

WP 6.1 Project Deliverable

Report on CAVE/HMD Installation



Project Number	IST-2000-29266
Project Title	Virtual Real Time Fire Emergency Simulator
Deliverable Type	Evaluation report
Deliverable Class	External

Deliverable Number	6.1.
Title of Deliverable	Report on CAVE / HMD Installation
Nature of the Deliverable	Report
Contributing WPs	WP 2, 4, 5, 7
Contractual Date of Delivery	01 Jan. 04
Actual Date of Delivery	09 Sep. 04
URL	www.virtualfires.org
Authors	Javier Urruzola, Gunther Lenz, Christian Redl, Thorsten Mai, Gert Svensson
Contact Details	European Virtual Engineering / EUVE Dr. Javier Urruzola Avda. de los Huetos, 79 01010 Vitoria / Spain Tel.: +34 945 21 46 46 Fax: +34 945 21 46 47 Email: jurruzola@euve.org

Abstract	Evaluation report of Virtual Fires System
Keywords	Validation, Evaluation, Software, Simulation

<u>1</u>	<u>INTRODUCTION.....</u>	<u>4</u>
<u>2</u>	<u>END USER EVALUATION PROCEDURE DESCRIPTION</u>	<u>4</u>
<u>3</u>	<u>QUALITY IN USE EVALUATION PROCEDURE DESCRIPTION.....</u>	<u>6</u>
3.1	INTRODUCTORY REMARKS	6
3.1.1	THE ISO VALIDATION FRAMEWORK.....	6
3.1.2	EVALUATION ACTIVITIES	7
3.1.3	INPUT TO THE EVALUATION PROCESS	7
3.1.4	OUTPUT OF THE EVALUATION PROCESS	7
3.2	EVALUATION REQUIREMENTS.....	8
3.2.1	PURPOSE OF THE EVALUATION	8
3.2.2	PRODUCT APPLICATION DOMAIN	8
3.2.3	GENERAL DESCRIPTION OF THE PRODUCT PURPOSE.....	8
3.2.4	QUALITY REQUIREMENTS.....	8
3.3	EVALUATION SPECIFICATION.....	9
3.3.1	SCOPE OF THE EVALUATION	9
3.3.2	SPECIFICATION OF MEASUREMENTS AND VERIFICATIONS	9
3.4	DRAFT EVALUATION PLAN	9
3.4.1	EVALUATION METHODS.....	9
3.4.2	MAPPING OF MEASURES	15
3.4.3	SCHEDULING (TIMETABLE) AND RESOURCE ALLOCATION (LOGISTICS)	17
<u>4</u>	<u>FOCUS GROUP TECHNIQUE DESCRIPTION</u>	<u>17</u>
<u>5</u>	<u>END USER EVALUATION RESULTS</u>	<u>18</u>
5.1	EFFECTIVENESS	18
5.2	PRODUCTIVITY	18
5.3	SATISFACTION	19
5.3.1	EVALUATION IN LYON	19
5.3.2	EVALUATION IN DORTMUND.....	21
5.3.3	EVALUATION IN STOCKHOLM	22
5.4	FOCUS GROUP	24
5.5	DIFFICULTIES ENCOUNTERED	28
<u>6</u>	<u>ANNEX 1. RESULTS OF ITEM CONSENSUAL ANALYSIS.</u>	<u>29</u>
6.1	EVALUATION IN LYON.....	29
6.2	EVALUATION IN DORTMUND.....	40
6.3	EVALUATION IN STOCKHOLM.....	51

1 Introduction

Workpackage 6, Evaluation and Validation, consists of two different tasks:

Task 6.1. End-user evaluation of the Virtual Fires System in the CAVE and PC versions.

Methodology: End-users take part in test cases to evaluate the quality in use of the system. This is complemented with the focus group technique to assess user needs and feelings.

Task 6.2. Functionality validation (validation against specifications).

Methodology: Developers test that code sticks to specifications and that functionality is appropriate.

Both tasks have been carried out in accordance with ISO. This deliverable contains the results of task 6.1.

