Design and Evaluation of a Web-based Training tool for the User Action Framework Explorer

Venkatramanan Balasubramanian

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
In
Industrial and Systems Engineering

Dr. Tonya L. Smith-Jackson, Ph.D, Co-Chair
Dr. H. Rex Hartson, Ph.D, Co-Chair
Dr. Robert C. Williges, Ph.D

Keywords: User action framework, UAF, training tool, web-based training, usability, human factors

May 10, 2002

Blacksburg, Virginia

DESIGN AND EVALUATION OF A WEB-BASED TRAINING TOOL FOR THE USER ACTION FRAMEWORK EXPLORER

Venkatramanan Balasubramanian

ABSTRACT

The proliferation of personal computing power and its associated affordability has seen an explosion in the use of computing and interactive systems of all kinds. This growth has brought with it an awareness among developers of interactive systems about the importance of user centered design and usability. In the Virginia Tech Usability tools lab, efforts are underway to develop tools and processes that assist in usability evaluation. This has led to the development of the User Action Framework (UAF) (Andre et al., 1999), a framework that forms the basis of several usability inspection tools like the Explorer, UPI (Usability Problem Inspector), UPC (Usability Problem Classifier) and the usability problem database. The UAF explorer is the tool in the toolkit, which allows users to explore and learn the structure of the UAF. The framework, in its final stages of development, is based on an interaction cycle derived from Norman's action model (1986). For its acceptance, the user action framework, which is based on human factors and cognitive concepts, needed a training program to make it accessible and understandable to the usability practioners in industry and academia.

This thesis addressed the following research activities: (1) Developing a web-based training tool for the User Action Framework explorer and (2) Evaluating a web-based training tool using various formative evaluation techniques and a final summative evaluation to measure effectiveness of the training, transfer of training, knowledge/skill acquisition and reaction to the training.

The summative evaluation used a pretest-posttest between subjects experimental design to determine the effectiveness of the training tool with the use of lecture-based training as a control group. The data collection included objective measures of performance and subjective measures through questionnaires and rating scales. The summative evaluation found no significant differences between gain scores on lecture-based training and web-based training under controlled conditions.

Opportunities for future research were identified and the training tool is expected to contribute to the efforts of the VT usability tools lab towards educating usability professionals and researchers alike on the usefulness of the User Action Framework and its associated tools. This work also seeks to proliferate the use of web based training methods as a valuable way to train remote learners on such developing frameworks and toolkits.

Acknowledgements

I would like to extend my gratitude to my advisors Dr. Tonya Smith-Jackson and Dr. Rex Hartson for their continued support and guidance. I have learnt much more than just courses from you in my years as a student and this work would have never been possible without you. I would also like to thank Dr. Robert C. Williges for his expertise and constant encouragement in his role on my thesis committee. Your dedication to helping students remains unparalleled.

I also thank my family for their continued love, concern, and support without which none of this would have been possible.

Dedication

I dedicate this work to my family, my greatest source of endless support and inspiration.

List of figures:

List of tables:

CHAPTER 1: INTRODUCTION

1.1 RATIONALE

With the coming of the digital era and the proliferation of personal computing power, the user is faced with numerous ways to do things. As the importance of involving the user in the design process gained credibility, a suitable terminology came into being. This concept of designing for the user and involving the user in the design process is termed user-centered design and has gained prominence in both industry and academia where efforts are underway to develop tools that assist in user centered design. A core component of user-centered design is designing for usability. Nielsen (1993) defines usability as an attribute that comprises but is not restricted to, effectiveness, efficiency, learnability, low error rate and satisfaction with the interface that is being evaluated. There currently exist a variety of usability evaluation techniques like usability testing methods, usability inspection methods and usability inquiry methods. These evaluation techniques present evaluators with data, both quantitative and qualitative. In the World Wide Web, which is just one of the few domains where usability has gained tremendous significance, it has been estimated that 65% of virtual shopping trips on the Web end up in failure (Souza, 2000) and usability has been cited as the primary reason for such aborted shopping attempts.

While usability as a field is exploding, it has created a demand for theory, unifying frameworks and technologies that help assess, quantify and address usability of various systems. The absence of a proper theoretical framework leaves many evaluations incomplete in that results are not organized in a proper way, preventing maximum utilization of collected data.

The User Action Framework (UAF) (Andre et al., 1999) serves to fill the gap and meet the demand for an integrating framework. It is a structured knowledge base of usability concepts, issues, and situations. The UAF is structured by the Interaction Cycle, an adaptation and extension of Norman's seven stage theory of action (Norman, 1986). The user action framework forms the underlying concepts built on a theory based framework of usability issues and concepts. Based on these concepts, development of a set of usability engineering support tools is currently underway. These tools include the UAF explorer, Usability Problem Inspector (UPI), Usability Problem Classifier (UPC) and the usability problem database. This set of tools aid classification of usability problems obtained during observation into various categories that bring out the causal factors of the problems in the interaction cycle. The User Action Framework (UAF) developed to serve as a foundation for a suite of structured methods and integrated tools, aims to get the most value from usability data by way of analysis, classification, storage and retrieval, reporting, and redesign (Andre et al., 1999). The UAF explorer tool is used to browse through the UAF and to help the practitioner learn about its content and structure.

Learnability and usability of various inspection techniques has often been a point of discontent among researchers and practitioners (Mack & Nielsen, 1994). The UAF and the explorer training tool, while not strictly inspection techniques, are tools that support inspection techniques. They therefore were developed with their own usability and learnability in mind. Being an 'in development' tool, the UAF has had limited exposure to the usability community and the complex cognitive structuring of the UAF has proved challenging to understand. This difficulty served as the primary motivation for this research. Complex cognitive structuring and lack of exposure necessitate the development of a training tool for the User Action Framework Explorer, which serves as the basis of understanding for further use of the UAF suite of tools.

The goal of this research was to develop a training tool to train a novice usability group with knowledge of HCI and usability concepts to demonstrate a complete understanding of the hierarchical structure and the underlying concepts of the User Action Framework (UAF). Based on that understanding, the trainees would be able to use the UAF Explorer tool to successfully navigate the UAF and classify usability problems according to its structure and use the other tools that form the UAF toolkit. One of the motivations for developing a training program is that a user of the UAF takes substantial time to learn how to navigate the UAF, without being proficient in its concepts and this adversely affects reaction to the tool. If the user were trained to understand these concepts, they would take less time to learn how to use the tools associated with the UAF, and use them effectively to classify usability problems. The research uses the systems approach model (Dick and Carey, 1996) for designing instruction and draws heavily on previous research conducted by the author and others wherein the first stages of the model were implemented. A needs analysis study conducted by the author (Balasubramanian et al., 2000; Appendix A) shows both a need for development of a training tool to address the problems with existing classroom training and a desire among usability practioners in the field on being trained on such a framework. The study revealed that training was considered necessary for the set of concepts on which the rest of the toolkit was based.

The work conducted in this research concentrated on the second half of the instructional design model as outlined in the systematic design of instruction and is elaborated in subsequent sections.

1.2 BACKGROUND

1.2.1 The User Action Framework

The User Action Framework is a structured knowledge base organized as a search or classification space of usability concepts, issues, and situations (Andre et al, 1999). Donald Norman (1986) in his work on developing a theory of action, which is described as a continuum of stages on the action/execution side and the perception/evaluation side of human-system interaction started efforts on forming a classification framework. It was also acknowledged by the author that the theory of action was incomplete and adaptable. The work went on to describe seven stages that are:

1. Establishing the goal
2. Forming the intention
3. Specifying the action sequence
4. Executing the action
5. Perceiving the system state
6. Interpreting the state
7. Evaluating the system state with respect to the goals and intentions

(Norman, 1986; pp 31-61)

Previous attempts at a classification framework include Cuomo and Bowen (1994), which also used Norman's (1986) theory of action model in the context of support for a direct manipulation framework. But this framework had a few downsides. For example, the integration and classification of data came after the evaluation and was not built into the method used by evaluators. This resulted in loss of the original information collected by the evaluators.

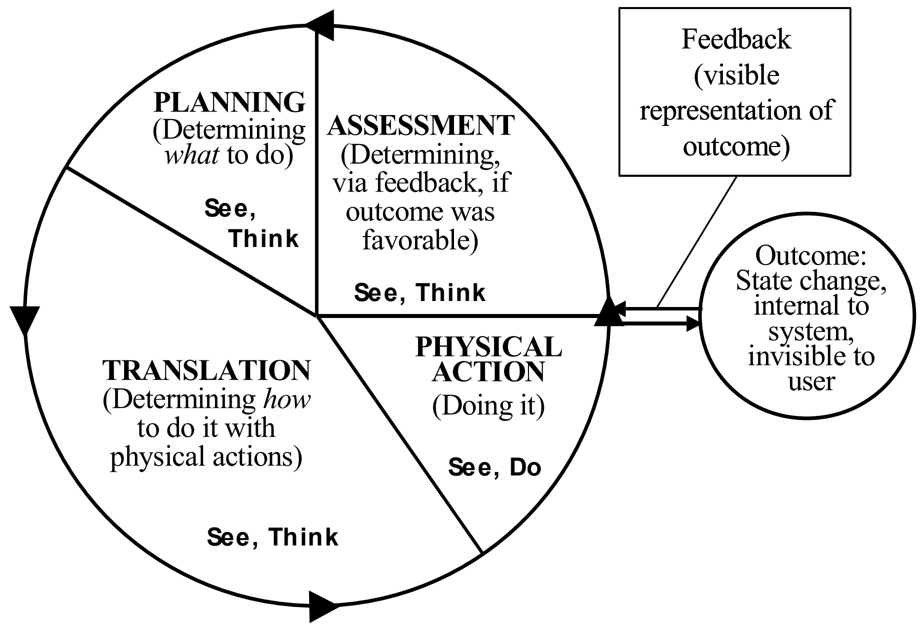
The theory-based UAF is structured by the Interaction Cycle, an adaptation and extension of Norman's seven stage theory of action. The user action theory consists of 5 stages with one system dependent stage. These stages, termed interaction activities, make up the highest level of the UAF tree and are:

1. Planning
2. Translation
3. Physical Actions
4. Outcome (System Interaction cycle)
5. Assessment
6. Independent (subject to change)

Adaptation of Norman's model to the user Action framework can be summarized as shown in **Table 1-1**:

Table 1-1: Adaptation of Norman's theory of Action to the User Action Framework

Norman's model	**Adaptation to the User action framework**	**Meaning in terms of usability issues**
Execution (Establishing the goal and forming the intention.)	Planning	Can user determine what to do?
Mapping	Translation	Can user determine how to do it (what physical actions)?
Physical activity (Executing the action)	Physical user action	Can user do actions (easily)?
Evaluation (Perceiving system state, interpreting state and evaluating state)	Assessment	Can user determine whether plan/actions worked?



[*Text and pictures adapted from the Usability tools lab at Virginia Tech, <http://miso.cs.vt.edu/~usab>]*

Fig: 1-1: Graphical Representation of the User Action Framework Stages

Fig 1-1 depicts a graphical representation of the various stages in the user interaction cycle as identified above.

Besides the above-illustrated stages, the user action framework includes another category called "Overall/independent". This category represents usability problems and issues that are general to the whole interaction design and does not arise just in one part of the interaction cycle. The various stages in the interaction cycle as outlined by Andre (1999) are explained in the subsequent sections.

1.2.1.1 Planning

Planning occurs when users determine what actions need to be taken and how to go about completing those actions. Planning includes all thinking by users to work out what to do and how to do it. Planning is concerned with the user's ability to understand the overall computer application in the perspective of work context, problem domain, environmental requirements and constraints. The primary focus is on the system model and metaphors, and the user's knowledge of system state and modalities. High-level planning includes user work goal decomposition across a hierarchy of plan entities: goals, task, and intentions, all expressed in cognitive problem-domain language.

1.2.1.2 Translation

The user translates intentions into plans for physical action(s). Translation is largely about cognitive affordances to support the user's ability to plan physical actions. The user draws on knowledge, experience, and cognitive affordances in the interaction design to determine, establish, or ascertain an action plan to carry out the intention. The activity here involves translating from the language of the problem domain to the language of actions upon user interface objects. For example, the intention to make a word bold in some text might result in a plan to drag-select the word and click on the "bold" button. The Translation sub-part is perhaps the most difficult for both designer and user, and accounts for a large proportion of usability problems observed in the field. Translation is purely cognitive- the user has formed a mental plan for actions, but has not yet done those actions. Translation issues also include effectiveness of content or cognitive affordance meaning (e.g., issues of clarity, completeness, error avoidance, consistency, and the ability to predict the effects of a given user action). Translation also includes preferences and efficiency issues (e.g., user control, task structure, number of steps, and short cuts).

1.2.1.3 Physical Action

Executing the planned actions is the focus of the Physical Action component of the Interaction Cycle. The Physical Action part of the Interaction Cycle is about perceiving objects to manipulate and manipulating objects. Perception has to do with the factors like but not restricted to noticeability, legibility, contrast, and timing. Object manipulation has to do with interaction complexity, input/output devices, interaction styles and techniques, manual dexterity, layout (Fitts' law), and physical disabilities.

1.2.1.4 Outcome (System Interaction Cycle)

As a result of physical user actions, the system computes an outcome, an internal and invisible state change. The system then produces a system response, providing feedback representing the outcome to the user. The only way a user can know something about the outcome is indirectly by way of the system response. Therefore, the outcome is primarily a system interaction cycle issue, but can have usability consequences. Outcome issues relating to usability problems include system automation, locus of control, and system errors.

1.2.1.5 Assessment

The Assessment component parallels the Translation component in that it has to do with presentation of feedback, meaning of feedback, and preferences and efficiency. However, these issues now apply to the feedback and how it supports the user's ability to gauge the outcome of physical actions. In addition to the usual issues of legibility, noticeability, timing, layout, grouping and presentation of feedback also involve complexity, clutter, consistency, organization of information displays, and presentation medium. Effectiveness of feedback content and meaning, as one might guess, depends on clarity, completeness, sufficiency, correctness, relevance, and consistency of feedback.

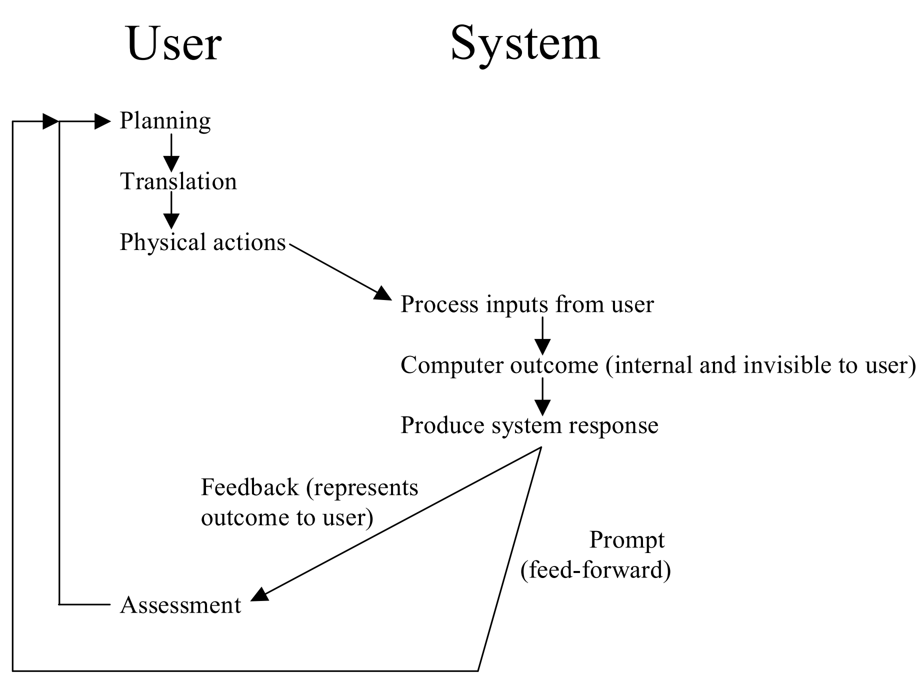


Fig 2-1: User and System interaction stages in the User Action framework

The division of the activities between the user and the system in the interaction cycle is shown in the above illustration. It is observed that Outcome is system dependent.

1.2.2 Design of Training

The aforementioned Systems Approach Model for Designing Instruction (Dick and Carey, 1996) is commonly used as a representative instructional design model, and for teaching instructional design to novice instructional designers. The model includes nine related processes. These processes are: (1) determine instructional goal, (2) analyze the instructional goal, (3) analyze learners and contexts, (4) write performance objectives, (5) develop assessment instruments, (6) develop instructional strategy, (7) develop and select instruction, (8) design and conduct the formative evaluation of instruction, (9) revise instruction, and (10) conduct summative evaluation (Dick and Carey, 1996; pp 5 - 7). The development of training for the user action framework explorer tool follows this design. The author of this work was part of a team that implemented the first stages of the outline proposed and the research as described in this thesis revolves around the actual development and evaluation of the tool. The results of these studies (Balasubramanian et al., 2000) are included in Appendix A. Figure1-3 below graphically illustrates the systems approach model for training explained above. The dotted line indicates the scope of this research.

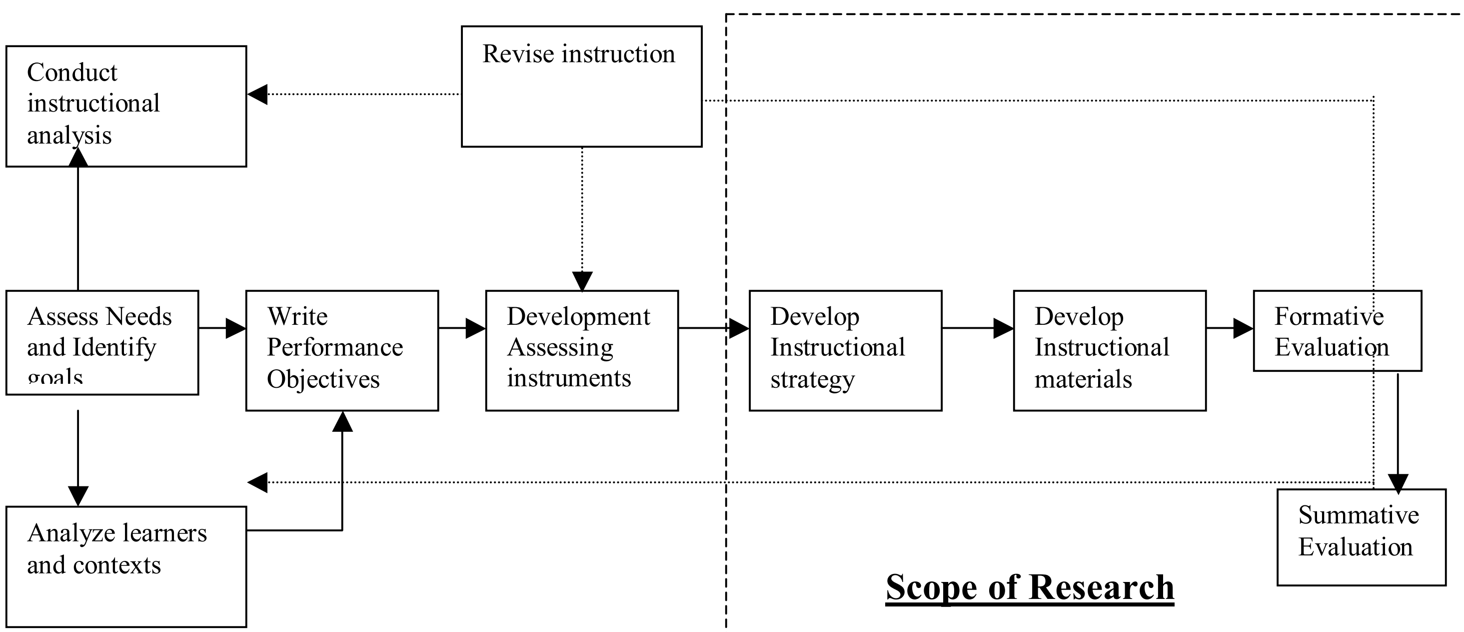


Fig 1-3: Systematic design of Instruction (Adapted from Dick and Carrey, 1996)

1.2.2.1 Computer aided instruction

Computer Aided Instruction (CAI) has students communicate with a program in the computer, which may provide "tutorial, drill and practice, or simulation and modeling exercises" (Hiltz, 1994, p. 21). More recently, the term 'intelligent tutoring systems' (Larkin & Chabay, 1992) has been used to describe CAI. Intelligent tutoring system or ICAI are systems that utilize the techniques of expert systems and artificial intelligence to help a person learn interactively (O' Lander, 1990). These systems have achieved enormous popularity among various communities. CAI can be delivered by different media like CDs and the World Wide Web (WWW). Currently the web based delivery of CAI or web based training or web based instruction, as it is known, is immensely popular due to the advantages offered by using the WWW as a medium. Web based instruction (WBI) can be defined as a hypermedia –based instructional program which utilizes the attributes and resources of the WWW to create a meaningful learning environment where learning is fostered and supported (Harasim, 1990). The history of computers in instruction can be traced from Web Based Training evolved directly from computer-based training (CBT). Computer based training was not Internet delivered, but was still delivered via the computer. Computer based training in turn evolved from the older text-based CAI (computer assisted instruction) linear programs that were popular in the 1970s until late 1980s. Studies have also shown web-based training as having numerous advantages over traditional computer-aided instruction. Some of the common advantages of WBT over other forms of CAI include:

- easy delivery of training to users
- easy updating of content (specifically as the UAF is a moving target)
- a shift from costly programming to interface and content design
- quicker turnaround of finished product
- controllable access
- availability of additional resources via links.

However, web based training also introduces disadvantages. These include:

- Limited formatting of content in current browsers
- Bandwidth/browser limitations that restrict instructional methodologies
- Access to server space and other infrastructure.

(Kilby, 1999)

Experimental results have also shown that students dislike reading from the screen and in some cases have printed out the parts of content to avoid having to read from the screen (Wade and Power, 1998). Despite these isolated drawbacks, Web based training environments are rapidly growing in popularity. There exist numerous resources devoted to web based training and educational multimedia and as computing costs decline it can be expected that these learning environments will make inroads into training at all levels.

Harasim (1990) states that the mere introduction of computer-mediated communication would by itself not improve learning and that design (or method) is crucial in ensuring the success of such an effort. The training tool developed- computer aided instruction modules will follow the widely adopted systems approach as outlined by Dick and Carey (1996). The approach includes a front-end analysis to find out about contextual factors and goals, a process for creating materials or instruction, and a feedback/evaluation mechanism to determine if the process is working. The scope of this thesis is limited to the latter 2 points as established by the approach. As mentioned earlier, this is because the author has already undertaken a needs analysis for the training program and the tool as part of an earlier project and these results are available in Appendix A.

1.2.2.2 Part task training

Part task training can be used to break the task into components that are practiced separately. Baldwin and Ford (1988) proposed that the complexity of the task and the relationship among the components determine the usefulness of whole and part task methods. Sequencing of learning activities in various stages can be used to ensure the trainee has developed capabilities necessary to proceed to the next part of the task. Training is designed such that the subtasks and goals are interrelated and the learners must learn each subtask to reach a new cognitive level. This lends reason for the use of part task training method because each subtask can be considered a module to be trained for the learners to achieve the instructional goal. It should however be mentioned that this is by no means justification of the use of the part task training approach for the UAF, but merely a reasonable assumption. Training content will be designed to train the learners on each of these subtasks (parts). The training content will be developed in such a way that the learners should have a thorough understanding of each module to proceed to the subsequent modules. Hence training content can also be considered progressive.

