attempt.....	60
Figure 27 Incorrect vs. Correct responses to Agent's questions on second attempt .....	60

## List of Tables

Table 1 Summary of explanation sessions.....	21
Table 2 Summary of Fact reference scores.....	22
Table 3 Information Content by participant.....	24
Table 4 Topic relevance ranking algorithm - initialization .....	27
Table 5 Topic relevance ranking algorithm – ranking.....	28
Table 6 Topic associations by Expert and the Lin and JcN algorithms.....	29
Table 7 Expert's Topic associations and corresponding ranking by Lin and JcN algorithms .....	29
Table 8 Overview of Document Explanation task.....	32
Table 9 Sample conversation - Verbose mode .....	37
Table 10 Sample conversation - Tailored mode .....	38
Table 11 Various document display strategies .....	39
Table 12 Participant interacting with Agent .....	40
Table 13 Comparison of AutoTutor and this work.....	41
Table 14 REALM scores and meaning.....	47
Table 15 Cross-tabulation of responses to Agent's questions on first attempt .....	59
Table 16 Cross-tabulation of responses to Agent's questions on second attempt.....	60

## **Abstract**

Research has shown that the comprehension of documents written by professionals for laypersons - such as research informed consent forms - is a challenging task for many due to low literacy level, lack of prerequisite domain knowledge, and the use of arcane, complex language that is above the individual's reading level.

Prior work has demonstrated that an embodied conversational agent can successfully explain research informed consent documents to participants, resulting in higher comprehension and satisfaction compared to self-study of the document. However, this prior work relied on scripted explanation dialogue, limiting the ability of the system to be deployed across a large number of studies.

In this thesis I describe a system that automatically and dynamically generates the explanation for a research informed consent document and delivers the explanation via an embodied conversational agent, given a machine-readable description of the document's contents and structure. This system is evaluated in an empirical study that compares two versions of the automated approach with self-study of the document. Results indicate that participants learned more when forced to hear more information from the agent, although they were less satisfied with this version of the system compared to a version that attempted to dynamically tailor information to the participant's prior knowledge.

# 1. Introduction

Documents are an integral part of our lives, being perhaps the most widespread method of information dissemination. Documents, however, are written only once, usually with a broad target audience in mind [36]. Thus, they may carry too much information for some individuals, or too little for others.

This is an especially serious problem in the health information domain, because the ramifications of misunderstanding on the part of the reader are serious. For example, the Institute of Medicine reports that nearly half of all American adults—90 million people—have difficulty understanding and acting upon written health information [4, 42, 44].

## 1.1. Research Informed Consent

Research Informed Consent is an important ethical and legal process, whose purpose is to provide participants or patients with the ability to make an educated and autonomous decision as to whether they want to participate in a research study. Indeed as Faden et al. indicate, “informed consent has less to do with the liability of professionals as agents of disclosure, and more to do with the autonomous choices of patients and subjects.” [20].

Research Informed Consent Documents are frequently used as the main vehicle of information in this process. Below is a sample document that was created based on a template for Genetic Research [18], with some sub-section headers added. The complete document is available in Appendix A (as the “Human Genetic Cell Repository” form).

<p><b>Human Genetic Cell Repository</b> Informed Consent Form</p> <p><b>Why am I being asked to take part in this research study?</b> You are being asked to participate in this study because you are an adult who knows if you or any family members have certain medical conditions or not.</p> <p><b>Why are you doing this research study?</b> The purpose of this study is to find if there are any relationships with certain medical conditions and genes.</p> <p><b>What will I be asked to do?</b> If you decide to take part in this study, a 5 ml (1 teaspoon) blood or skin biopsy sample will be taken to the NU Cell Repository, a research resource supported by the National Institutes of Genomics. The Repository collects, stores, and distributes cell cultures and DNA samples from people with many kinds of disorders, from unaffected family members, and from other healthy people. Your sample will be used to create a cell line that will be added to the Repository's collection. The purpose of this collection is solely to make specimens available for use in research and teaching and as standards in clinical genetics laboratories. Submission of your sample to the Repository may give scientists valuable research material that can help them to develop new diagnostic tests, new treatments, and new ways to prevent diseases.</p> <p><b>Where will this take place and how much of your time will it take?</b> This session will take place at Northeastern Institute for Medical Research and will take approximately one hour.</p> <p><b>Will there be any risk or discomfort to me?</b> The medical risks of providing these specimens are minimal. The risk for venipuncture is minor transient pain and slight possibility of infection. The risk for skin biopsy is mild local pain, slight bleeding, the possibility of a small scar, and slight possibility of infection.</p> <p><i>This is a fake consent document – it does not imply any actual commitments or contract.</i> Page 1 of 3</p>	<p><b>Will I benefit by being in this research?</b> There will be no direct benefit or payment to you for participating, but your sample may benefit the community at large or some particular group.</p> <p><b>Who will see the information about me?</b> The Repository will take measures to protect your privacy. Your blood or tissue specimen will be given a code number; your name will be removed, and the NU Cell Repository will not give out your name to the scientists who receive the samples. Some patient identification, such as age, sex, diagnosis, and race, will be made available to the Repository and scientists. There can be no absolute guarantee of confidentiality, however, and there is a small chance that some research may yield results that will have a negative impact on you, your family, other individuals, or groups. These may include insurability, employability, and/or family relationships.</p> <p>The Repository has obtained a Confidentiality Certificate from the Federal Government to help insure your privacy. This Certificate means that the Repository cannot be forced to tell people who are not connected with the study about your participation without your written consent.</p> <p><b>If I don't want to take part in the study, what choices do I have?</b> Your donation is voluntary, and if you choose not to participate there will be no penalty or loss of benefits to which you are entitled.</p> <p><b>What will happen if I suffer any harm from this research?</b> If you are injured as a result of your participation in this research, no special arrangements will be made for compensation or for payment for treatment.</p> <p><b>Can I stop my participation in this study?</b> Your sample will stay in the Repository indefinitely and you will not be able to withdraw it.</p> <p><i>This is a fake consent document – it does not imply any actual commitments or contract.</i> Page 2 of 3</p>	<p><b>Who can I contact if I have questions or problems?</b> If you have any questions or complications relating to collection of this specimen, you should contact Dr Brian Orlow at (617) 373-1234, who collected the specimen. If you have any questions about the Repository, you should contact Dr Michael Jack, Northeastern Institute for Medical Research, (617) 373-5678.</p> <p><b>Who can I contact about my rights as a participant?</b> If you have questions about your rights as a research subject you should call Dr Laura Silliman, (617) 373-7890.</p> <p><b>Will I be paid for my participation?</b> You will not be paid to participate in this research study.</p> <p><b>Will it cost me anything to participate?</b> Your donation is voluntary, and there are no costs to you for participating in this research study.</p> <p><b>Is there anything else I need to know?</b> Because researchers will not have access to your identity, it will not be possible to provide you or your physician with the eventual results of studies that might be performed using your specimen.</p> <p><b>Will this be used in any publication or teaching material?</b> It is possible that data resulting from use of your sample may eventually be used in a research publication. In that event, no individual's identifying information will be included, as this information will not be available to the researchers.</p> <p>(Signature) _____ Do Not Sign _____ (Date) _____</p> <p><i>This is a fake consent document – it does not imply any actual commitments or contract.</i> Page 3 of 3</p>
--	--	---

Figure 1 Sample Research Informed Consent Document

Some key elements of these documents are worth pointing out:

- An overview to the Research study being conducted
- An explanation of the Study Procedures
- An explanation of possible risks to the participant, and any possible benefits
- Information on privacy and confidentiality of the participant's information



- Information on whether and how the participant can withdraw from the study
- Contact information, in case the participant wishes to withdraw, needs to report anything, or has questions about his/her rights
- A place where the participant signs to acknowledge that they understood the Research Informed Consent document

It is generally advised that the Principal Investigator or an approved designee verbally explain a Research Informed Consent Document to a potential participant. The participant is encouraged to ask questions to clarify their understanding, and finally the participant is given as much time as they deem necessary to read the Research Informed Consent Document by themselves, before they decide whether to participate or not [1].

However there are many challenges in this ideal process. In particular, researchers (especially in clinical research) have limited capacity to spend enough time to explain the study properly, participants themselves may feel pressured to not ask too many questions, and researchers are not always consistent in following guidelines and the information they provide [8].

As an example, Dr. Michael Paasche-Orlow, a nationally recognized expert on health document readability, explained the Research Informed Consent Document shown in Figure 1 to a mock participant in a video-taped session at the Human-Computer Interaction lab of Northeastern University. In this session, Dr. Paasche-Orlow took approximately 25 minutes for the explanation, in the process asking the participant several questions to confirm his understanding, providing relevant background information, and answering questions the participant had. While being the ideal, this sort of “gold standard” procedure is followed in most real-world study scenarios.

The Research Informed Consent Document itself is therefore a critical part of this process, and it is critical that these can be understood by most, if not all, potential participants, irrespective of their educational backgrounds and literacy capabilities. Unfortunately though, some evidence suggests that an alarmingly high number of participants misunderstand Informed Consent documents [23].

## 1.2. Adult Health literacy

The National Literacy Act of 1991 defines literacy as “an individual's ability to read, write, and speak in English and compute and solve problems at levels of proficiency necessary to function on the job and in society, to achieve one's goals, and to develop one's knowledge and potential.” [32]. Within this context, health literacy is defined as “the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions” [42]. Both of these definitions show that working health literacy is critical to day-to-day function as well as making key health decisions.

Inadequate adult health literacy is a serious problem in the United States [44], and indeed the world over. In 2004, the Institute of Medicine reported that nearly half of all American adults, or 90 million people, have difficulty understanding and acting upon health information [1].

Individuals with low levels of health literacy have a difficult time functioning within the health care environment [5]. Even though the average reading level of patient materials related to health care has been 11<sup>th</sup> to 14<sup>th</sup> grade, even patients who read at the college level have been found to prefer medical information written at the 7<sup>th</sup> grade level [5]. There is thus a clear disparity between the readability level of medical information, and the abilities and preferences of the individuals who need to use this information.

Low health literacy is further associated with the risk of experiencing poorer health outcomes [1, 5], increased risk of hospitalization [4] and reduced rates of medication adherence [26]. It is also very interesting that properly developed material has been shown to help participants with inadequate health literacy or cognitive difficulties [20, 45], so the question really seems to be about how the disparities can be reduced.

This issue of limited health literacy among participants is a serious one that needs attention. Indeed, this issue is much broader, as even people with strong literacy skills may have trouble obtaining, understanding, and using health information [42].

### 1.3. Readability of Research Informed Consent Documents

Proponents of health information readability have been working towards guidelines for improving readability as well as objective metrics for measuring readability and creating standards [20].

Institutional Review Boards (IRB) of institutions conducting human participants research routinely publish guidelines and samples. However, a review by Paasche-Orlow et al. [46] of the readability of Informed Consent documents at the websites of 114 Medical schools in the USA found that while 61 provided specific guidelines on readability standards, the samples and templates provided by 92% of them did not even meet their own standards.

Thus it appears that the guidelines and standards, while certainly being an important step, have done little to actually cause significant changes to occur in the actual readability of the Research Informed Consent Documents.

### 1.4. Improving the Informed Consent process

Parallel with the movement to improve the readability of Informed Consent documents, is that of improving the process itself. Several key proposals have gained traction in recent times.

The use of multimedia to produce “Enhanced” Informed Consent documents has been demonstrated to be helpful in bringing patients with psychotic disorders to a comprehension score level comparable to normal control participants [20].

Testing the participant’s understanding periodically, rather than just asking “Do you understand?” is necessary to ensure that the participant has, in fact, understood the material. One way of doing this is to ask interactive questions while the Informed Consent document is being explained. One study reports that participants who received corrections after such interactive questions demonstrated higher post-consent comprehension rates [47] (although not statistically significant).

A complimenting strategy is the use of “Teach-back”, where the researcher asks the participant to explain key elements of the Informed Consent documents in their own words. This enables the researcher to identify areas of misunderstanding, and specially, common misconceptions (e.g: misconceptions like that the researcher will provide medical care free of charge for any injuries that may happen) [37].

Additionally, both reading a Consent Form together with the participant (also called “read aloud”) and repetition of a section as a response to the participant’s misunderstanding, were shown to increase participant understanding [47].

In the video-taped explanation (see section 1.1), Dr. Paasche-Orlow was observed using three of these four strategies; close-ended interactive questions (e.g.: “*What are they going to do with your genes, do you know?*”), more open-ended teach-back (e.g.: “*tell me what that means to you*” after explaining a section, “*tell me... if you can summarize in your own words, what this study is about*”), and a read-aloud protocol. While this sort of one-on-one explanation of an Informed Consent document to a participant is recommended [23], practical issues make this difficult.

## 1.5. Explanation of Research Informed Consent Documents by Embodied Conversational Agents

Over the past two decades, computer programs that can simulate a conversation with a person have been developed, further augmented by anthropomorphic embodiment and the capability to use synthesized speech to communicate with a user. They are also able to simulate human conversational behaviors like various body postures and the use of facial expressions and gestures. Such Embodied Conversational Agents (ECA) have been increasingly used as educational aids. Because a computer agent never gets tired or bored, it can engage in teaching tasks repeatedly and consistently.

There are a few attempts at applying Embodied Conversational Agent technology to improving the administration of Research Informed Consent Documents. Most notably, Bickmore et al. [8] extended their Conversational Agent system to explain a mock Research Informed Consent Document to participants and found that participants were

more satisfied with having the Agent explain the document to them, compared to having a researcher explain it to them or having to read it themselves.

Similar applications have been attempted for the purpose of training researchers too. Hubal and Day [31] developed an Agent that mimicked a research participant who asked questions from the researcher and was shown to be effective in training researchers.

The primary limitation of these applications, in particular that of Bickmore et al., is that these explanations rely on an explanation dialogue that needs to be created verbatim (“scripted”) *a priori*, and so requires a lot of time and effort.

Bickmore et al. make another interesting observation, where participants with inadequate health literacy asked more questions from the Agent than participants with adequate health literacy (although not statistically significant). In semi-structured interviews, participants with inadequate health literacy were also reported to indicate that participants felt comfortable in asking the Agent repeated questions.

## 1.6. Proposed Solution

Documents are a ubiquitous medium of information delivery, but because of various reasons, many individuals have difficulty in understanding these fully.

The use of Conversational Agents in explaining such documents has been demonstrated to be effective in educating individuals on these documents, and helping them to understand the documents better [6, 7, 8]. Developers of these systems need to spend a lot of time in preparing these explanations in advance though, and so their ability to be deployed and used across a large number of studies is limited.

However, previous research suggests that the task of explaining information can simply be based on what is known about the task domain. For example, Sibun et al [50] developed a system that could use a knowledge base to describe an apartment layout by mimicking some common strategies humans use. More recently, Bickmore et al. developed a Virtual Nurse that could explain to patients a tabular medical data in a document [6, 7] using a scripted dialogue augmented by patient-specific data that appeared on the document.

As indicated by Bickmore et al. [8], the ability to answer participants’ questions seems to be a feature that contributes to participant satisfaction and comprehension. The questions are also, however, part of the scripted conversation, and must be provided by the designers of the systems. In Bickmore’s system, possible responses that the participant can provide to the Agent, and any questions that they can ask the Agent, are represented as buttons on a touch screen. Since this relies on the participant’s initiative, the designers of these systems must make judgments about where in the conversation each question makes sense.

With these in mind, I propose an extension to this system that

- Uses the structure that is inherent in a document to guide an explanation of the document, so that *a priori* scripting is not required (or is minimal)
- Leverages a large knowledge base of possible questions that the participants can ask, and their relevant responses and seamlessly integrate these in to the conversation at the most appropriate points

In the sections that follow, I provide a more detailed look at these in turn.

## 1.7. Overview of Thesis

In this thesis is I attempt to answer three key questions:

- How well can a Conversational Agent perform as a learning guide, when provided with extremely simple capabilities for teaching?
- Can a Conversational Agent improve the Informed Consent experience for potential participants?
- Can we develop Conversational Agents that can be adapted very easily to explain new documents?

In particular, I present a new Agent system that can explain a Research Informed Consent Document without relying on scripted dialogue, and instead uses the structure of the document itself to guide the explanation. I also present an approach to linking these explanations with a large knowledge base of questions that participants can ask, and their responses, and evaluate an algorithm for determining when to present a question option in the dialogue.

The rest of this thesis provides background for my work, my own previous work, and finally the experimental setup that I designed. I then conclude with a discussion of the results obtained and their implications for future work.

Research Informed Consent documents are a significant challenge, particularly because

- Unlike certain other instructions (e.g.: prescriptions) they do not repeat and thus individuals are not able to develop familiarity with these concepts and procedures
- They refer to potentially complex procedures, and can contain many medical terms, and thus can be confusing
- The volume of information may be too great
- They are usually, intentionally or otherwise, phrased in a manner that protects the health service provider or researcher, from any possible ill-effects, and so may not be written in a manner favorable to the reader

Recent work by Bickmore et al. [6, 7] suggest that a Conversational Agent can be effective at explaining document to laypersons, and that participants would feel more satisfied with such an explanation as opposed to having to read a document themselves. More recently, Bickmore et al. showed similar results with a Conversational Agent [8] that explained Research Informed Consent documents, and so forms the motivation for the rest of this thesis.

## 2. Related work

In this section I discuss work towards automating the Informed Consent process and point out how Conversational Agents may fit this process.

### 2.1. Informed Consent Workflow Systems

While guidelines and readability standards exist for the authors of Informed Consent documents, adherence to these standards is still low [46], likely because of the manual effort required. Computer based systems could be a possible solution for improving this workflow, possibly incorporating feedback on issues like readability.

At least two such workflow systems are known to exist. The web-based ProtoType has been developed by a team of researchers from NIH [24, <http://prototype.cc.nih.gov/>]. While ProtoType appears to automate referencing and re-using prior protocols, one concrete advantage appears to be in dealing with adverse event notifications etc, which can be automatically sent out to collaborators. Other than these indirect advantages, it is not clear how this tool contributes to improving readability for the participant.

The Boston University School of Medicine also has a web-based system that provides authoring support for Informed Consent documents. [Paasche-Orlow, private communication].

### 2.2. Intelligent Tutoring Systems

Intelligent Tutoring Systems (ITS) use simulated dialogue to function as virtual tutors. Early examples like Anderson et al. [2] developed one of the earliest known examples that provided interactive learning environments for learners to engage in learning geometry and LISP programming.

The figure below shows their LISP programming environment, where the ITS gives a hint at the top, as well as a list of goals at the bottom.

I assume you are trying to test if the number argument is equal to 0. There is a special function for testing equivalence to 0. Try using that function instead of EQUAL.
CODE FOR fact
<pre>(defun fact (n)   (cond ((equal ) &lt;ACTION&gt;)         &lt;RECURSIVE-CASE &gt;))</pre>
GOALS
Write code to get the value of fact Code the terminating case *** Code the condition for the terminating case ***

Figure 2 Anderson et al.'s Intelligent Tutoring System for LISP programming

Using a variety of computational models, modern ITS have produced learning gains of between 0.3 and 1.0 standard deviations compared to students learning the same content in classrooms [27]. Modern ITS like AutoTutor employ a number of strategies, like feedback, pumping for information, hints, prompts, summaries, corrections and more [28]. They also employ complex models to represent the learner's current knowledge [16] [52].

There are at least two primary reasons cited for the pedagogical efficacy of ITS. First, the ITS allows one-on-one tutoring which is the most effective form of learning facilitation. Secondly, the ITS can sometimes model, or exemplify, good learning strategies [28].

There is also evidence to suggest that interactive tutorial dialogue of the kind offered by these ITS systems are not always reliably more effective than simply reading some material. For example, novice students learning physics were either given text meant for intermediaries, or the same content was presented by the Why2-Atlas and Why2-AutoTutor systems, and the novice students had bigger knowledge gains with the ITS than by reading themselves [40]. However, no such difference was found with intermediary students, suggesting that perhaps the strength of ITS systems is in fact making confusing material easier to understand for those who have the most difficulty with it.

### 2.2.1. Pedagogical Agents

Conversational Agents are virtual humans that function as ITS, simulating human conversational behavior, using speech, facial expressions (e.g.: smiling, head-nod to acknowledge or to show understanding), hand gestures (e.g.: emphasis, contrast or pointing) along with body postures [14]. These verbal and non-verbal gestures are used to improve the user experience through a number of mechanisms. In the health behavior change domain, these have been shown to increase adherence to treatment regimens [10].

Pedagogical Agents are the result of combining Intelligent Tutoring Systems with Conversational Agent technologies. Modern systems like AutoTutor are prime examples, featuring embodiment, mixed-initiated dialogue and user affect detection and response in addition to the Tutoring strategies pointed out earlier.

### 2.2.2. Representing teaching goals and achievements

For the sake of understanding how an ITS represents its teaching goals, and the learner's progress, the AutoTutor system will be taken as an example. A detailed explanation is presented in [27].

AutoTutor generally presents deep reasoning questions to learners and uses a variety of conversational moves to elicit the correct response from the learner. Some of these moves may be indirect (e.g.: a hint), or direct (e.g.: asserting the correct response), or somewhere in between (e.g.: prompts for a missing word). In order to do this, each question is associated with a number of expectations. At each turn of dialogue, AutoTutor attempts to get the learner to fulfill an expectation; when fulfilled, it selects the next expectation, until the question is answered. Since AutoTutor works with natural language text input (i.e.: typed on a keyboard by the learner), it makes use of Latent Semantic Analysis to match the learner's responses to expectations, and AutoTutor is reported to be almost as good as an expert in computer literacy in evaluating the quality of learner responses in tutorial dialogue.

AutoTutor's feedback to the learner works on three levels:

- Backchannel feedback (head nods, "uh-huh" after typing nouns etc.)
- Evaluative pedagogical feedback (facial expressions and intonation, for example, "not really" and shaking of the head to indicate negativity, and "right" and a quick head nod to indicate positivity).
- Corrective feedback (e.g.: corrections to misconceptions)

AutoTutor's tutorial dialogue is driven by a curriculum script, with a topic for each major deep reasoning question. Each topic is associated with a set of expectations, a set of hints and prompts for each expectation, a set of anticipated misconceptions and corrections and optional graphical content. Converting this content in to a cohesive dialogue is by means of a Dialogue Advancer Network (DAN).

## 2.3. Supporting technologies for teaching systems

A brief overview is presented here of two technologies in particular, that I feel have several interesting applications in teaching systems. Natural Language Generation and Automatic Text Summarization are two possible ways that could be used to help drive an automated explanation of a document.



### 2.3.1. Generating explanations

Natural Language Generation (NLG) is a branch of Natural Language Processing (NLP) that looks at generating natural language summaries and explanations of pieces of data or a source text. [47].

For example Sibun et al [50] that could describe houses and apartments by mimicking strategies humans use. This was accomplished by using a knowledge base of the objects in the living space, and traversing this knowledge base using a variety of strategies that were joined together by meta-strategies, where the strategies mimicked human strategies. An example strategy is to pick a room and an anchor point within the room, and then to sweep left and right from the anchor point and describe objects.

In general, the more structured the information is, the easier the task becomes, of explaining it. Bickmore et al.'s system provides some metastrategies for explaining a tabular document [8]. This however only applies to certain types of tables, and the other parts of the discharge booklet still require scripted explanations.

### 2.3.2. Automatic summarization of content

Spärck Jones discusses automatic (i.e.: computational) summarization of textual content in great detail [51]. A clear distinction is made between two very different high-level strategies: Extractive summarization essentially uses various algorithms to find the most representative sentences from a given input text. Abstractive summarization on the other hand, identifies and re-presents the source content.