2 End User evaluation procedure description

1.- The first objective of the evaluation sessions is to evaluate the "Quality in Use" of the Virtual Fires software. It was also decided to apply the user focus group technique described below to obtain more information about user needs.

2.- According to ISO/IEC 14598 and ISO 9126, "Quality in Use" means "the user's view of quality".

3.- Three types of end-users will evaluate the product in Lyon, Dortmund and Stockholm:

- Tunnel designers (including regulators) and tunnel operators (ventilation management staff)
- Fire brigade planning rescue scenarios
- Fire brigade getting familiar with a fire scenario

4.- End-users will evaluate three quality characteristics of the product:

Effectiveness

"The capability of the product to enable users to achieve specified goals with accuracy and completeness in a specified context of use".

Productivity

"The capability of the software product to enable users to expend appropriate amounts of resources in relation to the effectiveness achieved in a specified context of use. Relevant resources can include time to complete the task, the user's effort, materials or the financial cost of usage".

Satisfaction

"The capability of the software product to satisfy users in a specified context of use. Satisfaction is the user's response to interaction with the product, and includes attitudes towards use of the product."

5.- Evaluation technique:

A sample of users who are representative of a particular user group attempt to achieve representative task goals using the product in a simulated environment without any assistance other than that available in the real working environment. Users also fill in a satisfaction questionnaire and participate in a focus group session.

5.1.- End-users will be given a sheet of paper with several tasks and goals that they must carry out with the Virtual Fires software. For this purpose they will be helped by a technical facilitator.

5.2.-The evaluator will measure two types of data:

Task time: The total time taken by each user to complete the task (excluding any interruptions).

Task output: functionality located and correctly executed or not.

5.3.- Task output will be used to evaluate the effectiveness of Virtual Fires.

5.4.- The relationship between the task output and the resources expended (task time and resources cost) will be used to evaluate productivity of Virtual Fires.

5.5.- Focus group session.

5.6.- End-users will fill in a questionnaire for measuring usage satisfaction.

5.7.- The evaluator will use these questionnaires to measure the third type of data:

Satisfaction results: A completed satisfaction questionnaire like SUMI.

5.8.- Satisfaction of Virtual Fires will be measured by evaluating the SUMI questionnaire following a standard procedure.

6.- Evaluation session schedule

1.- General Description of the Virtual Fires project

2.- Technical description of the project

3.- Demonstration with the subway station case and the MtBlanc tunnel case

4.- Focus group session.

5.- Hands-on testing

6.- Filling in of SUMI questionnaire

6.- Testcases: tasks must be carried and goals must be achieved in a certain context of use or testcase. Two different testcases will be considered:

1) Dortmund subway station

2) MtBlanc

Tasks:

1.- Load a predefined scenario.

- 2.- View the events of the mission.
- 3.- View the boundary conditions of the scenario.
- 4.- Move through the scene (navigate in space).
- 5.- View the results of precalculated or online calculated cfd-data for a mission and select an appropriate visualisation method for the cfd-data.
- 6- Modify/add/delete a visualization method.
- 7.- Edit the boundary conditions and modify/add/delete/create events for the mission.

Goals:

For each task the goal is to locate in the Virtual Fires system the appropriate functionality and execute it correctly.

In order to make the test more interesting, the test case has been selected taking into account the interests of participants, i.e. MtBlanc for tunnel designers and Dortmund subway station for Dortmund fire brigade.

7.- Focus group session. It is based on four main questions:

What are your expectations of such a product?

What other products are the competitors of this one from your point of view?

What are the advantages of this new technology compared to traditional techniques?

How much would you be ready to pay for a product like this per year?

3 Quality in Use evaluation procedure description

3.1 *Introductory remarks*

3.1.1 The ISO validation framework

This section outlines the ISO-based framework for evaluation which will be applied in this project.

3.1.1.1 *Establishing evaluation requirements*

The first step is to establish what the purpose of the evaluation is.

The second step is to identify the types of products to be evaluated. There might also be additional constraints on the product that come into play here.

The third step in establishing the evaluation requirements is to define, in the light of the results of the first two steps, what characteristics of the object to be evaluated will decide whether it is a good object of its kind or not: e. g. defining the quality model; it is usual to start with quite vague, high level considerations, quality characteristics and sub-characteristics, and then to refine these down until we can arrive at quality attributes of the system which can be measured in some way.