1.2.3 Design of content

This section discusses two approaches to design of the actual training content: the minimalist instruction approach and accompanying research literature on it and the problem based learning approach.

1.2.3.1 Minimalist Instruction

Minimalism emphasizes the streamlining of instructional materials, and use of a task-oriented approach to determining and organizing the content of tutorials or other educational artifacts (Carroll and Rosson, 1990). The goal is to support users in accomplishing real and meaningful tasks quickly, and to allow them to take advantage of their existing task knowledge when learning about a new system. Minimalism as advocated by Carroll (1984) applied primarily to the design of instruction manuals. But subsequently work has been done on the design of online tutorials using a minimalist approach.

An aspect of minimalist training is the kind of hands- on practice provided in the training materials. Minimalist training advocates argue that people prefer active learning or learning by doing rather than by reading a manual and that hands-on practice may be essential for both learning and motivation in training. Charney et al. (1986), in studies on practice methods, found that methods involving problem-solving practice were superior to merely reading or typing worked-out examples

Studies by Carroll et al. (1987) support the effectiveness of the minimalist training approach compared to commercial tutorial materials. However the scope of application of minimalist instruction to the design of a training tool for the user action explorer is limited by the fact that minimalist instruction is exploratory in nature and takes time to study under experimental conditions. A limited minimalist approach is used in that numerous exercises for hands on practice will be included.

1.2.3.2 *Problem- based learning*

Problem-based learning (PBL), which originated in medical classification studies, is based on the premise that medical education would be more effective if medical facts were learned, not in isolation, but in the context of clinical cases. Therefore, in a PBL setting, students learn and are self-taught in the context of medical problems. Students collaborate, consult textbooks and other resources, and teach themselves with guidance and direction. (Berkson, 1993).

Problem-based learning has been extensively used in computer science education with success (Fekete et al., 1998). One characteristic element of problem-based learning is the division of learners into groups and this is difficult to implement under the scope of this thesis. While most of these studies used the definition of PBL to divide students into groups, this project, due to various constraints like time and space and the necessity to promote individual, rather than group learning as in PBL, did not use this approach.

The tasks are not tightly defined to exercise a single recently explained topic, but instead they are complex, involving ideas from many interacting disciplines. Trainees analyze numerous aspects of the scenario described, track down the information about possible causes, test to confirm or deny each explanation, treatment and their interaction with other parts of the whole situation. Studies have shown that PBL has some similarities with Discovery Learning, but PBL is more structured, with explicit learning objectives set by academics, and considerable amounts of facilitation provided to students in organized ways against the more trainee centered discovery learning approach (Baldwin, 1996).

Thus, while it may seem harder to incorporate technical material in a PBL subject, there is the prospect that whatever is introduced will be mastered more thoroughly, since students will actually use it in the context of a real task. This research will adopt a limited problem-based learning approach to the design of training and content. This is because the training tool is developed around usability problems identified in actual usability testing and this fits well into the PBL approach. However the training tool would cater to individual self- paced learning scenarios and not to group work as identified in the needs analysis (Appendix A). Hence the training tool developed for this study qualifies for a limited problem based learning approach.

1.2.4 Design of the instructional interface

Lohr (2000) in his work addresses the need for a procedural instructional interface design by describing a learner driven application of the instructional system design process to facilitate design of the interface elements within the interface. The work also identifies critical design guidelines based on graphic design, information design and instruction design. Furthering this work, Reigeluth and Nelson (1997) describe a paradigm of ISD where design and development follow an iterative series of ASEC (Analysis-synthesis-Evaluation-change) cycles. In his paper, Lohr discusses the implementation of the ASEC procedure to designing instructional interfaces.

The ASEC procedure consists of 4 main steps: Analysis, synthesis, evaluation and change as summarized below.

Table 1-2: ASEC Procedure for designing Instructional Interface

Phase of ASEC	**Description**
Analysis	This phase involves identification of environment, teacher and learner tasks that must be performed or facilitated by the interface. Tasks can be categorized as informational presentation and instructional practice.
Synthesis	Creation of interface elements to address needs identified in the analysis phase using design guidelines.
Evaluation	Testing the effectiveness of the interface in communicating with representative users.
Change	Identification of interface elements that require modification based on results of the evaluation phase. Redesign guided by interface design principles as in synthesis phase.

In his work, Lohr also brings out the presence of numerous design guidelines. Adhering to all of these requires very high design efforts, if at all it is possible. These include Nielsen's (1993) heuristics like simple and natural dialog, minimization of user memory load, consistency, feedback, shortcuts, prevention of errors, error messages, help and documentation. Galitz (1997) addresses principles for aesthetic appeal, clarity, compatibility, comprehensibility, consistency, efficiency, predictability, responsiveness etc. The above-mentioned are just examples of the existing guidelines.

Developing WBT systems demands a system of development, and a new look at instructional design of content. One approach that has been receiving increased attention recently is the use of the Rapid Application Development (RAD) model. It is considered an approach that can potentially reduce the time required to deliver high quality software. The RAD development lifecycle is designed to give much faster development and higher quality results than the traditional lifecycle. It is designed to take maximum advantage of powerful development software that has evolved recently. As in this case, using the RAD systems development for Internet courses involves the developer's use of GUI programming tools such as the Web authoring packages (Java or HTML editors or visual programming languages such as Visual Basic or Lingo (Authorware or Director) or Open Script (Toolbook or IconAuthor.)

In development of this work, which will follow the ASEC procedure, Macromedia Dreamweaver will be used as a web development tool.

1.2.5 Evaluation of the training tool

Evaluation is defined as "the systematic collection of descriptive and judgmental information necessary to make effective training decisions related to the selection, adoption, value, and modification of various instructional activities."
(Goldstein, 1993. p 181).

In this research, evaluation is seen as an integral part of the research both in evaluating the program to make changes if necessary and determining the effectiveness of the training.

This study proposed to conduct evaluation in two phases. The first phase, the formative evaluation, is an iterative process that is used to determine if the training program is operating as originally planned or if improvements are necessary before the program is implemented. During formative evaluation, data will be collected that can be used to revise the instruction to make it more efficient and effective. According to Dick and Carey (1996), formative evaluation helps to collect information during the development of training programs, which can be used to improve the effectiveness of the training. Formative evaluation work succeeds to the extent that it identifies priorities for redesign and refinement, and all the more to the extent that it provides concrete guidance as to how that redesign and refinement should be executed (Carroll and Rosson, 1995). The criteria for formative evaluation will center on accuracy, aptness of content and delivery style and flow of information through modules.

The formative evaluation techniques used will include:

(1) Subject Matter Expert (SME) evaluation

(2) One-on-One evaluation

(3) Non-intrusive individual evaluation

The second phase, the summative evaluation, is carried out to decide whether to maintain the existing instructional materials and also about adopting materials that may meet the instructional goals. Dick and Carey (1996) define summative evaluation as design of evaluation studies and data collection to verify effectiveness of instructional material with target learners. In order to perform a summative evaluation, an experimental design should be developed and data collected from target learners to verify the effectiveness of training materials in terms of actual learning and this forms the core of the experimental design outlined in a subsequent section. The evaluation criteria for summative evaluation will be the overall instructional goal and behavioral objectives as outlined in earlier sections and reaction towards and acceptance of web based training for the User Action Framework. The instructional goal for this research is set as follows:

'Given a set of usability problems, the trainee should be able to classify them according to the hierarchical structure of the UAF with at least 80% accuracy' (Balasubramanian et al., 2000).

Various data collection methods were used to identify performance objectives for evaluation of training (Balasubramanian et al., 2000). The criterion for acceptable performance was set as 80% accuracy. This is because this instructional goal provides for a reasonable, achievable goal. Since the skill to be developed is highly cognitive, 100% accuracy is seen as being impractical. By allowing for 20% or less errors, we will inculcate in trainees 'learning by errors', which is very useful and significant for highly cognitive tasks like understanding the UAF. Besides, allowing for errors helps train learners in error recovery and this becomes necessary as in practical situations users might make errors. The training system will allow for feedback that supports error recovery while training.

The summative evaluation method chosen is the field trial method. In this procedure, use of the tool with the target trainee population will be studied within an experimental setting. The studies will collect data that are both qualitative and quantitative in nature. This follows from observations made in studies regarding the nature of data when evaluating training. Mixing quantitative and qualitative methodologies results in high value when trying to study the richness and complexity of web based learning environments (Owston, 2000).

Nash et al., (1971) describe a variety of difficulties that hinder interpretation of data collected for evaluating CAI systems under effectively controlled conditions. A majority of students indicated that they believed their learning experience would be improved if the computer facilities were available within dedicated 'learning zones' such as college libraries or more private spaces. It is also observed that students strongly felt such monitoring was invasive and potentially damaging to the course experience (Wade and Power, 1998). The experiment will be designed to avoid these confounds.

The single most consistent result in numerous experiments investigating various diverse interests in research on use of computers, is that individual difference variables account for much more of the variance than any experimental manipulation. Studies also show that it is reasonable to expect that such variables will influence novice learning of computer systems (Nielsen, 1994).

Previous attempts to compare the efficacy of computer-assisted instruction to the traditional medium have produced mixed results (Lautar-lemay, 1993). However the conditions under which CAI is effective has not been properly investigated. This is because most of the studies carried out so far tend to study multimedia training or other forms of interactivity in training. This research seeks to compare the effectiveness of a traditional approach to a web based training approach. A few such examples are in a series of six investigations between 1990 and 1992. Adams (1992) showed that students using interactive web delivered multimedia had a fifty-five percent learning gain over students receiving traditional classroom teaching. Other important findings included that the students assimilated material sixty percent faster, and their long-term (30 day) retention was from twenty-five percent to fifty percent higher. However most of the above-mentioned studies focus more on the effect of interactive multimedia that is web delivered rather than web based text centric content.

While lecture-based training was found to be superior to self-study in enhancing learning (Davis and Davis, 1990), no significant difference was found between instruction-based and exploration-based training strategies (Davis and Bostrom, 1993). Behavior modeling, which combined lectures with hands-on experience, resulted in superior knowledge retention and learning compared to other training methods (Compeau and Higgins, 1995). In a recent review of empirical findings on hypermedia and learning outcomes (Dillon and Gabbard, 1998), it was observed that the use of hypermedia -based instructional systems in education had not produced significant learning gains. Two studies that were reviewed by Dillon and Gabbard are explained in further detail. Van der Berg (1991) developed a level of abstraction-structured test (LAST) covering statistics and hypothesis testing. The experiment, which used a lecture attending control group, found that the level of performance did not differ across instructional settings in objective testing. It was observed that when used as the only alternative to lecture based training it was evaluated subjectively the same way as the control group evaluated lectures. In another study, Blanchard (1990) examined community college students using a hypermedia system to learn MS-DOS. Blanchard found that in the experimental group a greater number of students passed the test and had higher mean scores and posted a significant increase in mean test scores between pretest and posttest. The following research attempts to evaluate the effectiveness of using web-based training to lecture based training by developing a training program for the UAF Explorer tool of the user action framework.

1.2.6 Problem Statement/Research objectives

Given the above background and previous studies, it is determined that the user Action Framework is a reliable tool for classifying usability problems (Andre, 1999) and that web-based training is rapidly becoming a preferred alternative to classroom training. However there is little on the effectiveness of web based training and comparing web based training to classroom training in transferring cognitive skills and concepts and this research seeks to develop bridge these gaps by developing and evaluating the effectiveness of a training tool aimed at academia and industry alike while underscoring the use of the UAF as a valuable post usability evaluation tool set.

Following the outline above, the research objectives, broadly stated, translate to:

Study 1:Formative Evaluation

- To apply training systems design principles to design a web based training tool for training users on the UAF explorer tool and formatively evaluate this tool.

Study 2: Summative Evaluation

- To evaluate and determine the effectiveness of learning of the web-based training and compare against the effectiveness of lecture based training.
- To measure and compare preferences, reactions and effectiveness of the web based training tool against classroom training.

1.2.7 Research hypotheses

The above mentioned research objectives in study 2 translate into the following research hypotheses:

- The web based training tool will be effective for training users to use the UAF explorer tool to classify usability problems. Effectiveness will be demonstrated by achieving 80% accuracy in classifying problems.
- The web-based training delivery system will be equivalent to or at least as good as classroom training delivery system as shown by significant differences on statistical tests conducted on accuracy scores and subjective ratings.

Support for these hypotheses will aid consideration of replacing classroom training with the web-based training tool.

CHAPTER 2: STUDY 1: FORMATIVE EVALUATION

2.1 SUBJECT MATTER EXPERT EVALUATION

2.1.1 Participants

Participants for the formative evaluation procedures included subject matter experts (SME) for the first phase, which is subject matter evaluation. Participants for the SME evaluation phase were remote participants well versed in the concepts of the UAF and involved with various stages of design and development of the UAF. Three SME's were used. These participants were selected from a pool of researchers and former researchers and developers at the Usability methods and tools lab. Each of the SME's had at least one year of experience studying the UAF and understood the nuances of the UAF.

2.1.2 Tools

For this phase of formative evaluation, data collection was executed using developed checklists, questionnaires and verbal protocol for the mentioned methods. The tools used and the data collected form these tools are outlined in **Table** 2-1.

Table 2-1: Tools and data collected in the SME evaluation procedure.

Subject Matter Expert Evaluation	
Tools	**Data collected**
Checklists/questionnaires and scratchpads	Responses on checklists and questionnaires and suggestions on scratchpads

The tools for the SME evaluation procedure are included in Appendix B.

2.1.3 Procedures

The SME evaluation procedure concentrated on using subject matter experts who helped to develop the content and those outside the project to check the accuracy, appropriateness and relevance of the content used. Since by nature of the subject, experts will also qualify as interface design experts, their feedback and suggestions on how to improve the instructional interface was also obtained. The SME evaluation was also designed to support last minute inclusions or exclusions to or from the training program. Three SMEs were used for this evaluation. Besides using a checklist (Appendix B) to evaluate the content on the training tool, the Subject matter experts also made numerous other corrections, notes and comments where necessary. Subject matter experts were encouraged to use any mode they think suitable to provide the feedback to the developers.

2.1.4 Data Analysis

The subject matter expert evaluation part of formative evaluation process yielded suggestions and refinements to the content for the training program. All the points regarding changes to the content were to be analyzed and accepted or rejected based on feasibility of proposed changes. However, inputs from the SME's mostly fell into one of two categories: minor changes for the current version and suggestions for the next version of the training tool (beyond the scope of this project). Therefore all changes pertinent to the current version were implemented. These mainly included revisions to content and addition of examples to illustrate specific concepts. The data collected in the SME phase of formative evaluation are shown in Table 2-2.

Table 2-2: Summary of data from Subject matter evaluation checklists

Question	SME #	YES	NO	Comments
Is information presented in relevant context?	1 2 3	X X X		2nd half of module 1 needs more work and more interactivity
Is the organization of information into relevant sets and subsets (modules)?	1 2 3	X X X		Inclusion of a review sub module at the end of each module might help in quick reference.
Does Feedback contain information about correctness of a response as well incorrectness of response, where necessary?	1 2 3	X	 X X	Inconsistency in presenting feedback to users. Feedback as feed forward exists in some modules and doesn't exist in other.
Do examples, contexts and applications progress from simple to complex, familiar to unfamiliar and concrete to abstract?	1 2 3	X X X		-
Does the instructional content maintain accuracy of facts, figures, data etc.?	1 2 3	X X X		-
Is there sufficient use of lists/tables and figures to summarize and encapsulate information?	1 2 3	X X	 X	Framework and hierarchical structure should be illustrated better. Maybe with a more interactive display of how classifying works.
Are the examples and the solutions used in the training accurate?	1 2 3	X X	 X	Examples need to be reliability tested to a point of high reliability before using them.

Question	SME #	YES	NO	Comments
Does the training tool provide a reasonable amount of interactivity to the learners?	1 2 3	X X 	 X	Within the constraints of a stand-alone text centric tool, interactivity is kept at a reasonable level.
Is repetitive practice provided to enable learners to smooth out the routine and if possible automate the skill?	1 2 3	X X X		-

Besides the use of checklists, the subject matter experts were encouraged to make notes and comments to provide inputs to the design of the training tool. These accounted for a major portion of the iteration after this stage to provide a refined training tool. The feedback from the SMEs was classified into one of the three following categories. These were:

- Feedback pertinent to current training tool.
- Feedback for future versions of the tool.
- Feedback for the UAF Explorer design team

Some of the comments pertinent to the study (category 1) are included in **Table** 2-3.

Table 2-3: Specific points from subject matter experts (with frequency >=2.)

SME positive comments	**SME negative comments**
Good navigation structure	Text centric.
Extensive coverage of training points provides depth in understanding concepts.	Overlap between background and practice sessions.
Feedback in popup windows is very useful.	More interactivity required explaining classification in the UAF.
Numerous examples and practice given crucial to learning.	Reliability problems within examples and exercises.

Redesign was limited to problems that at least 2 of the SME's had identified in their analysis of the training tool. Since this was in the initial stages of the study, time limitations did not impose any serious restrictions to redesign Most of the changes following this stage of formative evaluation were content specific and required revisiting the design of training content though some of the redesign also focused on the actual interface of the training tool. Some of the major changes to design that came up at the end of this round of formative evaluation are summarized in **Table 2-4** below:

Table 2-4: Design iteration 1: Major changes after the SME evaluation phase

Problem	**Solution in Design**
Lack of interactivity	Introduction of a UAF Explorer walkthrough page to enable trainees to understand classification better.
Overlap between background and practice sessions.	Altering the summative evaluation problem set and training content to introduce application of learnt concepts rather than direct reproduction.
Reliability problems within examples and exercises.	Replacing examples and exercises with instances based on results of reliability testing from previous studies.
More feedback	Extend feedback panes to reflect inaccurate answers

The SME evaluation also brought out a few system bugs that were platform specific and not related to the training interface or actual training content.

2.2 ONE – ON – ONE EVALUATION

Formative evaluation continued with a one-to-one evaluation of the training. This stage too involved data collection, but not formal experimental design. The number of subjects used for this procedure was 4 with one iteration run of evaluation-redesign. This would also dictate when evaluation was terminated and redesign began. The redesign in turn was limited by available resources for redesign as explained in detail in the following sections.

2.2.1 Participants

Representative participants were chosen from the spring 2001 Usability Engineering CS 5714 course. These were students who had training in usability engineering concepts but knew nothing about the UAF itself. The number of subjects used for this procedure was 4. Nielsen (1993) recommends the use of 4-5 participants for an evaluation stating redundancy in results from evaluation above that number.

2.2.2 Equipment

A schematic layout of the usability tools and methods lab where testing was carried out is shown below:

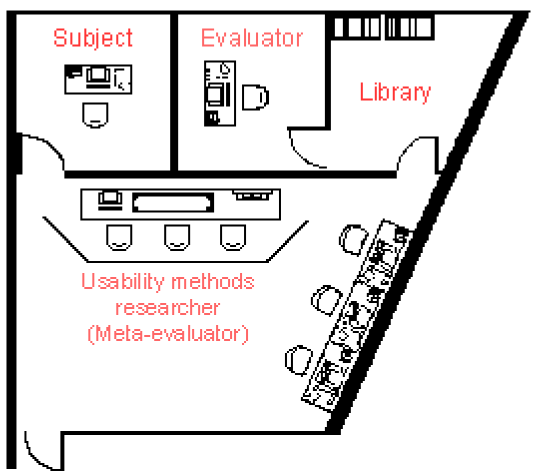


Fig 2-1: Schematic of the Usability Tools and Methods Lab

[Picture adapted from the Usability tools lab at Virginia Tech, http://miso.cs.vt.edu/~usab>]

The evaluator's area was equipped with soundproof walls, intercom to users room and video and audio editing facilities. The participant's area was equipped with a Mac and PC, video cameras and audio recording devices with a One-Way mirror separating user and evaluator. The computer used for the testing was equipped with a java enabled web browser. The computer itself was a Pentium based PC with the training tool preloaded on it. The observer's room was equipped with a viewing/recording facility and the lab setup is as shown in Figure 2-1. This same facility was also used for the non-intrusive individual evaluation phase that follows.

2.2.3 Tools

The tools used to aid data collection for the one-on-one evaluation procedure includes a questionnaire that is included in Appendix C. The critical incident observation method was used to identify and correct problem areas in design of content and interface. These tools and the data collected using them are summarized in **Table** 2-5.

Table 2-5: Tools and data collected in the One-on-One evaluation procedure.

One-on-one evaluation	
Tools	**Data Collected**
Critical incident method usability evaluation (with verbal protocol)	Critical incidents, descriptive data and feedback on content, organization and interface usability.
Questionnaires	Responses and ratings on user satisfaction and interface usability

2.2.4 Procedures

The participants were brought to the testing facility and were required to sign an informed consent form (Appendix D). This was an interactive session where the participant evaluated the training tool, both interface and content and talked to the evaluator while doing this. These participants were brought into sessions with the evaluator with recording of voice, facial expressions and onscreen interactions with the participants consent. The sessions lasted between 2 and 3 hours and were roughly predictive of the actual time taken for training under work settings. Most of the interactions with the evaluator remained informal with the participant picking apart problem areas while evaluating. This was followed by subjective ratings questionnaires to rate various components of the interface.

2.2.5 Data Analysis

In the one-on-one training evaluation, the data obtained were again mainly qualitative with the participant providing inputs mainly on the content of the training tool and the impact on the learner. As this was the first experience of the representative trainees with the training tool, the objectives of this stage largely remained clarity and impact. Clarity was mainly determined from descriptive data during participants think aloud and by observing critical incidents. A questionnaire was administered that measured the impact on the learner with emphasis on relevance, accomplishability and interest (Appendix C). A list of critical incidents observed in this stage is presented in Table 2-6.

Table 2-6: Critical incidents in the One-on-One evaluation phase. (Frequency>=2)

Critical incident category	Critical Incident Description	Frequency
Negative	Text centric nature of the training tool. Inability to concentrate for long periods of time to assimilate content in the training.	4
	Complexity of technical terms introduced, highly specific jargon (e.g.: cognitive affordance) especially towards the beginning of training.	2
	All users had problems with trying to move from the UAF training tool to the actual UAF.	4
	Conceptual clarity Issues: Presentation vs. content and meaning Physical affordance vs. cognitive affordance	3 3
	Trainees felt that additional examples and practice exercises would help concept assimilation and retention.	3
	User comments on readability of text on training tool	3
	Navigation within Module 4 which is the largest module with considerable scrolling	2
	At the end of module 2 most trainees thought that classification in the UAF was limited to classification at the highest level. Also highlights the lack of interest to read text.	3
Positive	One-line summaries of the different interaction activities in module 1.	2
	Section aiding classification at the highest level at the beginning of each module aids classification.	2
	User comments that Module 2 on highest-level classification greatly assists understanding.	3

Following the identification of critical incidents from user observations and verbal protocol a cost importance table was drawn up that would aid decision making regarding redesign for the problems identified in the evaluation phase. The method used for this follows the method recommended for analysis of data from evaluation by Hartson and Hix (2000).

The importance was drawn up based on the following simple principles, though not necessarily in the order listed:

- effect on trainee learning
- effect on training tool usability (independent of cost), include risk of not fixing
- frequency of occurrence

It must be noted that the importance was assigned independent of cost of fixes. This was factored into the next step before calculating priority ratios and ranks.

The decision rule for redesign following the one-on-one evaluation phase was based on availability of resources for redesign of the training. Redesign was considered only for critical incidents that had a frequency of at least 2 occurrences. It was decided that redesign would be carried out for a period of 24 person-hours and this would be the cut-off point for redesign decisions. This is illustrated in **Table 2-7** below:

Table 2-7: Cost-Importance Analysis for One-on-One evaluation.