Abstractive summarization is generally a supervised learning problem, which needs a large corpus at minimum to produce reasonable summaries, and there appear to be no readily usable public domain abstractive summarization tools. Because of this, abstractive summarization was not pursued further, but is noted here as a possible future direction.

In the first part of this thesis, I developed a Conversational Agent that explained documents by performing extractive summarization given a machine-readable source document (based on the document in Figure 1). As explained in section 3.5 however, this approach did not contribute to the ultimate vision of this thesis, but is again noted here as a possible future step.

One advantage of extractive summarization over abstractive summarization can be observed though. Especially in the health information domain, as is the case with this thesis, the summaries generated by extractive summarization systems contain verbatim text from the source, and so the only risk is of under-specifying information. With abstractive techniques however, it is possible that the summary conveys a message that a human participant may interpret in a substantially different manner.

## 2.4. Explanation of Research Informed Consent Documents by Conversational Agents

As explained previously, Conversational Agents have many applications in education, counseling, and advising, given their ability to emulate human behavior and form relationships.

In addition to previous work in document explanation [6, 7, 9], Bickmore et al. recently developed a Computer Agent to investigate the explanation of Research Informed Consent Documents [8] and is the most immediate precursor to this thesis and is discussed in further detail in the rest of this section.

Forming the basis of this system was an observational study of experts explaining Research Informed Consent Documents to potential participants. Pointing gestures at the document were found to be very common, and after analyzing 1994 expert utterances, 26% of utterances were reported to be accompanied by a hand gesture, and 90% of gestures to involve pointing at the document. These findings and previous findings were incorporated in to the BEAT toolkit [15], which generated appropriate non-verbal behavior (including document reference and pointing in this case) based on the discourse.

Further, initial mentions of part of a document were found to be more likely to be accompanied by a pointing gesture (43% vs. 19%). A reference to a page was found to be more predictive of a flat hand gesture, while a reference to a word or image was found to be predictive of pointing with a finger.

Finally, the expert was found to omit detail, and provide more scaffolding (i.e.: description of the document's structure) when explaining to a listener with low health literacy.

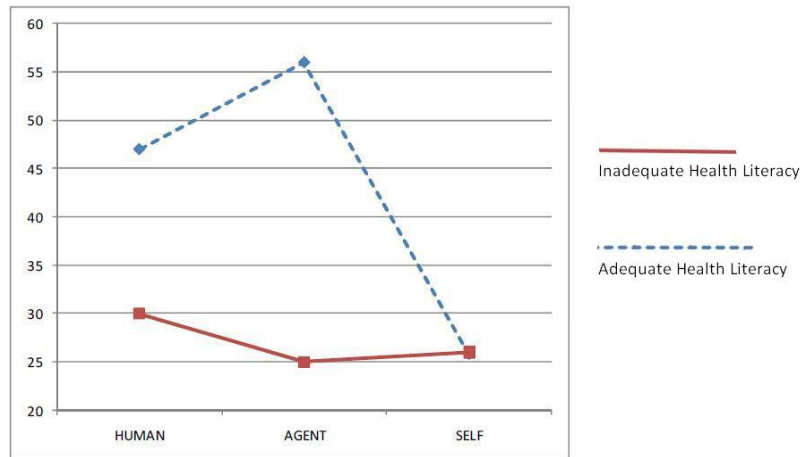
Two three arm randomized experiments were conducted, measuring two primary outcomes: post-intervention knowledge test and satisfaction in the experience. Both studies tested three conditions (self-reading vs. human expert explanation vs. Agent explanation).

The 18 participants in the first study were reported to be mostly college students with high levels of health literacy, and the study was thus deemed to have lesser ecological validity. The second study had 29 participants, 66% female, aged 28-91 (mean 60.2) with varying health literacy levels as measured by the Rapid Estimation of Adult Literacy in Medicine (REALM) instrument [19].

A significant interaction was reported between condition and literacy (Adequate vs. Inadequate), where participants with adequate health literacy showed significantly higher comprehension scores in the human expert and Agent conditions as compared with the Self-read condition, and no such differences existed among the participants with inadequate health literacy. Further, the participants with Inadequate health literacy as a

whole had lower comprehension scores when compared with the participants with Adequate health literacy.

The comprehension scores are demonstrated below (condition on X-axis, with dashed lines for the *Adequate health literacy* condition and solid lines for the *Inadequate health literacy* condition).



**Figure 3 Bickmore et al. Comprehension scores across Mode by health literacy level**

A significant main effect of condition is also reported, with participants being more satisfied in the Agent condition than with the Human condition.

This study stands as the first indication that Conversational Agents are a possible means of improving the Informed Consent process. The fact that participants felt more satisfied when the Agent explained the document than when they read it themselves or had a human expert explain it, confirms the initial suggestion made in this thesis that a Conversational Agent can improve the Informed Consent experience.

The interaction effect seen with the health literacy level of the participants is in fact an opportunity for new strategies to be explored to see if a Conversational Agent can facilitate learning among participants of varied health literacy levels.

## 2.5. Tailoring information to recipient

The concept of tailoring delivery of health information to a participant is very closely related to that of readability standards. Readability standards promote making the information easier to understand for the general audience as a whole; another possibility worth exploring is to see if the information can be tailored to be suitable for a particular person given some knowledge about that person.

Several studies support this view. Computer-tailored information [36] has been shown to increase fruit and vegetable intake and more changes in fat-related behavior as compared to non-tailored information [20], [54], [12].

This is even more critical when dealing with low health literacy individuals. Low health literacy Asthma patients have been associated with lower Asthma medication knowledge, and worse metered-dose inhaler (MDI) technique. A tailored one-on-one educational session showed improvements in knowledge and MDI techniques [45]. Interestingly, in spite of their low health literacy, tailored material and procedures were shown to help these patients in acquiring knowledge and skills, and thus similar improvements may be possible elsewhere.

### 3. Analysis of Document explanation by humans

#### 3.1. Introduction

In this section I discuss several preliminary analyses that I performed in order to design the Agent solution that was proposed earlier.

Most of this section is based on previous work by Bickmore et al. [9] which included the study of a human expert explaining health documents to volunteers.

This study involved video recorded explanations of a Research Informed Consent document to participants by two human experts. Nine of these participants provided written permission for the video recordings to be used for later research purposes, and of these the first expert's explanations ( $n = 5$ ) were chosen for further exploratory analysis.

#### 3.2. The explanation of documents by experts

##### 3.2.1. Observations of experts explaining documents

As the first step, the explanations in the chosen sessions were transcribed to text using the speech, phonetics and acoustics analysis application PRAAT [11].

During the study, the health literacy levels of the participants were measured using the Rapid Estimate of Adult Literacy in Medicine (REALM) instrument [19]. More details about the REALM instrument, and its scoring is provided in section 5.4.3.

Below is a summary of the sessions that were analyzed.

<b>Participant ID</b>	<b>REALM raw score</b>	<b>REALM bracket level</b>	<b>Session duration (seconds)</b>
102	17	3 <sup>rd</sup> grade or below	282
103	64	High school	336
108	62	High school	496
109	47	7 <sup>th</sup> to 8 <sup>th</sup> grade	368
111	1	3 <sup>rd</sup> grade or below	231

**Table 1 Summary of explanation sessions**

##### 3.2.2. Analysis of information content

In order to be able to compare the information that the expert presented to each participant, a list of key facts was created, containing all the facts that were mentioned in the Informed Consent document that was being explained. The explanations used a document very similar to that presented in Figure 1, with 25 sentences containing a total of 60 different facts. Appendix J lists the facts that were identified.

### 3.2.3. Analysis of the order of information

Once the facts were identified, I looked at the order in which each fact was introduced in the explanations, ignoring any facts which were not introduced. In each of the five explanation sessions, the expert proceeded to mention the facts in the same order as they were found in the document.

In other words, the expert's explanation followed the linear structure of the document from start to end. As pointed out, some facts from the document were never introduced in some or all of the sessions.

### 3.2.4. Analysis of information delivery

Once the facts were identified and the ordering was known, it was apparent that while some facts were explained to all participants (and some were likewise not presented to any), there were some facts which only some of the participants received.

It appeared that perhaps the expert decided consciously to refrain from referring to some facts with some participants. There are two possible explanations for this:

- either the expert believed the participant would not understand the fact
- or the expert believed the participant already knew or understood the fact.

Each session was compared against the 25 sentences and the 60 facts identified earlier. For each of the 60 facts identified, a "Fact Reference Score" is defined, where the Fact Reference Score of the  $i^{\text{th}}$  fact is 1 if the fact was mentioned fully, 0.5 if it was mentioned partially, or 0 otherwise.

Then, a Fact Reference Score is calculated for each of the 25 sentences in the source document, where the Fact Reference Score for the  $j^{\text{th}}$  sentence is the mean Fact Reference Score for all the facts that belong to the  $j^{\text{th}}$  sentence.

Finally, the Fact Reference Score of the session is defined as the mean Fact Reference Score of the 25 documents of the source document. All Fact Reference Scores (i.e.: at Fact, Sentence and session level) lie within the range 0..1, and are a simple indication of the amount of information that was conveyed to the participant.

The length of the expert's explanation was also calculated (in terms of words), including only utterances that contributed to any of the facts that were identified, and removing non-relevant words. Below this information is added to Table 1.

Participant ID	REALM raw score	REALM bracket level	Session duration (seconds)	Explanation length (words)	Session Fact Reference Score
102	17	1	282	572	0.732
103	64	4	336	657	0.750
108	62	4	496	697	0.697
109	47	3	368	488	0.669
111	1	1	231	524	0.739

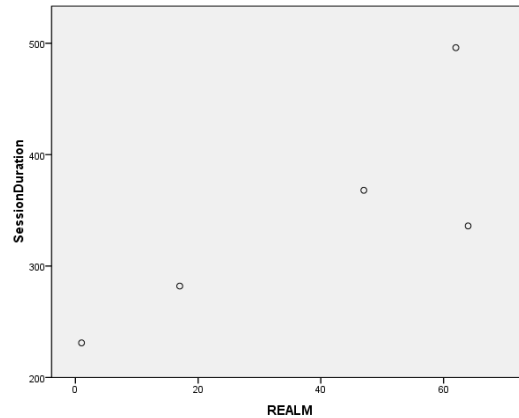
Table 2 Summary of Fact reference scores

Since each participant's health literacy level was known by way of their REALM scores, the test for Pearson's Product Moment Correlation Coefficient was performed between the REALM score and the mean fact reference scores. No significant correlation was found,  $r(4) = -0.315$ ,  $p > 0.05$ .

### 3.2.5. Analysis of explanation length

The next step was to test if the length of the expert explanation was dependent on the participant's health literacy level. To do this, all conversational fillers ("um", "er", "like" etc.), restarts, mistakes and casual conversation were removed from the conversation transcripts, and then the individual sentences and otherwise clearly separated phrases were listed.

I then performed a bivariate correlation test between the participant's REALM score and the duration of the session. A correlation approaching significance was found,  $r(4) = 0.812$ ,  $p < 0.1$ .



**Figure 4 Scatterplot for participant health literacy and session duration**

A similar correlation test between the participant's REALM score and the explanation length (see Table 2) was performed. No significant correlation was found.  $r(4) = 0.63$ ,  $p > 0.05$ .

Next, I mapped how each of these contributed to the 25 sentences that were identified on the source document. Contribution in this instance was measured in two ways. First, the number of words used to explain a given sentence (on the document) to a participant, indicate the amount of raw information present, and this is defined as *Words*. Secondly, the ratio between the number of words in a sentence on the document, and the number of words used to explain that sentence to the participant given an indication of the relative importance the expert assigned to that sentence, and this is defined as *Word Ratio*.

While this appears to be primitive, the belief was that this will give an indication of how much explanation effort went into each sentence and how this changed with the level of detail present in the source document.

This information is summarized below.

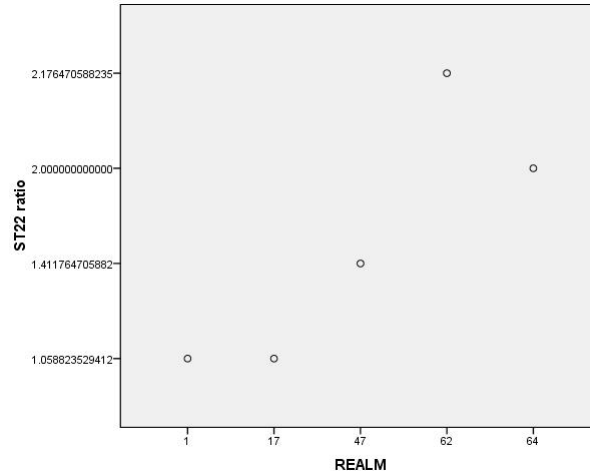
			Participant										
			102		103		108		109		111		
			Words	Word ratio	Words	Word ratio	Words	Word ratio	Words	Word ratio	Words	Word ratio	
Summary			1	41	41.00	44	44.00	32	32.00	27	27.00	40	40.00
Sentence	1	33	37	1.12	33	1.00	42	1.27	35	1.06	44	1.33	
	2	27	13.6	0.50	27.3	1.01	0	0.00	0	0.00	0	0.00	
	3	21	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	
	4	24	13.6	0.57	14.3	0.60	38	1.58	79	3.29	40	1.67	
	5	30	13.6	0.45	14.3	0.48	20	0.67	0	0.00	22	0.73	
	6	9	20	2.22	12	1.33	14	1.56	15	1.67	16	1.78	
	7	34	34	1.00	35	1.03	43	1.26	18	0.53	30	0.88	
	8	19	18	0.95	19	1.00	23	1.21	15	0.79	18	0.95	
	9	35	56	1.60	76	2.17	176	5.03	56	1.60	39	1.11	
	10	9	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	
	11	16	25	1.56	23	1.44	17	1.06	23	1.44	18	1.13	
	12	26	32	1.23	45	1.73	30	1.15	16	0.62	37	1.42	
	13	25	31	1.24	38	1.52	37	1.48	26	1.04	31	1.24	
	14	33	36	1.09	38	1.15	0	0.00	0	0.00	0	0.00	
	15	9	12	1.33	18	2.00	11	1.22	15.6	1.73	12	1.33	
	16	13	26	2.00	24	1.85	11	0.85	15.6	1.20	30	2.31	
	17	22	30	1.36	27	1.23	23	1.05	15.6	0.71	22	1.00	
	18	26	32	1.23	38	1.46	51	1.96	24	0.92	30	1.15	
	19	19	14	0.74	22.5	1.18	22.5	1.18	20.5	1.08	21	1.11	
	20	20	21	1.05	22.5	1.13	39.5	1.98	20.5	1.03	19	0.95	
	21	25	17	0.68	17	0.68	20	0.80	15	0.60	19	0.76	
	22	17	18	1.06	34	2.00	37	2.18	24	1.41	18	1.06	
	23	25	38	1.52	23.6	0.94	18	0.72	19	0.76	28	1.12	
	24	20	15	0.75	27.6	1.38	9	0.45	13	0.65	14	0.70	
	25	18	19	1.06	27.6	1.53	15	0.83	22	1.22	16	0.89	
REALM score			17		64		62		47		1		
REALM bracket			1		4		4		3		1		

**Table 3 Information Content by participant**

A bivariate correlation test was then performed, looking at the correlation between the word count of each sentence, and the REALM score, and similarly with the word ratio in each sentence and the REALM score.

The word ratio of only one sentence (sentence number 22 - “*My sample will stay in the Repository indefinitely and I will not be able to withdraw it.*”) correlated significantly with the participant’s REALM score. ( $r = 0.92$ ,  $p < 0.05$ ). It should of course be pointed out that only a single variable out of 25 correlated significantly with the REALM score, but observations with Dr. Paasche-Orlow (see 1.1) suggest that this was among the most important sentences in the entire document, so perhaps this is not entirely by chance either.





**Figure 5 Scatterplot for REALM score vs ST22 word ratio**

The final conclusion from this analysis was that there was no clear relationship between the relative importance the expert gave to any specific parts of the document and the participant's health literacy level.

### 3.2.6. Analysis of readability of sentences

In order to see if the readability level of each sentence had any impact on whether it was going to be mentioned or not, I calculated the Flesch-Kincaid readability score for each sentence, and then correlated this with the REALM scores.

The fact reference scores (see Table 2) for each session were then divided by the readability score for the corresponding sentence, and these values were correlated with the participant's REALM score.

No significant correlations were found between the mean fact reference scores (normalized by the readability level) for each session and the participant's REALM score.

Finally, to see if the readability level of the sentences on the document and the participants' REALM score were predictors of the explanation effort, a regression test was performed with the explanation word count as the dependent variable, and the Flesch-Kincaid Grade Level, Flesch-Kincaid Reading Ease score (where a higher score indicates easier reading) and participant REALM score as predictors.

The Flesch-Kincaid Grade Level of a sentence was found to be a significant predictor of the explanation word count,  $\beta = 3.776$ ,  $t(121) = 4.171$ ,  $p < 0.001$ . Further, the Flesch-Kincaid Reading Ease score was also a significant predictor of the explanation word count,  $\beta = 0.547$ ,  $t(121) = 3.869$ ,  $p < 0.001$ . The participant REALM score was not a significant predictor of the explanation word count,  $\beta = 0.063$ ,  $t(121) = 0.913$ ,  $p > 0.05$ .

The regression model accounted for a significant portion of variance in the dataset,  $R^2 = 0.138$ ,  $F(3, 121) = 6.352$ ,  $p < 0.001$ .

### 3.2.7. Summary of findings

The primary conclusion of this analysis is that the explanations followed the structure of the source document in a linear fashion.

With the small sample that was analyzed, however, it was not possible to explain how, given a participant's REALM score, to decide which facts to mention and which to avoid.

I also did not find any relationships between a participant's REALM score and the duration or length of the explanations that the expert provided. The Regression analysis showed that the readability of a sentence of the document was a reliable predictor of the effort the expert used to explain that sentence.

It was therefore decided that the best mimicry of explanation of a document should be one that linearly explains a document's structure. Later sections of this thesis discuss in greater detail our actual implementation.

## 3.3. Further observations of document explanation by experts

As referred to in section 1.1, Dr. Michael Paasche-Orlow, a noted national expert on readability of health documents, conducted a mock explanation of an Informed Consent while being videotaped in the Human-Computer Interaction lab at Northeastern University.

This session was not subject to detailed analysis, but several interesting observations were made. Firstly, Dr. Paasche-Orlow provided a lot of background information that was relevant to the form, but was not directly mentioned on the form itself. He also invited the participant to read the document together with him, and asked many questions from the participant, and provided clarifications when the participant's response seemed incorrect. The total explanation duration was approximately 25 minutes, and the content was essentially similar to that shown in Figure 1.

These observations were also taken as input in designing the final solution discussed here.

## 3.4. Providing domain knowledge

An important ability of an expert human tutor is the ability to provide context to a learning task, by way of background information, as observed in the previous section.

In particular, this thesis proposes that a simple way of providing the Conversational Agent with this ability would be to have a large knowledge base of background concepts that the Agent can elaborate upon. This requires that the Agent can estimate the relevance of the available concepts at each turn of its explanatory dialogue, and either guide the conversation to a concept tutorial, or offer it as a choice to the user.

### 3.4.1. Topic relevance ranking

In the initial part of this thesis, I also set out to develop and evaluate an algorithmic approach to rank the relevance of our background knowledge concepts at each point of the Agent’s dialogue.

Comparing the immediate content of the dialogue with the list of concept names was seen as a reasonable first attempt, and I used the JWordNetSim package [29], a port of WordNet::Similarity [48] for this purpose. JWordNetSim implements two algorithms, Jiang and Conrath’s algorithm (“JCn”) [33] and Lin’s algorithm [39], that both use the WordNet lexical database [22]. Although evidence suggests that LSA is superior to WordNet (for example in determining user likelihood to click a given link on a given information seeking task, see [35]), this was not evaluated due to a lack of known public domain LSA implementations.

The dialogue was broken down in to three levels, guided by logical hierarchies observed in documents:

- Sections are logical groupings within a document
- Sentences form Sections
- Terms are individual words that form Sentences, but do not include “noise” words or duplicates

Additionally, a background knowledge concept is defined as a “Topic”.

The following algorithm in pseudocode demonstrates the initialization:

```
terms <- extract-terms(Document)

for each Concept in background-concepts do

  for each Term in terms do
    similarity[Topic][Term] = calculate-similarity(Topic, Term)
  end

  for each Section in Document do
    similarity[Topic][Section] <-
      mean(similarity[Topic][Term2] | Term2 ∈ extract-terms(Section))

    for each Sentence in Section do
      similarity[Topic][Section] <-
        mean(similarity[Topic][Term3] | Term3 ∈ extract-terms(Sentence))
    end
  end
end
```

**Table 4 Topic relevance ranking algorithm - initialization**

Now, at each point of the dialogue, it is possible to rank the relevance of these Topics. The relevance score is defined such that whenever the score exceeds a certain threshold, it is reported as relevant. (note: a reliable threshold was not found).

The relevant Topic itself can be used by the Agent system in one of two different ways:

- Offer the Topic as a question that the user can ask the Agent (user initiated)
- Refer to the topic (“*Let me tell you about ... first*” – Agent initiated)

This however is outside the scope of this current study.

```
for current-section do
  terms <- extract-terms(current-section)
  topic <- argmax(mean(similarity[topic][term] | term ∈ terms)
end
```

**Table 5 Topic relevance ranking algorithm – ranking**

### 3.4.2. Empirical testing of Topic relevance ranking algorithm

In order to test if the Topic mappings produced algorithmically are in fact relevant, a simple experiment was conducted.

Three Research Informed Consent documents (see appendix A) were presented to an expert and the expert listed a number of background concepts as relevant to the documents. These concepts and a brief explanation text are present in appendix B. The list was then narrowed down to 12 items, and presented to the expert again the three documents, and the expert was asked to indicate the section of the document (if any) that each concept was most relevant to, by placing a sticker with the concept’s name on the document.

Next, Topic associations were generated to each section in three documents, using both the JCn and Lin algorithms and then the associations generated were compared against the expert’s associations. The JCn and Lin algorithms also produced a score indicating similarity (higher scores indicate greater similarity).

The associations produced by the expert, as well as the two algorithms are summarized below.