3.1.1.2 Specifying the evaluation

The first step here is to work out in concrete terms how the bottom line attributes are to be measured: e. g. selecting and validating a set of metrics. The next step is to decide on what counts as a good, bad or neutral value, and, in particular, what a satisfactory value will be for this evaluation: defining rating levels and overall assessment.

3.1.1.3 Producing the evaluation plan

Once the metrics have been decided upon, it is possible to plan the evaluation itself. Who will produce test materials if they are required? When will any tests be administered and by whom? How are the results to be reported and communicated?

3.1.1.4 Executing the evaluation

Once all the preparatory work has been done, the evaluation itself can be carried out. This will involve applying the metrics that have been decided upon, determining whether the values obtained fall into the good bad or indifferent range, and whether the values and combinations of values are satisfactory. The results will be combined to give an overall assessment of the product that has been evaluated, which will tell us whether the product is likely to meet the needs of the user whose requirements were articulated at the beginning of the creating the evaluation design.

3.1.2 Evaluation activities

The evaluation process comprises five activities:

- Establishment of evaluation requirements
- Specification of the evaluation based on the evaluation requirements and on the description of the product
- Design of the evaluation which produces an evaluation plan
- Execution of the evaluation plan
- Conclusion of the evaluation and delivery of the evaluation report.

3.1.3 Input to the evaluation process

- Product description.
- Product components

3.1.4 Output of the evaluation process

- evaluation records, including evaluation plan and records of evaluation actions,
- the draft evaluation report, including evaluation requirements, evaluation specification and synthesised evaluation results,
- the reviewed evaluation report.

The evaluation requirements, specification and plan are the intermediate products.

The evaluation records and evaluation report are the final products.

The evaluation requirements describe the objectives of the evaluation; in particular, quality requirements.

The evaluation specification defines all analyses and measurements to be performed on the product. The components of the product that will be analysed and measured are identified.

The evaluation plan describes operational procedures needed to implement the evaluation specification; in particular all the methods and tools to be used in the evaluation are described.

The evaluation records consist of the evaluation plan and a detailed account of actions performed by the evaluator while executing the evaluation plan; these records are kept by the evaluator.

The evaluation report contains evaluation requirements, the evaluation specification, results from the measurements and analyses performed

3.2 Evaluation requirements

3.2.1 Purpose of the evaluation

The purpose of the Evaluation and Validation workpackage of the Virtual Fires project is to evaluate whether the system developed fulfils the requirements of the intended end users.

3.2.2 Product application domain

The product application domain is double: computer aided engineering and computer aided training.

3.2.3 General description of the product purpose

The Virtual Fires system is a real time fire emergency simulator. It can be used for assessing the fire safety of tunnels, for training of rescue personnel and for planning rescue scenarios. It must also be able to replace some real fire tests.

The end users of the system are tunnel operators, government organizations involved in fire safety, and rescue organizations like fire brigades.

A more detailed description of the system can be found in Deliverable 2.4.

3.2.3.1 Type of product to be evaluated

Final product.

3.2.4 Quality requirements

3.2.4.1 Quality model

Because of the type of product and of the stage in the life cycle of the project, the appropriate quality model is Quality in Use (the user's view of quality.).

3.2.4.2 Quality characteristics

Effectiveness

"The capability of the product to enable users to achieve specified goals with accuracy and completeness in a specified context of use".

Level of evaluation: most stringent

Productivity

"The capability of the software product to enable users to expend appropriate amounts of resources in relation to the effectiveness achieved in a specified context of use. Relevant resources can include time to complete the task, the user's effort, materials or the financial cost of usage".

Level of evaluation: stringent. When compared to the use of traditional techniques, the Virtual Fires system will be handicapped by the fact that users have many years' experience in the use of traditional techniques. This experience cannot be equalled in the frame of this project.

Safety

"The capability of the software product to achieve acceptable levels of risk of harm to people, business, software, property or the environment in a specified context of use. Risks are usually a result of deficiencies in the functionality (including security), reliability, usability or maintainability".