Problem	Importance	Solution	Cost (In person-hours)	Priority Ratio	Priority Rank	Resolution	Cumulative Costs
Users had problems with trying to move from the UAF training tool to the actual UAF	5	Design and embed a visual aid that distinguishes UAF Explorer from training tool and shows transition. Incorporate the concept of a running explorer window that users can refer to. Consistently provide features throughout the training tool to support use of running explorer window.	6 hours	1250	1	Fix	6
Text centric content	3	Introduction of a review section at the end of each module and trim readable content throughout the tool.	5 hours	750	2	Fix	11
Conceptual clarity issues and need for more examples	4	Formulate and introduce at least 1 more worked example in each	6 hours	666	3	Fix	17

		module. Add one more example to all important distinguishers in the UAF like CA vs. PA					
UAF specific Technical Terms	2	Introduce a glossary section for quick reference to key terms	4 hours.	666	3	Fix	19
Readability of text	2	Make minor changes to font size and font type.	0.5 hours	400	5	Fix	21
Scrolling within large pages and navigation within modules.	2	Quick links to get to different sections of the page	2 hours	200	6	Fix	23
Module 2: Trainees thinking that classification was only at highest level	2	Include descriptive text and a complete classification example in module 1 to give trainees an idea of what classification is about.	2 hours	100	7	No fix	25
Total person hours for redesign efforts							**23**

COST CUTOFF: 24 person hours

It is seen from the above table that all changes resulting at the end of this evaluation stage were carried out as per the decision rules stated earlier. Table 2-8 below summarizes only the major changes after the second phase of formative evaluation.

Table 2-8: Design iteration 2: Major changes after the one-to-one evaluation phase.

Problem Area	Solution in design
Transition from the Training tool to the actual explorer	Developing a double window model to enable easily shifting between the UAF and the training tool
Classification using the UAF explorer unclear	Developing a short .avi file within the training that clearly illustrates how classification takes place in the UAF.
Too much text centric content	Sizing down text and adding extra examples to enable learning by problem solving

The data obtained on subjective ratings in this phase of evaluation is included in the results of the next phase. This is because the same tools (questionnaires and surveys) were used for collection of ratings data for subsequent formative evaluation methods and the results are treated together because of their small sample sizes.

2.3 NON-INTRUSIVE INDIVIDUAL EVALUATION

The last phase of formative evaluation was the non-intrusive individual evaluation phase. This was very similar in procedure to the one-on-one evaluation except for two major changes. In this phase, there was no interaction between the evaluator and a participant. The evaluator played the role of a silent observer. Another important difference was that this phase of evaluation used performance measures to measure the effectiveness of the training tool against a predetermined criterion. If this criterion of evaluation was achieved, then formative evaluation would be ended at that point.

2.3.1 Participants

Representative participants were again chosen from the spring 2001 Usability Engineering CS 5714 course. These were students who had training in usability engineering concepts but knew nothing about the UAF itself. Participants were selected for two iterations with 4 in the first run and 3 participants in the second.

2.3.2 Tools

The tool used to aid data collection for the non-intrusive individual evaluation procedure include questionnaires that are included in Appendix C. Tasks lists were used to assign participants learning tasks and classification problems and these served as performance measurement tools. These tools and the data collected using them are summarized in **Table** 2-9.

Table 2-9: Tools and data collected in the Non-intrusive individual evaluation procedure.

Non-intrusive individual evaluation	
Tools	**Data Collected**
Task list	Performance measures on classification accuracy.
Questionnaires	Responses and ratings on user satisfaction

2.3.3 Procedures

The testing followed the critical incident method for usability testing. The observer watched for and identified critical incidents in the users activities when the user performed a task and took notes on the user's thinking aloud. The observer could see and hear all the users activities from behind the one-way mirror in the observer's room but the user could not see the observer. These along with the usability questionnaire that participants answered at the end of the session provided valuable inputs to iteratively design the content and usability of the training tool. Care was taken to separate the interface components from the instructional components when assigning benchmark tasks. The benchmark tasks performed by the users are included in Appendix E. When deficiencies or weaknesses in the interface were discovered from testing, redesign was to be carried out. A criterion was pre-selected for formative evaluation and if this criterion was met then the design would be frozen and carried over to summative evaluation. The criterion is stated below:

The training tool will be effective as measured by at least 75% accuracy in classifying problems according to the hierarchical structure of the UAF.

The basis of selecting 75% as the criterion is explained based on the setup of the summative evaluation study. The summative evaluation was pre-designed to provide trainees with 2 sessions of 75 minutes of training each. Also the participants in the summative evaluation would have a week to finish the posttest. However for the formative evaluation, constraints on availability of subjects forced the use of single 2-3 hour-long sessions with no breaks between training or between training and testing. Taking these constraints into account it was decided that 75% presented a fair target as criterion for formative evaluation. The instructional goal, which was to be evaluated in summative evaluation, remains at 80% classification accuracy.

2.3.4 Data Analysis

The analysis of data consisted of simple quantitative calculations to determine if the set goal for formative evaluation was met. It was observed that the total time for completion of modules averaged over both the iterations was 109 minutes. Descriptive data obtained from formative evaluation is included in Appendix F. It can be seen that the lowest accuracy scores were obtained on question 7 on both the iterations. This could be attributed to the fact that this slightly complex problem involved problem extraction and this was not adequately dealt with in the training. Another observation was that many points relevant to the design of the UAF itself arose at this stage. These data will be submitted to the development team for their formative evaluation process. The data from formative evaluation is summarized in tables 2-10, 2-11 and 2-12 below:

Table 2-10: Time to complete task during formative evaluation.

Participant #	Total Time taken in mins
1	103
2	99.5
3	128
4	133
5	114.00
6	87.00
7	98.50

Mean(SD)	109.00(16.74)

Table 2-11: Accuracy scores on iteration 1 for non-intrusive individual evaluation.

Participant #	1	2	3	4	5	6	7
Score	83.00	80.00	92.00	76.00	74.0	84.00	78.00
Mean(SD)	**81.00 (6.02)**						

Table 2-12: Critical incidents observed in the Non-intrusive evaluation stage.

Critical incident category	Critical Incident Description	Frequency
Negative	Text centric content	2
	Loading times for heavy graphics and interactive content	2
	Use of JavaScript and frames.	2
	Over use of reviews at the end of sections	2
	User commented that the example of good design in the practice sessions caught them off guard.	2
Positive	Visual aid in classification	4
	Simple and effective navigation structure.	2
	Reviews at the end of sections	5
	Use of training tool's glossary of terms	5

The critical incidents and trainee comments that came out of this stage of formative evaluation are mostly optimization and refinement of the tool related. Any of the redesign that arose out of this stage were not central to training tool content or training itself. Optimization proceeded with making the training tool platform independent and reducing graphic file sizes to optimize download times on slower connections. It is also observed that some of the issues that arose consistently throughout the formative evaluation were text-centric nature of content for the training tool and the nature of terms introduced. These points merit a brief explanation. The User Action

Framework does introduce numerous terms and concepts that are sometimes more specific to trainees with certain skills and backgrounds, against others. But increasingly in industry, the lines between skills and roles adopted blur and this is inherent in a cross-disciplinary field like usability. It therefore seems that to clarify certain concepts and terms, a textual explanation is necessitated thus increasing the overall textual content in the training tool. It would be interesting to observe the results of a study concentrating solely on the use of an interactive tool with voice and visual depictions and see if this scores over text in trainee opinions.

Formative evaluation was terminated at this point and the evaluation phase moved into the summative phase as the criterion of 80% accuracy had been met.

Besides the above data, the formative evaluation evaluated, both in the one-on-one evaluation and the usability-testing phase, reaction to and satisfaction with the training tool. This was evaluated by means of questionnaires. Since this was done in both phases of formative evaluation (one-on-one and non-intrusive individual evaluation), the results are included in tables that combine data into a single n= 11[n = 4 for one-on-one evaluation, n= 7 for non-intrusive individual evaluation. The data from the impact questionnaires have been summarized in the **Table** 2-13.

Questions	**Mean(SD) rating for n = 11** **1** Strongly Disagree — **5** Strongly Agree
There are clear navigational aids that tell me where I am.	4.45 (0.68)
It is easy to find the information I need using the training tool.	4.00(0.77)
I like using the UAF training tool.	3.72(0.78)
The training tool has all the functions and capabilities I expect it to have	3.54(0.52)
I am able to complete my tasks easily using the interface.	4.72 (0.46)
The UAF explorer makes it easy for me to organize my thoughts about a usability problem.	3.54 (0.68)

Table 2-13: Ratings on impact of and satisfaction with training tool.

The participants were required to fill out a subjective ratings questionnaire on the ease of use and other features of the training tool. These questions mainly targeted user response to the user interface component of the training tool but also contained questions designed to measure overall reaction to the training tool (Appendix C). The data from these qualitative analyses were carried out to obtain an idea about the response to and satisfaction with the tool and the opinion of representative trainees from the targeted population. No specific data analysis was performed on this data.

Some of the comments on open-ended questions that were asked of participants after the conclusion of the tests provided more information about the problem areas identified. These sessions were viewed as a chance to understand participant's perception of satisfaction with the tool. Some issues that were missed during the regular evaluation came to light when measuring the subjective ratings of participants. The results for these questions are summarized in **Tables** 2-14 and 2-15.

Table 2-14: Subjective ratings on interface of training tool.

Questions	Mean(SD) rating for n =11
Reading characters on screen Hard -2 2 Easy	1.04(1.15)
Organization of the Information Confusing -2 2 Very clear	1.31(0.78)
Sequence of screens Confusing -2 2 Very clear	1.77(0.41)
Use of popup windows Unhelpful -2 2 Helpful	1.63(0.45)

Table 2-15: Subjective ratings of reaction to training and interface in formative evaluation.

Overall reaction	Mean(SD) rating for n = 11
Terrible-1 5- Wonderful	3.81(0.98)
Difficult- 1 5- Easy	3.63(0.80)
Frustrating- 1 5- Satisfying	3.63 (0.92)
Dull-1 5- Stimulating	3.45 (0.68)

For example, in the one-on-one evaluation, interacting with participants revealed that the use of a particular font for the text in the training tool made reading characters on screen difficult. This is reflected in the relatively low rating that on screen reading received from participants. However changes were made to this in the next design iteration and led to subsequently better feedback on readability in the next stages of evaluation.

2.4 RESULTS

At the end of the rounds of formative evaluation, a refined training tool with adequate functional features was ready for summative evaluation in which numerous trainees would use the tool to learn the UAF explorer tool. Participants expressed satisfaction with the tool and numerous features of the training tool were appreciated. The formative evaluation process proved invaluable in the construction of the training tool for the summative evaluation process and eventually, deployment of the training tool.

Figure 2-2 below summarizes the changes that took place after each round of formative evaluation:

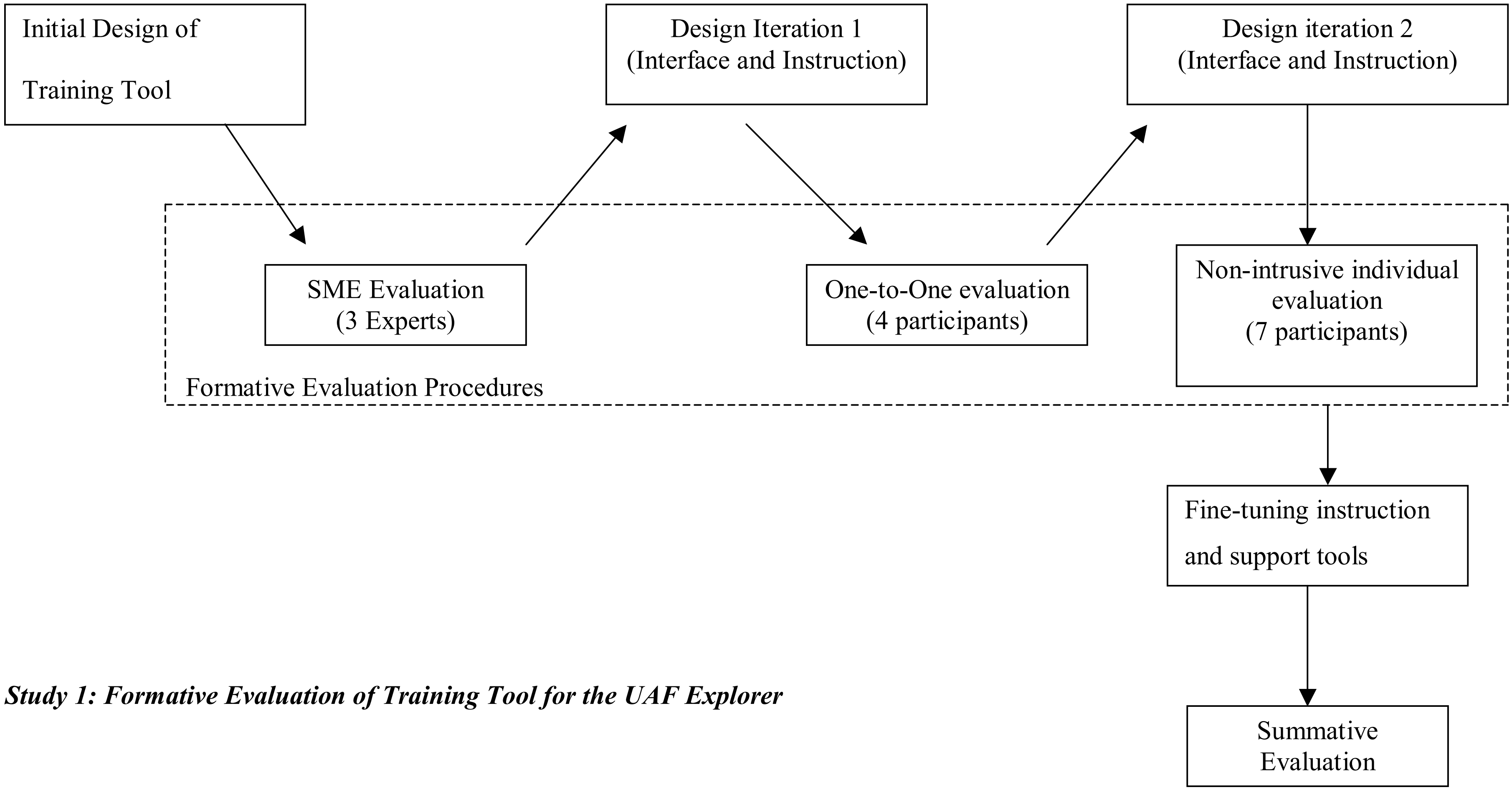


Fig. 2-2: Study 1: Formative Evaluation

CHAPTER 3: SUMMATIVE EVALUATION

3.1 PARTICIPANTS

Participants were selected from a pool of graduate students in the Computer Science and Industrial and Systems Engineering departments at Virginia Tech. Assigning participants to groups was randomized. McKeachie et al (1986) stressed that a novice trainee's prior knowledge in the domain and cognitive skills are of crucial importance in an instruction situation. These participants will have no prior experience using or learning the UAF, but have prior knowledge on usability concepts that forms the training domain. These volunteers were selected from the graduate level course CS 5614- Usability Engineering and are highly representative of the novice usability group. The prerequisites to be met were verified by screening potential participants by means of a questionnaire (Appendix G) to ensure that they met the following requirements:

- At least one year experience using a PC and an email server application.
- Able to write and speak fluent English.
- No graduate-level courses in human factors engineering with high exposure to the UAF.
- Prior use of a Java enabled browser.
- Prior exposure to usability concepts

Participants were identified in the study using a coding scheme. All subject information and data sheets were kept confidential through the course of the study and data analysis.

3.2 EXPERIMENTAL DESIGN

An experimental pretest/posttest between subjects design was used to evaluate whether participants performed better on the web-based training system or a classroom training system. The dependent variable for the study was training effectiveness and preference ratings for the training tools. The between subjects design randomly assigned volunteers to one of two groups, the experimental group and the control group, with the classroom training group serving as the control group. Lack of experimental control on time spent on the UAF (frequency of exposure) other than the scheduled training sessions provided the researcher with only moderate control over exposure for each group, but this cancelled out between the two groups.

The design has been illustrated below:

Group 1	(R)	O1	X1	O2
Group 2	(R)	O1	X2	O2

Where:

R: Random assignment of participants.

O1 : Observation or pretest

O2 : Observation or posttest

X1 : Training condition 1 – Classroom training

X2 : Training condition 2 - Web based training

It must be noted that the use of traditional training as a control group and basis of comparison could potentially introduce various confounding effects and these were avoided by building controls in to the experiment. Other studies (Pane et al., 1996; Wiedenbeck et al., 1995) that used a treatment group as a control group or have totally avoided the use of a control group have adopted similar approaches to reduce confounding effects.

This pretest-posttest design measured the performance of the user with the UAF Explorer Tool before and after the training program, resulting in data that can be analyzed to see if there was improvement after the participant completes the training program. Since the participants were randomly selected and assigned to the level of training, most internal threats to validity were controlled. If the experiment was not carefully designed, it could have been difficult to interpret the data because apparent improvement of performing a task might have resulted from other factors like sensitization to the area due to pretest or intuition rather than by training alone (Goldstein, 1996). Since a control group was used, the fact that the improvement of the task performance is due to other factors that are external can be addressed. Because groups received the same pretest and posttest, sensitization was equal in both groups and therefore this equally affected the data from the posttest of both groups. Because of the random selection of participants, differences between the groups were very unlikely, so homogeneity between the groups was almost ensured. Nonetheless, the pretest verified if both groups have the same knowledge, skills and ability levels. Ideally the pretest results of both groups would be similar, this would mean that the two groups had relatively the same knowledge or background about the factors that affect their performance.

The Pretest would measure the level of skills, knowledge and abilities that trainees have about usability and about the User Action Framework (UAF) concepts and classification. Trainees might never have heard of the UAF before, but knowledge about related things and intuition could also have affected the performance with the system. So a pretest will determine the trainee performance with the system counting on intuition and knowledge prior to training. The entry behaviors for the trainees defined for this project were knowledge of usability concepts. These knowledge, skills and abilities of the trainees were measured to see if the groups are homogenous.

After receiving the pretest, one of the groups received the web based training tool. The other group, the control group, received the classroom based training program. The pretests were compared to ensure that both groups started out at the same level of skill and knowledge. The posttest had classification problems to measure the impact of the training program. The classification measured skills when using the UAF Explorer tool (Refer to Appendix H). Pretest results of the trained group were compared with the results of the posttest of the same group to determine the impact of the training program. Then posttests of both the treatment group and the control group were compared to establish which of these was more effective.

The groups and their characteristics are summarized in **Table** 3-1.

Table 3-1: Similarities and differences between control and experimental groups

Characteristic	**Control group**	**Experimental group**
Sample size	16	15
Assignment of subjects	Randomized	Randomized
Testing	Pretest-training-posttest	Pretest-training-posttest
Location	Classroom	PC lab
Training	Lecture-based	Web-based

3.3 TOOLS

The tools used focused primarily on the pretests and posttests for the summative evaluation and are summarized **Table** 3-2.

Table 3-2: Tools and data collected in the summative evaluation procedure

Tools	**Data Collected**
Pretest-posttest	Accuracy scores
Likert type scale questionnaires	Subjective satisfaction rating scores

The pretests and posttests for the summative evaluation procedure are included in Appendices G and H. The questionnaires are included in Appendices I and J.

3.4 EQUIPMENT

Participants in the experimental group used a Pentium based PC running windows 2000 or a Mac running Mac OS or higher versions and a 17-inch CRT display operating at 1280x1024 resolution displays the interface. All computers were equipped with a java enabled web browser (Netscape or Internet Explorer versions 3.0 and above). The facility used for training the experimental group was the PC lab at Togersen Hall. The control group had an instructor using teaching aids like a notebook PC equipped with a java enabled web browser and a projector.

3.5 PROCEDURES

Once recruited and verified to meet all screening requirements, participants were assigned to one of two treatment conditions that are classroom training and web based training. The participants were presented with a written overview of the study, its purpose, objectives, and the requirements of their participation. The overview was verbalized and included in the participants informed consent form. If in agreement with the terms participants were asked to sign it. Samples of these informed consent forms are included in Appendix I. Training took place in two 75-minute sessions over three days. The sessions for the experimental group were conducted at the PC lab while the control group underwent classroom training simultaneously. This controlled the time spent on training for both groups, as the duration of instruction sessions for both groups was the same. At the end of the training sessions the trainees undertook a posttest to obtain data that measures effectiveness of the training program. For the same durations, the control group underwent training on the UAF Explorer at the same time in classroom settings under the supervision of an instructor with regular class teaching aids and methods. Both groups received the same amount of time (approximately one week) to complete the posttests and had access to class notes during this period. Both groups also filled out subjective ratings of experience questionnaires that will be used as qualitative data (Refer to Appendices K and L).

A simple scoring system was used to score the tests according to the number of levels traversed in the UAF to accurate classification. The highest level was accorded more importance than the next level and so on. The lower levels, which have the highest ambiguity within the UAF, carried the least points. The scoring key for the tests is included in Appendix J along with the solutions for the pretests and the posttests.

3.6 DATA ANALYSIS

The summative evaluation procedure involved a two-group pretest-posttest design to evaluate the overall effectiveness of the training, as described in the previous section. Data collected from summative evaluation includes performance measures, learner attitudes and preferences. The principle performance measure was the gain score, which is computed from the pretest-posttest scores for participants in each group.

The improvement (gain) from pretest to posttest can be computed for each participant by subtracting each person's pretest score from their posttest score -

Gain = posttest - pretest

A positive gain score indicates that the posttest score was greater than the pretest score; a negative gain score indicates that the posttest score was less than the pretest score. In this experiment the dependent variable was learning so it was expected that successful treatment would lead to higher posttest scores. The gain score should therefore ideally be positive.

The gain score controls for individual differences in pretest scores by measuring the posttest score relative to the each person's pretest score. But, a gain score analysis does not control for the differences in pretest scores between the two groups and this is accounted for by performing a t- test on pretest scores between the two groups. The mean gain for the WBT group was **23.73** while the mean gain for the classroom-based training group was **14.25.**

Data obtained are included in Appendix M and summarized in **Table** 3-3.

Table 3-3: Pretest – posttest data from summative evaluation

Group	Number of Participants	**Mean (SD) Pretest score**	**Mean(SD) Posttest Score**	**Gain score**
EXPERIMENTAL GROUP	15	57.13(15.66)	80.8(9.15)	**23.73(16.45)**
CONTROL GROUP	16	64.12(15.16)	78.375(13.63)	**14.25(15.39)**

The pretest and posttest scores for both groups have been illustrated and summarized in the Figures 3-1, 3-2 and 3-3.