Section	Expert	Genetic				Expert	RED				Expert	NSF			
		Lin	Lin-score	JCn	JCn-score		Lin	Lin-score	JCn	JCn-score		Lin	Lin-score	JCn	JCn-score
1	Certificate of Confidentiality	1.94	Certificate of Confidentiality	1.48			Voluntariness	1.91	Voluntariness	12.00		Privacy	1.57	Certificate of Confidentiality	1.38
2	Randomization	1.34	Alternatives	1.35		Randomization	Certificate of Confidentiality	1.96	Certificate of Confidentiality	1.62		No benefits	1.92	No benefits	2.05
3	Certificate of Confidentiality	1.46	Randomization	12.00	Best care		Certificate of Confidentiality	1.89	Certificate of Confidentiality	12.00		Certificate of Confidentiality	1.56	No benefits	1.59
4	Certificate of Confidentiality	1.66	Randomization	1.37			Certificate of Confidentiality	1.74	Randomization	1.45		Certificate of Confidentiality	1.53	Certificate of Confidentiality	1.21
5	Risk	1.68	Certificate of Confidentiality	12.00	Risk		Certificate of Confidentiality	1.42	Certificate of Confidentiality	8.00	Risk	No benefits	1.56	Certificate of Confidentiality	12.00
6	No benefits	2.16	No benefits	12.00	No benefits		No benefits	2.40	No benefits	12.00	No benefits	No benefits	2.11	No benefits	12.00
7	Privacy, Confidentiality, Certificate of Confidentiality	1.52	Best care	3.75	Confidentiality		Certificate of Confidentiality	1.82	Randomization	1.38	Confidentiality	Certificate of Confidentiality	1.65	Certificate of Confidentiality	1.34
8	Voluntariness	2.42	No benefits	6.00	Alternatives		Alternatives	2.59	Alternatives	12.00	Alternatives	Best care	1.80	Alternatives	12.00
9	Treatment	2.02	Randomization	12.00	Treatment	Voluntariness, Withdrawal	No benefits	2.02	Randomization	12.00	Treatment	No benefits	2.02	Randomization	12.00
10	Withdrawal	1.44	Best care	1.36	Withdrawal		Certificate of Confidentiality	1.38	Privacy	6.00	Withdrawal	No benefits	1.95	Privacy	6.00
11	No benefits	1.48	No benefits	1.39			Certificate of Confidentiality	1.67	Certificate of Confidentiality	1.40		Certificate of Confidentiality	1.50	Certificate of Confidentiality	1.29
12	No benefits	1.39	Alternatives	1.38			Certificate of Confidentiality	1.70	Certificate of Confidentiality	1.38		Certificate of Confidentiality	1.45	Alternatives	1.29
13	Certificate of Confidentiality	1.94	Confidentiality	1.84			Certificate of Confidentiality	2.10	Certificate of Confidentiality	1.84	Lose eligibility	No benefits	1.99	Lose eligibility	12.00
14	Certificate of Confidentiality	1.65	Privacy	12.00			Certificate of Confidentiality	1.91	Certificate of Confidentiality	1.63		Certificate of Confidentiality	1.91	Confidentiality	1.83
15	No benefits	1.52	Alternatives	1.89	Lose eligibility	Best care	1.81	No benefits	12.00						
16	Certificate of Confidentiality	1.58	Alternatives	1.92											

**Table 6 Topic associations by Expert and the Lin and JCn algorithms**

Further, I compared the ranking that the two algorithmic approaches gave to the association made by the expert, theorizing that if the algorithmic associations were ranked sufficiently high, it would still be useful. The results are summarized below.

Form	Section	Expert assigned Topic	Lin's Algorithm		JCn Algorithm	
			Rank of Expert's Topic	Rank difference	Rank of Expert's Topic	Rank Difference
Genetic	5	Risk	11	10	12	11
Genetic	6	No benefits	1	0	1	0
Genetic	7	Privacy	5	4	9	8
Genetic	7	Confidentiality	10	9	3	2
Genetic	7	Certificate of Confidentiality	1	0	4	3
Genetic	8	Voluntariness	6	5	4	3
Genetic	9	Treatment	2	1	9	8
Genetic	10	Withdrawal	7	6	8	7
RED	3	Randomization	4	3	5	4
RED	3	Best care	3	2	7	6
RED	5	Risk	12	11	12	11
RED	6	No benefits	1	0	1	0
RED	7	Confidentiality	10	9	10	9
RED	8	Alternatives	1	0	1	0
RED	9	Treatment	11	10	9	8
RED	10	Voluntariness	5	4	6	5
RED	10	Withdrawal	7	6	8	7
RED	15	Lose eligibility	8	7	10	9
NSF	5	Risk	12	11	12	11
NSF	6	No benefits	1	0	1	0
NSF	7	Confidentiality	9	8	9	8
NSF	8	Alternatives	2	1	1	0
NSF	9	Treatment	11	10	9	8
NSF	10	Voluntariness	7	6	4	3
NSF	10	Withdrawal	10	9	9	8
NSF	13	Lose eligibility	6	5	1	0
Average rank difference			5.26923077		5.34615385	

**Table 7 Expert's Topic associations and corresponding ranking by Lin and JCn algorithms**

Out of 45 possible Topic mappings (in 16, 15 and 14 sections in the Genetic, RED and NSF documents, respectively), the Expert and the Lin algorithm agreed only in 5 instances (11%), and the Expert and the JCn algorithm agreed only in 6 instances (13%), both being slightly above chance levels (8.33%). Further, the Lin and JCn algorithms showed agreement in 22 instances (48%).

The Lin and JCn algorithms were in agreement for the Certificate of Confidentiality topic the most number of times (13 instances, 20% of total), and the “No benefits” topic the next (5 instances, 11% of total).

The mean Lin-score for items in agreement between the Expert and the Lin algorithm was 2.16, compared to 1.73 for the items not in agreement. Similarly, the mean JCn-score for items in agreement between the Expert and the JCn algorithm was 12, compared to 4.96 for the items not in agreement.

Based on Table 7, the relative ranking that the two algorithms gave to the Topic associations made by the expert were also evaluated. The expert’s associations were ranked, on average, in the top six of the twelve possible associations.

These results are therefore not very promising, and this algorithm could not be refined to a level sufficient enough for use in the Agent system. Some experts suggest that approaches based on the WordNet lexical database, as is the case with the Lin and JCn algorithms suffer from excessive noise [Hafner, personal communication], and this could have led to the results seen here.

Three approaches for future work in this area are suggested: first, instead of calculating the word similarity between the text of a document section and the Topic’s title, the Topic’s text itself could be used. Secondly, and possibly complimenting the first, Latent Semantic Analysis (LSA) may provide a better measure of semantic similarity, as demonstrated in the AutoTutor system [27]. The third possibility is the use of Machine Learning techniques on large corpora of Research Informed Consent documents that are already associated with Topics, and use these associations to predict associations for new documents. Given that a lot of Research Informed Consent documents are very similar in structure to each other, this approach seems the most promising.

### 3.5. Conversational Agents explaining Research Informed Consent Documents – a first attempt

During the first phase of this thesis the applicability of extractive text summarization for the purpose of explaining Research Informed Consent documents was explored (see also 2.3.2).

The first document in Appendix A was used as input for this test. This document had several sections of text, and the Agent system attempted to provide a summary of each section by extracting content accounting for at most 40% of the content (in terms of word count) of the section. However, 12 of the 16 sections consisted of a single sentence, and so the extraction would always return the full-text of the section (i.e.: the single sentence). Thus, this approach was abandoned.

Abstractive summarization was also considered, but on one hand, no public-domain abstractive summarization tools were found, and on the other hand, because of the arguments presented in section 2.3.2 this was not considered to be an essential element in this study.

Abstractive summarization does however offer one very interesting manipulation that could be done: if a clear relationship can be found between a participant's health literacy level (again, using the REALM measure) and the ideal level of detail level for the participant, the summarization algorithm could be parameterized based on the required detail level, offering fine-grained freedom in tailoring the detail level.

## 4. Computational Model for Research Informed Consent document explanation

### 4.1. Background of Research Informed Consent Document explanation task

As discussed in previous sections, I propose that a Conversational Agent that can explain a health document to a participant can use a variety of strategies to help the participant understand the document. Key strategies include:

- Make the system natural to interact with, by simulating a conversation
- Displaying the document so that the participant can orient themselves within the document
- Pointing out the section that the current dialogue relates to
- Where possible, allow the user to control the level of detail
- Explain the document in a logical order
- Provide any background information that may help the participant understand part of the document better
- Ask interactive questions to establish the participant's understanding, and provide corrections if needed (as explained in 1.4)

### 4.2. Describing the Document Explanation task

The task of Document Explanation by an Agent is as follows:

<pre>For a given document   display a representation of the document to help orient the user   create an index of high-level sections within the document   for each section     tell the user what the current section is     rank the relevance of background knowledge concepts     for all concepts that have not been introduced to the conversation yet       attempt to introduce it     provide the information in the current section, either as the content     itself, or as a summary       if summary provided, allow user to request additional detail     attempt to check the user's understanding of the current section       if understanding is not adequate, review the section a second time</pre>
--

Table 8 Overview of Document Explanation task

### 4.3. Implementation of the Document Explanation system

#### 4.3.1. Producing Informed Consent documents

I reasoned previously that we wanted to add structure to an Informed Consent document, and that structured content would become input to a Conversational Agent system. It is therefore a logical extension to design a system that can produce this structured content, and I decided to model it based on our observations of existing authoring tools for study



protocols and Informed Consent documents. This tool is still not functional, but some of its envisioned features are:

- Users can create sections in a given document. Each section becomes a titled paragraph in the Informed Consent document.
- Each section can be annotated with concepts from a knowledge base that the Agent can use for providing background knowledge
- Each section can be further annotated with simple comprehension questions that the Agent can use to establish the user's understanding, and if needed, provide additional information or repetitions.
- With each section, the author is currently requested to provide a simpler explanation. This can be extended in the future so that the system provides a simple explanation that the author can review and, if necessary, modify.
- With the main text being written by the author, readability statistics are calculated so (per-section as well as per-document) that the author can verify that readability guidelines are being met.
- While I implemented a simple document rendering algorithm using Java 2D in the first phase of this thesis, this functionality was removed for the experimental setup given the complexity of document layout. In the experiment, word processing tools were used to produce bitmap outputs that the Agent system displays, and then produced the structured document that the Agent system uses to generate dialogue.

#### 4.3.2. Informed Consent documents as input

Documents have an inherent structure to them, making explanation of a document easier than other tasks, like explaining the layout of a living room [50]. However, inferring the structure of a document is itself not trivial, so a structured document was instead chosen as input.

In doing this, I started developing a system that authors of Informed Consent documents can use to produce this structured content (see section 4.1 for details). XML was chosen for this purpose (see Appendix I for the Document Type Definition).

This file describes each of the sections for the document, annotated with the background knowledge concepts that relate to each section, as well as any comprehension check questions that could be asked.

The document also has an index of references to the document at each section and each utterance, so that the Agent could point at the document.

#### 4.3.3. A non-scripted Conversational Agent

The dialogue engine developed by Bickmore et al., called RAGServer, and its relevant client components, were used for the development of this system. RAGServer uses a visual dialogue scripting environment called ScriptBuilder. ScriptBuilder converts conversational scripts written in a custom language to Java source files that the RAGServer system uses for generating dialogue and responding to user input.

The RAGServer system in turn is a complete Conversational Agent management system that controls dialogue flow, handling of volatile and persistent information and controlling of the actual client system that hosts and displays the Agent.

Dialogue in this system is scripted in files that model a Hierarchical Transition Network (HTN). While the scripting process allows a great deal of flexibility of actions that the Agent can carry out, it also means that a significant authoring effort is needed for all new content that the Agent can converse about.

With this in mind, I set out to create a Conversational Agent that did not rely on pre-scripted interactions, but could use a document as a guide for creating a dialogue.

#### 4.3.4. Document explanation strategies

Extensive modifications were made to Bickmore et al.'s system to implement what I call metastrategies: high-level strategies for various document explanation strategies.

Since our observation study revealed that the expert proceeded through the document in a linear manner, I too chose to use such a linear metastrategy in this study. This linear metastrategy realizes two slightly different explanation strategies that each proceeds in a linear manner starting from the first section to the last.

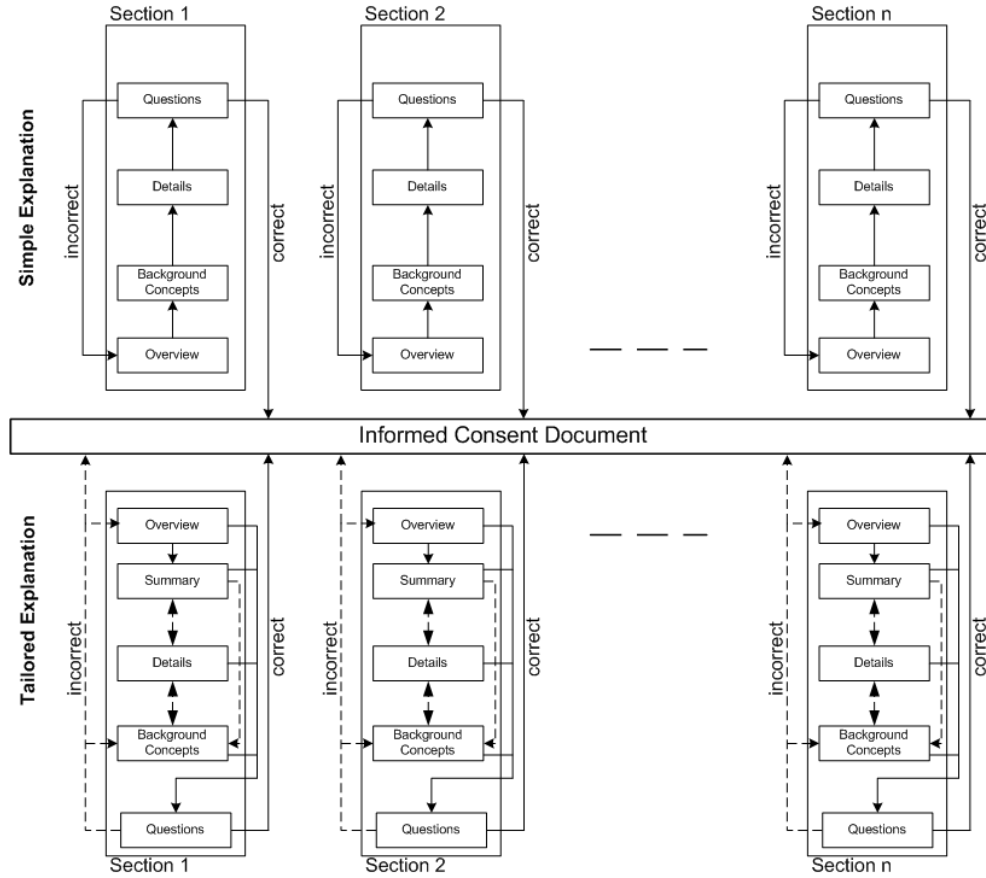
Based on prior observations, I further incorporated three features within these explanation strategies:

- base the explanation on simulating a collaborative reading of a shared document
- provide background information at the beginning of a logical section
- pose simple questions to the user at the end of logical sections, and provide correction or review of content if necessary

Incorporating these features, two possible document explanation strategies were produced:

- **Verbose (Simple) explanation** traverses all content available from within each section. Thus it offers a brief overview and then any related background concepts before offering the primary content. This is followed by a comprehension check, and if the user provides an incorrect response, a second attempt is made. If the comprehension check fails again, a correction is offered and the explanation moves to the next available content.
- **Tailored explanation** skims each section, allowing the user to control the detail level and background concepts. Background concepts are offered as questions that the user can ask the Agent, thus adding a more conversational aspect to the dialogue. If a comprehension check fails, the user has the option of reviewing the content a second time, reviewing the relevant background concepts only, or moving on to the next content.

The diagram below shows the flow of the dialogue, and sample transcripts are shown in 4.4.



**Figure 6 Comparison of document explanation strategies.**

Note that dashed lines indicate optional paths of the dialogue, and solid lines indicate mandatory paths.

#### 4.3.5. Enabling traversal of documents

In addition to support for a linear metastrategy, I introduced several actions that describe what users can perform at any given moment.

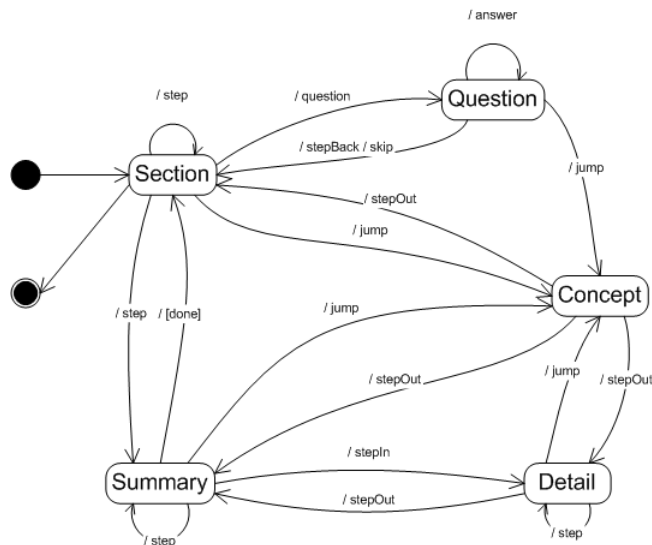
I first introduce several levels of the conversation:

- **Section:** The Agent is providing a very brief overview of the section.
- **Summary:** Sub-level of Section, where the Agent is providing a summarized explanation of a section.
- **Detail:** Sub-level of Section, where the Agent essentially reads the section along with the user.
- **Concept:** The Agent is providing some background information for a given section.
- **Question:** The Agent is asking a comprehension check question from the user, or responding to the user's answer.

Below is a list of the actions that were implemented.

- **step** – move to the next available unit of discourse, which could be the next utterance of the current section, any comprehension checks for the current section, or the next section.
- **stepIn** – elaborate on the current unit of discourse. If we are currently at a section summary, this produces the section detail.
- **stepOut** – move to a higher unit of discourse. If we are currently at a Concept level, move back to the Section.
- **stepBack** – an implicit action, when the conversation moves back to the beginning of the current section (because the user wanted to, or because the Agent determined that the section should be repeated).
- **jump** – save the current context within the document, and move to another discourse unit (e.g.: Concept), with the intention of returning and resuming later. This is parameterized with the Concept to jump to.
- **answer** – given a question, conveys the user’s response back to the system. This is parameterized by the answer given.

A state graph summarizes the levels of the conversation, and the actions that trigger these transitions.



**Figure 7 Statechart for Document Explanation**

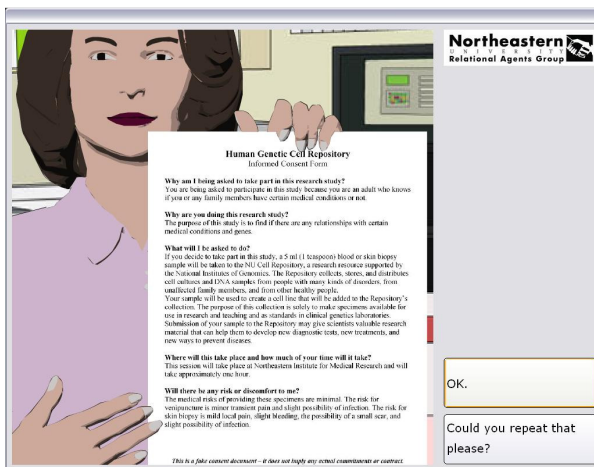
#### 4.4. Example conversation transcript

Agent: Just let me know when you're ready.	<introduction>
User : OK.	
Agent: This Informed Consent form is about a study that we are inviting you for.	
User : OK.	
Agent: Let's take a few minutes, and go over it together.	
User : OK.	
Agent: This explains why we're inviting you for this study.	<section>
User : OK.	
Agent: You are being asked to participate in this study because you are an adult who knows if you or any family members have certain medical conditions or not.	
User : OK.	
...	
Agent: The risk for skin biopsy is mild local pain, slight bleeding, the possibility of a small scar, and slight possibility of infection.	<sentence>
User : OK.	
Agent: Here are some possible benefits to you.	
User : OK.	
Agent: But first, let me tell you a little bit about benefits in research studies.	<concept>
User : OK.	
Agent: Some people are comfortable being in research studies even though they will not benefit directly, because they want to help others in the future who may benefit.	
User : OK.	
Agent: This study is like that, and it will not help you directly, but it may help scientists find out ways to diagnose and treat certain medical conditions in the future.	
User : OK.	
Agent: Okay, let's look at the form again.	<return to form>
User : OK.	
Agent: There will be no direct benefit or payment to you for participating, but your sample may benefit the community at large or some particular group.	
User : OK.	
Agent: Before we move on, I would like to ask you a quick question to make sure I am being clear.	<section end>
User : OK.	
Agent: Can you tell me how you will benefit from this study?.	<question>
User : 2.	<answer>
Agent: That is correct, well done!.	<feedback>
User : OK.	
...	
Agent: Can you tell me how the Certificate of Confidentiality helps you?.	<question>
User : 2.	<answer>
Agent: I think we should review what we just talked about.	<stepback>
User : OK.	
Agent: This part is about how we treat information about you.	

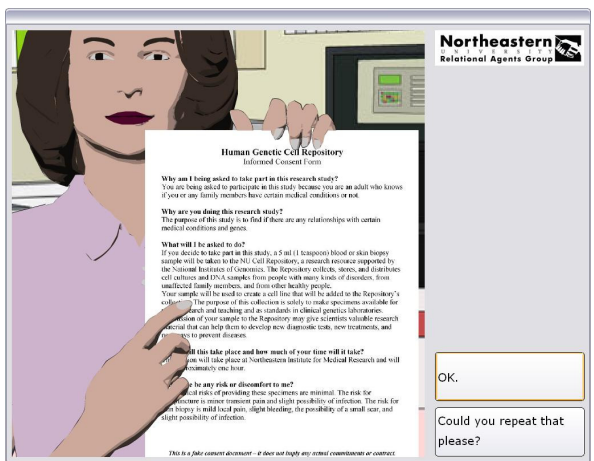
Table 9 Sample conversation - Verbose mode

In the Verbose mode, the Agent proceeds through the document, and chooses when to talk about a background concept, and then returns to the document. While asking questions, if the response is correct, it is acknowledged. If not, the agent says “*I think we should review what we just talked about*” and reviews the section.

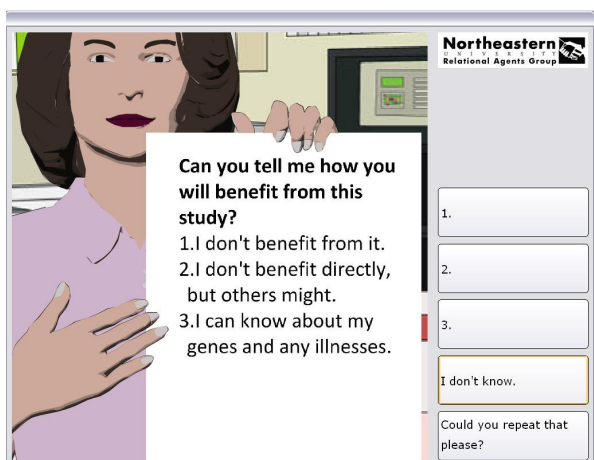




Reference to a section with a flat hand for orienting the user.



Reference to a sentence with the finger.

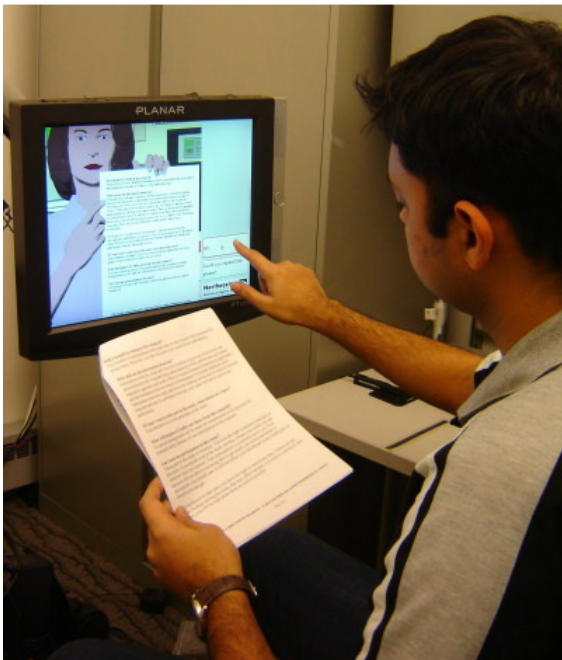


Displaying comprehension questions on screen. The Agent asks the question verbally, but the responses are only shown on the screen. The user selects a button (1-3) or can say “I don’t know”.

Table 11 Various document display strategies



Test participant interacting with the Agent. The Agent is deployed on a computer with a touchscreen display.



Close up of participant interacting with the Agent.

**Table 12 Participant interacting with Agent**



#### 4.5. Parallels with Intelligent Tutoring Systems

While this implementation is more of a conversational agent, there are clear similarities with Intelligent Tutoring Systems. Below is a summary of these similarities, compared in particular with the AutoTutor system (see section 2.2.1 for details).