Level of evaluation: none. No significant risks or harms can be foreseen for people operating/using the system.

Satisfaction

"The capability of the software product to satisfy users in a specified context of use. Satisfaction is the user's response to interaction with the product, and includes attitudes towards use of the product."

Level of evaluation: medium. When compared to the use of traditional techniques, the Virtual Fires system will be handicapped by the fact that users have many years' experience in the use of traditional techniques. This experience cannot be equalled in the frame of this project. In addition, the project is expected to improve in this area once it succeeds commercially.

3.3 Evaluation specification

3.3.1 Scope of the evaluation

The components of the product to be evaluated are those developed during the project Virtual Fires as they are defined in Deliverable 2.4.

3.3.2 Specification of measurements and verifications

For the purpose of specifying measurements and verifications, the "Usability and Quality in Use" evaluation module in Annex D of ISO/IEC 14598-6:2001 has been taken as starting point with a few modifications.

3.4 Draft evaluation plan

3.4.1 Evaluation Methods

The evaluation module in Annex D of ISO/IEC 14598-6:2001 is for measuring quality in use as specified ISO 9126-1 and using the principles of ISO 9241-11. It is based on the MUSiC methodology.

It gives the principles for evaluating quality in use by evaluating the results of using the product, with a representative sample of users carrying out representative tasks in a simulated environment.

Effectiveness, productivity and satisfaction are evaluated according to their definitions in ISO/IEC 9126-1.

3.4.1.1 Level of evaluation

This evaluation procedure can provide accurate estimates of three characteristics of quality in use. The degree of accuracy depends on how closely the context of evaluation simulates the context of use, and on the number of users in each user group who are evaluated. At least eight users should be evaluated in a realistic context of use to obtain reliable results.

3.4.1.2 Technique

A sample of users who are representative of a particular user group attempt to achieve representative task goals using the product in a simulated environment without any assistance other than that available in the real working environment. Users also fill in a satisfaction questionnaire.

3.4.1.3 Applicability

The evaluation module is appropriate for any product which forms part of an system with which users interact to achieve task goals.

It can be applied during development, acquisition or operation, for quality assurance or validation. During development the evaluation module may be used to evaluate early prototypes using only 3 or 4 users, to obtain an indication of whether quality in use targets are likely to be met. During acquisition the evaluation module can provide assurance that the product is appropriate for the intended working environment. During operation, the evaluation module can establish baseline values against which to compare future products, and can indicate which product attributes may need to be improved.

3.4.1.4 Terms and definitions

Context of use: The users, tasks, equipment (hardware, software and materials), and the physical and social environments in which a product is used. [ISO 9241-11]

User: The person who interacts with the product. [ISO 9241-11]

Goal: An intended outcome. [ISO 9241-11]

Task: The activities required to achieve a goal. [ISO 9241-11]

3.4.1.5 Inputs and metrics

3.4.1.5.1 Input for the evaluation

3.4.1.5.1.1 Product component: working prototype

A working prototype is evaluated (including executable code and user documentation – the figure of the technical facilitator compensates for the lack of documentation in this case).

3.4.1.5.1.2 Product information: context of use

A definition of the intended contexts of use of the product is required, including the essential characteristics and capabilities of the intended user groups, their goals and tasks and the intended technical and support environment.

The contexts of use will be defined according to the information provided in Deliverables 2.4 and 7.1.

Context of use: the users, tasks, equipment (hardware, software and materials), and the physical and social environments in which a product is used.

Context of use 1: Tunnel road scenario (PC)

This is a fire scenario following the French guide for specific study of danger in road tunnel.

User profiles

Tunnel designers (including regulators) and tunnel operators (ventilation management staff)

In this example some calculations are done for determining the progress of the smoke and toxic gas. The use of the projected Virtualfires system is obvious for a quicker answer and high possibilities of changes in assumptions and men actions. The Virtualfires system will have a friendly user interface as to act quickly on the scenario, giving the capacity of visualisation of the effects of the changes.

Tasks:

The users will collect together and/or separately in order to analyze a certain fire scenario: the Mont Blanc case in D7.1..