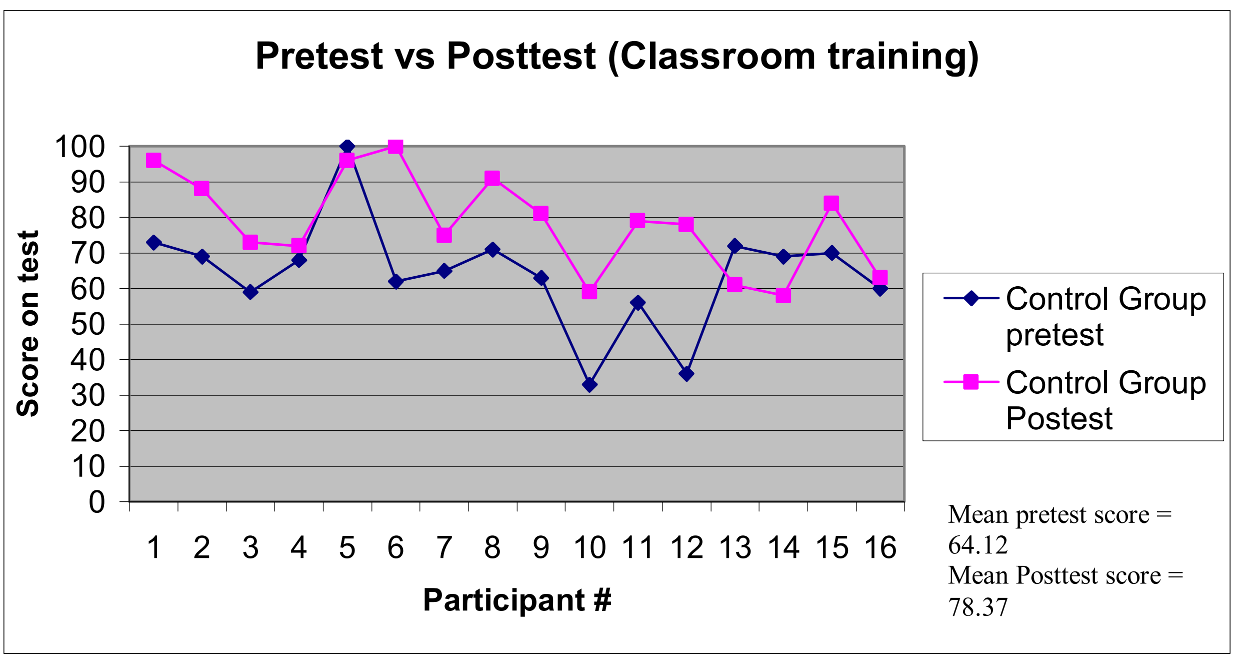


Fig 3-1: Pretest vs Posttest scores for classroom training (control) group

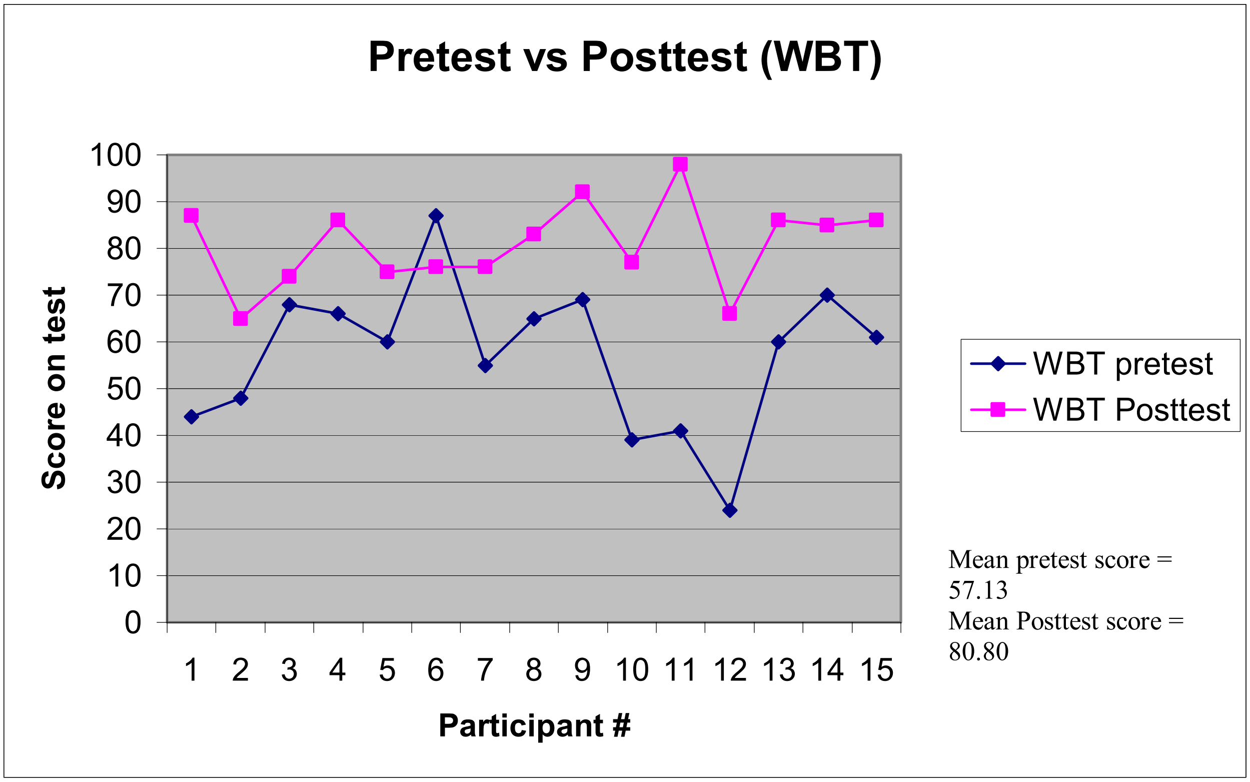


Fig 3-2: Pretest vs Posttest scores for web-based training group

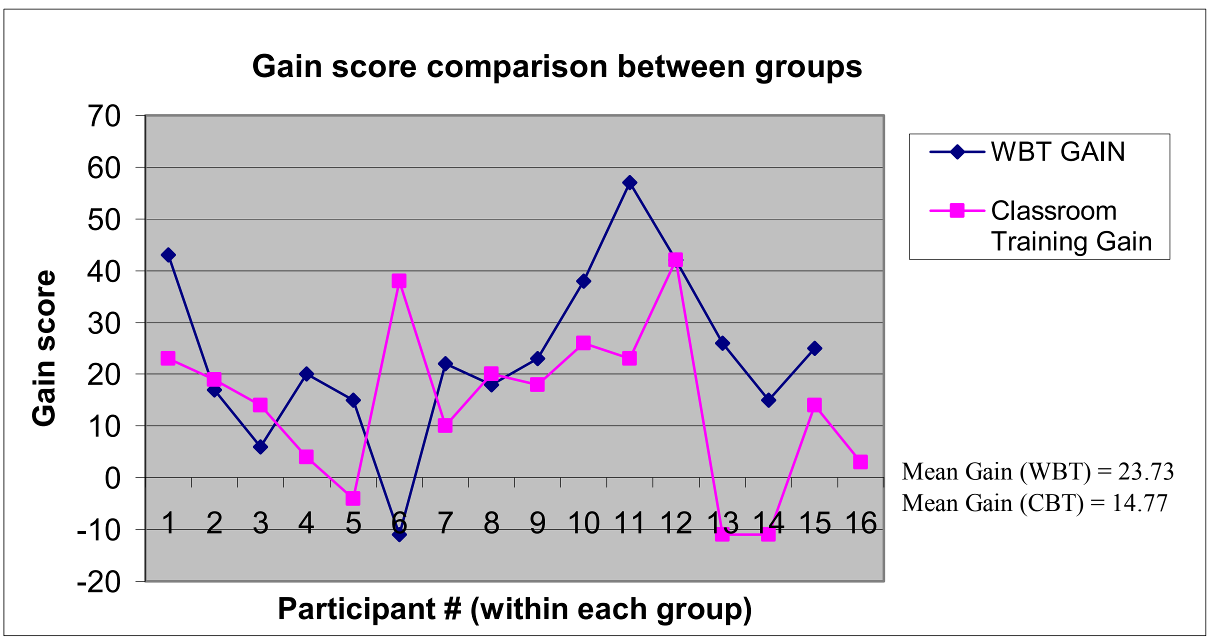


Fig 3-3: Gain score comparisons between groups

The graph shows a higher gain score for students in the experimental group for most data points. To ascertain if the training tool was effective, t- tests were preformed.

A student's t- test was performed to ascertain any differences between the two groups at the pretest stage. The results of the unpaired t-Test indicated $t\,(21) = 0.358$, $p = 0.158$ for a two-tailed $\alpha = 0.05$. This p value of 0.158 (two- tailed for $\alpha = 0.05$) leads us to infer that there are no significant differences between the groups at the pretest level.

A paired student's t-test was within each group between the pretest and posttest to ascertain if there indeed was learning. For the control group the results indicated: $t(15) = 3.7036$, $p = 0.0021$ for a two-tailed $\alpha = 0.05$. This shows a significant difference at $\alpha = 0.05$ confirming that there is a difference between the scores obtained at the pretest level and the posttest for the control group.

Similarly, within the experimental group between the pretest and posttest the results of this t-test indicated $t\ (14) = 5.5859$, $p < 0.0001$ for a two-tailed $\alpha = 0.05$.

The two tailed $p<0.0001$ shows a significant difference at $\alpha = 0.05$ confirming that there is a difference between the scores obtained at the pretest level and the posttest for the experimental group.

A student's t-test was performed between the gain scores of the experimental group and the control group to see if there was a significant difference. The results of this t-test indicated $t\ (21) = 1.658$, $p = 0.1081$ for a two-tailed $\alpha = 0.05$.

This p value of 0.108 (two- tailed for $\alpha = 0.05$) leads to a failure to accept the null hypothesis, i.e. there is no significant difference between the gains for the groups.

The results from the summative evaluation address the training goal of achieving 80% accuracy in classification for the experimental group as was shown in **Table** 3-3. The mean score on the posttest for the experimental group was 80.80 with a mean gain of 23.73. This satisfies the research hypothesis that states that the training tool will be effective as indicated by accuracy scores on the posttest. It is seen that the gain is significant and given that numerous controls were imposed on the experimental design, it can be concluded that the gain in accuracy scores was due to the training. It is worth mentioning that the gain score for the experimental group is higher than the gain for the control group. It is also observed that both groups have isolated instances of negative gains between pretests and posttests.

Subjective data collected includes ratings on satisfaction and attitude and impact on learners. The main criteria that were measured include satisfaction with the training and impact on learner attitude. The data obtained from the questionnaires (Appendices 9 and 10) are summarized in the **Tables** 3-4 and 3-5.

Table 3-4: Subjective ratings in summative evaluation

Construct	**Question**	**Mean Rating(SD) (1- strongly disagree 5-srongly agree)**	
		Web based training (n = 14)	**Classroom Training (n = 15)**
Impact	I will continue to use the UAF when required to classify usability problems.	3.25(0.42)	3.40(0.50)
	I feel better prepared to classify usability problems accurately using the UAF explorer after going through the training tool.	3.46(0.77)	3.50(0.77)
	I understand the underlying concepts of the UAF and can apply them confidently.	2.82(1.09)	2.96(1.04)
	Mean (SD)	**3.17(0.32)**	**3.28(0.28)**
Satisfaction	The UAF training has the required capabilities.	2.60(1.0)	2.80(1.08)
	The UAF training satisfies all my expectations from training/ a training tool	2.32(0.92)	3.18(0.67)
	Mean (SD)	**2.46(0.19)**	**2.99(0.26)**

To test for significant differences between the ratings received by each group, Mann-Whitney U tests were conducted for the ordinal data obtained. These tests show no significant differences between the two groups for each of the measured criteria ($U = 3.0$, $p = 0.70$) for $\alpha = 0.05$.

Similarly Mann -Whitney U tests were conducted to measure differences, if any, between data from other questions as shown in **Table** 3-5. These tests showed no significant differences between ratings for either group on questions measuring reaction to training. While higher ratings were observed for classroom training, there was no significant difference between the ratings for the web based training group. Further results of this test are summarized in **Table** 3-5.

Table 3-5: Results of Mann-Whitney U tests on subjective ratings

Question	**Mean(SD) Rating**		**Results of MW test** ($\alpha = 0.05$)	**Interpretation**
	Classroom Training n = 15	Web based Training n = 14		
How would you rate the overall training you received for the UAF? 0 – Poor 10 - Excellent	6.13(1.45)	5.42(1.45)	p = 0.511	No significant difference
How would you rate your skill on using the UAF explorer to classify usability problems? 1- Poor 4- Excellent	2.33(0.72)	2.07(0.61)	p = 0.537	No significant difference

Though not part of the research objectives or hypotheses, data were also collected about trainee's preference for media from the pretest questionnaires where the participants rated their preference of training media. These data were used to see if preference had an effect on performance. The data are presented in detail in Appendix M.

The relevant data are also summarized in the **Table** 3-6. The first sub table indicates the performance of participants who preferred classroom training and the second sub-table indicates the performance for those who preferred web training.

Table 3-6: Preference vs. Performance data

Preferred	Mean (SD) Score on classroom training (for n = 12)	Mean (SD) Score on web training (for n=11)
Class	79.41 (14.71)	79.90(7.58)

Preferred	Mean (SD) Score on classroom training (for n = 5)	Mean (SD) Score on web training (for n=3)
Web	77.6 (10.78)	82(16.52)

An analysis (as described below) of preference vs. performance data collected showed no significant difference between participants who preferred one treatment and got the other treatment. This analysis compared participants within each group that preferred a medium and got that treatment to those who preferred a medium and did not receive that treatment.

The null hypothesis assumes that there is no significant difference between groups that preferred class and got class training and those that preferred classroom training but got web training. A t-test indicated that: $t\,(21) = 0.099$, $p = 0.9218$for a two-tailed $\alpha = 0.05$. For $p<0.05$ level of significance, the two tailed p value of 0.9218 (for $\alpha = 0.05$) leads us to fail to reject the null hypothesis.

Similarly, within the other group, the null hypothesis assumes that there is no significant difference between groups that preferred web-based training and got web-based training and those that preferred web-based training but got classroom training. A t-test indicated that $t\,(21) = 0.464$, $p = 0.658$ for a two-tailed $\alpha = 0.05$.
For $p<0.05$ significance, the obtained two tailed p value of 0.658 leads us to fail to reject the null hypothesis.

The tests show no significant difference within the groups with preference satisfied and the groups whose preference was not satisfied. While this was not hypothesized in the research goals, data available warranted an examination to see if preference had an effect on performance. The data were analyzed within each treatment group to observe effects within the groups only and no other interactions were examined.

3.7 RESULTS

The summative evaluation seeks to evaluate the actual effectiveness of the developed training tool before deployment. By meeting the instructional goal of achieving 80% mean on accuracy scores for classification in the summative evaluation, the effectiveness of web-based training for training users on the UAF Explorer was ascertained. Gain scores between pretests and posttests conducted for a control group and an experimental group measured effectiveness. Both groups showed a positive gain with significant differences between the pre -posttests within each group. However between the control group and the experimental group there was no significant difference in gain scores.

Subjective data were obtained for trainee reaction to the training and impact on the learners. Tests on this data also revealed no significant differences between the ratings for each group. Both of these results supports the use of the constructed web-based training tool to train users on the UAF Explorer tool.

CHAPTER 4: CONCLUSIONS AND FUTURE RESEARCH

4.1 CONCLUSIONS

The overall purposes of this study were to design a training tool for the UAF Explorer using the systematic design of instruction approach and quantitatively evaluate its effectiveness. Design of the tool involved designing the content, the interface and iterative formative evaluation studies to optimize the training tool developed. The first part of the research concentrated on actually developing the functional training tool. This was followed by 3 rounds of formative evaluation. The effectiveness of the tool was then evaluated by using a pretest-posttest design.

Trainees often found the use of the PBL approach very important in learning to classify problems using the hierarchical structure of the UAF. This might validate earlier studies (Fekete et al., 1998) that support the use of PBL in computer science and other sciences. The formative evaluation process is used to evaluate and address problems if any at an early stage in the design process (Hix & Hartson , 1993). This proved to be invaluable in designing as many problems were encountered and effectively addressed early to produce a tested and refined training tool for the subsequent study. The training tool met the objectives set for formative evaluation after 4 rounds of testing on the original tool as shown in Figure 2-2.

Summative evaluation centered on measuring performance to ascertain the effectiveness of both training media. Gain scores between pretests and posttests conducted for a control group (classroom based training) and an experimental group (Web based training) were used as measures of effectiveness. Both groups showed a positive gain with significant difference between the pre and posttests within each group. The instructional goal of achieving 80% mean on accuracy scores for classification was met. However between the control group and the experimental group there was no significant difference in gain scores. Qualitative data were obtained to gauge trainee reaction to the training and impact on learner and no significant differences were observed between the ratings for each group.

Based on this research, it is observed that web based training is as effective as classroom training and provides a cost effective alternative to classroom training and other training media.

While the results obtained show a higher gain for the web-based training tool and higher scores on the posttests, generalizing these results to include all comparisons of hypermedia based training to traditional training is not warranted. Past studies (Blanchard, 1990; Compeau and Higgins, 1995; Dillon and Gabbard, 1998) fail to conclusively point to either as being the better medium, except under the conditions specific to that study. The same applies to this study and within the constraints of the study, web based training can be concluded to be effective for training novice users on the user action framework explorer.

It is also seen that data obtained do not indicate that performance has no correlation with preference as measured before the experiment. The use of web based training and short tutorials for training in skills and toolkits should assume significance and present a chance to use the reach of the web for creating an awareness of such tools and skills.

These results somewhat differ from earlier studies that show that independent variables collectively influence novice learning of computer systems (Nielsen, 1994; Pennington, 1981). But earlier works treated independent variables as a whole and tried to measure their effect on the measured variable. However this research seeks only to address the effects of one such variable, preference of training medium on actual performance. A well-designed study seeking to establish effects of preference on actual learning performance has been suggested in the future research section to study such interactions.

This research also supports the use of a web based training tool for the User Action Framework Explorer tool as a cost effective training medium. Given that effectiveness matches the effectiveness of classroom training, the inherent cost – effectiveness and how the world-wide web lends itself to distance learning to a distributed audience, it is hoped that the web tool gradually phases out classroom training for the purpose of educating researchers and students on the use of the UAF toolkit. The combination of cost-effectiveness and training effectiveness lends immense credibility to the proliferation of web-based training tools for specific purposes. The concluded research illustrates this point well as the target audience of usability professionals and students represent a demographic that uses the web extensively (Appendix- A) and have no infrastructure/bandwidth and other high initial capital costs that is normally inherent to the use of web-based systems as a training medium. This also presents the developers with enormous flexibility and reach into the research community in terms of exposure for the UAF. Based on this research, it is observed that web based training is as effective as classroom training and provides a cost effective alternative to classroom training and other training media. Also worth mentioning is that from the data analysis stage onwards potential trainees and usability engineers have consistently expressed interest in understanding the use of such a toolkit and the contributions of such a framework to the usability community.

4.2 FUTURE RESEARCH

Interactive fully functional training tool with synchronous and asynchronous communication and compare to the static version.

The subjective ratings after the posttests brought out the need for a more streamlined training tool, which allowed for communication between the trainees and the instructor if necessary. While that was beyond the scope of this study, it is definitely an interesting topic as the UAF by nature of its structure is difficult to understand and classification has in the past been time consuming and complicated. Therefore it would be good to have some form of communication for clarifications that the students would have in order to simulate a classroom experience. This would also enable comparing various levels of interactivity in training.

Investigating emerging technologies in training and testing.

As early as 1995, three emerging training techniques training through Internet, Performance Support Systems, and Videoconferencing had been identified (Fitzgerald and Cater-Steel, 1995). At the present time, research is ongoing to validate these proposed technologies. However as training requirements evolve, so will such techniques and tools and this should be seen as an evolutionary process. The role of new and emerging technologies should be investigated with regard to training needs. There exist numerous Rapid Application Development (RAD) tools like camtasia and Macromedia Director that make designing and developing tutorials with high multimedia content and

A/V support easy. These create training programs in the form of portable tutorial files that can be viewed by anyone, anywhere with a PC and required software. This presents interesting opportunities to satisfy training demands for toolkits like the UAF that adds value to regular data analysis by extending the reach of such skills to remote users. Besides, camtasia has a role to play in remote usability evaluation as it allows evaluators to record screen actions and with audio. Efforts are currently underway to examine such techniques.

Alternative approaches to training development.

This research concentrated on developing a training program using the Dick and Carey systems approach model for designing instruction with training content developed partly on the problem based learning approach coupled with a few facets of the minimalist design approach. Efforts can be made to study the efficacy of each of these content development models separately.

Other research interests could include studying the effects of variables like preference, setting and the interactions of these with learning and training effectiveness. An attempt was made to briefly examine this topic in this research but an entire study devoted to this might yield some interesting results.

The following section addresses future research specific to the efforts of the Usability Methods and Tools lab.

Explore a more valid framework for higher reliability using the UAF.

The existing structure of the User action framework, by its nature, has a lot of ambiguity in the lower levels. While the scoring of the tests were done taking into account that reliability decreases at the lower levels, steps could be taken to cement the foundations for accurate and inaccurate answers by having a more valid extraction-classification process. Some of the ambiguity involving wording the usability problems can be addressed by exploring the option of integrating the UAF with a problem identification-extraction mechanism that lets the usability practioners identify the problem from a scenario themselves and then classify it.

Development of the UAF

Through the various stages of evaluation, data were obtained that identified issues with the structuring and working of the UAF itself. Efforts to address these issues might include an attempt at eliminating ambiguity in wording, especially towards the lower levels and an information architecture approach to designing the framework. Besides there are issues associated with the user interface of the UAF but formative evaluation conducted on the UAF during the course of this study might have since corrected some of the problems with the UAF.

Development of the UAF toolkit built around existing tools with integrated training.

Inherent to the UAF is a thorough understanding of human factors concepts without which the use of the UAF is limited. While the UAF is useful as a classification framework, the acceptance of it might hinge on the toolkit that is built around usability and understandability of the underlying concepts. The web based training tool is an effective practice tool that can aid many future learners become effective usability professionals. Leveraging on this, an integrated UAF toolkit that incorporates training with other tools like the UPC, UPI and the explorer needs to be developed and this would make the UAF toolkit very useful to usability practitioners.

REFERENCES

Adams, G. (1992) Why Interactive? Multimedia and Videodisc Monitor, 10,(3), pp. 20-24.

Albanese, M. A., & Mitchell, S. (1993). Problem-based learning: A review of literature on its outcomes and implementation issues [published erratum appears in Academia Medica 1993 Aug;68(8):615] . Academia Medica, 68(1), pp. 52-81.

Andre, T. S., Hartson, H. R., Williges, R. C. & Van Rens,L.(1999)The User Action Framework: a theory-based foundation for inspection and classification of usability problems. In Proceedings of the 8th International Conference on Human-Computer Interaction (HCI International '99): Ergonomics and User Interfaces. Vol. 1, pp. 1058-1062.

Andre, T. S., Williges, R. C., & Hartson, H. R. (1999). The effectiveness of usability evaluation methods: Determining the appropriate criteria. In Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting , pp. 1090-1094. Santa Monica, CA: Human Factors and Ergonomics Society.

Balasubramanian, V., Sridharan,S., Rao,S., Ferriera, R. (2000) User Action Framework Explorer Training Program. Unpublished manuscript. <http:// hci.ise.vt.edu/williges/ise5634/uaf.pdf> Last modified on: December 13th,2000

Baldwin, D. (1996). Discovery learning in computer science. In ACM SIGCSE Technical Symposium on Computer Science Education , pp. 222–226.

Baldwin, D. & Ford (1988). Transfer of training: A review and directions for future research. Personnel Psychology, *41*, 63- 105.

Berkson , L. (1993). Problem-based learning: Have the expectations been met?

Acad Med, 68(10 Suppl.), S79-88.

Blanchard, D. (1990). A Computer Aided Instruction program: Teaching the Disk Operating System to community college students. Unpublished thesis, Center for advancement of education. Nova university.