	AutoTutor	This work
Dialogue content	Based on Curriculum Script	Dynamically generated from structured Informed Consent document.
Knowledge representation	Questions and Topics in Curriculum Script	Concepts of background knowledge, and comprehension questions for Sections
Handling knowledge units	Latent Semantic Analysis (LSA) used to determine coverage	Word Similarity algorithm to associate Concepts with current text (using WordNet [22], and JWordNetSim [29] and the JCN and Lin algorithms [33], [39]). Topics were manually associated for our empirical study.
Teaching goals	Set of Topics, each with a set of Expectations	Set of Document Sections, some of which have a set of Concepts and Comprehension questions
Authoring content	Lesson Planner and authoring tools	Informed Consent document authoring tool
Embedded multimedia content	Graphics associated with topics	On-screen representation of Informed Consent document pages and question pages.
Measure learner comprehension	LSA used to measure overlap between expectation and learner response	Question with multiple responses, and learner response as a choice.
Feedback	Varied (see 2.2.1 for details).	Spoken acknowledgement of correct or incorrect responses.
Correction mechanisms	Varied (see 2.2.1 for details).	Repetition of previous content.
Dialogue Management	Dialogue Advancer Network	Dynamic extensions to RAGServer system

**Table 13 Comparison of AutoTutor and this work**

In essence, this work is a step towards augmenting the RAGServer dialogue engine with some simple implementations of the features commonly seen in modern ITS systems.

## 4.6. Realizing the Document Explanation task

At the beginning of this chapter, a list of strategies was presented for the document explanation task. A final review is presented here explaining how those strategies will be realized.

### 4.6.1. Facilitate a natural method of interaction

Humans engage in inter-personal conversations on daily basis, and therefore the method of interacting with a Conversational Agent is expected to feel quite natural. The Agent uses synthesized speech to “talk” to the participant, and the participant’s responses appear as buttons (see Table 11). A touchscreen is used so that participants can simply tap their finger on the button that represents what they want to say (see Table 12).

Some features were however removed to minimize their effect on the Satisfaction that participants were asked to report. For example, the Agent does not introduce herself or engage in any introductory conversation, uses the same neutral facial expression throughout, and does not address the participant by their name, all of which have been shown to also increase the participants’ perception of the quality of the interaction [10].

### 4.6.2. Displaying the document

The input from the machine-readable files can be rendered in to a bitmap that the Agent system is able to load and display as previously done by Bickmore et al. [6, 7, 8].

### 4.6.3. Pointing out the section that the current dialogue relates to

Following previous experiences by Bickmore et al [6, 8, 9] the Agent will use two kinds of hand gestures to refer to the document. When the Agent refers to a Section within a document, the Agent will use the whole hand to gesture towards the section, but when the Agent is reading a sentence from the document, the Agent will point at the sentence with the finger.

### 4.6.4. Where possible, allow the user to control the level of detail

In section 2.5 of this thesis, I proposed that it might be useful to change the detail level of the explanation as a simple way of tailoring the content for a user. I also proposed that abstract summarization could be used in the future to provide fine-grained control over the detail level presented.

For the purpose of this study, this was simplified, where instead the machine-readable input has an optional summary for the text of most sections. Thus the summary provides an essentially “low-detail” condition, and the full text of the section provides a “high-detail” condition. Together with the StepIn and StepOut pair of actions defined (see 4.3.5), this allows the user to control the amount of detail.

#### 4.6.5. Explain the document in a logical order

Following the observations made earlier, the document will be explained by following its structure in a linear manner.

#### 4.6.6. Provide any background information that may help the participant

In section 3.4 an approach was presented for allowing the Agent to judge the relevance of background knowledge concepts. For the purposes of this study though, I decided to instead base the presentation of background concepts on those suggested by the expert. The expert suggestions were further modified slightly, so that out of the 12 concepts identified, 4 concepts would be referred to in each of the three documents included in this study.

#### 4.6.7. Ask interactive questions and provide necessary corrections

Interactive questions were introduced earlier as a possible improvement to the Informed Consent process. However, given that the Agent system deals with user contributions to the dialogue as buttons being clicked, such questions need to be closed-ended (see Table 11 for an example) . Therefore, the questions associated with a section were always multiple-choice, and were rendered on to bitmaps and displayed on screen, with corresponding responses available as buttons.

If the participant's response is incorrect, the Agent could offer to review the current section, or even possibly force it. It was determined that allowing the participant to make two attempts is sufficient, and if both attempts are unsuccessful, the Agent will provide a correction and move on to the next section.

## 5. Evaluation of document explanation system

### 5.1. Background

In this section I discuss an empirical evaluation of the Agent system for explaining Research Informed Consent Documents.

The objectives here are two-fold: primarily I am interested in seeing if a Conversational Agent could be developed that would not rely on pre-scripted dialogue, and would instead explain a document based on an annotated machine-readable representation of the document. Further, two different possible explanation strategies were introduced in section 4.3.4, and it would be very useful to see if these strategies would have different levels of effectiveness.

Secondly, I am interested in studying if the mere presence of the Agent improves the quality of the Informed Consent process. Conversely, current research suggests many approaches for improving the Informed Consent process, and I am interested in seeing if any of those can be incorporated in to the Agent, and study if that leads to improvements to the process again. I therefore propose that a dialogue where the user has greater control over the level of detail in the explanation would be an interesting comparison against a control condition of a simple, yet rigid and verbose dialogue.

### 5.2. Hypotheses

As expressed above, the primary purpose of this thesis is to study if a Conversational Agent that does not rely on pre-scripted dialogue is effective at helping a participant understand a Research Informed Consent Document better, and if such an Agent would make the participants feel more satisfied with the process.

The secondary purpose is to examine if the above Conversational Agent can tailor the detail level of its explanation, and if doing so would mark the participants feel more satisfied with the process.

With this in mind, the study hypotheses are framed thus:

H<sub>1a</sub>: When participants have a Conversational Agent explain a document to them, they will tend to have higher levels of comprehension than when they read a document by themselves.

H<sub>2a</sub>: When participants have a Conversational Agent explain a document to them, they will tend to feel more satisfied with the Informed Consent process than when they read a document by themselves.

H<sub>3a</sub>: When participants have a Conversational Agent tailor the detail level of the explanation of a document, they will feel more satisfied than when the Conversational Agent that presents a verbose fixed dialogue.

### 5.3. Experimental design

#### 5.3.1. Methods

A conversation agent system was developed to test these approaches (tailored explanation vs. verbose explanation) and I compared these against the baseline case of asking the participant to read the document themselves. Section 4.3.3 provides more detail on the implementation of this system.

20 volunteer participants were asked to participate in an experiment to test these hypotheses. An amendment was approved by Northeastern University's IRB to an already approved protocol, allowing us to carry out this study (IRB# 07-02-11).

Each participant was asked to complete three sessions, with random orderings of content (i.e.: document) and mode (self-read, Verbose Agent or Tailored Agent). Comprehension and Satisfaction measures were administered after each session.

#### 5.3.2. Materials

Three Informed Consent documents were used as source material in this study. The participants were asked to read one by themselves, while the Conversational Agent explained the other two; one with a simple but rigid and verbose explanation, (VERBOSE) and one with a tailored explanation (TAILORED). The ordering of these conditions, as well as the assignment of documents to conditions, were counterbalanced and randomized.

The three documents had some core concepts common to all of them (e.g.: voluntariness), but each document also had 4 more concepts that are significantly unique to them, and the Conversational Agent attempted to provide some background information on these concepts while it was explaining the document. The choice of these background concepts was based on the expert associations reported in 3.3, but was then re-assigned so that for the sake of this study, each document would have an equal number of concepts associated with it.

Each document was followed by a verbal comprehension test, which was largely teach-back in nature. Each comprehension test consisted of 3 general questions common to all documents (e.g.: "*What is the primary purpose of the study?*") and 3 questions specific to the document (see appendix E) and one general feedback question about the interaction.

The Conversational Agent system utilized a touch-screen to display the animated Agent and allow the participants to respond by touching buttons on the screen representing various responses. The Agent used synthesized speech to talk to the participant, and also displayed a representation of the document and referred to it by using pointing gestures. While a participant was interacting with the Agent, the researcher left the room.

### 5.3.3. Research Design

I utilized a 3-condition counterbalanced within-subjects design, to test the two different strategies for document explanation by a Conversational Agent against a participant reading it themselves, for a 3 (document) x 3 (condition) nested study design.

#### **Independent variables**

- Condition (self-read, Verbose explanation or Tailored explanation)

#### **Dependent variables**

- Comprehension (measured by a questionnaire based on the BICEP instrument [19] – see 5.4.1 for details)
- Participant's attitude towards the agent (in the Verbose or Tailored explanation conditions, measured with a 6-item self-report measure – see 5.4.2 for details)

#### **Covariates**

- Participant's general health literacy level (measured by the REALM instrument [53] – see 5.4.3 for details)
- Participant's Need for Cognition level
- Other socio-demographics

### 5.3.4. Explanation strategies

The two agent conditions will involve two different approaches for explanation of Informed Consent documents.

1. **Tailored agent-explanation:** A Conversational Agent will explain an Informed Consent document. The Agent will start with a low level of detail, and participants can ask for additional information relating to background knowledge concepts. Throughout each section, the user will have the ability to ask questions from the Agent pertaining to background concepts that are relevant to that section. At the end of some sections, the Agent will administer some questions to the user. If the user's response to this is incorrect, the Agent will suggest that they review the background concepts and the section one more time before moving on, but will allow the user to make the choice.
2. **Verbose agent-explanation:** A Conversational Agent will explain an Informed Consent document. The explanation of the document will be fixed, but at the beginning of each section of the document, the agent will explain any background concepts relevant to that section. At the end of some sections, the Agent will administer some questions to the user. If the user's response to this is incorrect, the Agent will attempt to explain the background concepts and the section one more time before moving on.

## 5.4. Measures

### 5.4.1. Primary outcome measure - Comprehension

Comprehension is proposed as the primary outcome measure, and will be used to test for hypothesis  $H_{1a}$ . The comprehension tests used will be based on a comprehension questionnaire that was used earlier by Bickmore et al. [7, 8], itself based on the Brief Informed Consent Evaluation Protocol (BICEP) instrument, and was modified slightly. Since three different Research Informed Consent Documents were to be used in the study, a comprehension questionnaire was created for each document.

Each questionnaire consisted of 11 questions; the first four are questions about the participant's perceptions of the quality of the Informed Consent process. Questions 5, 6 and 9 are general questions about benefits, risks and the primary purpose of the study, respectively (the respective answers however, are specific to each document). Question 11 is an open-ended prompt for the participant's impressions, and questions 7, 8 and 10 are specific to the document. In order to reduce carry-over effects, a 60-second cognitive task was administered before the start of sessions 2 and 3, where participants were asked to name as many animals, and vegetables as possible, respectively.

The comprehension tests were verbally administered in a “closed-book” manner.

### 5.4.2. Secondary outcome measure – Satisfaction

In order to test for hypothesis  $H_{2a}$ , participant satisfaction in the Informed Consent process was measured in two ways. The Comprehension questionnaire itself had a question “*How satisfied were you with the Informed Consent process?*” with a 1-7 rating scale, that was administered verbally.

Secondly, for interactions that involve the Agent, a separate 6-item survey was administered on paper (see “Instructor Evaluation Questionnaire” on Appendix H).

### 5.4.3. Participant's Health Literacy

Participant Health Literacy was expected to be a covariate, and therefore was measured. The Rapid Estimation of Adult Literacy in Medicine (REALM) instrument [53] has been used in previous studies by Bickmore et al., and will be used in this study as well.

REALM is a list of 66 medical words that the participants read aloud, and scoring is made based on correct pronunciation. The raw REALM score, in the range of 0 – 66 is then translated in to bracket levels as indicated below.

Raw REALM score range	REALM bracket level	REALM bracket
0-18	1	3 <sup>rd</sup> grade and below
19-44	2	4 <sup>th</sup> to 6 <sup>th</sup> grade
45-60	3	7 <sup>th</sup> to 8 <sup>th</sup> grade
61-66	4	High school

Table 14 REALM scores and meaning

#### 5.4.4. Need for Cognition

Comprehension relies on many things, with the cognitive effort of the learner being a key predictor of learning. The 18-item reduced Need for Cognition instrument [13] measures a participant's likelihood to engage in and enjoy effortful cognitive activities.

#### 5.4.5. Participant Demographics

Basic participant demographics, including gender, age, educational levels and ethnicity were also collected.

Since the study deals with the Informed Consent process, it was theorized that participants' prior research study participation could have an effect on their perceptions, and so additional questions were added, asking about their research study participation in the past.

### 5.5. Results

#### 5.5.1. Participants

20 volunteer participants were recruited for the experiment, and compensated for their time with \$15. With the exception of one participant, who dropped out after two sessions, all participants completed all sessions. All participants were debriefed at the end of the experiment.

Some of the participants were recruited by posting flyers at Marcus Garvey Apartments, at 44 John Eliot Sq, Roxbury, MA 02119, and their sessions were conducted at the same location. The other participants were recruited via Craigslist Boston, and the study sessions conducted at the Human-Computer Interaction Laboratory at Northeastern University. I conducted all the sessions. The participants' demographic information is presented below.

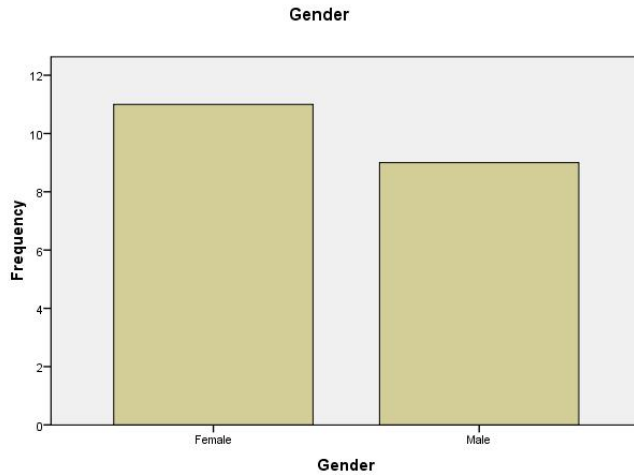
All study protocols and activities were part of a protocol approved by the Institutional Review Board of Northeastern University (IRB# 07-02-11).

A few observations stand out: Health Literacy, measured by the REALM instrument [18] shows a strong ceiling effect, and the majority of participants are regular computer users. Three participants (15%) reported never having participated in a research study, and 16 participants (80%) reported having participated in at least one research study in 2009.



### Participant demographics

Participants were asked to report their gender, date of birth, ethnicity and marital status. Age was calculated from this information.



The distribution of participants by Gender was roughly equal.  
( $\chi^2(1) = 0.2, p > 0.05$ )

Figure 8 Participants by gender

Participants' age was in the range 22-85 years, with the mean being 46.3 years.

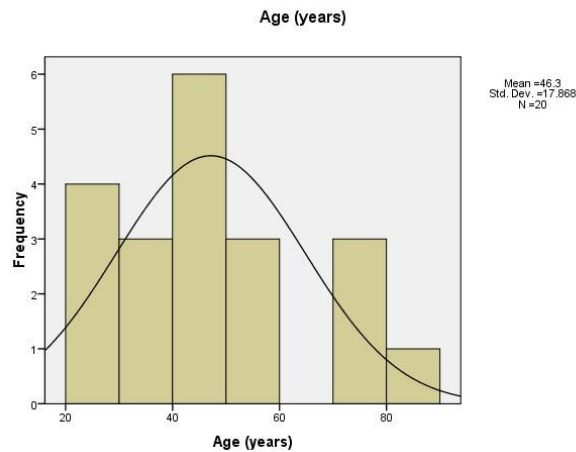
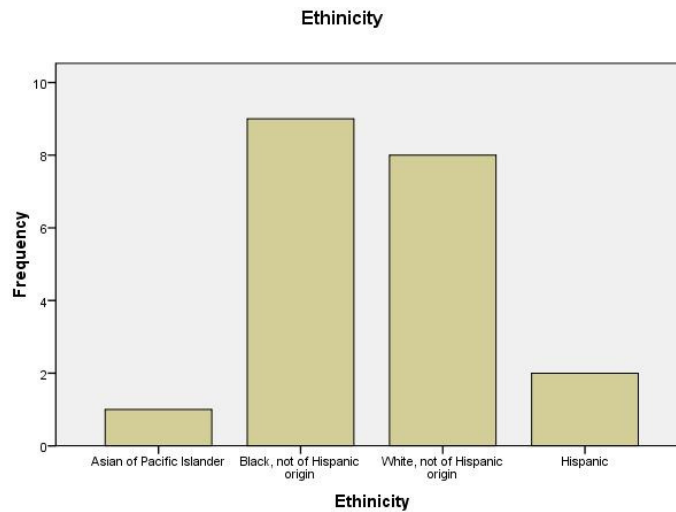


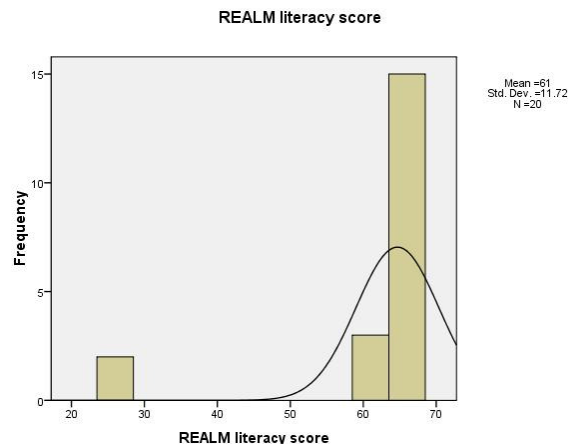
Figure 9 Age distribution



**Figure 10 Participant ethnicity**

### **Participant's health literacy and health background**

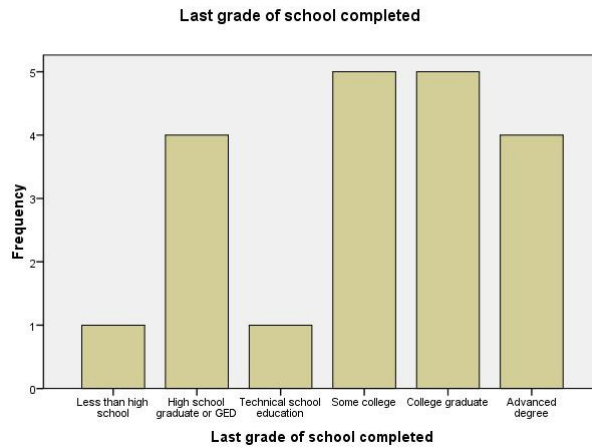
Some questions were asked to gauge participants' health literacy, background and habits. Height and weight were gathered, and converted to BMI values, and subsequently categorized according to NHLBI guidelines [41]. Health literacy was assessed using the REALM instrument [18] (see section 5.4.3 for more details).



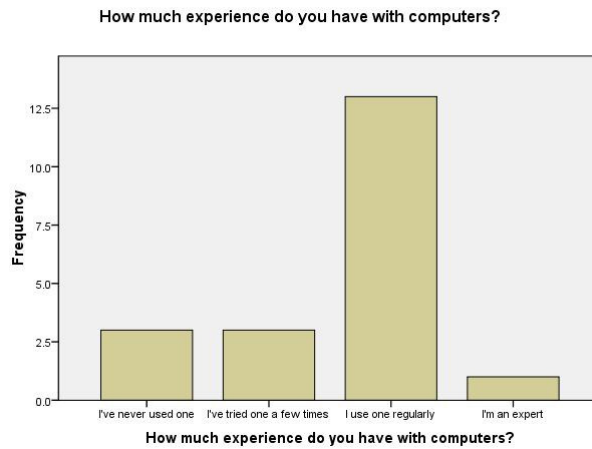
**Figure 11 Health Literacy level results**

### **Education and previous computer experience**

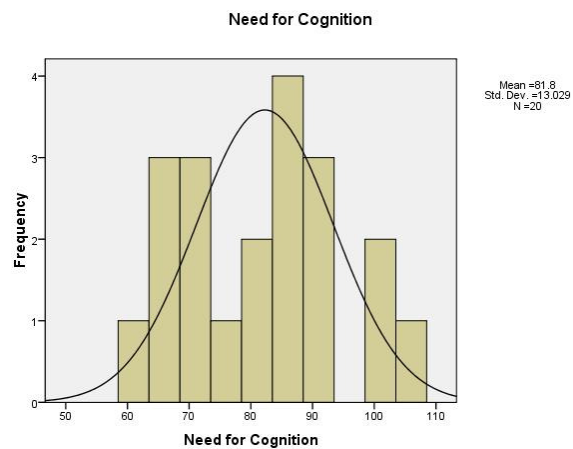
In addition to self-reported data about education and previous experience and familiarity with computers, participants were also asked to fill the 18-item shortened Need for Cognition survey [13] (see section 5.4.4 for details).



**Figure 12 Last grade of school completed**



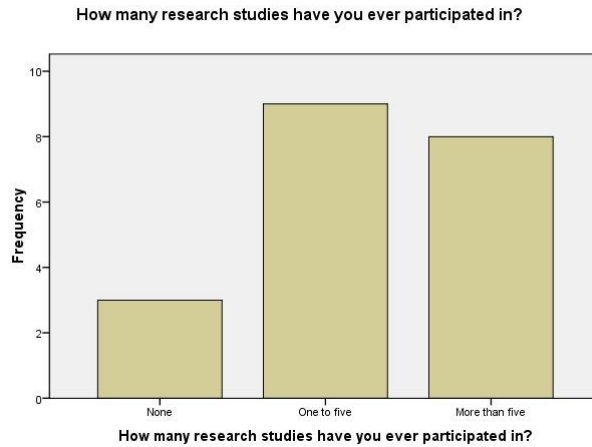
**Figure 13 Experience with Computers**



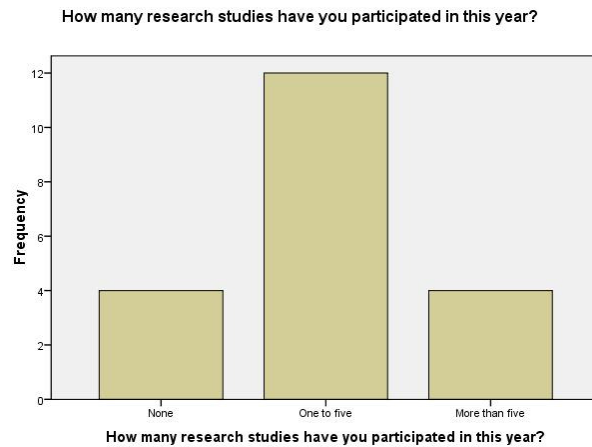
**Figure 14 Need for Cognition for participants**

### Previous exposure as research participants

Three participants reported having never participated in research studies, and 16 participants reported participating in some research study in 2009.



**Figure 15** Prior participation in research studies

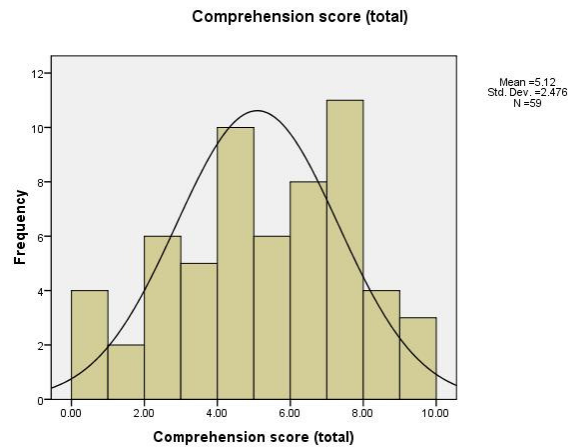


**Figure 16** Participation in research studies in 2009

### 5.5.2. Scoring Comprehension tests

Verbal responses to comprehension were transcribed, omitting conversational artifacts that were deemed irrelevant to the comprehension result (restarts, disfluencies [43], non-relevant speech etc.). The resulting responses were then compared against a Scoring Manual (see Appendix F). Each question has one or more relevant facts, and each fact is scored 0 to 2 (inclusive), where correct responses get 2, partially correct responses 1 and incorrect, incomplete or irrelevant responses get 0 (i.e.: no negative scoring). Each question's score is then the mean of the fact scores, and an overall comprehension result is calculated as the mean of the question scores.