- 1.- Load a predefined scenario (=tunnel+objects+boundary conditions) and a mission for this scenario. Mont Blanc and Mont Blanc car.
- 2.- View the events of the mission. Mont Blanc car.
- 3.- View the boundary conditions of the scenario. Initial BC of the outlet.
- 4.- Move through the scene (navigate in space). From entrance to exit.
- 5.- View the results of precalculated or online calculated cfd-data for this mission and select an appropriate visualisation method for the cfd-data: smoke density with volume rendering (playback)
- 6.- Modify a visualization method. Change resolution of volume rendering to 0.8 and the colour map to 3.
- 7.- Edit the boundary conditions and modif events for the mission: For the fresh air inlet at time 190 s. set velocity to 0.9 m/s and temperature to 300K.

Goals:

For each task the goal is to locate in the Virtual Fires system the appropriate functionality and execute it correctly.

Equipment and environment

For Tunnel designers (including regulators) and tunnel operators (ventilation management staff) the PC screen version is the best-adapted version.

Context of use 2: Subway station (PC)

This is a fire scenario in a subway station.

User profiles

Fire brigade planning rescue scenarios

The action of firemen is defined by the accurate location and direction of the fire fighting systems used (nozzles, turbo-ventilators, etc.) as well as the exact time when they are used.

The VIRTUALFIRE simulator will show the efficiency of the fire attack, that is the distribution of gas temperature showing if the escape routes are safe or not. Other informations as lethal zone (toxic gaz) are also given.

Tasks:

The users will collect together and/or separately in order to analyze a certain fire scenario: the Dortmund subway station M1 case in D7.1..

- 1.- Load a predefined scenario (=station+objects+boundary conditions) and a mission for this scenario. Subway station (Dortmund).
- 2.- View the events of the mission. Subway nozzles.
- 3.- View the boundary conditions of the scenario. Initial BC portal west.
- 4.- Move through the scene (navigate in space). From viewing all to the yellow train.
- 5.- View the results of precalculated or online calculated cfd-data for this mission and select an appropriate visualisation method for the cfd-data: smoke density with volume rendering (playback)
- 6- Modify a visualization method. Change resolution of volume rendering to 0.8 and the colour map to 3.
- 7.- Edit the boundary conditions and modify events for the mission: set nozzle 1 at 200 seconds to speed 10 and state "on".

Goals:

For each task the goal is to locate in the Virtual Fires system the appropriate functionality and execute it correctly.

Equipment and environment

For the fire brigade planning rescue scenarios the PC screen version is the best-adapted version for their training exercises.

Context of use 3: Subway station (CAVE RT)

This is a fire scenario in a subway station simulated in a highly immersive virtual reality environment (CAVE) with real time interactive calculation capabilities.

User profiles

Fire brigade getting familiar with a fire scenario

The interactions with other actors are:

- radio connection to the control centre to reverse or to stop the ventilation,

and the job they have to do can be summarized by:

- walking (with oxygen masks) in a very smoky environment without any visibility
- extinguishing of the fire with different devices

The VIRTUALFIRE simulator will show the efficiency of the fire attack, that is the distribution of gas temperature showing if the escape routes are safe or not. Other informations as lethal zone (toxic gaz) are also given.

Tasks:

The users will collect together and/or separately in order to analyze a certain fire scenario: the Dortmund subway station M1 case in D7.1..

- 1.- Load a predefined scenario (=station+objects+boundary conditions) and a mission for this scenario. Subway station (Dortmund).
- 2.- View the events of the mission. Subway nozzles.
- 3.- View the boundary conditions of the scenario. Initial BC portal west.
- 4.- Move through the scene (navigate in space). From viewing all to the yellow train.
- 5.- View the results of precalculated or online calculated cfd-data for this mission and select an appropriate visualisation method for the cfd-data: smoke density with volume rendering (playback)
- 6- Modify a visualization method. Change resolution of volume rendering to 0.8 and the colour map to 3.
- 7.- Edit the boundary conditions and modify events for the mission: set nozzle 1 at 200 seconds to speed 10 and state "on".