Bostrom, R. P., Olfman, L. & Sein, M. K. (1988). The importance of individual differences in end-user training: The case for learning style. Proceedings of the ACM SIGCPR conference on Management of information systems personnel, pp. 133 – 144.

Campbell.D.T and Stanley J. C (1966). Experimental and Quasi-experimental designs for research; Houghton Mifflin Company

Carroll, J.M. (1984). Minimalist Training. Datamation Magazine. November. pp. 125-136.

Carroll, J. M., Rosson, M. B. (1990) Human Computer Interaction Scenarios as Design Representation. Proceedings of the Hawaii International Conference on System Sciences, Los Alamitos CA IEEE Computer Society Press, pp. 555-561**.**

Carroll,J.M. & Rosson,M.B.(1995). Managing evaluation goals for training. In *Communications of the ACM* 38,(7), pp. 40 – 48.

Carroll,J.M., Bellamy,R.K.E. &Rosson,M.B.(1995). *Smalltalk scaffolding: a case study of minimalist instruction.* In Conference proceedings on Empowering people: Human factors in computing system: special issue of the SIGCHI Bulletin,Seattle: Washington, pp. 423-430.

Charney,D.H. and Reder, L.M. (1986). Designing interactive tutorials for computer users. Human- Computer Interaction, (2) , pp. 297-317.

Compeau, D. and Higgins, C. Application of Social Cognitive Theory to Training for Computer Skills. Information Systems Research, Vol. 6 (2), June 1995, pp. 118-143.

Cuomo, D. L., & Bowen, C. D. (1992). Stages of user activity model as a basis for user-system interface evaluations. In Proceedings of the Human Factors Society 36th Annual Meeting (pp. 1254-1258). Santa Monica, CA: Human Factors and Ergonomics Society

Davis, S.A. and Bostrom, R.P. (1993). Training end users: an experimental investigation of the roles of the computer interface and training methods. MIS Quarterly, (17), pp. 61-85.

Davis, D. L. & Davis, D. F. (1990). The effect of training techniques and personal characteristics on training end users of information systems. Journal of Management Information Systems, 7(1), 93-110.

Dick, W. and Carey, L. (1996). The Systematic Design of Instruction, Fourth Edition. Addison Wesley Longman, Inc. New York.

Dillon, A. (2000). Designing a better learning environment with the web: problems and prospects Cyberpsychology and behavior 3 (1). Mary Ann Liebert inc

Fekete, A. Greening, T. & Kingston, J. (1998). Conveying technical content in a curriculum using problem based learning. In Proceedings of the third Australasian conference on Computer science education.

Galitz,W.O. (1997). The essential guide to user interface design: an introduction to gui design principles and techniques. New York .NY. Wiley computer publishing

Goldstein, I.L. (1993). Training In Organizations- Needs Assessment, Development, and Evaluation, 3rd Edition. Brooks/Cole Publishing Company. Pacific Grove, California

Goodman, A.& Street, S. (1998). Some experimental evidence on the educational value of interactive Java applets in Web-based tutorials. In Proceedings of the third Australasian conference on Computer science education. Brisbane, Australia.

Hartson, H.R., Andre, T.S., Williges, R.C. & Rens, L.V. The User Action Framework: A theory-based foundation for inspection and classification of usability problems. In Proceedings of the 8th International Conference on Human-Computer Interaction (HCI International '99): Ergonomics and User Interfaces, Vol. 1. Mahway, NJ: Lawrence Erlbaum Associates, 1058-1062.

Hartson , H.R. (2000). CS-5714 : Usability Engineering course pack.

Harasim , L. (1990). On-line education: Perspectives on a new environment. New York: Praeger.

Hiltz, R. S. (1994). The virtual classroom. Norwood, NJ: Ablex Publishing Corporation.

Hix, D., & Hartson, H. R. (1993). Developing user interfaces: Ensuring usability through product and process. New York: John Wiley & Sons, Inc.

Kilby,T. Advantages and Disadvantages of web based training http://www.filename.com/wbt/pages/advdis.htm Last viewed on : June 10th 2001

Larkin, J. H., & Chabay, R. W. (1992). Computer-assisted instruction and intelligent tutoring systems: Shared goals and complementary approaches. Hillsdale, NJ: Lawrence Erlbaum Associates.

Lautar-Lemay, C. (1987). Comparison of Computer-Assisted-Learning With Traditional Lecture and Reading Assignment. In Proceedings of the International Conference on Computer Assisted Learning in Post-Secondary Education. 293-294 Calgary, Canada .

Lohr, L.L. (2000). Designing the instructional interface. Computers in human behavior 16(1) (pp 161-182)

Lee A.Y., Gillan D. J., Upchurch E.E. & Melton, J. S. (1995). Is multimedia-based training effective?: yes and no. Conference companion on Human factors in computing systems pp. 111 – 112.

Mack, R. L., & Nielsen, J. (1994). Executive summary. In Nielsen, J., & Mack, R. L. (Eds.), Usability inspection methods, 1-23. New York, NY: John Wiley & Sons. Reprinted in Baecker, R. M., Grudin, J., Buxton, W. A. S., & Greenberg, S. (1995), Readings in human-computer interaction: Toward the year 2000, 170-181. 2d ed. San Francisco, CA: Morgan Kaufmann Publishers.

Mckeachie, W.J., Pintrich, P.R., Lin, Y., Smith, D.A.F. y Sharna, A, R. (1986). Teaching and Learning in the College Classroom. Michigan: NCRIPTAL, The University of Michigan.

Nash, A.N., Muczyk, J.P. & Vettori, F.L. (1971). The relative practical effectiveness of programmed instruction. Personnel Psychology, (24), pp. 397-418

Nielsen, J. (1993). Usability engineering. Boston: Academic Press.

Nielsen, J. and Levy. J (1994). Measuring usability: preference vs. performance. Communications of the ACM 37 (4) pp 66 – 75.

Norman, D.A. (1986). Cognitive engineering. In D. A. Norman & S. W. Draper (Eds.), User centered system design: New perspectives on human-computer interaction (pp. 31-61). Hillsdale, NJ: Lawrence Erlbaum Associates.

O' Lander, R. (1990). Use of intelligent computer aided instruction in the classroom. In Proceedings of the 1990 ACM annual conference on Cooperation

Owston, R. D. (2000). Evaluating web based learning environments : strategies and insights.Cyberpsychology and behavior 3 (1). Mary Ann Liebert inc.

Pane, J.P., Corbett,A.T. & John,B.E**.** (1996). Assessing dynamics in computer-based instruction. Conference proceedings on Human factors in computing systems pp. 197-204.

Reigeluth, C. M. & Nelson, L. M. (1997). A new paradigm of ISD?. Educational media and technology yearbook (22), pp. 24-35.

Souza, R. K. (2000). The best of retail site design. Cambridge, MA: Forrester Research Inc.

User Action Framework Explorer Tool. http://hemlock.cs.vt.edu Last viewed on : July 20th 2001.

Usability Methods and Tools Lab at Virginia Tech. http://miso.cs.vt.edu/~usab
Last viewed on: July 20th 2001

Van der berg, S. & Watt, J. (1991). Effects of educational setting on student responses to structured hypertext. Journal of computer based instruction. 18(4) 118-124.

Wade, V.P. & Power, C.(1998). Evaluating the design and delivery of WWW based educational environments and courseware. Proceedings of the 6th annual conference on the teaching of computing/3rd annual conference on integrating technology into computer science education on Changing the delivery of computer science education 243 – 248.

Wiedenbeck, S., Zila, P.L., McConnell, D.S. : End-User Training: An Empirical Study Comparing On-Line Practice Methods. CHI-1995. 74-81

Williges, R. C. & Hartson, H. R. (1986). Human-computer dialogue design and research issues. In R. W. Ehrich & R. C. Williges (Eds.), Human-computer dialogue design. Amsterdam: Elsevier. 367-376.

APPENDICES

Appendix A: Needs Analysis

Data Collection

To collect information that will help determine various requirements of training program, a combination of three data collection procedures were used. Questionnaire for untrained incumbents (usability practitioners), subject matter expert interviews and user surveys for graduate students (who have taken the usability engineering course) were selected as optimal methods to collect data. An effort is made in this section to explain these procedures in detail.

Questionnaires

It was decided to administer a questionnaire to untrained incumbents, the usability practitioners in the real field. They are the main trainee class for this project and also the main user class of User Action Framework. However, since our subjects have no exposure to UAF, a word document giving a brief introduction about UAF was also sent with the questionnaire. They have had experience in the usability field and encountered real-world problems in the same. The questionnaire consisted of 11 open-ended questions and was sent out to 10 usability practitioners. The subjects had varying backgrounds in fields related to usability engineering and were selected from User Works, Inc, a consultancy firm specializing in usability engineering. The questionnaire was designed to collect information about users' training requirements. The questionnaires were sent via e-mail since our subjects were remotely located.

Subject matter expert interviews

Three subject matter experts were interviewed to obtain information about some basic aspects about User Action Framework. The subject matter experts were identified as the people involved in the development of the UAF tools. They have an "in-depth" knowledge of the UAF concepts, which would be necessary for the development of the tools. Keeping the questions) open ended ensured that they would be able to give us input on issues that we may have not considered as yet and yield more information than an objective questionnaire. The answers and comments were recorded using a tape recorder and transcribed to get the most important details.

User survey

Fifteen graduate students who have taken usability engineering course at Virginia Tech participated in a Likert type scale survey (Appendix 4). The course provided the students an exposure to the UAF concepts. They were asked to classify some usability problems according to the structure of UAF in their homework assignments. The survey was useful to collect information about users' skills and preferences.

Results

Questionnaires

Table 1 summarizes the major findings from the questionnaires given out to ten usability practitioners:

Questions	**Answers**
What is your academic and professional background?	• MS – Human Factors Engineering., Focus on HCI • MS, PhD – Experimental Psychology • MS – Technical Communication • 3+ years experience in usability evaluation (Including usability testing and heuristic evaluation)
What is your prior experience with post-usability testing procedures like data classification, data analysis etc?	• Informal classification of data on a project- specific basis • Using spreadsheets to classify data • Classification according to severity ratings (e.g., Low, Medium, High) and make global level design recommendations
Have you had exposure to any tool for	• No

performing post usability testing data analysis?	• Yes - most seem to be too laborious for the results returned
Have you ever received formal instruction in conducing a post-usability data classification, analysis etc? If yes, please explain briefly	• No • Referred to several other reports and created own style of classification and reporting
Do you do post-usability testing procedures alone or in teams? • Alone	• Alone • In teams • Both
How do you feel about being trained on understanding a knowledge base, to be able to use a tool with real world usability data?	• Would welcome it • Time and money expended versus the return should be considered
What do you expect to learn out of a training program like this for the UAF?	• To consistently classify usability problems • Vocabulary, tools, and standards shared by usability practitioners • Underlying framework and context for use, procedures for use, limitations and constraints, future directions, time estimates for use
Preference for training medium	• Self paced instructions with manual available • One-on-one

	• Lecture / classroom
Given the scope of the UAF, what do you see as constraints (personal and organizational) towards being trained on the UAF?	• Time • Money • Acceptance among usability practitioners • Accommodating new hires i.e. to get people up-to speed

Table 1: Findings from Questionnaires

Subject matter expert interviews

Presented below in table 2 is a summary of the results collected from interviewing the subject matter experts. The questions asked during the interview have been given in appendix 3.

Table 2: Findings from interviews

Questions	**Answers**
How did you learn to use the UAF?	• Existing training. (CS 5714 course) • Self-learning by trial and error
Do you think this training was adequate?	• No • Yes
Is the UAF easy to learn?	• Background dependent. • No. Background independent. • In its current version, no comments.
What background would you consider necessary to understand the UAF?	• Human factors/cognitive psychology background would help more than anything else • Even brief exposure to HF basics would

	help • Strongly favors HF background
What are the difficulties that you encountered when learning to use the UAF? (Problem areas)	• Navigating the UAF tree • Navigation- the interface does not support easy navigation • Differentiating between similarly worded causative factors at the lowest level
Following from your difficulties, what do you think should be done to avoid them?	• High-level descriptions that goes deeper into the UAF tree • Rewording key identifiers in the UAF to avoid ambiguity • Large number of carefully worded examples
What do you see as constraints to a training program for the UAF?	• Still in development and hence not yet in a stable form • Introduces a lot of new terminology that might not be too simple to understand
Comments/suggestions	• Walking though the decision-making process and Norman's model would help understand the UAF • Having a large problem base to work with while training should increase accuracy in classification

	• Providing for instantaneous feedback when solving example problems

1.4.3 User Survey

The following graph (Fig 1) shows the skill level classification of the respondents. As mentioned earlier, fifteen graduate students who have taken usability-engineering course at Virginia Tech participated in this survey.

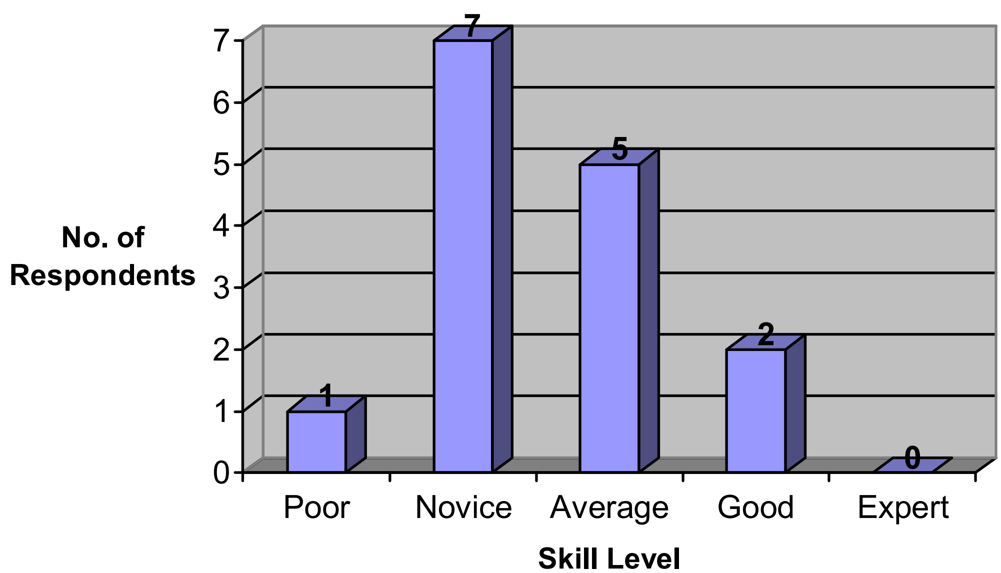


Figure 1: Skill level classification

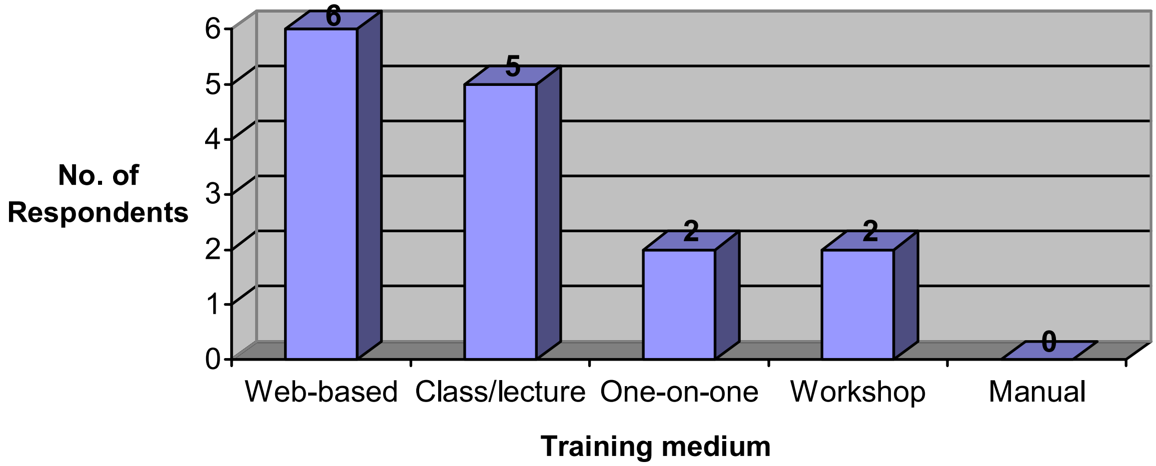
Results of the survey are given in Table 3. Results are presented as average score for each statement

Table 1: Results from user survey

Statement	Average Value	Corresponding label

The concept behind the UAF is easy to learn	2.2	Disagree
I had trouble traversing the UAF tree	4.6	Strongly agree
What do you think about the structure in which the UAF is organized?	3.1	OK
Have you received any training at all to use the UAF other than what was taught in the class?	1.3	None
The language used in the UAF is too technical even for people with some usability knowledge to understand	3.3	Neither disagree nor agree
The UAF introduces a lot of new terminology (e.g., Translation, Affordance etc)	3.9	Agree
Training is necessary to effectively use the UAF	4.2	Agree

Fig 2 shows the respondents' preference towards effective training medium to distribute the training content:



(source: Balasubramanian et al., 2000 URL: http://hci.ise.vt.edu/williges/ise5634/uaf.pdf)

Appendix B: Checklist for SME evaluation

Instructions: please complete the following questions by choosing the appropriate choices.

- Is information presented in relevant context?

  YES _______ NO ________

  Comments:

- Is the organization of information into relevant sets and subsets (modules)?

  YES _______ NO ________

  Comments:

- Does Feedback contain information about correctness of a response as well incorrectness of response?

  YES _______ NO ________

  Comments:

- Do examples, contexts and applications progress from simple to complex, familiar to unfamiliar and concrete to abstract?

  YES _______ NO ________

  Comments:

- Are measures of performance congruent with goals and objectives of instruction and target learner characteristics?

YES _______ NO ________

Comments:

- Does the instructional content maintain accuracy of facts,figures,data etc.?

YES _______ NO ________

Comments:

- Is there sufficient use of lists/tables and figures to summarize and encapsulate information?

YES _______ NO ________

Comments:

- Is repetitive practice provided to enable learners to smooth out the routine and if possible automate the skill?

YES _______ NO ________

Comments:

In the space provided below, indicate if any, features on the UAF that warrants inclusion to or exclusion from the training.

Inclusions:

Exclusions:

Appendix C: Questionnaire for formative evaluation

Instructions: Please indicate how strongly you disagree or agree to the statements by circling using the scale provided. Try to respond to all the items.

There are clear navigational aids that tell me where I am.

Strongly Disagree	Disagree	Neither Disagree nor Agree	Agree	Strongly Agree

It is easy to find the information I need using the training tool.

Strongly Disagree	Disagree	Neither Disagree nor Agree	Agree	Strongly Agree

I like using the UAF training tool system.

Strongly Disagree	Disagree	Neither Disagree nor Agree	Agree	Strongly Agree

The training tool system has all the functions and capabilities that I expect it to have.

Strongly Disagree	Disagree	Neither Disagree nor Agree	Agree	Strongly Agree

I am able to complete my tasks easily using the interface.

Strongly Disagree	Disagree	Neither Disagree nor Agree	Agree	Strongly Agree

The UAF Explorer tool makes it easier for me to organize my thoughts about a usability problem.

Strongly Disagree	Disagree	Neither Disagree nor Agree	Agree	Strongly Agree

Instructions:

Please indicate your subjective opinions on the interface using the number scale provided. Try to respond to all the items.

Reading characters on screen

Hard	-2	-1	0	1	2	Easy

Organization of the Information

Confusing	-2	-1	0	1	2	Very clear

Sequence of screens

Confusing	-2	-1	0	1	2	Very clear

Use of terms throughout tasks

Inconsistent	-2	-1	0	1	2	Consistent

Use of popup windows

Unhelpful	-2	-1	0	1	2	Helpful

Overall reaction to the interface

Terrible	1	2	3	4	5	Wonderful
Difficult	1	2	3	4	5	Easy
Frustrating	1	2	3	4	5	Satisfying
Dull	1	2	3	4	5	Stimulating

Using the UAF Explorer will make understanding usability concepts

harder										easier
0	1	2	3	4	5	6	7	8	9	10

End of questionnaire. Thank you.

Appendix D: Informed consent for Formative Evaluation.

INFORMED CONSENT FORM

Title of Project: Evaluation of Training Tool for the UAF Explorer
Principal Investigators: Venkatramanan Balasubramanian, Dr. Rex H. Hartson

PURPOSE OF PROJECT

You are invited to participate in an evaluative study to evaluate a web based training tool for the User Action Framework explorer tool.

INFORMATION

This activity will require about one session of 2-3 hours. These sessions will take place at the Usability methods and Tools lab at 102,McB.

Sessions will be videotaped. We will use the tapes to gather required important information you provide during the session.

In each session, you will be given a task list and asked to read through training modules and complete tasks associated with those scenarios. Computers will be used for each task. Our goal is NOT to evaluate YOUR skills, but to determine how the design of instruction and information can be improved.

RISKS

There are no physical or emotional risks associated with this research project.

BENEFITS

At the end of this session, you will be provided with a copy of this form. At the bottom of this form, you will find contact information that can be used to contact the principal investigators after the research has been completed in order to receive information about the results.

CONFIDENTIALITY

The information gained in this research project will be kept strictly confidential. At no time will the researchers release the results of the study to anyone other than individuals working on the project without your written consent. You will be identified only by a 3-digit study code. Data will be stored securely and will be made available only in the context of research publications and discussion. No reference will be made in oral or written reports that could link you to the data nor will you ever be identified as a participant in the project. When clips from video-tapes are used for presentation, your faces and name-references will be edited so that your confidentiality is protected

COMPENSATION

For your participation you will be compensated at $7.50 an hour.

FREEDOM TO WITHDRAW

You are free to withdraw from this study at any time without penalty.

APPROVAL

This research project has been approved, as required, by the Institutional Review Board for Research Involving Human Subjects at Virginia Polytechnic Institute and State University and by the Department of Industrial and Systems Engineering.

PARTICIPANT'S RESPONSIBILITIES

It is very important that you keep the activities and information discussed confidential, since others will be participating in this research. You will be encouraged to think aloud during the course of the study.

QUESTIONS

If you have questions, or do not understand information on this form, please feel free to ask them now.

PARTICIPANT'S PERMISSION

I have read and understand the Informed Consent and conditions of this project. I have had all questions answered. I hereby acknowledge the above and give my voluntary consent for participation in this project.

If I participate, I may withdraw at any time without penalty. I agree to abide by the rules of this project.

Signature ____

Date ____

CONTACT

If you have questions at any time about the project or the procedures, you may contact the principal investigators, Drs. Rex Hartson at 231-6270, hartson@vt.edu or Tonya Smith-Jackson at 231-4119, smithjack@vt.edu .

If you feel you have not been treated according to the descriptions in this form, or your rights as a participant have been violated during the course of this project, you may contact Dr. David Moore, Acting Chair of the Institutional Review Board Research Division at 231-5281.

Appendix E: Task list for Small group evaluation/Usability testing in Formative Evaluation

Instructions: You are allowed to move between modules as and when you wish as long as you complete the practice questions marked out here.

Task 1: Read module 1 thoroughly

Task 2: Read module 2 thoroughly and attempt practice session Q's 1- 10.

Task 3: Read module 3 and attempt practice sessions questions. See the solution only after you have tried classifying.

Solution for practice session 1:

Solution for practice session 2:

Task 4: Read module 4 thoroughly and attempt practice sessions. See the solution only after you have tried classifying.

Solution for practice session 1:

Solution for practice session 2:

Task 5: Read module 5 and attempt practice sessions. See the solution only after you have tried classifying.

Solution for practice session 1:

Solution for practice session 2:

Task 6: Read module 7 thoroughly and attempt practice sessions. See the solution only after you have tried classifying.