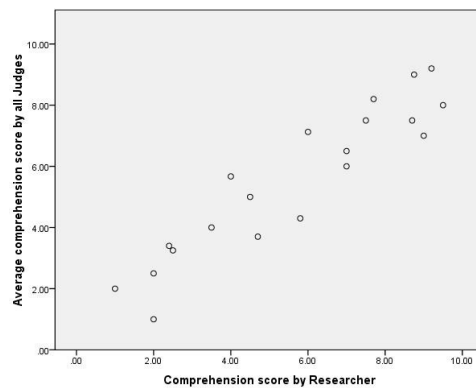
The distribution of the Comprehension scores (i.e.: for all sessions) was roughly normal.



**Figure 17 Distribution of all Comprehension scores**

Three random subsets of transcribed participant responses were given to three judges along with the Scoring Manual, and they were asked to assign comprehension scores. The scores given by the judges were then tested against the scoring that was performed. The summary of scores given by each judge is presented in Appendix G.

The Comprehension scores used in this analysis correlated strongly and significantly with the Comprehension scores given by first judge ( $r(7) = 0.949$ ,  $p < 0.001$ ), the second judge ( $r(7) = 0.924$ ,  $p < 0.001$ ) and the third judge ( $r(7) = 0.878$ ,  $p < 0.005$ ). Thus it was determined that the Scoring Manual (see Appendix F) was unambiguous and that the Scoring was reliable and consistent.



**Figure 18 Scatter-plot of comprehension scores by Researcher and Judges**

### 5.5.3. Testing for order effects – Comprehension

A repeated measures Analysis of Variance (ANOVA) was performed with the Comprehension scores for the three documents. No significant within-subjects effects were found,  $F(2, 36) = 1.60$ ,  $p > 0.05$ .

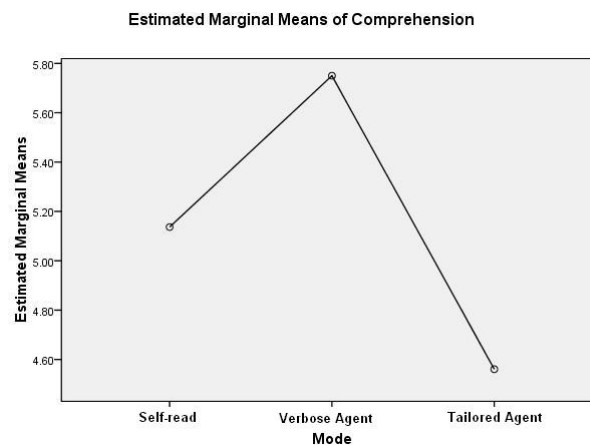
Similarly, a repeated measures ANOVA was performed with the Comprehension scores in the first, second and third sessions for each participant. No significant within-subjects effects were found,  $F(2, 36) = 1.07$ ,  $p > 0.05$ .

Based on these two results, effects of the documents and the order of presentation were disregarded from further analysis.

### 5.5.4. Hypotheses testing – Comprehension

A repeated measures analysis of variance (ANOVA) was the performed to assess comprehension scores across the three modes of the session (Self-read, Verbose Agent and Tailored Agent).

The Verbose Agent explanation mode was found to outperform the other two modes (self and tailored agent explanation) in terms of comprehension scores.



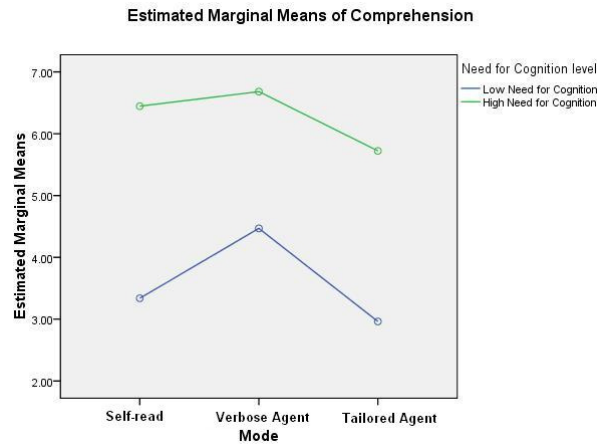
**Figure 19 Means for Comprehension scores by Mode of presentation**

Post-hoc tests confirmed that a statistically significant difference exists between the Comprehension scores in the Verbose Agent ( $M=5.75$ ) and the Tailored Agent ( $M=4.561$ ),  $p < 0.05$ .

A difference was also observed between the Comprehension scores in the Self-read mode ( $M=5.137$ ) and the Tailored Agent mode (4.561), approaching significance,  $p < 0.1$ .

I then dichotomized the Need for Cognition scores in to High and Low at the mean ( $M=81.8$ ,  $N=20$ , Standard deviation = 13.03). 12 participants were categorized as “High” and 8 as “Low”. The repeated measures ANOVA was performed again with the Need for

Cognition level as a between-subjects factor, and a significant between-subjects effect was observed.  $F(1, 17) = 9.19, p < 0.01$ .



**Figure 20 Comprehension scores and Need for Cognition across Modes**

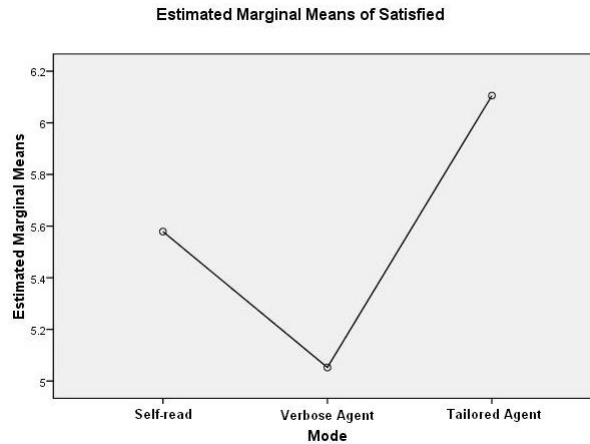
Participants with high Need for Cognition outperformed their low Need for Cognition counterparts in the comprehension tests, regardless of mode.

#### 5.5.5. Hypotheses testing – Satisfaction

Participant satisfaction was measured in two ways. After each session, a modified version of the BICEP instrument [53] was administered verbally. The first four questions measured the participants' perceptions of the amount of information (too little to too much), their likelihood to sign it (extremely unlikely to extremely likely), the amount of pressure they felt (no pressure to extreme pressure) and their satisfaction (extremely unsatisfied to extremely satisfied).

A repeated measures ANOVA was again performed with each of these values as dependent variables. While interesting trends are seen, no significant differences were found for the amount of information, the likelihood to sign and the amount of pressure felt within the three modes of presentation.

Participant satisfaction has an interesting trend that is approaching statistical significance,  $p < 0.1$ , where it appears that participants were most satisfied with the Tailored agent-explanation, and least with the Verbose agent-explanation.

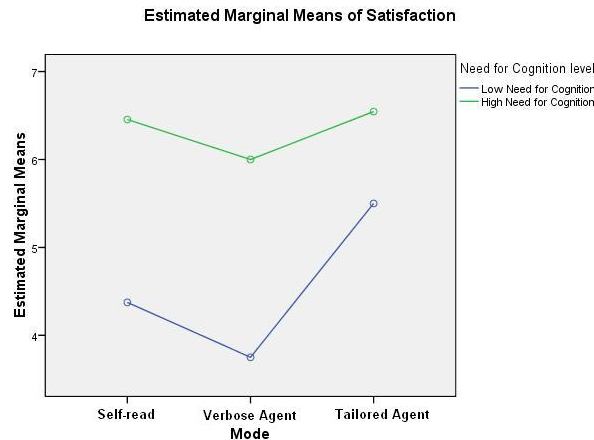


**Figure 21 Satisfaction scores by Mode of session (approaching significance)**

Post-hoc tests confirmed that the difference in participants' self-report Satisfaction scores in the Verbose mode ( $M=5.05$ ,  $N=19$ ) and the Tailored mode ( $M=6.11$ ,  $N=19$ ) was approaching significance,  $p < 0.1$ .

Secondly, after each session with the Agent, participants also filled out a 6-item questionnaire (See "Instructor Evaluation Questionnaire" in Appendix H) on paper. This questionnaire had high internal consistency (Cronbach's  $\alpha = 0.9$ ) and was combined in to a single Satisfaction score. A paired T-test was performed to see if the Satisfaction scores were different in the Verbose Agent ( $M=5.71$ ,  $N=18$ ) and Tailored Agent ( $M=5.88$ ,  $N=18$ ) modes. No significant differences were found,  $p > 0.05$ .

When the participants' Need for Cognition level was added as a between-subjects factor to the repeated measures ANOVA, the Need for Cognition level had a significant between-subjects effect on Satisfaction,  $F(1, 17) = 13.083$ ,  $p < 0.005$ .

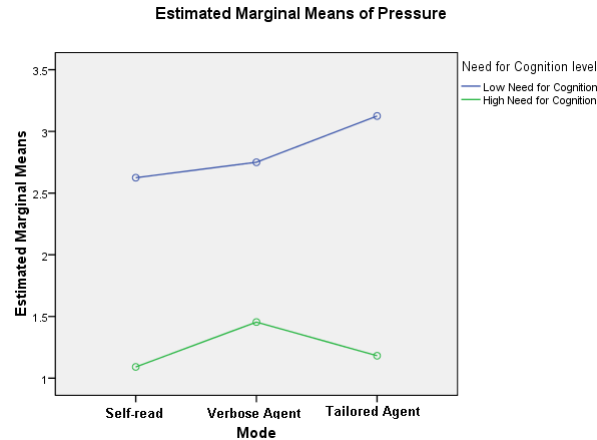


**Figure 22 Effect of Need for Cognition on Satisfaction across Modes**

High Need for Cognition participants reported higher levels of satisfaction than the low Need for Cognition participants.

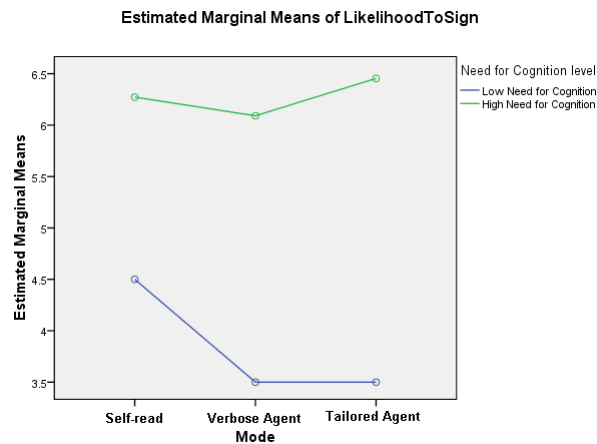


A similar test was performed on self-reported Pressure to sign the Informed Consent form. A significant between-subjects effect was found, where participants with low Need for Cognition generally felt more pressure in all three modes, and participant with high Need for Cognition generally reported feeling very little pressure (rated 1 to 7),  $F(1, 17) = 6.015, p < 0.05$ .



**Figure 23 Effect of Need for Cognition on self-reported perceived Pressure**

Self-reported likelihood to sign the Informed Consent form too exhibited a significant between-subjects effect on Need for Cognition, where participants with high Need for Cognition generally reported a higher likelihood to sign than participants with low Need for Cognition,  $F(1, 17) = 14.459, p = 0.001$ .



**Figure 24 Effect of Need for Cognition on self-reported Likelihood to Sign**

Need for Cognition did not have any significant between-subjects effects on self-reported amount of information received,  $F(1, 17) = 2.616, p > 0.05$ .

#### 5.5.6. Other tests

##### **Prior research study participation and Need for Cognition**

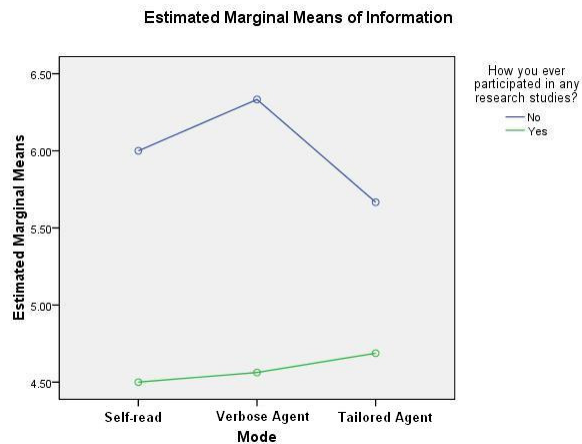
Participants were asked if they had participated in any research studies ever, and if they had participated in any research studies in 2009. The former was then dichotomized in to a “Prior research study participation” variable.

A one-way ANOVA was then performed to test if the Need for Cognition scores of participants who had prior research study participation experience was different from those who had not participated in research studies previously. No significant difference was found,  $F(1, 18) = 0.15$ ,  $p > 0.05$ . However, it must be pointed out that only 3 participants (15%) reported having never participated in research studies.

##### **Prior research study participation and outcome measures**

“Prior research study participation” was added as a between-subjects factor and a repeated measures ANOVA was performed with Comprehension and Satisfaction as dependent variables. No significant between-subjects effects were found.

When the amount of information participants reported as receiving was tested similarly, a significant between-subjects effect of prior research participation was found,  $F(1, 17) = 10.343$ ,  $p < 0.005$ . Participants who reported having never participated in research tended to rate the information received in all three modes as being close to too much (4 = just right, 7= too much), while other participants reported a level slightly above “just right”.



**Figure 25 Means of "How much information did you get?" by Mode and prior research participation**

Participants' self-reported pressure felt to sign the Informed Consent document, likelihood to sign the Informed Consent document and Satisfaction in the Informed Consent process were also tested similarly, but no significant effects were found of prior research participation.

### Conversation features and outcome measures

I extracted statistics about the conversation from the system logs. From this, I noted the number of conversation turns, the number of times the participant requested more detail, the number of times the participant requested background information and the number of times the participant reviewed a section after getting a question wrong.

A paired T-test confirmed that the number of turns of dialogue in the Verbose Agent-explanation mode ( $M=144.1$ ,  $N=19$ ) was significantly higher than the mean number of turns of dialogue in the Tailored Agent-explanation mode ( $M=87.53$ ,  $N=19$ ), paired  $t(18) = 8.883$ ,  $p < 0.001$ .

The number of turns of dialogue was also found to correlate negatively with participant's satisfaction with the Agent,  $r(36) = -0.34$ ,  $p < 0.05$ . This is an interesting finding, since it provides what could explain the difference in Satisfaction levels between the two Agent modes, and seems to suggest that our background concept explanations are a little too lengthy for most participants.

No significant correlations were found between the number of turns of dialogue and participant's perceptions of amount of information, self-reported likelihood to sign or self-reported perceived pressure.

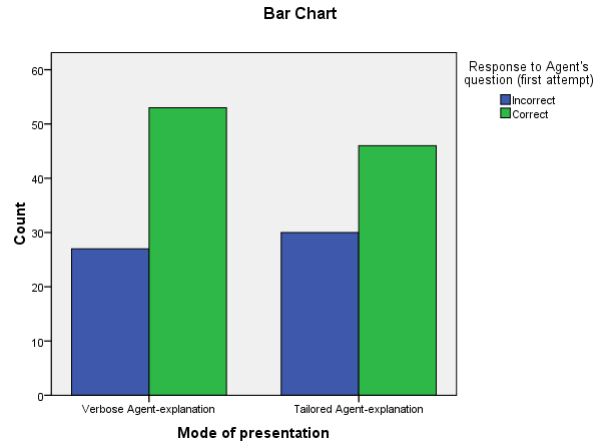
I also tested to see if participants responded differently to the agent's comprehension questions in the two agent modes. Participants' responses to the agent's questions were recorded as either incorrect or correct on the first attempt. Where participants made a second attempt, it too was recorded as either incorrect or correct.

A Chi-squared test for independence was performed on each of these, and the cross-tabulations are below.

		Response to Agent's question (first attempt)		Total
		Incorrect	Correct	
Mode of presentation	Verbose Agent	27	53	80
	Tailored Agent	30	46	76
Total		57	99	156

**Table 15 Cross-tabulation of responses to Agent's questions on first attempt**

No significant difference was found between the distribution of correct and incorrect responses on the first attempt between the Verbose Agent and Tailored Agent modes,  $\chi^2(1) = 0.551$ ,  $p > 0.05$ .



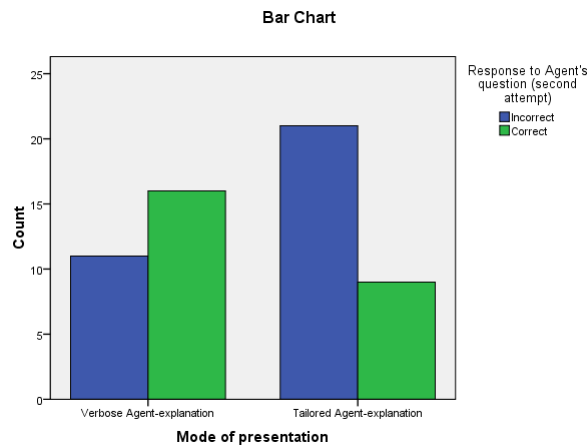
**Figure 26 Incorrect vs. Correct responses to Agent's questions on first attempt**

Where participants were asked a comprehension question by the agent for a second time, a similar test was performed. The Cross-tabulation is below.

		Response to Agent's question (second attempt)		Total
		Incorrect	Correct	
Mode of presentation	Verbose Agent	11	16	27
	Tailored Agent	21	9	30
Total		32	25	57

**Table 16 Cross-tabulation of responses to Agent's questions on second attempt**

A significant difference was found in the distribution of Incorrect vs. Correct responses on the second attempt,  $\chi^2(1) = 4.941$ ,  $p = 0.05$ .



**Figure 27 Incorrect vs. Correct responses to Agent's questions on second attempt**

The verbose agent-explanation mode had more correct responses and less incorrect responses on the second attempt.

## 5.6. Qualitative review

After each participant's study session, a semi-structured interview was conducted. Some of the questions asked are listed in the Study Protocol in Appendix D.

Ten of the participants (50%) indicated a preference for the Agent, attributing this to the ease of the interaction, the additional information, the thoroughness of the Agent and the fact that they paid more attention when the Agent was explaining the document.

Five of the participants (25%) on the other hand, indicated a clear preference for reading the document themselves, and felt that the Agent either slowed them down, kept them less focused, or found the Agent's additional information confusing.

Seven of the participants (35%) further mentioned noticing a clear difference between the explanations provided by the Verbose and Tailored agents.

Some interesting comments are listed below:

- “The second one [Verbose Agent] was more detailed, and like I said, more redundant...”, “... the intention [Verbose Agent] was good, but it added too much...”, (41 year old Male)
- “She [Tailored Agent] did a little more paraphrasing, but she didn't read word-for-word and that was confusing.... I don't think that's good when you're talking about a Consent Form... it should be word-for-word.” (48 year old Female)
- Did the Agent help you to understand the documents better? “When she provided additional information, yes, but just by reading it, no.” (25 year old Female)
- “Best would be a combination where you could get additional information if you wanted, but skip sections if you wanted to.” (25 year old Female)
- “It was much more thorough with the Agent... she kind of categorized areas...” (25year old Female)
- “The only advantage [with the Agent] is that she asked you a couple of questions, so it made me, you know, if I got the answer wrong, she'd repeat it, so I think in that respect it's good.” (51 year old Female)

## 5.7. Discussion

In this thesis, I presented the hypothesis that participants would exhibit higher Comprehension scores in both of the Agent conditions when compared with the Self-read condition. The results show partial support for this hypothesis, and the Verbose Agent-explanation mode was shown to be associated with significantly higher comprehension scores, compared with the Tailored Agent mode.

In addition to the comprehension test that was used to test the former hypothesis, the Agent too administered some simple comprehension tests. These too showed an interesting result, where participants got a similar number of questions correct and

incorrect on their first attempt, but on the second attempts, the Verbose Agent mode had more correct responses than the Tailored Agent mode.

All this points to the fact that the Verbose agent-explanation's strategy of providing background information and high-detail level provided additional learning benefits for the participants. However, the majority of participants in this study were in the highest Health Literacy bracket level, so the efficacy of this system with individuals with low health literacy remains unknown.

Another interesting effect was seen where participants with high levels of Need for Cognition outperformed their low Need for Cognition counterparts in comprehension tests. This provides support to the argument that the comprehension scores rely also on the participant's initiative for learning. In this sense, the Agent becomes indeed a facilitator of learning for the participant. Again, as mentioned earlier, this may turn out to be valid only for participants with high levels of Health Literacy such as those found in this sample.

It was further hypothesized that the Tailored Agent-explanation would lead to higher levels of satisfaction than the Verbose Agent-explanation condition. One of the two measures of Satisfaction used approaches significance, and the other measure had no significant results.

A possible explanation is that the Verbose agent-explanation mode was far too lengthy and repetitive. The fact that participants were not significantly more satisfied with the Verbose agent, as opposed to reading the document themselves, is also evidence that the Agent either tried to give too much information, or took too much control of the conversation. The higher comprehension scores in the Tailored Agent mode than the Self-read mode (and approaching statistical significance) however is very promising.

Some of these findings also hint at limitations with this thesis. The primary limitation of this work is that the majority of participants had high health literacy (as measured by the REALM instrument). Secondly, the implemented document explanation strategies were based on observational studies with very few participants. The expert observed in the primary observational study (see 3.2) was someone with experience in administering Informed Consent for research studies, but was not a true "Expert" in health information communication. Thirdly, several results approaching statistical significance levels suggest that the study is under-powered, and that perhaps more participants should have been recruited. Finally, the "Tailoring" that was actually implemented was very simple and based on user-initiative, so its effect may have been minimized.

A few more limitations were introduced in to the new Agent system. As an example, this system does not yet deal well enough with branching logic. Branching in conversations has many important implications like facilitating question-answer dialogue instead of the largely monologue nature of the current Agent. In fact, several participants referred to how it became too repetitive when "OK" was the only response that could be given to the agent.

## 6. Conclusions and future work

### 6.1. Conclusions

In this thesis I presented a new approach for building Conversational Agent systems that does not rely on pre-scripted dialogue. In empirical tests, this new Conversational Agent was found to be effective in facilitating higher levels of comprehension in participants than they achieved by reading a Research Informed Consent document themselves.

I finally review the research questions that I posed at the beginning and attempt to judge the contribution of this thesis on each of those questions.

The first question I posed was how well a Conversational Agent can perform as a learning guide, when provided with extremely simple capabilities for teaching. I have successfully implemented a new Agent system that has extremely simple teaching capabilities, where the traversal of a document represents teaching goals, and this is supported by linkages to a knowledge base and some questions that the Agent can ask from the participant. This Agent was then tested in a real-world scenario, and it was demonstrated to help the participants comprehend the document better than when participants read the documents themselves.

Secondly, I asked if a Conversational Agent can improve the Informed Consent experience for potential participants. Many participants were quite vocal about how much they enjoyed having the Agent explain the document, and statistical tests showed trends, but no conclusive results.

Further, the Agent only used a very minimal set of non-verbal behaviors, such as head nods and pointing. There are many other features that could be added, for example, varied facial expressions, greetings, farewells and other casual conversation, which might contribute to increased satisfaction for the participants. This, however is an open question.