Goals:

For each task the goal is to locate in the Virtual Fires system the appropriate functionality and execute it correctly.

Equipment and environment

For the fire brigade getting familiar with a fire scenario, the real time CAVE version is the best-adapted version for their training exercises.

3.4.1.5.2 Supporting information: context of evaluation

The context of evaluation coincides with the context of use except for some additional details that can be found in the description of the end user evaluation procedure above.

3.4.1.6 Data elements

3.4.1.6.1 Task time

The total time taken by each user to complete the task (excluding any interruptions).

3.4.1.6.2 Task output

A concrete representation of the results of the task produced by each user (eg data, a paper record or the user's response to a questionnaire).

3.4.1.6.3 Satisfaction results

A completed satisfaction questionnaire like SUMI.

3.4.1.6.4 Difficulties encountered

It is usually appropriate to provide additional qualitative data identifying the problems encountered by users which caused difficulties with quality in use. Suggestions may also be included for changes to the product which would improve quality in use.

3.4.1.7 Metrics and measures

3.4.1.7.1 Effectiveness

Effectiveness is a measure of whether users can achieve their goals accurately and completely. It does not take into account how the goals were achieved, only whether they were achieved.

Effectiveness should be measured by the extent to which the goals of the task have been achieved. One possible metric is the percentage of users who wholly achieve their goals. If goals can be partially achieved (e.g. by incomplete or sub-optimum results) then a more appropriate metric would be the average goal achievement, scored on a scale of 0 to 100% based on specified criteria. In some cases the percentage of users making uncorrected critical errors may be important.

3.4.1.7.2 Productivity

Productivity relates the level of effectiveness achieved to the quantity of resources expended. Efficiency is generally assessed by the mean time taken to achieve the task. Efficiency may also relate to other resources (e.g. total cost of usage), or be relatively unimportant (e.g. for some consumer applications).

Task time is the general measure of efficiency. It assumes that the less time a user takes to complete a task, the less resources the task takes, and the better the product is. Measuring efficiency as effectiveness/time gives a measure of work rate, and is useful when comparing different products for the same user group and task.

3.4.1.7.3 Satisfaction

Satisfaction is an assessment of the users' reaction to using the product and should be measured using a standardized questionnaire (SUMI).

3.4.1.8 Interpretation of results

3.4.2 Mapping of measures

3.4.2.1.1 Effectiveness

Effectiveness is measured as a percentage. The criteria for effectiveness will depend on the evaluation level and the nature of the business objectives.

Only a 100% result for all subgoals will be considered satisfactory.

3.4.2.1.2 Productivity

Productivity can be measured by task time or effectiveness/task time. The criteria for efficiency will depend on the evaluation level and the nature of the business objectives.

Relevant rating levels here are:

The cost and time spent in training with the Virtual Fires System compared with traditional training systems.

3.4.2.1.3 Satisfaction

Criteria for satisfaction can be set by comparison with previous results for related products, or by comparison with a supplied database of industry standard norms for the questionnaire.

3.4.2.1.4 Interpretation of measures

3.4.2.1.5 Accuracy

All metrics reported should give mean values and the standard error of the mean. Any claims of differences between values should state the probability that the difference did not occur by chance.

3.4.2.1.6 Interpretation

Each measure should be interpreted in relation to the requirements for quality in use in a specific context of use. It is not generally meaningful to combine efficiency and satisfaction scores.

3.4.2.2 Reporting

A report should have the following contents:

- The purpose of the product: what the product is for, what it is intended to do for its users.
- The objectives of the evaluation, and any specific target values for the measures to be taken.
- The essential characteristics and capabilities that are expected of the user groups that are being evaluated.
- Context of evaluation: the conditions under which the tasks were performed, and any known differences between the evaluated context and the expected context of use.
- Evaluation design: the user groups, the tasks given, any other independent variables, and the measures taken.

- Procedure: sequence of events.
- Task Instructions.
- Results (which should include graphical representation).
- Difficulties encountered by users and suggestions for changes to the product (optional).
- Interpretation.