Solution for practice session 1:

Solution for practice session 2:

Task 7: Read module 8 and attempt practice session. See the solution only after you have tried classifying.

Solution for practice session 1:

Appendix F: Data from Formative Evaluation

Table A-1: Time taken for task completion during formative evaluation.

Task	Time taken for participant #			
	1	2	3	4
Read module 1	20	22	19	20
Read module 2 and attempt the practice session.	14	18.5	20.5	18
Read module 3 and attempt the practice session.	22	18.5	24	27
Read module 4 and attempt the practice session.	21	17	20	26
Read module 6 and attempt the practice session.	16	18	23	28
Read module 7 and attempt the practice session.	10	5.5	6.5	14
Total Time	**103**	**99.5**	**128**	**133**

Table A-2: Accuracy scores on iteration 1 for small group evaluation.

Question #	Accuracy score for participant #				Mean
	1	2	3	4	
1	10	10	9	7	9
2	8	7	10	6	7.75
3	10	8	10	7	8.75
4	5	6	9	8	7.0
5	8	10	8	7	8.25
6	10	10	9	9	9.50
7	6	7	8	6	6.75
8	10	8	10	9	9.25
9	7	8	10	8	8.25
10	8	6	9	9	8.0
Total	**83**	**80**	**92**	**76**	
Average	**82.75**				

Table A-3: Time taken for task completion during formative evaluation.

Task	Time taken for participant #		
	1	**2**	**3**
Read module 1	21.5	16	18
Read module 2 and attempt the practice session.	17	20	16.5
Read module 3 and attempt the practice session.	20	16	19
Read module 4 and attempt the practice session.	25	14.5	22
Read module 6 and attempt the practice session.	18	14	15
Read module 7 and attempt the practice session.	12.5	6.5	8
Total Time	**114**	**87**	**98.5**

Table A-4: Accuracy scores on iteration 2 for small group evaluation

Question #	Accuracy score for participant #			Mean
	1	**2**	**3**	
1	8	10	6	8
2	7	7	9	7.6
3	10	8	8	8.6
4	6	8	10	8.0
5	7	9	8	8.0
6	6	10	8	8.0
7	4	8	6	6.0
8	10	8	7	8.34
9	8	8	8	8.0
10	8	8	8	8.0
Total	**74**	**84**	**78**	
Average	**78.6**			

Table A-5: Summary of mean accuracy scores on questions.

Question #	Mean(SD) scores averaged over both iterations (n =4+3)
1	8.50(1.0)
2	7.67(1.1)
3	8.67(0.75)
4	7.50(1.04)
5	8.12(0.89)
6	8.75(0.95)
7	6.37(1.49)
8	8.80(0.80)
9	8.12(0.69)
10	8.00(0.57)

Appendix G: Pretest for classroom training group and web-based training group.

Instructions

1. This is a test used only for a study and should not take more than an hour or so to complete. We know that you have not yet covered this topic in class but we wanted to get an idea of whether people can use this tool without any given training at all. Don't worry if you can't answer or if you are not confident of your answers. Your answers will not count against your grade. Please use the following URL: http://hemlock.cs.vt.edu/training
to classify the questions under the structure of the UAF as discussed in class.

2. The solution to a problem is the path shown in the topmost part of the web page when you have gone to the lowest level. These are clickable links in case you need to go back to other parts of the UAF. Fill in an alternate solution only if you are unsure if your first choice or can't decide between the two choices.

3. Keep explanations brief and to the point. Explanation is the reason why you chose that path to classify the usability problem using the Explorer.

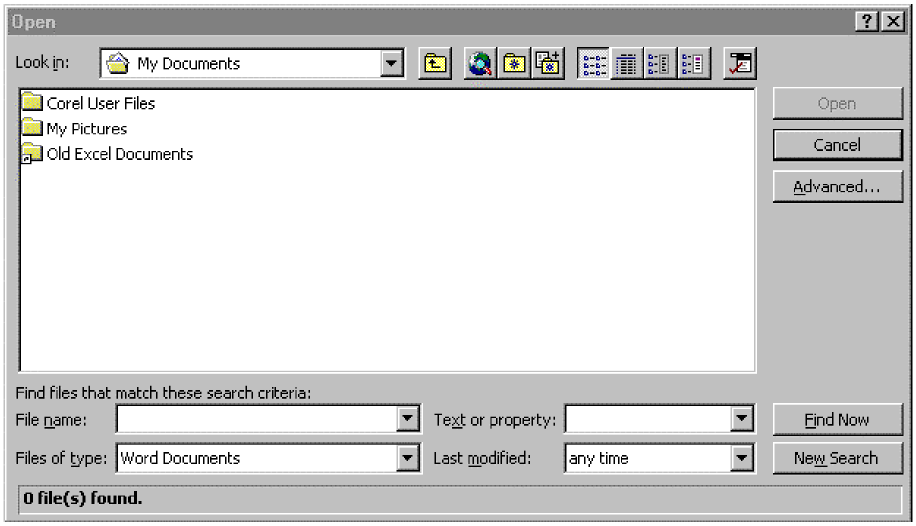
4. ***After*** you finish the problem set, we would like you to complete a short questionnaire that seeks some information about you. Thanks for participating in this study.

Reminder: You will ***not*** be graded on accuracy of solutions for this assignment, but on whether you attempt these questions using the UAF and turn it in.

Please write down your name and student ID number at the top of this sheet and return all the sheets together with the questionnaire.

1. The Advanced button

The Open dialogue box of word has a set of buttons on the upper right hand corner that are labeled :Open, Cancel and Advanced..



These appear to be a set of buttons directly related to the Open function. The lower portion of the dialogue box is about the Find function, which can be needed in conjunction with the Open function. Offering the Find as part of the Open function is good design because it gives an affordance to a likely related task. This problem, though, is about the button, up in the Open group, labeled "Advanced." It turns out that this button actually goes with the Find function. When you click on it, it leads to the "Advanced Find…" dialogue box. This can be confusing, disrupting, and distracting.

Solution (first choice):

Alternate Solution:

Explanation:

2. ***Resetting the screensaver timeout***

A user did not want to be bothered by the screen saver kicking in during lulls in computer usage during the day. So he decided to reset the timeout on the screensaver. This user pulled up the control panel for Display and clicked on the screen saver tab and saw a screensaver field labeled "Wait", followed by a text box containing "15", followed by "minutes". Wanting to set it for at least and hour and a half before the screensaver activated, he selected the "15" and typed over it with "90", and clicked on OK. The dialogue box disappeared, and he returned to his work. He was a little surprised sometime later when the screensaver activated after only 60 minutes. It turns out that this feature is limited to 60 minutes.

But this case isn't about that problem. This problem is about the fact that, after performing the task, the user wasn't aware that he had entered an erroneous (not accepted) value. After he clicked on OK, the dialogue box apparently disappeared and no indication was given that the input wasn't fully accepted.

Solution (first choice):

Alternate Solution:

Explanation:

3. ***Doesn't like the colors***

After performing several tasks during a usability testing session, a user voluntarily commented that she thought the whole interface suffered from bad use of colors. In some places the colors were too bright and hard to look at for very long. In other cases, the colors for different objects were too close to each other and difficult to distinguish.

<u>**Solution (first choice):**</u>

<u>**Alternate Solution:**</u>

<u>**Explanation:**</u>

4. **User wants to reuse document numbers in a document retrieval system.**

A user of a personal document retrieval system has been deleting numbered documents. The user now wants to reuse the old document numbers, but the system does not allow this.

<u>**Solution (first choice):**</u>

Alternate solution:

Explanation:

5.***Doesn't understand spreadsheets***

A relatively new user of spreadsheets thought he knew at least the basics. However, he was occasionally surprised and confused about certain features such as the automatic re-calculation of formulae in some cells when values here were changed in other cells. It became clear that he didn't know how to approach his assigned tasks because he really didn't understand enough about how spreadsheets work.

Solution (first choice):

Alternate Solution

Explanation:

6. ***Message Tone***

A user of a database application gets this message five seconds after invoking a command to do a database transaction: "Still processing your request. Guess I'm a little slow today." This problem is just about the tone of the message.

Solution (first choice):

Alternate Solution

Explanation:

7. **No-Fill Choice not obvious**

In PowerPoint, the user wanted to remove any fill color from a rectangle she had drawn. The fill color tool on the graphics tool bar has a "no fill" choice at the top, but it does not look like an active button, even when you put the cursor over it. The user who wanted no fill, went by that button, not knowing it could be selected, and used the white fill choice in the palette below, but that was the wrong choice and had to be undone.

Solution (first choice):

Alternate solution:

Explanation:

8. **Icon too small**

In a CAD/CAM program, the icon to print a page is very small and located too close to other icons. As a result expert users, who are usually working very fast, often click on the wrong icon and lose time having to recover.

Solution (first choice):

Alternate Solution:

Explanation:

9. **Exit label**

The label on a button for dismissing a dialogue box says, "Exit." The user thinks this is probably the way to leave the dialogue box, but is hesitating because of the uncertainty and a little intimidation caused by the word "Exit." The user is worried that it might cause an exit from the whole system, possibly losing some work.

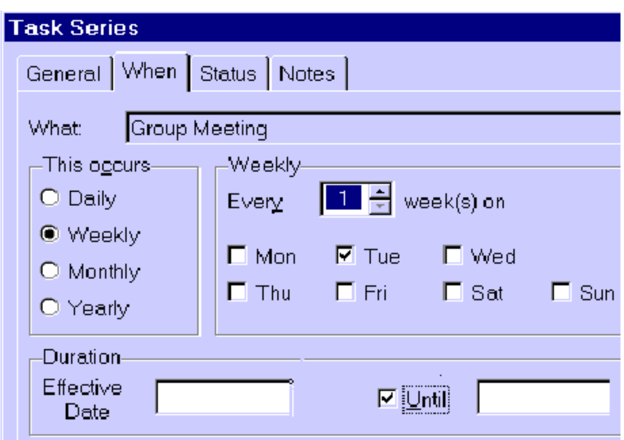
Solution (first choice):

Alternate solution:

Explanation:

10. ***Data entry format missing***

User is filling out an on-line form and gets to a field for a date, but there is no indication about what format to use. The user tries something and gets an error message and then is able to correct it and get the system to recognize the date.



.

Solution (first choice):

Alternate Solution:

Explanation:

Thank you for participating in this evaluation of the training tool for the User Action Framework. Please answer the following questions as clearly as you can.

1. Name: ____

2. *Gender*: Male Female *Academic Level:* Masters PhD

3. Major of Study: _____________________________________

4. How many university-level courses have you taken that have addressed human factors, HCI or other usability concepts such as evaluation methodologies or other usability techniques. *(Please check one*):

None

1 - 2

3- 5

5+

Besides academic courses, do you have any practical experience in usability evaluation or human factors/HCI concepts. If yes, state the nature and duration of your experience.

YES NO

____

____

5. Based on the limited exposure you have had on the UAF, how would you rate your skill on classifying problems using the UAF at this stage? *(Circle your choice)*

Poor Fair Good Excellent

6. What is the frequency of your use of: (Please circle your choice)

Computers

Very Rarely | Rarely | Sometimes | Frequently | Very frequently

Web browsers

Very Rarely | Rarely | Sometimes | Frequently | Very frequently

7. Rank order your preference of medium for delivery of training for acquiring computer based skills (like learning a software package etc.). Please use 1, 2, 3, and 4 to rank your preferences. Use 1 for "most preferred", 2 and 3 as your next respective preference ranks, and 4 as "least preferred". Please use only 1 number for each training delivery system and use each number only once.

Web based training ____

Classroom training ____

Paper based/Instruction manual ____

Audio/Video training ____

8. Have you used ever used an on-line instructional tool (ex. training tool, on-line course)?

Yes No

If yes, please describe this tool:

____

____

9. Do you think it would be difficult to learn the structure and function of the User Action Framework Explorer tool? Why or why not?

Thanks for completing this questionnaire. If you have further questions, you can contact the researcher at: vbalasub@vt.edu.

Appendix H: Posttest for classroom training and web-based training groups

Instructions

5. Your answers will count against your grade. Please use the following URL: http://hemlock.cs.vt.edu/training to classify the questions under the structure of the UAF as discussed in class.

6. The solution to a problem is the path shown in the topmost part of the web page when you have gone to the lowest level. These are clickable links in case you need to go back to other parts of the UAF. Fill in an alternate solution only if you are unsure if your first choice or can't decide between the two choices.

7. Keep explanations brief and to the point. Explanation is the reason why you chose that path to classify the usability problem using the Explorer.

4. Note that the version of the UAF explorer you used will have changed from the last study. The changes are minor and designed to make using the explorer easier. It includes color-coding, increasing font sizes and so on.

Reminder: You **will** be graded on accuracy of solutions for this assignment.

- **Please fill in your name and student ID number at the top of this sheet and email the document to** vbalasub@vt.edu**and** dtessend@vt.edu **.**
- **In the subject of the email, include: 'CS5714 - HW#3'**

1. ***Message Tone***

A user of a database application gets this message five seconds after invoking a command to do a database transaction: "Still processing your request. Guess I'm a little slow today." This problem is just about the tone of the message.

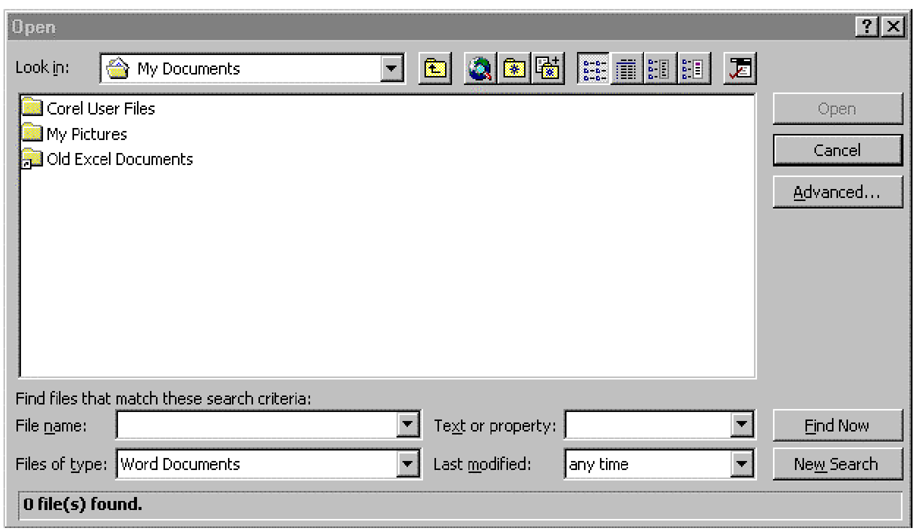
Solution (first choice):

Alternate Solution

Explanation:

2. The Advanced button

The Open dialogue box of word has a set of buttons on the upper right hand corner that are labeled :Open, Cancel and Advanced..



These appear to be a set of buttons directly related to the Open function. The lower portion of the dialogue box is about the Find function, which can be needed in conjunction with the Open function. Offering the Find as part of the Open function is good design because it gives an affordance to a likely related task. This problem, though, is about the button, up in the Open group, labeled "Advanced." It turns out that this button actually goes with the Find function. When you click on it, it leads to the "Advanced Find..." dialogue box. This can be confusing, disrupting, and distracting.

Solution (first choice):

Alternate Solution:

Explanation:

3. **No-Fill Choice not obvious**

In PowerPoint, the user wanted to remove any fill color from a rectangle she had drawn. The fill color tool on the graphics tool bar has a "no fill" choice at the top, but it does not look like an active button, even when you put the cursor over it. The user who wanted no fill, went by that button, not knowing it could be selected, and used the white fill choice in the palette below, but that was the wrong choice and had to be undone.

Solution (first choice):

Alternate solution:

Explanation:

4. **Icon too small**

In a CAD/CAM program, the icon to print a page is very small and located too close to other icons. As a result expert users, who are usually working very fast, often click on the wrong icon and lose time having to recover.

Solution (first choice):

Alternate Solution:

Explanation:

5. ***Resetting the screensaver timeout***

A user did not want to be bothered by the screen saver kicking in during lulls in computer usage during the day. So he decided to reset the timeout on the screensaver. This user pulled up the control panel for Display and clicked on the screen saver tab and saw a screensaver field labeled "Wait", followed by a text box containing "15", followed by "minutes". Wanting to set it for at least and hour and a half before the screensaver activated, he selected the "15" and typed over it with "90", and clicked on OK. The dialogue box disappeared, and he returned to his work. He was a little surprised sometime later when the screensaver activated after only 60 minutes. It turns out that this feature is limited to 60 minutes.

But this case isn't about that problem. This problem is about the fact that, after performing the task, the user wasn't aware that he had entered an erroneous (not accepted) value. After he clicked on OK, the dialogue box apparently disappeared and no indication was given that the input wasn't fully accepted.

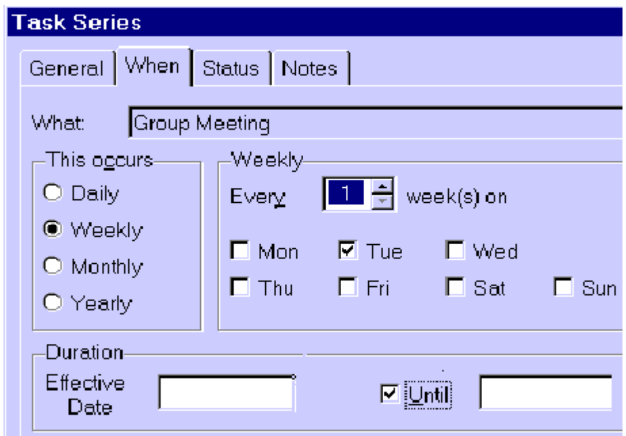
Solution (first choice):

Alternate Solution:

Explanation:

6. ***Data entry format missing***

User is filling out an on-line form and gets to a field for a date, but there is no indication about what format to use. The user tries something and gets an error message and then is able to correct it and get the system to recognize the date.



.

<u>Solution (first choice):</u>

<u>Alternate Solution:</u>

<u>Explanation:</u>

7. ***Doesn't like the colors***

After performing several tasks during a usability testing session, a user voluntarily commented that she thought the whole interface suffered from bad use of colors. In some places the colors were too bright and hard to look at for very long. In other cases, the colors for different objects were too close to each other and difficult to distinguish.

Solution (first choice):

Alternate Solution:

Explanation:

8. ***User wants to reuse document numbers in a document retrieval system.***
A user of a personal document retrieval system has been deleting numbered documents. The user now wants to reuse the old document numbers, but the system does not allow this.

Solution (first choice):

Alternate solution:

Explanation:

9. ***Exit label***
The label on a button for dismissing a dialogue box says, "Exit." The user thinks this is probably the way to leave the dialogue box, but is hesitating because of the uncertainty and a little intimidation caused by the word "Exit." The user is worried that it might cause an exit from the whole system, possibly losing some work.

Solution (first choice):

Alternate solution:

Explanation:

10. ***Doesn't understand spreadsheets***

A relatively new user of spreadsheets thought he knew at least the basics. However, he was occasionally surprised and confused about certain features such as the automatic re-calculation of formulae in some cells when values here were changed in other cells. It became clear that he didn't know how to approach his assigned tasks because he really didn't understand enough about how spreadsheets work.

Solution (first choice):

Alternate Solution

Explanation:

Appendix I : Informed consent for Summative Evaluation

Virginia Polytechnic Institute and State University

Informed Consent for Participants of Investigative Projects

Title of Research Project: Evaluation of Training for the UAF Explorer tool.

Investigators:

VenkatRamanan Balasubramanian, Dr.Tonya Smith Jackson, and Dr.Rex Hartson

Phone No. (Off): (540) 231 4119

Location of Investigation: Blacksburg, Virginia

Purpose of Study:

This study is being conducted to evaluate a training program design for the User Action Framework Explorer Tool and measure the effectiveness of web based training tool to train people to classify usability problems based on the UAF.

Principal investigators

Venkatramanan Balasubramanian

Dr. Tonya L. Smith Jackson, PhD

Dr.Rex Hartson, PhD

Procedure

If you decide to participate in this study, the test administrator will describe the procedure to you verbally during the time of the experiment. A brief description is as follows: the participant will have to come to one of two labs that will serve as training locations. Each participant will have to come in for 1 to 3 training sessions over a week's time and each session will last approximately 75 minutes. At the end of all the sessions the participant will take a post-test and answer a questionnaire to help evaluate the effectiveness of the training tool.

Rights

Any published data will not reveal your identity. You will be assigned a Participant Identification Number so that we may keep all data collected confidential. If you decide to participate, you are free to withdraw your consent at any time without penalty.

If you have any questions, we expect you to ask via email or telephone. The test administrators will be happy to answer them. Your electronic or real signature indicates that you have read and understand the above information that your questions have been answered to your satisfaction, and that you have decided to participate based on the information provided on this form.

Risks

No specific risks have been identified though it can be mentioned that the risks encountered in this evaluation will be minimal and are typical of the everyday risks encountered by personal computer use.

Benefits

Your participation in this project will provide information that may be used to improve the usability and effectivenessof the training tool for the UAF. No guarantee of benefits has been made to encourage you to participate. You may receive a synopsis summarizing this research when completed. Please leave a self-addressed envelope with the experimenter if you wish a copy of the results to be sent to you.

Confidentiality

Confidentiality of personal information identifying the subject will be maintained. This information will be excluded from the test data by the assignment of a participant code at the time of the administration of the test and will only be made available for official purposes.

Approval of Research

This research has been approved, as required, by the Institutional Review Board for projects involving human subjects at Virginia Polytechnic Institute and State University, and by the Department of Industrial and Systems Engineering.

Participant's rights

Any published data will not reveal your identity. Your participation in this evaluation is voluntary. If you choose not to participate in this evaluation, or later wish to withdraw from any portion of it, you may do so without penalty.

Compensation.

Students of the CS5714 class participating in this experiment will receive extra credit towards their coursework and a chance to complete training on the UAF, which forms part of their normal curriculum.

Cautions

a. The test participant may withdraw from the study at any time without any resulting consequences.

b. The number of subjects involved in the study will be approximately 30 to 40.

c. No precautions are required to be observed by the subject either before or after the study.

Points of contact

Venkatramanan Balasubramanian:

Email: vbalasub@vt.edu

phone: (540) 961 9091 (Home)

Dr.Tonya L. Smith Jackson

smithjack@vt.edu

Phone: (540) 231 4119 (office)

Disposition of Volunteer Agreement Affidavit

The principal investigator will retain the original signed Volunteer Agreement Affidavit and forward it to the chair of the Human Use Committee after the investigation. An electronic copy will be provided to the volunteer by the test administrator.

Your signature indicates (1) that you have read the information on this form, (2) that you have been given the opportunity to ask questions and those questions have been answered to your satisfaction, and (3) that you have decided to participate based on the information provided on this form.

I have read and understand the Informed Consent and conditions of this project. I have had all my questions answered. I hereby acknowledge the above and give my voluntary consent for participation in this project. If I participate, I may withdraw at any time without penalty.

PRINTED NAME OF PARTICIPANT **DATE**

______________________________ ______________

SIGNATURE OF PARTICIPANT

____

NAME OF TEST ADMINISTRATOR SIGNATURE

______________________ ___________________________

If you feel you have not been treated according to the descriptions in this form, or your rights as a participant have been violated during the course of this project, you may contact Dr. David Moore, Acting Chair of the Institutional Review Board Research Division at 231-5281.