Finally, I asked if we can develop Conversational Agents that can be adapted very easily to explain new documents. The evidence from this study confirms that this is indeed possible. Unlike earlier systems where the Agent system had to be re-compiled and re-deployed for even a minor change, the new Agent has a single deployment unit which takes a document as input and traverses the document, explaining it.

## 6.2. Future work

The results observed in this research are very promising. Some possible future research areas are discussed below.

- The primary lesson from this research is that a balance must be struck between providing all the explanation that the participants need, and the right degree of flexibility.
- One of the underlying premises of this thesis is that we can leverage the structure inherent to a document to make the document explanation task easier. It might therefore seem counterintuitive that this thesis used a machine-readable representation of a document, instead of a document itself. It is however easy to understand why this was a necessary first step. A logical next step therefore, is to enable the explanation of any arbitrary document, by inferring its structure.
- This study used an extremely simple 12-item knowledge base that was linked to the document by an expert's annotation. The testing of the Topic relevance ranking algorithm too confirms that even such a simple knowledge base is not trivial to use dynamically. However this also presents many opportunities for evaluating different strategies for utilizing larger knowledge bases, which might in fact make the relevance ranking task easier.
- Abstractive text summarization has an interesting application, where the detail-level of the summary might be parameterized based on some knowledge about the participant (e.g.: health literacy level). This way, fine-grained control could be achieved for tailoring the document.
- The results from this study seem to suggest that repeating the same content twice (as was done when the participant's response to a question by the agent was incorrect) is actually effective, since most of the participants answered correctly on the second attempt. While some previous studies in improving Informed Consent comprehension have also suggested this procedure (e.g.: [47]), this warrants further study.
- An authoring tool for Research Informed Consent documents could provide an environment where an author could quickly and easily create new source material that would allow the Agent to explain new Informed Consent documents. It could also incorporate some of the current guidelines, like for example showing the author statistics about the readability of the each and every section of the document as well the whole document itself.



## 7. Acknowledgements

I gratefully acknowledge the guidance of my advisor, Prof. Timothy Bickmore, whose pioneering work in addition provided the basis for much that is reported here. Dr. Michael Paasche-Orlow of the Boston University School of Medicine provided valuable input on the problems with health information documents, and with some of the study design. I am especially grateful to Prof. Carole Hafner and the rest of my thesis committee for their advice and guidance, and also for patience and understanding as I struggled to meet deadlines.

I am also indebted in particular to Laura Pfeifer, Langxuan Yin, Margaret McElduff, Lindsey Hollister, Thomas Brown and Francisco Crespo who conducted previous user studies which I have used here, as well as their volunteer participants. Prof. Timothy Bickmore and Laura Pfeifer were instrumental in an earlier version of document explanation systems that this work is based on. Dr. Donna Byron provided advice on some of the linguistic aspects of the study. Dan Schulman helped me with some of the statistics. I am also grateful to all members of the Relational Agents Group of the College of Computer and Information Science at the Northeastern University for the much appreciated feedback during the various points of this study, and especially the three judges (Dan, Laura and Langxuan) who scored some of the comprehension tests for comparison. Brian and Rachel along with Prof. Agnes Chan and the MS committee and the rest of the College of Computer and Information Science at Northeastern University too helped in numerous ways.

I am grateful also towards all the study participants, and everyone who helped me setup and run the study sessions, including Krissy Mainello and the Human Subject Research Protection Office at Northeastern University and Estella Saraceni and the staff at Marcus Garvey Apartments in Roxbury, MA.

A word of thanks goes out to all my friends, colleagues and family, especially Hardik, Shahzad, Keshi, Chris, Therese, Pradeep, Ira, Neville, Rukshan, Oshani, Chaamari, Sepalika, Dan, Laura, James, Lazlo, Tom, Jenna, Juan, Dolphy, Prasanna, Zoe, Liang, Jon, Krissy, Donna, Jim, Luc, Francesco, the Redemptorist Fathers and Brothers at the Basilica of Our Lady of Perpetual Help and everyone else in Boston and Sri Lanka who helped me remain sane during an incredibly challenging time.

Finally, this is dedicated especially to my parents, Charles and Felicia, my sister Nilmini, and Asha, my strength and my wife – your sacrifices brought me to the beginning of this thesis, and your dedication to your lives and missions inspired me to persevere when I was tempted to give up.

## 8. References

1. Agency for Health Research and Quality. The AHRQ Informed Consent and Authorization Toolkit for Minimal Risk Research. AHRQ Publication No. 09-0089-EF, September 2009, Agency for Healthcare Research and Quality, Rockville, MD. <http://www.ahrq.gov/fund/informedconsent/>
2. Anderson, J. R., Boyle, C. F., and Reiser, B. J. (1985). Intelligent tutoring systems. *Science*, 228(4698):456-462
3. Baker, D. W., Parker, R. M., Williams, M. V., Clark, W. S., and Nurss, J. (1997). The relationship of patient reading ability to self-reported health and use of health services. *American journal of public health*, 87(6):1027-1030.
4. Baker, D. W., Parker, R. M., Williams, M. V., and Clark, W. S. (1998). Health literacy and the risk of hospital admission. *Journal of general internal medicine*, 13(12):791-798.
5. Berkman ND, DeWalt DA, Pignone MP, Sheridan SL, Lohr KN, Lux L, Sutton SF, Swinson T, Bonito AJ. Literacy and Health Outcomes. Evidence Report/Technology Assessment No. 87 (Prepared by RTI International-University of North Carolina Evidence-based Practice Center under Contract No. 290-02-0016). AHRQ Publication No. 04-E007-2. Rockville, MD: Agency for Healthcare Research and Quality. January 2004.
6. Bickmore, T. W., Pfeifer, L. M., and Jack, B. W. (2009). Taking the time to care: empowering low health literacy hospital patients with virtual nurse agents. In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*, pages 1265-1274, New York, NY, USA. ACM.
7. Bickmore, T., Pfeifer, L., & Paasche-Orlow, M. K. (2007). Health Document Explanation by Virtual Agents. *Intelligent Virtual Agents '07*.
8. Bickmore, T. ., Pfeifer, L. ., and Paasche-Orlow, M. . (2009). Using computer agents to explain medical documents to patients with low health literacy. *Patient Education and Counseling*, 75(3):315-320.
9. Bickmore, T., Pfeifer, L., and Yin, L. (2008). The Role of Gesture in Document Explanation by Embodied Conversational Agents. *International Journal of Semantic Computing* Vol. 2, No. 1, 47-70.
10. Bickmore, T. W. and Picard, R. W. (2005). Establishing and maintaining long-term human-computer relationships. *ACM Trans. Comput.-Hum. Interact.*, 12(2):293-327.

11. Boersma, P. & Weenink, D. (2008). Praat: doing phonetics by computer (Version 5.0.38) [Computer program]. Retrieved November 1, 2008, from <http://www.praat.org/>
12. Brug, J., Glanz, K., Van Assema, P., Kok, G., and van Breukelen, G. J. (1998). The impact of computer-tailored feedback and iterative feedback on fat, fruit, and vegetable intake. *Health education & behavior : the official publication of the Society for Public Health Education*, 25(4):517-531.
13. Cacioppo, J. T., Petty, R. E., and Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of personality assessment*, 48(3):306-307.
14. Cassell, J., Sullivan, J., Prevost, S., and Churchill, E. F. (eds) (2000). *Embodied conversational agents*. MIT Press
15. Cassell, J., Vilhjálmsón, H. H., and Bickmore, T. (2001). Beat: the behavior expression animation toolkit. In *SIGGRAPH '01: Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 477-486, New York, NY, USA. ACM.
16. Conati, C., Gertner, A., and Vanlehn, K. (2002). Using bayesian networks to manage uncertainty in student modeling. *User Modeling and User-Adapted Interaction*, 12(4):371-417.
17. Corbett, A. T., Koedinger, K. R., & Anderson, J. R. (1997). Intelligent tutoring systems. In M. G. Helander, T. K. Landauer, & P. V. Prabhu (Eds.), *Handbook of human-computer interaction* (pp. 849-874). Amsterdam: Elsevier Science B. V.
18. Coriell Institute for Medical Research. Informed Consent form template for Genetic Research (Retrieved December 11<sup>th</sup>, 2009, from <http://ccr.coriell.org/Sections/Support/NIGMS/Model.aspx?PgId=216>)
19. Davis, T. C., Long, S. W., Jackson, R. H., Mayeaux, E. J., George, R. B., Murphy, P. W., Crouch, M. A. (1993). Rapid Estimate of Adult Literacy in Medicine: A Shortened Screening Instrument. *Journal of Family Medicine* 25(6):391-5.
20. Dunn, L. B., Lindamer, L. A., Palmer, B. W., Golshan, S., Schneiderman, L. J., and Jeste, D. V. (2002). Improving understanding of research consent in middle-aged and elderly patients with psychotic disorders. *American Journal of Geriatric Psychiatry*, 10(2):142-150.
21. Faden, R. R., Beauchamp, T. L., and King, N. M. P. (1986). *A history and theory of informed consent*. Oxford University Press.
22. Fellbaum (1998). *WordNet: An Electronic Lexical Database (Language, Speech, and Communication)*. The MIT Press.

23. Flory, J. and Emanuel, E. (2004). Interventions to improve research participants' understanding in informed consent for research: A systematic review. *JAMA*, 292(13):1593-1601.
24. Gallin, J. I. (2006). Taming clinical research's paper tigers. *Clinical Cancer Research*, 12(14):4131.
25. Gans, K. M., Risica, P. M., Strolla, L. O., Fournier, L., Kirtania, U., Upegui, D., Zhao, J., George, T., and Acharyya, S. (2009). Effectiveness of different methods for delivering tailored nutrition education to low income, ethnically diverse adults. *The international journal of behavioral nutrition and physical activity*, 6:24+.
26. Gazmararian, J. A., Kripalani, S., Miller, M. J., Echt, K. V., Ren, J., and Rask, K. (2006). Factors associated with medication refill adherence in cardiovascular-related diseases: a focus on health literacy. *Journal of general internal medicine*, 21(12):1215-1221.
27. Graesser, A. C., VanLehn, K., Rosé, C. P., Jordan, P. W., and Harter, D. (2001). Intelligent tutoring systems with conversational dialogue. *AI Mag.*, 22(4):39-51.
28. Graesser, A. C., Jeon, M., and Dufty, D. (2008). Agent technologies designed to facilitate interactive knowledge construction. *Discourse Processes: A multidisciplinary journal*, 45(4):298-322.
29. Greenwood, M. (1997). JWordNetSim (version 1.0.0) [Computer program]. Retrieved October 26, 2009, from <http://nlp.shef.ac.uk/result/software.html>.
30. Hawkins, R. P., Kreuter, M., Resnicow, K., Fishbein, M., and Dijkstra, A. (2008). Understanding tailoring in communicating about health. *Health Educ. Res.*, 23(3):454-466.
31. Hubal, R. C. and Day, R. S. (2006). Informed consent procedures: an experimental test using a virtual character in a dialog systems training application. *J. of Biomedical Informatics*, 39(5):532-540.
32. Irwin, P. M. (1991). National literacy act of 1991: Major provisions of p.l. 102-73. crs report for congress. Technical report.
33. Jiang, J. J. and Conrath, D. W. (1997). Semantic similarity based on corpus statistics and lexical taxonomy. In *International Conference Research on Computational Linguistics (ROCLING X)*, pages 9008+.
34. Kandula, S. and Zeng-Treitler, Q. (2008). Creating a gold standard for the readability measurement of health texts. *AMIA ... Annual Symposium proceedings / AMIA Symposium*. AMIA Symposium, pages 353-357.

35. Kaur, I. and Hornof, A. J. (2005). A comparison of lsa, wordnet and pmi-ir for predicting user click behavior. In CHI '05: Proceedings of the SIGCHI conference on Human factors in computing systems, pages 51-60, New York, NY, USA. ACM.
36. Kreuter, M. W., Strecher, V. J., and Glassman, B. (1999). One size does not fit all: the case for tailoring print materials. *Ann Behav Med*, 21(4):276-283.
37. Kripalani, S., Bengtzen, R., Henderson, L. E., and Jacobson, T. A. (2008). Clinical research in low-literacy populations: using teach-back to assess comprehension of informed consent and privacy information. *IRB*, 30(2):13-19.
38. Lothian, N. Classifier4j (Version 0.6) [Computer program]. Retrieved November 1, 2008, from <http://classifier4j.sourceforge.net/index.html>
39. Lin, D. (1998). An information-theoretic definition of similarity. In Shavlik, J. W. and Shavlik, J. W., editors, *ICML*, pages 296-304. Morgan Kaufmann.
40. Matthews, D. E., VanLehn, K., Graesser, A. C., Jackson, G. T., Jordan, P., Olney, A., and Rosa, A. C. (2007). When are tutorial dialogues more effective than reading? *Cognitive Science*, 31(1):3-62.
41. National Heart Lung and Blood Institute. Calculate your Body Mass Index. Retrieved December 10, 2009, from <http://www.nhlbisupport.com/bmi/>
42. Nielsen-Bohlman, L., Panzer, A. M., Kindig, D. A., Institute of Medicine. (2004). *Health literacy : a prescription to end confusion*. National Academies Press.
43. Oviatt, S. (1994). Predicting and managing spoken disfluencies during human-computer interaction. In *HLT '94: Proceedings of the workshop on Human Language Technology*, pages 222-227, Morristown, NJ, USA. Association for Computational Linguistics.
44. Paasche-Orlow, M., Parker, R., Gazmararian, J., Nielsen-Bohlman, L., and Rudd, R. (2005). The prevalence of limited health literacy. *Journal of General Internal Medicine*, 20(2):175-184.
45. Paasche-Orlow, M. K., Riekert, K. A., Bilderback, A., Chanmugam, A., Hill, P., Rand, C. S., Brancati, F. L., and Krishnan, J. A. (2005). Tailored education may reduce health literacy disparities in asthma self-management. *American Journal of Respiratory and Critical Care Medicine*, 172(8):980-986.
46. Paasche-Orlow, M. K., Taylor, H. A., & Brancati, F. L. (2003). Readability Standards for Informed-Consent Forms as Compared with Actual Readability. *N Engl J Med* 2003;348:721-6.

47. Palmer, B. W., Cassidy, E. L., Dunn, L. B., Spira, A. P., and Sheikh, J. I. (2008). Effective use of consent forms and interactive questions in the consent process. *IRB*, 30(2):8-12.
48. Pedersen, T., Patwardhan, S., and Michelizzi, J. (2004). Wordnet::similarity: measuring the relatedness of concepts. In *HLT-NAACL '04: Demonstration Papers at HLT-NAACL 2004 on XX*, pages 38-41, Morristown, NJ, USA. Association for Computational Linguistics.
49. Reiter, E., & Dale, R. (2000). *Building Natural Language Generation Systems*. Cambridge University Press.
50. Sibun, P., Huettner, A. K., & McDonald, D. D. (1988). Directing the generation of living space descriptions. *Proc. of the 12th conference on Computational linguistics*. 626-629.
51. Spärck Jones, K. (2007). Automatic summarising: The state of the art. *Information Processing & Management*, 43(6):1449-1481.
52. Suebnukarn, S. and Haddawy, P. (2007). Comet: A collaborative tutoring system for medical problem-based learning. *IEEE Intelligent Systems*, 22(4):70-77.
53. Sugarman, J., Lavori, P. W., Boeger, M., Cain, C., Edson, R., Morrison, V., and Yeh, S. S. S. (2005). Evaluating the quality of informed consent. *Clinical trials* (London, England), 2(1):34-41.
54. van Keulen, H. M., Mesters, I., Brug, J., Ausems, M., Campbell, M., Resnicow, K., Zwietering, P. J., van Breukelen, G., van Mechelen, W., Severens, J. L., and de Vries, H. (2008). Vitalum study design: Rct evaluating the efficacy of tailored print communication and telephone motivational interviewing on multiple health behaviors. *BMC public health*, 8:216+.

**Appendix A – source Informed Consent documents used in the study**  
Research Informed Consent Document 1 – Genetic Cell Repository

**Human Genetic Cell Repository**  
Informed Consent Form

**Why am I being asked to take part in this research study?**

You are being asked to participate in this study because you are an adult who knows if you or any family members have certain medical conditions or not.

**Why are you doing this research study?**

The purpose of this study is to find if there are any relationships with certain medical conditions and genes.

**What will I be asked to do?**

If you decide to take part in this study, a 5 ml (1 teaspoon) blood or skin biopsy sample will be taken to the NU Cell Repository, a research resource supported by the National Institutes of Genomics. The Repository collects, stores, and distributes cell cultures and DNA samples from people with many kinds of disorders, from unaffected family members, and from other healthy people.

Your sample will be used to create a cell line that will be added to the Repository's collection. The purpose of this collection is solely to make specimens available for use in research and teaching and as standards in clinical genetics laboratories.

Submission of your sample to the Repository may give scientists valuable research material that can help them to develop new diagnostic tests, new treatments, and new ways to prevent diseases.

**Where will this take place and how much of your time will it take?**

This session will take place at Northeastern Institute for Medical Research and will take approximately one hour.

**Will there be any risk or discomfort to me?**

The medical risks of providing these specimens are minimal. The risk for venipuncture is minor transient pain and slight possibility of infection. The risk for skin biopsy is mild local pain, slight bleeding, the possibility of a small scar, and slight possibility of infection.

*This is a fake consent document – it does not imply any actual commitments or contract.*

**Appendix A – source Informed Consent documents used in the study**  
Research Informed Consent Document 1 – Genetic Cell Repository

**Will I benefit by being in this research?**

There will be no direct benefit or payment to you for participating, but your sample may benefit the community at large or some particular group.

**Who will see the information about me?**

The Repository will take measures to protect your privacy. Your blood or tissue specimen will be given a code number, your name will be removed, and the NU Cell Repository will not give out your name to the scientists who receive the samples. Some patient identification, such as age, sex, diagnosis, and race, will be made available to the Repository and scientists. There can be no absolute guarantee of confidentiality, however, and there is a small chance that some research may yield results that will have a negative impact on you, your family, other individuals, or groups. These may include insurability, employability, and/or family relationships.

The Repository has obtained a Confidentiality Certificate from the Federal Government to help insure your privacy. This Certificate means that the Repository cannot be forced to tell people who are not connected with the study about your participation without your written consent.

**If I don't want to take part in the study, what choices do I have?**

Your donation is voluntary, and if you choose not to participate there will be no penalty or loss of benefits to which you are entitled.

**What will happen if I suffer any harm from this research?**

If you are injured as a result of your participation in this research, no special arrangements will be made for compensation or for payment for treatment.

**Can I stop my participation in this study?**

Your sample will stay in the Repository indefinitely and you will not be able to withdraw it.

*This is a fake consent document – it does not imply any actual commitments or contract.*



**Appendix A – source Informed Consent documents used in the study**  
Research Informed Consent Document 1 – Genetic Cell Repository

**Who can I contact if I have questions or problems?**

If you have any questions or complications relating to collection of this specimen, you should contact Dr Brian Orlow at (617) 373-1234, who collected the specimen. If you have any questions about the Repository, you should contact Dr Michael Jack, Northeastern Institute for Medical Research, (617) 373-5678.

**Who can I contact about my rights as a participant?**

If you have questions about your rights as a research subject you should call Dr Laura Silliman, (617) 373-7890.

**Will I be paid for my participation?**

You will not be paid to participate in this research study.

**Will it cost me anything to participate?**

Your donation is voluntary, and there are no costs to you for participating in this research study.

**Is there anything else I need to know?**

Because researchers will not have access to your identity, it will not be possible to provide you or your physician with the eventual results of studies that might be performed using your specimen.

**Will this be used in any publication or teaching material?**

It is possible that data resulting from use of your sample may eventually be used in a research publication. In that event, no individual's identifying information will be included, as this information will not be available to the researchers.

(Signature) ----- Do Not Sign -----

\_\_\_\_\_

(Date) \_\_\_\_\_

*This is a fake consent document – it does not imply any actual commitments or contract.*

**Appendix A – source Informed Consent documents used in the study**  
Research Informed Consent Document 2 – Re-hospitalization reduction study

**Cardiopulmonary re-hospitalization reduction study**  
Informed Consent Form

**Why are you doing this research study?**

This study is being done to determine how best to prepare patients for discharge from the hospital.

**Why am I being asked to take part in this research study?**

You are being asked to participate in this study because you are an adult preparing to be discharged from the hospital. You will be one of approximately 814 subjects to be asked to participate in this study.

**What will I be asked to do?**

If you agree to participate, we will ask you to read a list of 66 words (REALM) to assess your health literacy level and ask you questions about your demographic information. Then you will have a 50% chance of having a nurse discharge advocate spending about 20 minutes with you to carefully review information that might be helpful to you as you leave the hospital. This information will then be used by an interactive, computer-based patient-education program before discharge. In this group, you also receive telephone calls from a computerized health education and monitoring system beginning after you leave the hospital and lasting until you see your doctor. There is a 50% chance of having the routine hospital discharge procedures. In both groups, you will receive a phone call 30 days after you go home from the hospital to ask about your health since leaving the hospital and to ask if you have been back to the hospital.

**Where will this take place and how much of my time will it take?**

All or part of the research in this study will take place at Northeastern University Medical Center.

**Will there be any risk or discomfort to me?**

The only risk is the possibility of some embarrassment in talking with the nurse about your health while in the hospital.

There may be unknown risks/discomforts involved. Study staff will update you in a timely way on any new information that may affect your health, welfare, or decision to stay in this study.

*This is a fake consent document – it does not imply any actual commitments or contract.*

**Appendix A – source Informed Consent documents used in the study**  
Research Informed Consent Document 2 – Re-hospitalization reduction study

**Will I benefit by being in this research?**

The benefits of participating in this study may be that you are better prepared for going home. However, you may not receive any benefit from participating.

**Who will see the information about me?**

Information from this study and from your medical record may be reviewed and photocopied by the Food and Drug Administration (FDA) and/or state and federal regulatory agencies such as the Office of Human Research Protection as applicable, and the Institutional Review Board of Northeastern University Medical Center. Information from this study and from your medical record may be used for research purposes and may be published; however, your name will not be used in any publications.

**If I don't want to take part in the study, what choices do I have?**

Your alternative is to not participate in the study.

**What will happen if I suffer any harm from this research?**

No special arrangements will be made for compensation or for payment for treatment solely because of your participation in this research.

**Can I stop my participation in this study?**

Taking part in this study is voluntary. You have the right to refuse to take part in this study. If you decide to be in the study and then change your mind, you can withdraw from the research. Your participation is completely up to you. Your decision will not affect your being able to get health care at this institution or payment for your health care. It will not affect your enrollment in any health plan or benefits you can get.

If you choose to take part, you have the right to stop at any time. If there are any new findings during the study that may affect whether you want to continue to take part, you will be told about them as soon as possible.

*This is a fake consent document – it does not imply any actual commitments or contract.*

**Appendix A – source Informed Consent documents used in the study**  
Research Informed Consent Document 2 – Re-hospitalization reduction study

**Who can I contact if I have questions or problems?**

The investigator or a member of the research team will try to answer all of your questions. If you have questions or concerns at any time, or if you need to report an injury while participating in this research, contact Brian Fernando at (617) 373-9876 during the day and the YourHealthNet attending physician at 617-373-8475 after hours.

**Who can I contact about my rights as a participant?**

You may obtain further information about your rights as a research subject by calling the Office of the Institutional Review Board of Northeastern University Medical Center at 617-373-9182.

**Will I be paid for my participation?**

You will not be paid to participate in this research study.

**Will it cost me anything to participate?**

There are no costs to you for participating in this research study.