3.4.2.3 Application procedure

3.4.2.3.1 Resources Required

- An evaluator with skill or expertise in human factors evaluation.
- The minimum effort for the evaluators is approximately: 3 person days planning, 2 person days for the evaluation, and 2 person days for analysis and reporting.

3.4.2.3.2 Evaluation instructions

The purpose of the evaluation is to help readers of the report decide whether the product will have quality in use for their particular users, tasks and working environment. The design of the evaluation should be based on a simulation of the intended usage environment.

The evaluation and the results should be reported in sufficient detail to enable readers to be able to judge the relevance of the results to the needs of their own users, tasks and working environments.

The following guidelines will help ensure that the evaluation procedure is as close to real world usage as possible:

- The evaluation report needs to make it clear which users, tasks and working environments the product is intended for, and the extent to which these characteristics were actually simulated in the evaluation.
- The task instructions should tell users what they need to achieve, without giving any clues about which product features to use.
- To be representative of real world usage, the evaluation situation should be as natural as possible. This may mean simulating distracters and other working conditions. The evaluators should be as unobtrusive as possible (preferably observing remotely in another room).
- During the evaluation participants should not be asked to think aloud. *This condition does not apply to a highly collaborative environment where a facilitator is needed.*
- Participants should not be given any hints or assistance, other than by mechanisms available to real users (such as documentation or a telephone help desk). *This condition does not apply to a highly collaborative environment where a facilitator is needed.*
- Data should be obtained from sufficient users in each category for the sample of users to be representative of the intended user group. Given the typical variability of participants in an evaluation, it has been found that for consistent results it is best to evaluate at least 8 participants from each user group. *This condition can be difficult to fulfil due to the limited availability of some users like authorities, operators and designers.*

- It should be possible for the measures taken to be used to establish acceptance criteria or to make comparisons between products. This means that the measures should be counting items of known value. *Rating levels and acceptance criteria have been already defined.*

3.4.3 Scheduling (timetable) and resource allocation (logistics)

3.4.3.1 Scheduling

According to the description of work, the evaluation process will go on for about two months and it will comprise the following tasks:

- Evaluation of CAVE and HMD installation
- General validation

The plan presented in this document is considered a draft produced at the request of the reviewers. It is clear that defining the final evaluation plan is a task that requires the participation of the whole consortium.

3.4.3.2 Resource allocation

Each evaluation session will be carried out in the presence of an evaluator.

The PC system will also demand the participation of a technical facilitator who will carry out the commands of users and will collaborate with them to get the most out of the system capabilities.

In the case of the CAVE system two technical facilitators will be needed: one controlling the system out of the CAVE and the other inside the CAVE to carry out the commands of users and to collaborate with them to get the most out of the system capabilities

In addition, the hardware and software necessary for the CAVE version and the PC version should be available for each session.

About eight users from the different end users groups will take part in each scenario.

However, a lower number of end users from the authorities/operators/designers group will be used because of their limited availability.

All the CAVE tests will take place in Stockholm because of the particular incidences that have affected this setup.

The PC tests will take place at the partners locations depending on the availability of users to validate the system.

4 Focus group technique description

A number of methods are currently available to assist in developing and testing applications. Group testing is a mixture of both focus groups and usability, it delivers the best of both worlds and provides an efficient way to gather important data on users and their behavior.

In general, a focus group is a session in which current or potential customers gather to discuss issues and concerns about the features of the product presented to them. This somewhat informal technique helps assess user needs and feelings both before interface design and long after implementation.

A focus group usually comprises six to nine users, along with a moderator who keeps the group focused on issues and concerns related to the features of a user interface. Focus groups often bring out users' spontaneous reactions about the product, and ideas for both new features and marketing positioning.

Focus groups are easy to organize and useful for judging subjective measures. However, their success depends heavily on the quality of the moderator, and they only present the customers' descriptions of what they do - not the way they actually work with the product.

Usability testing is the observation and analysis of user behavior while they are using a product or product prototype to achieve a goal. Conducting a usability test is a great way to refine an ongoing design or lay the groundwork for a future offering. The purpose of usability testing is to find out practical information about how users actually use a proposed product or service. Usability tests make it possible to get user feedback on specific features that are of particular interest to the designer or developer.