Appendix J: Solutions and Scoring key for posttests

Q1. **Message tone**

Solution:

Assessment> Issues about feedback > Content, meaning of feedback> Preferences and Efficiency (about feedback content) > Style of feedback content > Wording, vocabulary> Anthropomorphism

Scoring : 4>2>2>1>1>0

Q2: **Advanced button**

Solution:

Translation>Presentation (of cognitive affordance) >layout and grouping by function>grouping related functions, objects and controls together by task

Scoring: 4> 3> 2> 1

Q3 : **No fill choice not obvious**

Solution:

Translation> Existence>existence of Cognitive affordance>existence of cognitive affordance to show HOW to manipulate an object

Scoring : 4>2>2>2

Q 4: **Icon too small**

Solution:

Physical Actions> Manipulating objects> Physical layout> Proximity (closeness) of object as a factor in ability to manipulate reliably

Scoring: 4>2>2>2>

Q 5: **Resetting the screensaver timeout**

Solution:

Assessment> Issues about feedback > Existence > Existence of necessary or desirable feedback > Error message needed but missing

Scoring: 4>2> 2>1>1

Q 6: **Data entry format missing**

Solution(first choice):

Translation> Existence> Existence of a cognitive affordance > Existence of field data format information to guide data entry

Scoring: 5>2>2>1

Q 7: **Doesn't like the colors**

Solution: *Independent > Overall style*

Scoring: 7> 3

Q 8: **User wants to reuse document numbers in a document retrieval system.**

Solution:

Translation> Existence> Existence of a way > Existence of feature- functionality and physical affordance.

Scoring: 5>2>2>1

Q 9: **Exit label**

Solution:

Translation> Content, meaning of cognitive affordance > Clarity, precision, predictability of meaning > precise use of words> labels for buttons, menus > Clearly labeled exits

Scoring 4>2>2>1>1>0

Q 10: **Doesn't understand spreadsheets**

Solution:

Planning> User's model of a system > User awareness of overall application concept(how it works)

Scoring: 5>3>2

Appendix K: Posttest questionnaire for classroom training group

How would you rate the training your received for the UAF?

Excellent Poor

0 1 2 3 4 5 6 7 8 9 10

I felt better prepared to classify usability problems accurately using the UAF explorer after going through the training tool.

Strongly Disagree Disagree Neither Disagree nor Agree Agree Strongly Agree

I understand the underlying concepts of the UAF and can apply them confidently.

Strongly Disagree Disagree Neither Disagree nor Agree Agree Strongly Agree

The UAF training tool has the needed capabilities

Strongly Disagree Disagree Neither Disagree nor Agree Agree Strongly Agree

Based on the exposure you have had on the UAF and the training, how would you rate your skill on using the UAF explorer to classify usability problems now?

Poor Fair Good Excellent

Now that you've had some experience with a web based training tool, rank order your choice of training medium for acquiring a computer-based skill. Please use 1,2,3 and 4 to rank your preferences. Use 1 for "most preferred" and 4 for "least preferred".

Web based training ______

Classroom training ________

Paper based/instruction manual ______________

Audio/video training _______________

List a few positive points about the UAF training you received

List a few negative points about the UAF training you received

End of questionnaire. Thank you!

Appendix L: Posttest questionnaire for web-based training group

Last Name OR Student ID number:

How would you rate the training your received for the UAF?

Excellent Poor

0 1 2 3 4 5 6 7 8 9 10

I felt better prepared to classify usability problems accurately using the UAF explorer after going through the training.

Strongly Disagree Disagree Neither Disagree nor Agree Agree Strongly Agree

I understand the underlying concepts of the UAF and can apply them confidently.

Strongly Disagree Disagree Neither Disagree nor Agree Agree Strongly Agree

The UAF training met my expectations and provided me with good training to classify problems.

Strongly Disagree Disagree Neither Disagree nor Agree Agree Strongly Agree

Based on the exposure you have had on the UAF and the training, how would you rate your skill on using the UAF explorer to classify usability problems now?

Poor Fair Good Excellent

I was motivated to use the UAF Explorer to classify usability problems and understand its use.

Strongly Disagree Disagree Neither Disagree nor Agree Agree Strongly Agree

After your experience with training for the UAF Explorer, rank order your choice of training medium for acquiring a computer-based skill. Please use 1,2,3 and 4 to rank your preferences. Use 1 for “most preferred” and 4 for “least preferred”.

Web based training ______
Classroom training ________
Paper based/instruction manual ______________
Audio/video training _______________

List a few positive points about the UAF training you received

List a few negative points about the UAF training you received

End of questionnaire. Thank you!

Appendix M: Data from summative evaluation

Table A-6: Pretest – posttest data from summative evaluation

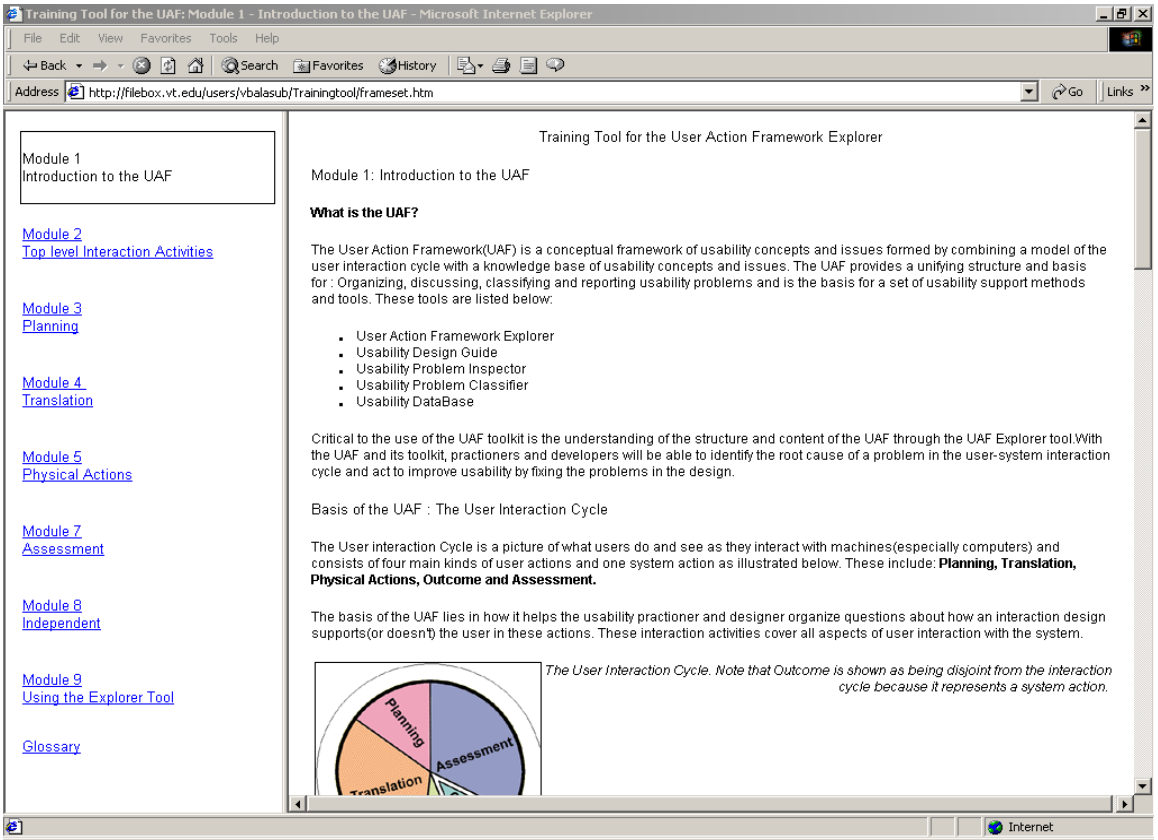
Group	Participant #	Pretest score	Posttest Score	Gain score
EXPERIMENTAL GROUP	1	44	87	43
	2	48	65	17
	3	68	74	6
	4	66	86	20
	5	60	75	15
	6	87	76	-11
	7	55	76	22
	8	65	83	18
	9	69	92	23
	10	39	77	38
	11	41	98	57
	12	24	66	42
	13	60	86	26
	14	70	85	15
	15	61	86	25
MEANS		**57.13**	**80.8**	**23.73**
SD		**15.66**	**9.15**	**16.45**
CONTROL GROUP	1	73	96	23
	2	69	88	19
	3	59	73	14
	4	68	72	4
	5	100	96	-4
	6	62	100	38
	7	65	75	10
	8	71	91	20
	9	63	81	18
	10	33	59	26
	11	56	79	23
	12	36	78	42
	13	72	61	-11
	14	69	58	-11
	15	70	84	14
	16	60	63	3
MEANS		**64.12**	**78.375**	**14.25**
SD		**15.16**	**13.63**	**15.39**

Table A-7 : Preference vs. Performance data

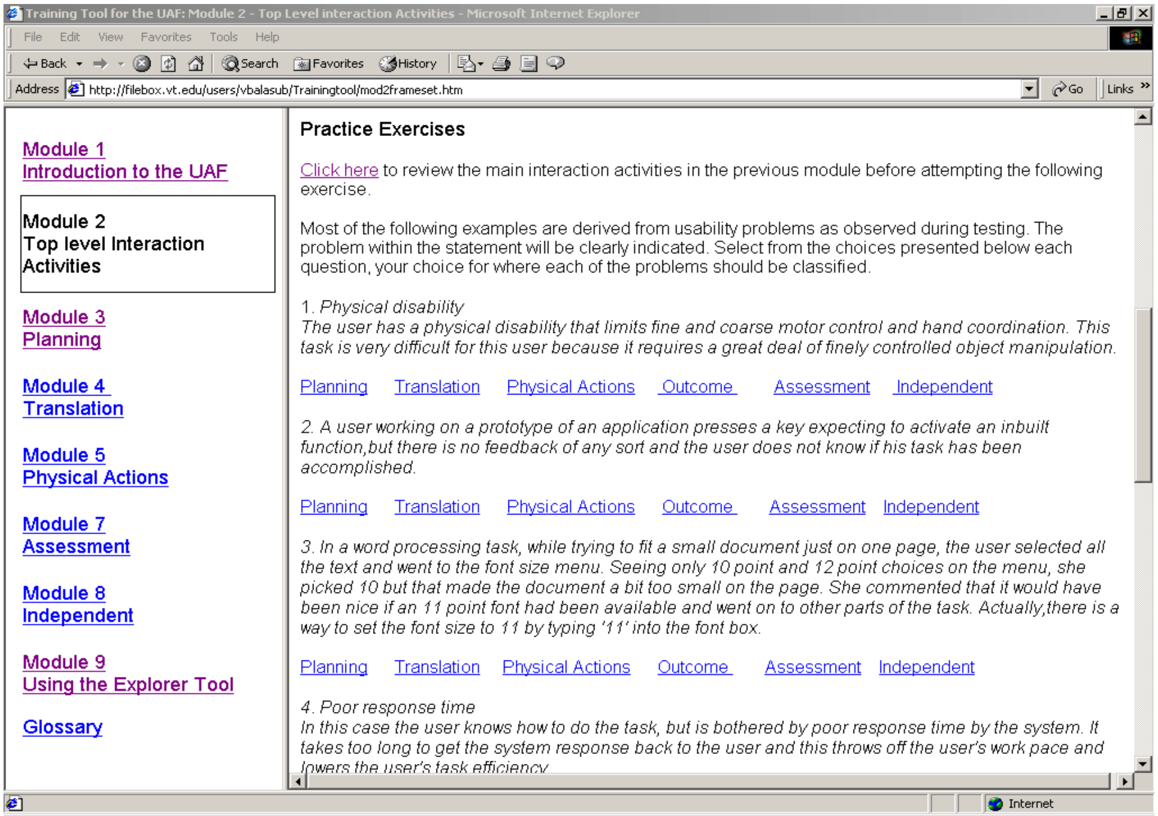
Preferred	Score on classroom training	Score on web training	Preferred	Score on classroom	Score on web
Class	96		Web	72	
Class	87		Web	91	
Class	88		Web	78	
Class	73		Web	84	
Class	96		Web	63	
Class	100		Web		83
Class	75		Web		98
Class	81		Web		65
Class	59				
Class	79				
Class	61				
Class	58				
Class		74			
Class		86			
Class		75			
Class		76			
Class		76			
Class		92			
Class		77			
Class		66			
Class		86			
Class		85			
Class		86			