**Is there anything else I need to know?**

By consenting to participate in this study you do not waive any of your legal rights. Giving consent means that you have heard or read the information about this study and that you agree to participate. You will be given a copy of this form to keep. If at any time you withdraw from this study you will not suffer any penalty or lose any benefits to which you are entitled.

The investigator may decide to discontinue your participation without your permission because he/she may decide that staying in the study will be bad for you, or the sponsor may stop the study.

(Signature) ----- Do Not Sign ----

\_\_\_\_\_

(Date) \_\_\_\_\_

*This is a fake consent document – it does not imply any actual commitments or contract.*

**Appendix A – source Informed Consent documents used in the study**  
Research Informed Consent Document 3 – Long-term human-computer relationships

**Long-Term Human-Computer relationships study**

**Informed Consent Form**

**Why am I being asked to take part in this research study?**

You are being asked to participate in this study because you are a healthy, English speaking adult who owns a home computer with an Internet connection.

**Why are you doing this research study?**

The purpose of this research is to find ways to help people stick with exercise programs over many months or years.

**What will I be asked to do?**

If you decide to take part in this study, we will first ask you some questions about your background and health status. If we find any areas of health concern, we may ask you to obtain a note from your physician before participating in the study. After this, we will show you how to use a pedometer, which is a small battery-powered device that counts the number of steps you walk each day. We will then show you how to install and use software that we will give you to install on your home computer (at your request, a member of the research staff may come to your home to help you with this). Once you have gone home and installed the software, you will begin using the pedometer every day and running the software every day for 10 minutes. During these sessions you will have a short conversation with an animated exercise advisor and answer some additional questions. You will continue this every day as long as you like, until either: (1) you notify us that you no longer wish to be in the study; (2) you have stopped using the software for two consecutive weeks, or; (3) we notify you that the study has ended (expected to be no earlier than April, 2011).

**Where will this take place and how much of my time will it take?**

The initial study session will take place either in the Human-Computer Interaction Laboratory at Northeastern University and will take approximately one hour. After that, it is expected to take approximately 10 minutes of your time each day that you choose to run the software on your home computer.

*This is a fake consent document – it does not imply any actual commitments or contract.*

## **Appendix A – source Informed Consent documents used in the study**

### **Research Informed Consent Document 3 – Long-term human-computer relationships**

#### **Will there be any risk or discomfort to me?**

The primary risk from the study is injury or other health problems from beginning an exercise program. However, the software will only ask you to perform walking as a form of exercise, and only in a way that is consistent with the current national guidelines for physical activity. In addition, you will give up approximately one hour of your time today and 10 minutes every day you choose to use the software. There is a possibility that some of the questions you will be asked might make you feel uncomfortable. In that case, you are free to refuse to answer any question. The study may also have risks that are not now known. You will be told if new information becomes available that may affect your willingness to participate.

#### **Will I benefit by being in this research?**

The software will provide information to you about the benefits of walking and exercise in general, and many tips for walking and for staying motivated. In addition, your participation may help the investigators learn to build new computer systems that will help older adults get more exercise.

#### **Who will see the information about me?**

Information from this study may be used for research purposes and may be published. However, your name will not be used in any publications. In rare instances, authorized people may request to see research information about you and other people in this study. This is done only to be sure that the research is done properly. We would only permit people who are authorized by organizations such as Northeastern University or the federal government to see this information.

#### **If I don't want to take part in the study, what choices do I have?**

You have the option to not participate in the study.

#### **What will happen if I suffer any harm from this research?**

No special arrangements will be made for compensation or for payment for treatment solely because of your participation in this research.

*This is a fake consent document – it does not imply any actual commitments or contract.*

## Appendix A – source Informed Consent documents used in the study

### Research Informed Consent Document 3 – Long-term human-computer relationships

**Can I stop my participation in this study?**

Your participation in this research is completely voluntary. You do not have to participate if you do not want to. Even if you begin the study, you may quit at any time. If you do not participate or if you decide to quit, you will not lose any rights, benefits, or services that you would otherwise have.

**Who can I contact if I have questions or problems?**

If you have questions or concerns at any time, or if you need to report an injury while participating in this research, contact Timothy Brown at (617) 373-1234.

**Who can I contact about my rights as a participant?**

If you have any questions about your rights as a participant, you may contact Thomas Ring, Coordinator, Human Subjects Research Protection, Division of Research Integrity, 413 Lake Hall, Northeastern University Boston, MA 02115 tel. 617-373-4567. You may call anonymously if you wish.

**Will I be paid for my participation?**

You will be paid \$1 every time you successfully use the software (while connected to our server over the Internet), up to once a day between now and 4/1/2011, paid in monthly installments mailed to you. If you do not use the software for two weeks you will lose your eligibility to continue in the study and will not receive further compensation.

**Will it cost me anything to participate?**

There are no costs to you for participating in this research study.

(Signature) ----- Do Not Sign -----

\_\_\_\_\_

(Date) \_\_\_\_\_

*This is a fake consent document – it does not imply any actual commitments or contract.*

## **Appendix B – Background knowledge base**

The following texts were used as background knowledge concepts in the study. Concept titles, in Bold face, were used in creating the associations. Question forms, in Italic face, represent the question that the user could ask (in the Tailored Agent mode) to get the concept information.

### **Randomization**

*Can you tell me about randomization?*

In this study, half of the people will get treated one way and half of the people will be treated another way. To be specific, half of the people will work with a computer system to learn about how to take care of themselves when they go home, and half of the people will get the regular discharge process for Northeastern University Medical Center. You cannot decide if you will be put in one group or the other group, and even I cannot decide this. The decision of which group you would go into is made by random chance. This is like flipping a fair two-sided coin. Half of the time it comes up heads and half of the time it comes up tails. After you decide to join the study, we will tell you which group you are in, but you will not know this before you decide. Even after you join, you are free to quit if you do not like the group you are in, or if you change your mind.

### **Voluntariness**

*What does voluntary participation mean?*

You do not have to be in this study if you do not want to. Even if you choose to be in the study, you can still stop at any time. You are free to be in the study or not, and nothing bad will happen if you refuse to be in the study, and you will not be penalized. If you do not want to be in the study, no one will treat you differently, the care from your doctor will not change, and you will not lose any benefits.

### **Risk**

*Do I have to worry that all these things will happen to me?*

I want to tell you about possible risks, because there is a very small chance that they could happen. Some of the things that could go wrong are possibilities of pain and infection in the collection today. Another thing is that some of the questions we ask could make you sad or upset. Someone could also find out that you were in the study, and learn something about you that you did not want others to know. Also, you could have a legal problem if you told us about a crime such as child abuse, that we are required to report. But you do not have to worry about all of these things happening to you. We are simply telling you that these things could happen, so that you can decide if you want to participate or not.

### **No benefits**

*Why should I take part even if there are no benefits to me?*

Some people are comfortable being in research studies even though they will not benefit directly, because they want to help others in the future who may benefit. This study is like that, and it will not help you directly, but it may help scientists find out ways to diagnose and treat certain medical conditions in the future.

### **Treatment**

*If I am harmed, who will pay for the treatment?*

If you are injured while you are taking part in this study, we can arrange for your treatment but you will be responsible for paying for this. However, we cannot compensate you for any of these, because there is no compensation fund for injuries that are related to research in this study.

### **Best care**

*Does being in the study guarantee the best care for me?*

Being in the study does not mean that you will get special care. For example, I told you that in this study, you might get a computer-based discharge. We hope that it is a little better than the usual discharge process, but we don't know that yet, which is why we're inviting you to take part. Even if you go through the usual discharge process, you will get all the help you need to prepare you to go home. Taking part also does not mean that your doctors or nurses will treat you any differently, and you will continue to get the same kind of care in the future.



## **Appendix B – Background knowledge base**

### **Privacy**

*Who will know that I took part in this study?*

When we store information about you taking part in this study and what you do today, we will make sure that your name is not linked to your information. We will do this by giving your information a code number, and removing your name. Your name will not appear in any of our reports or publications. However, there might still be some information, like your gender, age or race, that someone could use to identify you.

### **Confidentiality**

*Who will see the information about me?*

The only people allowed to see your answers will be the people who work on the study and people who make sure we run our study the right way. Your survey answers, health information, and a copy of this document will be locked in our files. We will not put your answers into your medical record. When we share the results of the study, for example in research journals, we will not include your name. We will do our best to make sure no one outside the study will know you are a part of the study. There are several institutional groups that make sure that that we follow ethical, legal, and quality standards for research. These groups are there to make sure we run our study correctly, and so we sometimes need to share your information with them. In addition, Federal and state agencies can ask to see your information too. Some of these agencies are the US Department of Health and Human Services, the National Institutes of Health, and the Massachusetts Department of Public Health.

### **Alternatives**

*What if I don't want to participate?*

You do not have to be in this study if you do not want to. You will not lose any rights or benefits you already have. Your doctor and the hospital staff will continue to care for you as before.

### **Certificate of Confidentiality**

*What is a Certificate of Confidentiality?*

We have obtained a special certificate from the federal government, called a Certificate of Confidentiality. This protects us against compulsory legal demands, such as court orders and subpoenas, for identifying information about a research participant like you. This adds an additional layer of protection for your privacy.

### **Lose eligibility**

*Can someone else remove me from the study?*

There are some cases where you might be removed from the study by the researchers. For example, if they decide that it is dangerous for you to participate, or if you are not able to follow the study procedures properly, you might be removed from the study. For studies that involve payments, you may be paid for the time you have already spent in the study.

### **Withdraw**

*What happens if I say yes, but change my mind later?*

You can stop being in the study at any time. You will not be penalized, and you will not lose any benefits. You can tell us to stop using and sharing health information that can be traced to you. If you also want us to stop, you have to tell us in writing, by contacting the people whose names are on this form. If you stop, the care you get from your doctor will not change. Because we will use a code number with your information that cannot be linked to your name, even if you decide to withdraw, we will not be able to remove your sample and some of your information from the repository.

## Appendix C – Informed Consent form for participants of study

Northeastern University, College of Computer and Information Science  
Timothy W. Bickmore, Ph.D.  
Agent-Enhanced Document Explanation

### Consent to Participate in a Research Study

We are inviting you to take part in a research study. This form will tell you about the study, but the researcher will explain it to you first. You may ask this person any questions that you have. When you are ready to make a decision, you may tell the researcher if you want to participate or not. You do not have to participate if you do not want to. If you decide to participate, the researcher will ask you to sign this statement and will give you a copy to keep.

#### Why am I being asked to take part in this research study?

You are being asked to participate in this study because you are an English speaking adult.

#### Why are you doing this research study?

The purpose of this research is to learn how people explain documents to one another.

#### What will I be asked to do?

If you decide to take part in this study, we will first ask you some questions about your background and then we will give you three documents that we would like you to read and understand. You will be asked to read one of the documents by yourself, and an animated computer character will explain the other two to you. Before you interact with the animated character, the researcher will first show you how to use the computer system, and then leave the room. Your activity during the experiment may be videotaped, but you may request that the tape segment be erased after we analyze it. After you have had time to understand each document, the researcher will ask you some questions about it.

#### Where will this take place and how much of my time will it take?

The study will take place either in the Human-Computer Interaction Laboratory at Northeastern University. It will take approximately one hour.

#### Will there be any risk or discomfort to me?

You will give up approximately one hour of your time. There is a possibility that some of the questions you will be asked might make you feel uncomfortable. In that case, you are free to refuse to answer any question. In addition, the study may have risks that are not now known. You will be told if new information becomes available that may affect your willingness to participate.

#### Will I benefit by being in this research?

You will receive no direct benefit from your participation in this study. However, your participation may help the investigators learn to build new computer systems that will be able to explain documents.

**APPROVED**

NU IRB# 07-02-11  
VALID: 11/24/09  
THROUGH: 3/15/10

## Appendix C – Informed Consent form for participants of study

### Who will see the information about me?

Information from this study may be used for research purposes and may be published. However, your name will not be used in any publications. All videotapes will be stored in a locked cabinet, accessible only to the researchers, and will be destroyed at the end of the study unless you give explicit permission for them to be used for teaching purposes. In rare instances, authorized people may request to see research information about you and other people in this study. This is done only to be sure that the research is done properly. We would only permit people who are authorized by organizations such as Northeastern University or the federal government to see this information.

### If I don't want to take part in the study, what choices do I have?

You have the option to not participate in the study.

### What will happen if I suffer any harm from this research?

No special arrangements will be made for compensation or for payment for treatment solely because of your participation in this research.

### Can I stop my participation in this study?

Your participation in this research is completely voluntary. You do not have to participate if you do not want to. Even if you begin the study, you may quit at any time. If you do not participate or if you decide to quit, you will not lose any rights, benefits, or services that you would otherwise have.

### Who can I contact if I have questions or problems?

If you have questions or concerns at any time, or if you need to report an injury while participating in this research, contact TIMOTHY BICKMORE at (617) 373-5477.

### Who can I contact about my rights as a participant?

If you have any questions about your rights as a participant, you may contact Nan C. Regina, Coordinator, Human Subjects Research Protection, Division of Research Integrity, 413 Lake Hall, Northeastern University Boston, MA 02115 tel. 617-373-7570. You may call anonymously if you wish.

### Will I be paid for my participation?

You will be paid \$15 for participating in this laboratory session.

### Will it cost me anything to participate?

There are no costs to you for participating in this research study.

### Is there anything else I need to know?

- You must be at least 21 years old to participate.
- You will be one of approximately 96 subjects to be asked to participate in this study.

**APPROVED**  
NU IRB# 07-02-11  
VALID: 11/24/09  
THROUGH: 3/13/10

## Appendix C – Informed Consent form for participants of study

I agree to take part in this research.

\_\_\_\_\_  
Signature of person agreeing to take part

\_\_\_\_\_  
Date

\_\_\_\_\_  
Printed name of person above

\_\_\_\_\_  
Signature of person who explained the study to the  
participant above and obtained consent

\_\_\_\_\_  
Date

**APPROVED**

NU IRB# 07-02-11  
VALID: 11/24/09  
THROUGH: 3/13/12



## Appendix D – Study protocol

### Document Explanation - Protocol

Hi, and welcome. We are developing an animated health advisor agent that we would like you to test out for us. The agent's name is Laura, and she is going to try and explain to you some health documents, and we would like to know how well she can teach something new. Laura looks like this [Demonstrate screenshot], and will talk to you using speech and animation. When it is your turn to say something to her, a list of the things you can say will appear on the right side of the screen, and you can tap your finger on that button.

We will ask you to read three health documents. Laura will try to explain two of those documents, and we will ask you to read the other one by yourself. Those health documents will also talk about research studies and being in hospital, but that is all fake, make believe stuff. Once you are done reading a document, I will ask you a few questions based on it, so I'd like you to pay close attention. We just want to know how well Laura can help you with those documents.

The whole session will about an hour and you will be paid \$15 for your participation. Is this something you are interested in doing? [if no, thank and dismiss].

Now I will go over a few things that you need to understand about the study.

- You are doing this under your own free will and may drop out at anytime without penalty.
- You have not committed to anything yet and you are not required to do anything by being here right now.
- You also need to know that the nature of this study is experimental, and though the risk is minimal, it is research, and may have negative outcomes or no benefit to you at all.
- The possible benefit to you from the study is that you will get a chance to learn some health related information.
- The risks from the study are minimal.

This is our consent form, on last page you will find a consent for use of video form. It is separate from the other consent and just gives us permission to use very short clips of videos for educational purposes, if you do not give us permission for this you can still participate in the study by signing the other consent.

I want to tell you again, the documents you will read in the study are also Consent forms, but they are all fake ones about studies that do not exist.

Please take as long as you need to go over the consent and make a decision. Let me know if you have any questions. [Give subject as much time as they need to read consent forms and make decision.]

## Appendix D – Study protocol

OK, great. Please sign the written consent on the last page.  
[Collect consent forms.]

Thank you. Now I have a few more forms for you to fill out.

[Give Background/Demographic forms]

- Ok, now it's time for your first session:
  - o Please remember the form I am giving you now is a fake, from a study that does not exist. We are only asking you to read it [with Laura] just so we can know how to make these documents easier to understand. We want to see how we can improve documents list this, and we just want your opinion on that. You do not have to sign it, and you will not have to do any of the things it says, but pay attention, because I will ask you some questions about it later.
  - o [if self] please take as much time as you need to read this Consent Form. When you are done, let me know, and I'll ask you a few questions about it.
  - o [if agent] here is the agent [start agent]. She will explain the document, and maybe ask you a few questions. I'm going to step outside while you talk with her, and you can let me know when you're done.
  - o When done, take the informed consent document back. Administer Comprehension questions verbally. [if agent] administer the Instructor Evaluation questionnaire on paper.
  - o Do Cognitive task
  - o [repeat for second and third interactions]
- Great, that's the end of the study.
- Now I'd like to just ask you a few general questions to know what you thought about it.
  - o Can you compare your experience of reading the document yourself, and the agent trying to explain it to you? Was one of them better than the other?
  - o Did you think the Agent helped you to understand the forms I gave you?
  - o Did you see a difference between the two conversations you had with the agent? What were they?
  - o Was one conversation more helpful than the other? Did you like any of those conversations more than the other?
  - o How did you feel about the agent asking you questions?

There are a few things I am required to tell you about this study. We were studying two different ways for Laura to explain documents to someone, and comparing this with someone reading a document by themselves.

- Feel free to ask any questions about the experiment at this time. [Wait for response.]
- Your help has been greatly appreciated, and will aid the Human-Computer Interaction Laboratory in the construction of new software that will ultimately be used to assist people in making

## Appendix D – Study protocol

positive changes to their health behavior. [Answer any questions.]

Okay, this last form, is a receipt to prove that you are a real person, that actually participated in the study and allows you to be paid. [Pay subject.]

---

### **Cognitive Task**

[SAY] "Now I want you to name things that belong to another category: Animals. You will have one minute. I want you to tell me all the animals you can think of in one minute. Ready? Begin."

Start timer as you say "Begin". Write actual responses as legibly as possible on the Worksheet.

Stop the procedure at 60 seconds. One prompt ("Tell me all the animals you can think of.") is permitted if the participant makes no response for 15 seconds or expresses incapacity (e.g., "I can't think of any more."). It is also permissible to repeat the instruction or category if the subject specifically requests it.

Next, read the instructions for the Vegetables category (worksheet follows this instruction page):

[SAY] "Now I want you to name things that belong to another category: Vegetables. You will have one minute. I want you to tell me all the vegetables you can think of in one minute. Ready? Begin."

Start timer as you say "Begin". Write actual responses as legibly as possible on the Worksheet.

Stop the procedure at 60 seconds. One prompt ("Tell me all the vegetables you can think of.") is permitted if the participant makes no response for 15 seconds or expresses incapacity (e.g., "I can't think of any more."). It is also permissible to repeat the instruction or category if the participant specifically requests it.

## Appendix E – Comprehension questionnaires

### **Informed Consent Comprehension - Human Genetic Cell Repository**

**Remember – these questions are only about the make believe consent form that you just went through.**

1. Did you get all the information you needed to make a good decision about participating in the study?  
Was it too much; just right; too little?
2. If you were in a real situation, would you have signed the consent form to participate in the study?  
<show scale>
3. Did you feel any pressure to participate in the study?  
<show scale>
4. Were you satisfied with the informed consent process?  
<show scale>
5. What are the benefits to you of participating in the study? Can you think of any benefits? Can you tell me what those benefits are?
  - No benefits to me
  - might benefit community
  - future benefits
6. What are the risks to you of participating in the study? Can you think of any risks? Can you tell me what those risks are?
  - some pain
  - some possibility of infection
  - some possibility of scarring
  - bleeding
7. If you get injured while your blood sample is being taken, what kind of medical care do you expect the researchers to provide? Who would you expect to pay for this care?
  - Researchers will provide medical care I need / only basic care
  - I may have to pay for treatment
8. If a gene is identified in your sample that puts you at risk for a serious disease, how will you be notified?
  - I will not be notified
  - because my name is not linked to my sample
9. What is the primary purpose of the study?
  - collect genetic material
  - find the relationship between genes and diseases
10. If, after today, you decide that you do not want your sample to be used for any more research, what can you do?
  - Nothing – sample cannot be withdrawn / in the repository forever
  - because my name is not linked to my sample
11. I have been trying to learn about your impressions of the informed consent process that you just now went through. Is there anything else you would like to tell me about it?



## Appendix E – Comprehension questionnaires

**Informed Consent Comprehension - Cardiopulmonary re-hospitalization reduction study**  
**Remember – these questions are only about the make believe consent form that you just went through.**

1. Did you get all the information you needed to make a good decision about participating in the study?  
Was it too much; just right; too little?  
<show scale>
2. If you were in a real situation, would you have signed the consent form to participate in the study?  
<show scale>
3. Did you feel any pressure to participate in the study?  
<show scale>
4. Were you satisfied with the informed consent process?  
<show scale>
5. What are the benefits to you of participating in the study? Can you think of any benefits? Can you tell me what those benefits are?  
- I may be better prepared to [be discharged / to go home]
6. What are the risks to you of participating in the study? Can you think of any risks? Can you tell me what those risks are?  
- embarrassment talking to a nurse about my health  
- possible unknown risks
7. There is a chance that you might go through something other than the routine discharge procedures if you continue with this study. Can you tell me the chance that you'll go through the other set of procedures, and what that will be like?  
- 50-50 / 50% / equal chance  
- other is : 20 minute conversation with discharge advocate  
    : computer-based discharge  
    : computerized telephone calls
8. If you refuse to take part in this study, will the medical care you get here change? If so, how will it change? Will your doctors or nurses treat you differently?  
- My [treatment | care] will not change  
- My doctors and nurses will not treat me differently
9. What is the primary purpose of the study?  
- learn how best to prepare patients for discharging
10. Can the researchers discontinue your participation in this study, even if you decide to participate? If so, why would they do that?  
- Yes, if they decide that taking part is bad for [me/my health]  
- if the sponsor stops the study / funding stops
11. I have been trying to learn about your impressions of the informed consent process that you just now went through. Is there anything else you would like to tell me about it?

## Appendix E – Comprehension questionnaires

**Informed Consent Comprehension - Long-Term Human-Computer relationships study**  
**Remember – these questions are only about the make believe consent form that you just went through.**

1. Did you get all the information you needed to make a good decision about participating in the study? Was it too much; just right; too little?

<show scale>

2. If you were in a real situation, would you have signed the consent form to participate in the study?

<show scale>

3. Did you feel any pressure to participate in the study?

<show scale>

4. Were you satisfied with the informed consent process?

<show scale>

5. What are the benefits to you of participating in the study? Can you think of any benefits? Can you tell me what those benefits are?

- learn about walking and exercise
- be motivated to exercise

6. What are the risks to you of participating in the study? Can you think of any risks? Can you tell me what those risks are?

- injury or health problems from beginning to exercise
- some questions might make me uncomfortable

7. Can you tell me under what conditions your participation in this study might end, without you specifically asking to stop?

- if I do not talk to the Agent for two weeks
- study ends

8. Can you tell me how often we are asking you to use this system, and how much time you will have to spend when you do use it?

- use the system everyday
- 10 minutes per day

9. What is the primary purpose of the study?

- help people stick with exercise programs for many months or years

10. If you get injured while taking part in exercise, what kind of medical care do you expect the researchers to provide? Would you expect to pay for this care?

- no medical care will be provided
- I will have to pay for treatment

11. I have been trying to learn about your impressions of the informed consent process that you just now went through. Is there anything else you would like to tell me about it?

## Appendix F – Scoring Manual for Comprehension tests

Instructions: For each participant, check their transcribed answer against the scoring sheet. The sheet lists the questions (Q5 – Q10), and for each, lists one or more facts that make up a response. For each fact, give the participant 2 points if they substantially mention it, or 1 point if they deserve partial credit with regard to the fact. Incorrect responses, or non-responses get 0 (i.e.: no negative scoring). The sheet will calculate the average for each question (range 0 – 2). Enter this value in the main spreadsheet.

### Genetic Study - Scoring sheet

	Q05		Q06					Q07		Q08		Q09		Q10	
Participant ID	No benefits to me	Some benefits to community / future	some pain	possibility of infection	scarring	bleeding	insurability / employability / future impact of info	Basic care provided	I will have to pay	won't be notified	my name not linked to sample	Collect blood / skin / genetic material	find relationship between genes and diseases	Can't withdraw sample	name not linked to sample
1001															
1002															
1003															
1004															
1005															
1006															
1007															
1008															
1009															
1010															
1011															
1012															
1013															
1014															
1015															
1016															
1017															
1018															
1019															
1020															

### Re-hospitalization reduction study - Scoring sheet

Participant ID	better prepared for discharge	Embarrassment	unknown risks	50-50 / 50% / equal chance	20 min conv with	computer discharge	Computerized telephone calls	Treatment doesn't change	No one treats me differently	Learn how best to prepare patients for discharge	If detrimental	if funding stops
1001												
1002												
1003												
1004												
1005												
1006												
1007												
1008												
1009												
1010												
1011												
1012												
1013												
1014												
1015												
1016												
1017												
1018												
1019												
1020												

## Appendix F – Scoring Manual for Comprehension tests

### Long-term human-computer relationships study - Scoring sheet

	Q05		Q06		Q07		Q08	Q09		Q10	
Participant ID	start exercise routine	learn / get tips / be motivated	injury	other	2 weeks inactive	funding stops	10 minutes everyday	exercise program	long-term maintenance	no compensation / minimal treatment	I would pay for costs
1001											
1002											
1003											
1004											
1005											
1006											
1007											
1008											
1009											
1010											
1011											
1012											
1013											
1014											
1015											
1016											
1017											
1018											
1019											
1020											

## Appendix G – Comparison of scoring

Participant ID	Content	Scores				
		Researcher	Judge1	Judge2	Judge3	Judges' average
1001	Genetic	7.7		8.2		8.2
1002	Genetic	7.7				
1003	Genetic	6.8				
1004	Genetic	6.3				
1005	Genetic	9.2	9.2			9.2
1006	Genetic	5.3				
1007	Genetic	5.8		4.3		4.3
1008	Genetic	3.7				
1009	Genetic	6.2				
1010	Genetic	4				
1011	Genetic	3.5	4			4
1012	Genetic	0.5				
1013	Genetic	2.4				
1014	Genetic	2.4	2.9		3.9	3.4
1015	Genetic	4.7		3.7		3.7
1016	Genetic	1			2	2
1017	Genetic	4.2				
1018	Genetic	8.7			7.5	7.5
1019	Genetic	5				
1020	Genetic	1.4				
1001	RED	4.5				
1002	RED	6		7	7.25	7.13
1003	RED	7.5				
1004	RED	4.5				
1005	RED	7.75				
1006	RED	7.5	7.5			7.5
1007	RED	6				
1008	RED	5				
1009	RED	5.75				
1011	RED	3				
1012	RED	2		2	3	2.5
1013	RED	2.75				
1014	RED	4				
1015	RED	7.5				
1016	RED	0				
1017	RED	5				
1018	RED	8.75	9			9
1019	RED	4	5	5	7	5.67
1020	RED	3.5				
1001	NSF	8				
1002	NSF	7				
1003	NSF	6.5				
1004	NSF	9	7			7
1005	NSF	8.5				
1006	NSF	6.5				
1007	NSF	6				
1008	NSF	7		6		6
1009	NSF	4.5				
1010	NSF	4.5	5		5	5
1011	NSF	2		1		1
1012	NSF	0.5				
1013	NSF	2.5		2	4.5	3.25
1014	NSF	3				
1015	NSF	7			6.5	6.5
1016	NSF	0.5				
1017	NSF	7				
1018	NSF	7				
1019	NSF	9.5	8			8
1020	NSF	4				

## Appendix H - Other study forms

### Questionnaire 1

Page 1 of 2

Respondent ID# \_\_\_\_\_

Date: \_\_\_\_\_

Please take a moment and answer a few questions about yourself:

**Date of Birth:** \_\_\_\_\_

**Sex:** M / F

**Height:** \_\_\_\_\_

**Weight:** \_\_\_\_\_

**Do you smoke:** Y / N

**Ethnic Background (check one):**

American Indian or Alaskan Native \_\_\_\_\_  
Asian or Pacific Islander \_\_\_\_\_  
Black, Not of Hispanic Origin \_\_\_\_\_  
White, Not of Hispanic Origin \_\_\_\_\_  
Hispanic \_\_\_\_\_

**Marital Status (check one):**

Single \_\_\_\_\_  
Married \_\_\_\_\_  
Divorced/Widowed \_\_\_\_\_

**Last grade of school completed (check one):**

Less than high school (0-8) \_\_\_\_\_  
Some high school \_\_\_\_\_  
High school graduate or GED \_\_\_\_\_  
Technical school education \_\_\_\_\_  
Some college \_\_\_\_\_  
College graduate \_\_\_\_\_  
Advanced degree \_\_\_\_\_

**Occupation:** \_\_\_\_\_

## Appendix H - Other study forms

### Questionnaire 1

Page 2 of 2

**How often do you read books (check one):**

Never \_\_\_\_\_  
Less than once a week \_\_\_\_\_  
Once a week \_\_\_\_\_  
A few times a week \_\_\_\_\_  
Every day \_\_\_\_\_

**How much experience do you have with computers (check one):**

I've never used one. \_\_\_\_\_  
I've tried one a few times. \_\_\_\_\_  
I use one regularly. \_\_\_\_\_  
I'm an expert. \_\_\_\_\_

**How do you feel about using computers (check one):**

I don't like them. \_\_\_\_\_  
They're OK. \_\_\_\_\_  
They can be useful. \_\_\_\_\_  
I love playing with them. \_\_\_\_\_

**How comfortable are you using a computer mouse (check one):**

Not comfortable \_\_\_\_\_  
Somewhat comfortable \_\_\_\_\_  
Very comfortable \_\_\_\_\_

**What would you rather use to enter information into a computer (check one):**

A keyboard and mouse \_\_\_\_\_  
A touch screen \_\_\_\_\_

**How would you prefer written information be given to you by a computer (check one):**

As written text \_\_\_\_\_  
Spoken \_\_\_\_\_

**How many research studies have you ever participated in? (check one):**

None \_\_\_\_\_  
One to five \_\_\_\_\_  
More than five \_\_\_\_\_

**So far in 2009, how many research studies have you participated in? (check one):**

None \_\_\_\_\_  
One to five \_\_\_\_\_  
More than five \_\_\_\_\_

## Appendix H - Other study forms

Questionnaire 2		Page 1 of 3						
Respondent ID#	_____							
Date:	_____							
I would prefer complex to simple problems.								
extremely uncharacteristic	•	•	•	•	•	•	•	extremely characteristic
I like to have the responsibility of handling a situation that requires a lot of thinking.								
extremely uncharacteristic	•	•	•	•	•	•	•	extremely characteristic
Thinking is not my idea of fun.								
extremely uncharacteristic	•	•	•	•	•	•	•	extremely characteristic
I would rather do something that requires little thought than something that is sure to challenge my thinking abilities.								
extremely uncharacteristic	•	•	•	•	•	•	•	extremely characteristic
I try to anticipate and avoid situations where there is a likely chance I will have to think in depth about something.								
extremely uncharacteristic	•	•	•	•	•	•	•	extremely characteristic
I find satisfaction in deliberating hard and for long hours								
extremely uncharacteristic	•	•	•	•	•	•	•	extremely characteristic



## Appendix H - Other study forms

Questionnaire 2								
Page 2 of 3								
I only think as hard as I have to.								
extremely uncharacteristic	•	•	•	•	•	•	•	extremely characteristic
I prefer to think about small, daily projects than long-term ones.								
extremely uncharacteristic	•	•	•	•	•	•	•	extremely characteristic
I like tasks that require little thought once I've learned them.								
extremely uncharacteristic	•	•	•	•	•	•	•	extremely characteristic
The idea of relying on thought to make my way to the top appeals to me.								
extremely uncharacteristic	•	•	•	•	•	•	•	extremely characteristic
I really enjoy a task that involves coming up with new solutions to problems.								
extremely uncharacteristic	•	•	•	•	•	•	•	extremely characteristic
Learning new ways to think doesn't excite me very much.								
extremely uncharacteristic	•	•	•	•	•	•	•	extremely characteristic

## Appendix H - Other study forms

Questionnaire 2 Page 3 of 3								
I prefer my life to be filled with puzzles that I must solve.								
extremely uncharacteristic	•	•	•	•	•	•	•	extremely characteristic
The notion of thinking abstractly is appealing to me.								
extremely uncharacteristic	•	•	•	•	•	•	•	extremely characteristic
I would prefer a task that is intellectual, difficult and important to one that is somewhat important but doesn't require much thought								
extremely uncharacteristic	•	•	•	•	•	•	•	extremely characteristic
I feel relief rather than satisfaction after completing a task that required a lot of mental effort.								
extremely uncharacteristic	•	•	•	•	•	•	•	extremely characteristic
It's enough for me that something gets the job done; I don't care how or why it works.								
extremely uncharacteristic	•	•	•	•	•	•	•	extremely characteristic
I usually end up deliberating about issues even when they do not affect me personally.								
extremely uncharacteristic	•	•	•	•	•	•	•	extremely characteristic

# REALM Questionnaire

Respondent ID# \_\_\_\_\_

Date: \_\_\_\_\_

## List 1

fat \_\_\_\_\_  
 flu \_\_\_\_\_  
 pill \_\_\_\_\_  
 dose \_\_\_\_\_  
 eye \_\_\_\_\_  
 stress \_\_\_\_\_  
 smear \_\_\_\_\_  
 nerves \_\_\_\_\_  
 germs \_\_\_\_\_  
 meals \_\_\_\_\_  
 disease \_\_\_\_\_  
 cancer \_\_\_\_\_  
 caffeine \_\_\_\_\_  
 attack \_\_\_\_\_  
 kidney \_\_\_\_\_  
 hormones \_\_\_\_\_  
 herpes \_\_\_\_\_  
 seizure \_\_\_\_\_  
 bowel \_\_\_\_\_  
 asthma \_\_\_\_\_  
 rectal \_\_\_\_\_  
 incest \_\_\_\_\_

## List 2

fatigue \_\_\_\_\_  
 pelvic \_\_\_\_\_  
 jaundice \_\_\_\_\_  
 infection \_\_\_\_\_  
 exercise \_\_\_\_\_  
 behavior \_\_\_\_\_  
 prescription \_\_\_\_\_  
 notify \_\_\_\_\_  
 gallbladder \_\_\_\_\_  
 calories \_\_\_\_\_  
 depression \_\_\_\_\_  
 miscarriage \_\_\_\_\_  
 pregnancy \_\_\_\_\_  
 arthritis \_\_\_\_\_  
 nutrition \_\_\_\_\_  
 menopause \_\_\_\_\_  
 appendix \_\_\_\_\_  
 abnormal \_\_\_\_\_  
 syphilis \_\_\_\_\_  
 hemorrhoids \_\_\_\_\_  
 nausea \_\_\_\_\_  
 directed \_\_\_\_\_

## List 3

allergic \_\_\_\_\_  
 menstrual \_\_\_\_\_  
 testicle \_\_\_\_\_  
 colitis \_\_\_\_\_  
 emergency \_\_\_\_\_  
 medication \_\_\_\_\_  
 occupation \_\_\_\_\_  
 sexually \_\_\_\_\_  
 alcoholism \_\_\_\_\_  
 irritation \_\_\_\_\_  
 constipation \_\_\_\_\_  
 gonorrhea \_\_\_\_\_  
 inflammatory \_\_\_\_\_  
 diabetes \_\_\_\_\_  
 hepatitis \_\_\_\_\_  
 antibiotics \_\_\_\_\_  
 diagnosis \_\_\_\_\_  
 potassium \_\_\_\_\_  
 anemia \_\_\_\_\_  
 obesity \_\_\_\_\_  
 osteoporosis \_\_\_\_\_  
 impetigo \_\_\_\_\_

SCORE	
List 1	_____
List 2	_____
List 3	_____
Raw	_____
Score	_____

## Appendix H - Other study forms

### Instructor Evaluation Questionnaire 1

Respondent ID# \_\_\_\_

Date: \_\_\_\_

Please answer the following questions about the computer instructor who just explained the document to you:

**Write an 'X' on each line (on one of the dots):**

How **satisfied** are you with the **instructor**?

not at all    •    •    •    •    •    •    •    very satisfied

How **satisfied** are you with the **instructional experience**?

not at all    •    •    •    •    •    •    •    very satisfied

How much would you like to **continue** working with the instructor?

not at all    •    •    •    •    •    •    •    Very much

How much do you **trust** the instructor?

not at all    •    •    •    •    •    •    •    very much

How much do you **like** the instructor?

not at all    •    •    •    •    •    •    •    very much

How knowledgeable was the instructor?

not at all    •    •    •    •    •    •    •    very  
knowledgeable

## Appendix H - Other study forms

### Instructor Evaluation Questionnaire 2

Respondent ID# \_\_\_\_\_

Date: \_\_\_\_\_

Please answer the following questions about the computer instructor who just explained the document to you:

**Write an 'X' on each line (on one of the dots):**

How **satisfied** are you with the **instructor**?

not at all    •    •    •    •    •    •    •    very satisfied

How **satisfied** are you with the **instructional experience**?

not at all    •    •    •    •    •    •    •    very satisfied

How much would you like to **continue** working with the instructor?

not at all    •    •    •    •    •    •    •    Very much

How much do you **trust** the instructor?

not at all    •    •    •    •    •    •    •    very much

How much do you **like** the instructor?

not at all    •    •    •    •    •    •    •    very much

How knowledgeable was the instructor?

not at all    •    •    •    •    •    •    •    very  
knowledgeable

## Appendix H - Other study forms

Informed Consent Scales	
Respondent ID# _____	
Date: _____	
<b>Question 1</b>	
On a scale of 1-7, how much information did you get? With 1 being too little information, 4 being just right, and 7 being too much information.	
To little	• • • • • • • too much
<b>Question 2</b>	
On a scale of 1-7, how likely would you have been to sign it? With 1 being extremely unlikely, and 7 being extremely likely.	
Extremely unlikely	• • • • • • • extremely likely
<b>Question 3</b>	
On a scale of 1-7, how much pressure did you feel? With 1 being no pressure, and 7 being extreme pressure.	
No pressure	• • • • • • • extreme pressure
<b>Question 7</b>	
On a scale of 1-7, how satisfied were you? With 1 being extremely unsatisfied, and 7 being extremely satisfied.	
extremely unsatisfied	• • • • • • • extremely satisfied

## Appendix I – Source file DTD

The below XML Document Type Definition (DTD) describes the schema of the machine readable representation that was used for Research Informed Consent documents.

```
<?xml version="1.0" encoding="UTF-8" ?>

<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="detail">
    <xs:complexType mixed="true" />
  </xs:element>

  <xs:element name="file">
    <xs:complexType>
      <xs:attribute name="page" type="xs:NMTOKEN" use="required" />
      <xs:attribute name="path" type="xs:string" use="required" />
    </xs:complexType>
  </xs:element>

  <xs:element name="files">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="file" maxOccurs="unbounded" />
      </xs:sequence>
    </xs:complexType>
  </xs:element>

  <xs:element name="informedConsent">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="source" />
        <xs:element ref="title" />
        <xs:element ref="synopsis" />
        <xs:element ref="sections" />
        <xs:element ref="files" />
        <xs:element ref="locations" />
      </xs:sequence>
    </xs:complexType>
  </xs:element>

  <xs:element name="location">
    <xs:complexType>
      <xs:attribute name="x" type="xs:NMTOKEN" use="required" />
      <xs:attribute name="y" type="xs:NMTOKEN" use="required" />
      <xs:attribute name="ref" type="xs:NMTOKEN" use="required" />
      <xs:attribute name="page" type="xs:NMTOKEN" use="required" />
    </xs:complexType>
  </xs:element>

  <xs:element name="locations">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="location" maxOccurs="unbounded" />
      </xs:sequence>
    </xs:complexType>
  </xs:element>

  <xs:element name="overview">
    <xs:complexType mixed="true" />
  </xs:element>

  <xs:element name="question">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="responses" />
      </xs:sequence>
      <xs:attribute name="correction" type="xs:string" use="required" />
      <xs:attribute name="correct" type="xs:NMTOKEN" use="required" />
    </xs:complexType>
  </xs:element>
</xs:schema>
```

```

    <xs:attribute name="form" type="xs:string" use="required" />
    <xs:attribute name="page" type="xs:NMTOKEN" use="required" />
  </xs:complexType>
</xs:element>

<xs:element name="questions">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="question" />
    </xs:sequence>
  </xs:complexType>
</xs:element>

<xs:element name="response">
  <xs:complexType>
    <xs:attribute name="text" type="xs:string" use="required" />
    <xs:attribute name="id" use="required">
      <xs:simpleType>
        <xs:restriction base="xs:NMTOKEN">
          <xs:enumeration value="0" />
          <xs:enumeration value="1" />
          <xs:enumeration value="2" />
          <xs:enumeration value="3" />
        </xs:restriction>
      </xs:simpleType>
    </xs:attribute>
  </xs:complexType>
</xs:element>

<xs:element name="responses">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="response" maxOccurs="unbounded" />
    </xs:sequence>
  </xs:complexType>
</xs:element>

<xs:element name="section">
  <xs:complexType>
    <xs:choice>
      <xs:element ref="detail" />
      <xs:element ref="overview" />
      <xs:element ref="questions" />
      <xs:element ref="topics" />
    </xs:choice>
    <xs:attribute name="title" type="xs:string" use="required" />
    <xs:attribute name="desc" type="xs:string" use="required" />
  </xs:complexType>
</xs:element>

<xs:element name="sections">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="section" maxOccurs="unbounded" />
    </xs:sequence>
  </xs:complexType>
</xs:element>

<xs:element name="source">
  <xs:complexType>
    <xs:attribute name="url" type="xs:NMTOKEN" use="required" />
  </xs:complexType>
</xs:element>

<xs:element name="synopsis">
  <xs:complexType mixed="true" />
</xs:element>

<xs:element name="title">

```



```
<xs:complexType mixed="true" />
</xs:element>

<xs:element name="topic">
  <xs:complexType mixed="true" />
</xs:element>

<xs:element name="topics">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="topic" />
    </xs:sequence>
  </xs:complexType>
</xs:element>
</xs:schema>
```

## Appendix J – Fact List

Following is the list of 60 facts identified in the Informed Consent document used in the observational study. (See also “Human Genetic Cell Repository” document in Appendix A).

[Sentence 01]

Fact01: subject-consent-(collect-SAMPLE)  
Fact02: SAMPLE = (5ml of blood)-or-(skin biopsy)  
Fact03: SAMPLE-submit-REPOSITORY  
Fact04: REPOSITORY = NU Cell Repository  
Fact05: REPOSITORY-supported-by-(National Institute of Genomics)

[Sentence 02]

Fact06: REPOSITORY-(collect, store, distribute)-SPECIMENTYPE  
Fact07: SPECIMENTYPE = cell culture, DNA sample  
Fact08: SPECIMENTYPE-from-PEOPLE  
Fact09: PEOPLE = people with many kinds of disorders, unaffected family members, other healthy people

[Sentence 03]

Fact10: SAMPLE-create-CELL\_LINE  
Fact11: REPOSITORY-collect-CELL\_LINE

[Sentence 04]

Fact12: collect-purpose-(SPECIMEN-available-RESEARCH\_TEACHING\_STANDARD)  
Fact13: RESEARCH\_TEACHING\_STANDARD = research, teaching, standards in CLINICAL\_LAB  
Fact14: CLINICAL\_LAB=clinical genetics laboratories

[Sentence 05]

Fact15: SAMPLE-valuable-help-AIM  
Fact16: AIM = develop new diagnostic tests, new treatments, new ways to prevent disease

[Sentence 06]

Fact17: REPOSITORY-steps-protect-privacy

[Sentence 07]

Fact18: SAMPLE-(give code number, remove name, not give name to scientists)

[Sentence 08]

Fact19: SOME\_INFO-available-(repository, scientists)  
Fact20: SOME\_INFO = age, sex, diagnosis, race

[Sentence 09]

Fact21: no-absolute-guarantee-confidentiality  
Fact22: small-chance-result-impact-negative-AFFECTED\_PEOPLE  
Fact23: AFFECTED\_PEOPLE=me, my family, other individuals, other groups

[Sentence 10]

Fact24: result-impact-IMPACT\_AREAS  
Fact25: IMPACT\_AREAS=insurability, employability, (family relationships)

[Sentence 11]

Fact26: REPOSITORY-obtain-CONFIDENTIALITY\_CERTIFICATE  
Fact27: CONFIDENTIALITY\_CERTIFICATE=Confidentiality certificate from federal government  
Fact28: CONFIDENTIALITY\_CERTIFICATE-insure-privacy

[Sentence 12]

Fact29: CONFIDENTIALITY\_CERTIFICATE-prevent-FORCED\_DISCLOSURE

Fact30: FORCED\_DISCLOSURE=forced to tell EXTERNALS about participation  
WITHOUT\_WRITTEN\_CONSENT  
Fact31: EXTERNALS=people not connected with study  
Fact32: WITHOUT\_WRITTEN\_CONSENT=without written consent

[Sentence 13]

Fact33: no-DIRECT\_BENEFIT-participation  
Fact34: sample-benefit-BENEFICIARIES  
Fact35: DIRECT\_BENEFIT=direct benefit or payment  
Fact36: BENEFICIARIES=community at large, some particular group

[Sentence 14]

Fact37: RESEARCHERS-no-access-identity  
Fact38: SPECIMEN-result-not-(possible to provide ME\_PHYSICIAN)  
Fact39: ME\_PHYSICIAN=me or my phsyician

[Sentence 15]

Fact40: risk-provide-specimen-minimal

[Sentence 16]

Fact41: (venipuncture)-risk-BLOOD\_COLLECTION\_RISK  
Fact42: BLOOD\_COLLECTION\_RISK=minor transient pain, slight possibility of infection

[Sentence 17]

Fact43: (skin biopsy)-risk-SKIN\_BIOPSY\_RISK  
Fact44: SKIN\_BIOPSY\_RISK=minor local pain, slight bleeding, the possibility of a small scar, slight possibility of infection

[Sentence 18]

Fact45: participating-injury-NO\_COMPENSATION\_TREATMENT  
Fact46: NO\_COMPENSATION\_TREATMENT=no special arrangements for compensation or for payments for treatments

[Sentence 19]

Fact47: study-data-publishable

[Sentence 20]

Fact48: publish-no-identifying-information  
Fact49: RESEARCHERS-no-access-identity

[Sentence 21]

Fact50: donation-voluntary  
Fact51: no-participation-NO\_PENALTY\_LOSS  
Fact52: NO\_PENALTY\_LOSS=no penalty or loss of benefits entitled

[Sentence 22]

Fact53: SAMPLE-indefinitely-in-REPOSITORY  
Fact54: SAMPLE-withdraw-impossible

[Sentence 23]

Fact55: questions-comments-STUDY-CONTACT\_1  
Fact56: CONTACT\_1=Dr. Brian Orlow (617)-373-1234

[Sentence 24]

Fact57: questions-REPOSITORY-CONTACT\_2  
Fact58: CONTACT\_2=Dr. Micheal Jack Northeastern Institute for Medical Research (617) 373-5678

[Sentence 25]

Fact59: questions-rights-(research-subject)-CONTACT\_3  
Fact60: CONTACT\_3=Dr. Laura Silliman (617) 7890