Appendix N: Screenshots of the web based Training tool

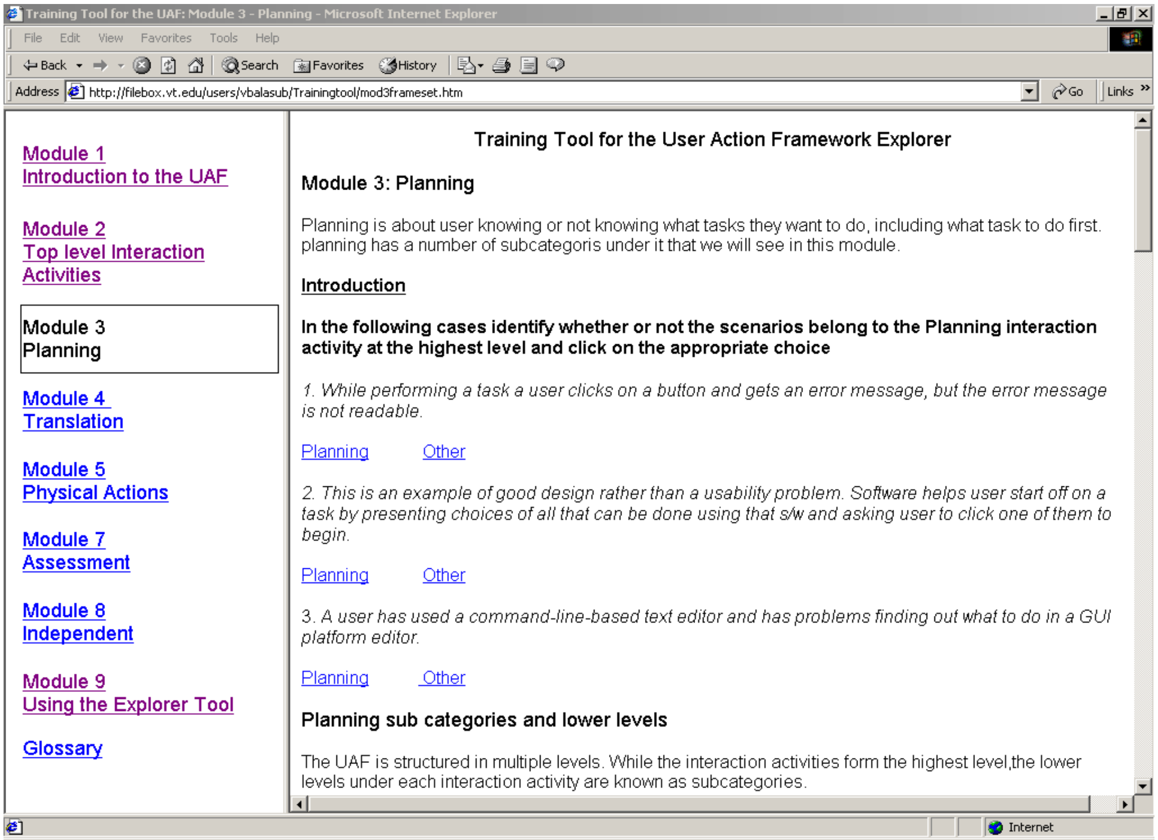
Module 1: Introduction to the UAF



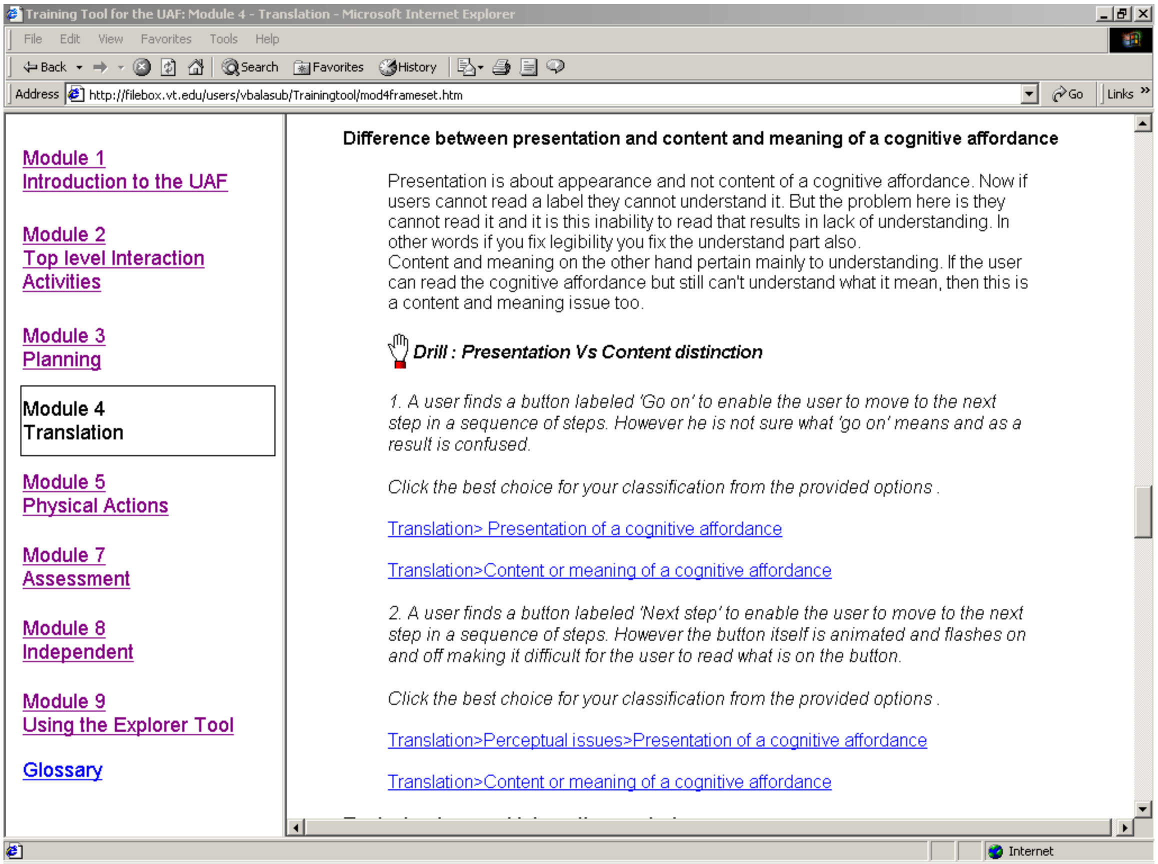
Module 2: Interaction Activities

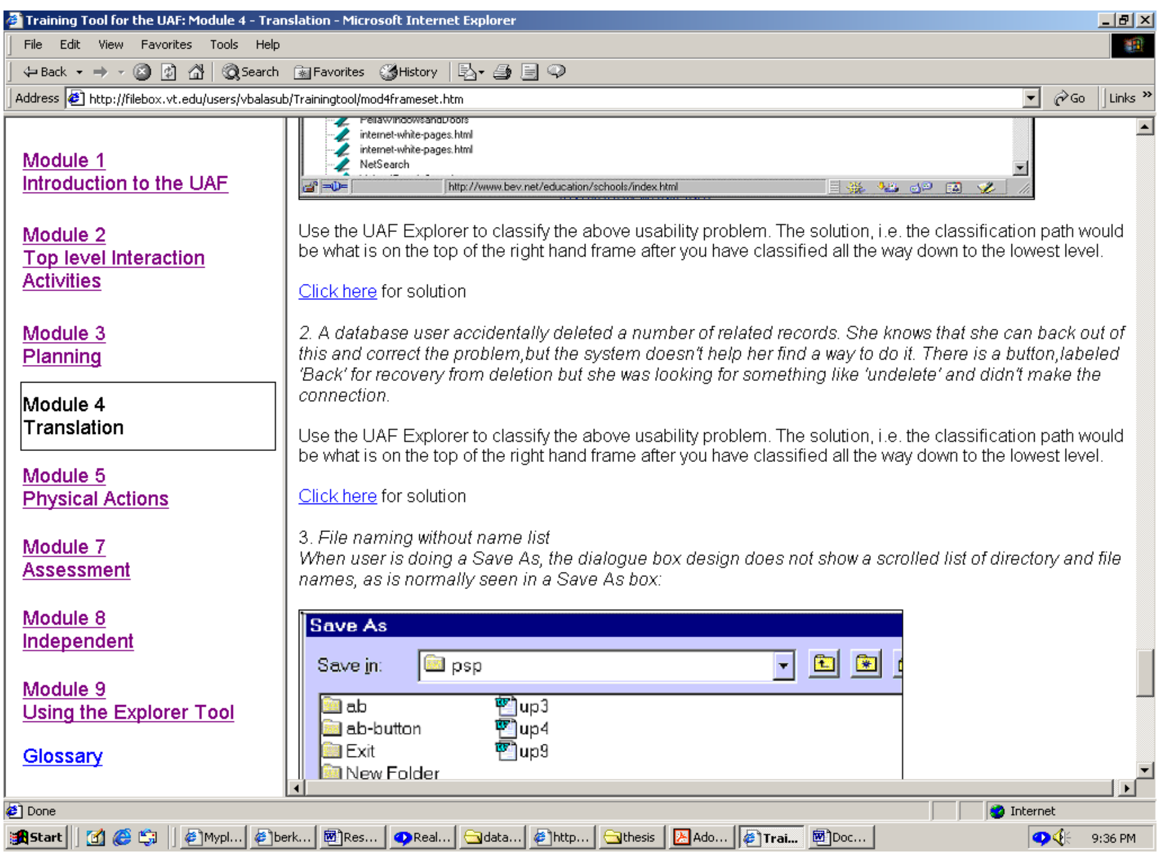


Module 3: Planning

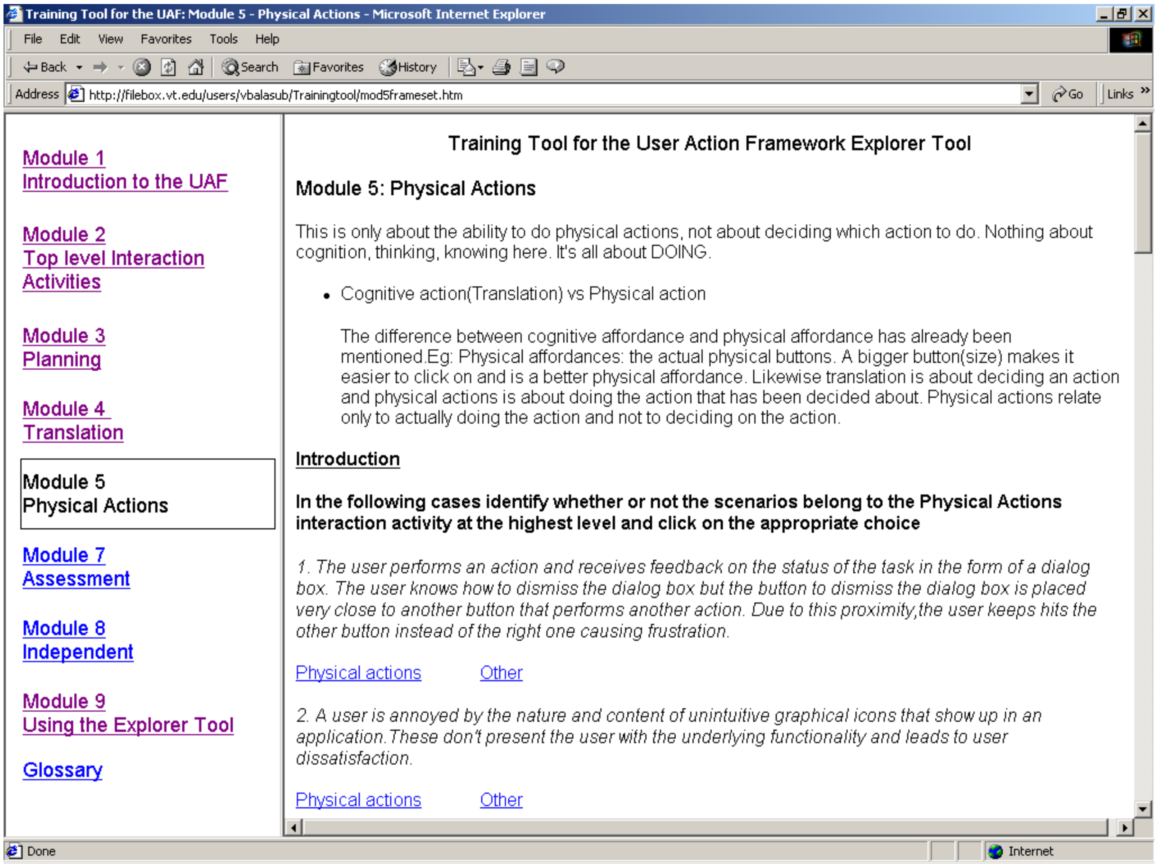


Module 4: Translation

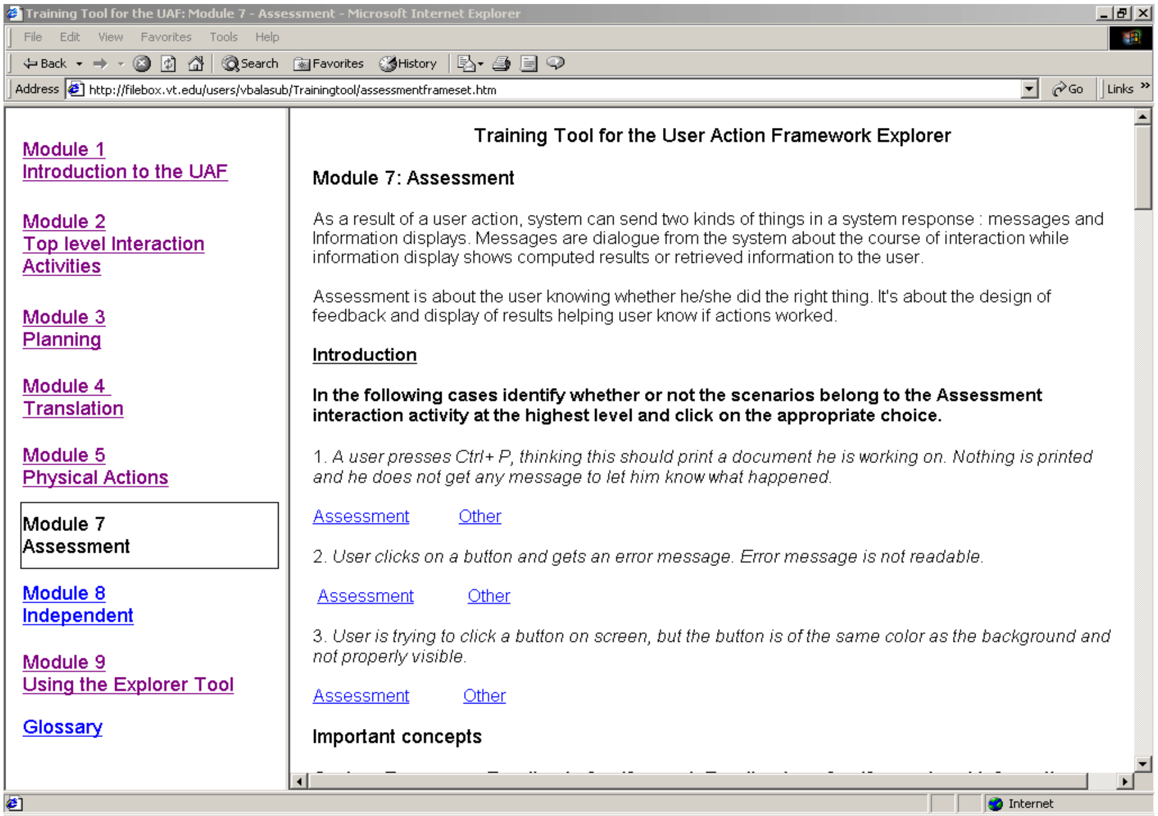

Training Tool for the UAF: Module 4 - Translation - Microsoft Internet Explorer
Address http://filebox.vt.edu/users/vbalasub/Trainingtool/mod4frameset.htm
Module 1
Introduction to the UAF
Module 2
Top level Interaction Activities
Module 3
Planning
Module 4
Translation
Module 5
Physical Actions
Module 7
Assessment
Module 8
Independent
Module 9
Using the Explorer Tool
Glossary
Difference between presentation and content and meaning of a cognitive affordance
Presentation is about appearance and not content of a cognitive affordance. Now if users cannot read a label they cannot understand it. But the problem here is they cannot read it and it is this inability to read that results in lack of understanding. In other words if you fix legibility you fix the understand part also.
Content and meaning on the other hand pertain mainly to understanding. If the user can read the cognitive affordance but still can't understand what it mean, then this is a content and meaning issue too.
Drill : Presentation Vs Content distinction
1. A user finds a button labeled 'Go on' to enable the user to move to the next step in a sequence of steps. However he is not sure what 'go on' means and as a result is confused.
Click the best choice for your classification from the provided options .
Translation> Presentation of a cognitive affordance
Translation>Content or meaning of a cognitive affordance
2. A user finds a button labeled 'Next step' to enable the user to move to the next step in a sequence of steps. However the button itself is animated and flashes on and off making it difficult for the user to read what is on the button.
Click the best choice for your classification from the provided options .
Translation>Perceptual issues>Presentation of a cognitive affordance
Translation>Content or meaning of a cognitive affordance
Internet


Training Tool for the UAF: Module 4 - Translation - Microsoft Internet Explorer
Address http://filebox.vt.edu/users/vbalasub/Trainingtool/mod4frameset.htm
Module 1
Introduction to the UAF
Module 2
Top level Interaction Activities
Module 3
Planning
Module 4
Translation
Module 5
Physical Actions
Module 7
Assessment
Module 8
Independent
Module 9
Using the Explorer Tool
Glossary
Use the UAF Explorer to classify the above usability problem. The solution, i.e. the classification path would be what is on the top of the right hand frame after you have classified all the way down to the lowest level.
Click here for solution
2. A database user accidentally deleted a number of related records. She knows that she can back out of this and correct the problem,but the system doesn't help her find a way to do it. There is a button,labeled 'Back' for recovery from deletion but she was looking for something like 'undelete' and didn't make the connection.
Use the UAF Explorer to classify the above usability problem. The solution, i.e. the classification path would be what is on the top of the right hand frame after you have classified all the way down to the lowest level.
Click here for solution
3. File naming without name list
When user is doing a Save As, the dialogue box design does not show a scrolled list of directory and file names, as is normally seen in a Save As box:
Save As
Save in: psp
ab
ab-button
Exit
New Folder
up3
up4
up9
Done
Internet
9:36 PM

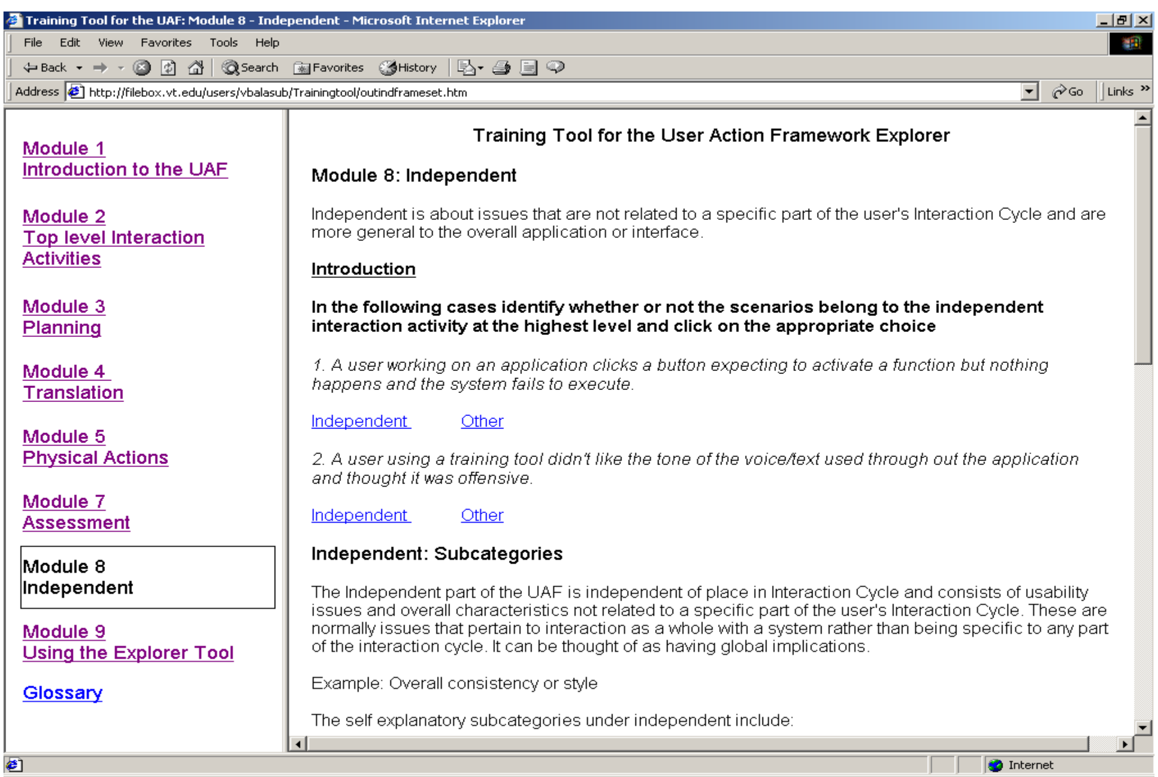
Module 5: Physical Actions



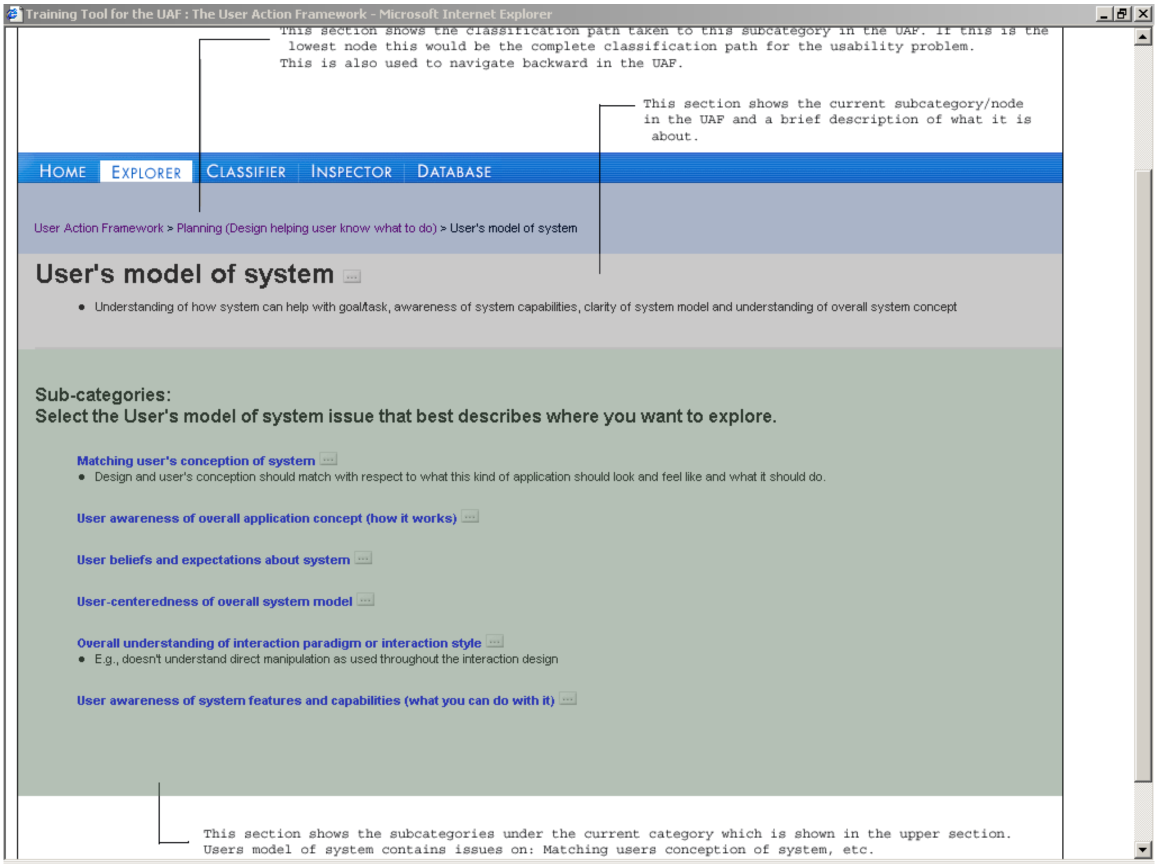
Module 7 : Assessment

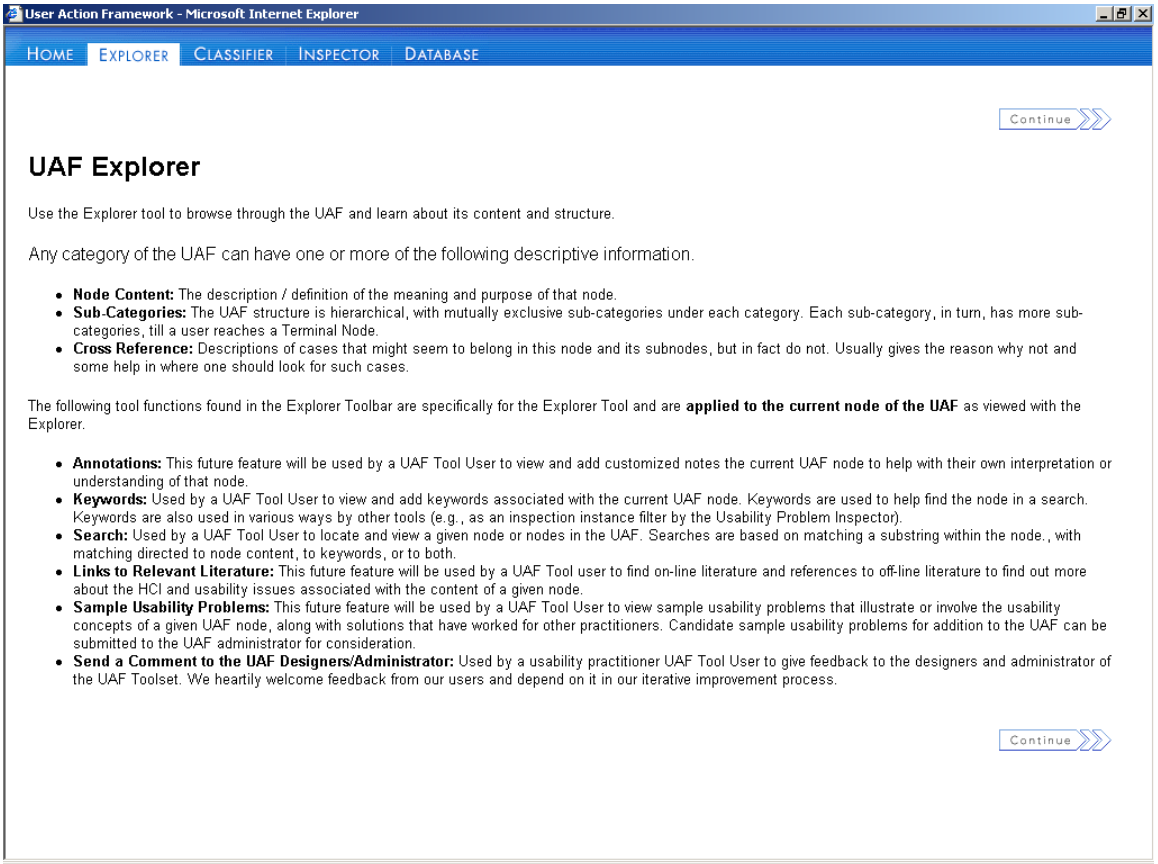


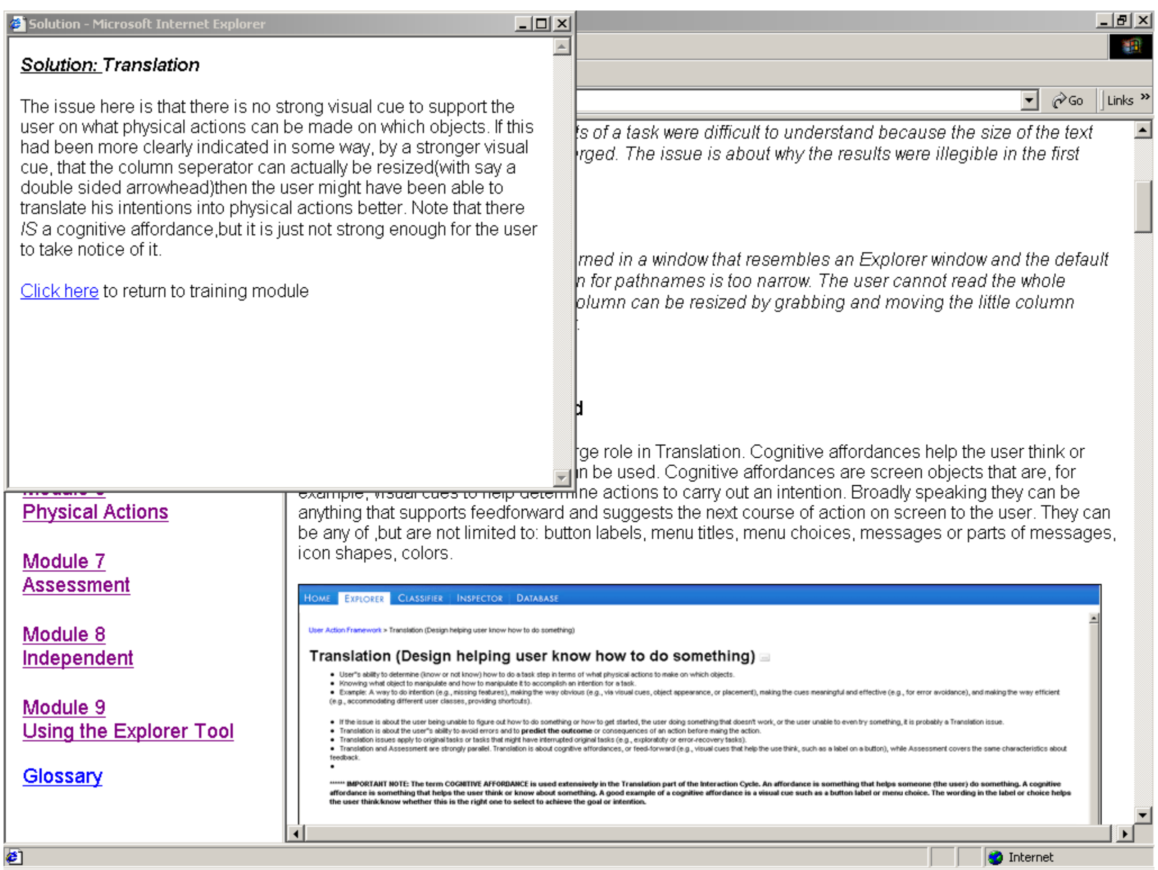
Module 8:Independent



Other:


Training Tool for the UAF : The User Action Framework - Microsoft Internet Explorer
This section shows the classification path taken to this subcategory in the UAF. If this is the lowest node this would be the complete classification path for the usability problem. This is also used to navigate backward in the UAF.
This section shows the current subcategory/node in the UAF and a brief description of what it is about.
HOME
EXPLORER
CLASSIFIER
INSPECTOR
DATABASE
User Action Framework > Planning (Design helping user know what to do) > User's model of system
User's model of system
Understanding of how system can help with goal/task, awareness of system capabilities, clarity of system model and understanding of overall system concept
Sub-categories:
Select the User's model of system issue that best describes where you want to explore.
Matching user's conception of system
Design and user's conception should match with respect to what this kind of application should look and feel like and what it should do.
User awareness of overall application concept (how it works)
User beliefs and expectations about system
User-centeredness of overall system model
Overall understanding of interaction paradigm or interaction style
E.g., doesn't understand direct manipulation as used throughout the interaction design
User awareness of system features and capabilities (what you can do with it)
This section shows the subcategories under the current category which is shown in the upper section. Users model of system contains issues on: Matching users conception of system, etc.


User Action Framework - Microsoft Internet Explorer
HOME
EXPLORER
CLASSIFIER
INSPECTOR
DATABASE
Continue
UAF Explorer
Use the Explorer tool to browse through the UAF and learn about its content and structure.
Any category of the UAF can have one or more of the following descriptive information.
Node Content: The description / definition of the meaning and purpose of that node.
Sub-Categories: The UAF structure is hierarchical, with mutually exclusive sub-categories under each category. Each sub-category, in turn, has more sub-categories, till a user reaches a Terminal Node.
Cross Reference: Descriptions of cases that might seem to belong in this node and its subnodes, but in fact do not. Usually gives the reason why not and some help in where one should look for such cases.
The following tool functions found in the Explorer Toolbar are specifically for the Explorer Tool and are applied to the current node of the UAF as viewed with the Explorer.
Annotations: This future feature will be used by a UAF Tool User to view and add customized notes the current UAF node to help with their own interpretation or understanding of that node.
Keywords: Used by a UAF Tool User to view and add keywords associated with the current UAF node. Keywords are used to help find the node in a search. Keywords are also used in various ways by other tools (e.g., as an inspection instance filter by the Usability Problem Inspector).
Search: Used by a UAF Tool User to locate and view a given node or nodes in the UAF. Searches are based on matching a substring within the node., with matching directed to node content, to keywords, or to both.
Links to Relevant Literature: This future feature will be used by a UAF Tool user to find on-line literature and references to off-line literature to find out more about the HCI and usability issues associated with the content of a given node.
Sample Usability Problems: This future feature will be used by a UAF Tool User to view sample usability problems that illustrate or involve the usability concepts of a given UAF node, along with solutions that have worked for other practitioners. Candidate sample usability problems for addition to the UAF can be submitted to the UAF administrator for consideration.
Send a Comment to the UAF Designers/Administrator: Used by a usability practitioner UAF Tool User to give feedback to the designers and administrator of the UAF Toolset. We heartily welcome feedback from our users and depend on it in our iterative improvement process.
Continue


Solution - Microsoft Internet Explorer
Solution: Translation
The issue here is that there is no strong visual cue to support the user on what physical actions can be made on which objects. If this had been more clearly indicated in some way, by a stronger visual cue, that the column seperator can actually be resized(with say a double sided arrowhead)then the user might have been able to translate his intentions into physical actions better. Note that there IS a cognitive affordance,but it is just not strong enough for the user to take notice of it.
Click here to return to training module
Physical Actions
Module 7
Assessment
Module 8
Independent
Module 9
Using the Explorer Tool
Glossary
anything that supports feedforward and suggests the next course of action on screen to the user. They can be any of ,but are not limited to: button labels, menu titles, menu choices, messages or parts of messages, icon shapes, colors.
Translation (Design helping user know how to do something)
Internet