A focus group is not a usability test and vice versa. A focus group can be very effective at getting opinions, but it has nothing to say about performance. According to usability experts Jakob Nielsen, the proper role of focus groups is not to assess interaction styles or design usability, but to discover what users want from the system. The opposite can be applied to usability testing: It is an excellent technique for evaluation, but inadequate for assessing user needs. The evaluator felt that it made sense in this case to complement the quality in use evaluation with the focus group technique to help with the dissemination, exploitation and improvement of the product.

5 End user evaluation results

Refer to the previous sections for a description of: the purpose of the product, the objectives of the evaluation, context of use and context of evaluation, evaluation design and procedures.

5.1 Effectiveness

All end users managed to perform the tasks assigned to them, so the value of the effectiveness metrics in this case is 100%.

It can be argued that the tasks assigned were not very complicated but this is coherent with the type of product being evaluated and the very stringent level of evaluation proposed.

What is indicated here is that the Virtual Fires system works satisfactorily in all the functionality that offers to end users.

The question arises whether this functionality is enough for end users but this question is better answered by the results of the focus group sessions. In this respect, the Virtual Fires system shows some exclusive functionality in graphical representation and lacks a catalogue of validation cases like the one that a CFD package like Fluent can show.

5.2 Productivity

The average time expended in completing the tasks assigned was 11' with peak values of 9' and 12'. This time is in line with the time invested on dealing with a standard menu-based software application.

A more relevant productivity measure is the time needed to set up a case with the ICE-PC (one week) combination compared to the time needed to do the same with a established CFD package like Fluent (4 weeks). The assessment of the Virtual Fires system in this case is excellent.

Another relevant indicator is cost. In this case, the cost of the ICE-PC package and the cost of a Fluent professional licence is approximately the same (18000 euros), so the evaluation in this case is satisfactory.

5.3 Satisfaction

5.3.1 Evaluation in Lyon

Profile Analysis

Scale	UF	Ucl	Medn	Lcl	LF
Global	66	62	57	52	48
Efficiency	80	68	60	52	44
Affect	75	66	61	55	47
Helpfulness 64	56	53	49	43	
Control	62	57	53	49	45
Learnability 67	55	50	45	39	

Note:

The Median is the middle score when the scores are arranged in numerical order. It is the indicative sample statistic for each usability scale.

The Ucl and Lcl are the Upper and Lower Confidence Limits. They represent the limits within which the theoretical true score lies 95% of the time for this sample of users.

The UF and LF are the Upper and Lower Fences. They represent values beyond which it may be plausibly suspected that a user is not responding with the rest of the group: the user may be responding with an outlier.

Individual User Scores

User	Globa	Effic	Affec	Helpf	Contr	Learn	
1	56	58	66	50	56	52	Etienne
2	60	69	62	54	51	63	Faure RM
3	58	62	56	57	58	48	Ponticq Xavier
4	44	38	47	49	44	48	Rahmani (GEC)
5	66	68	65	63	55	57	Michelle (G)
6	54	56	59	51	50	44	Cottaz

Any scores outside the interval formed by the Upper and Lower Fences are potential outliers. The user who produced an outlier is indicated in the right hand column. The initial letter of the scales in which outliers are found are indicated in parentheses.

Item Consensual Analysis

Detailed results of Item Consensual Analysis are collected in Annex1.

Interpretation

All average values are over 50, except learnability, which is 50, so usability is poor in this respect. Affect and Efficiency are good (values of 60 and 61 respectively). Helpfulness and control, both with values of 53 are improvable.

Affect	61
Efficiency	60
Control	53
Helpfulness	53
Learnability	50

Learnability refers to the early phase in which the use of the software starts to be familiar and to the quality of documentation, help files, etc. A low value of learnability shows indicates that the software is initially hard and that it is easy to forget how to do things. Users doubt that they will manage to control the program. A low value of learnability can also reflect that end users have not received enough training before starting the evaluation or that they did not have access to manuals, etc.

Helpfulness refers to messages and help functions. A low index of Helpfulness indicates that the software does not help in an optimal way, it does not offer information in a consistent way or does not help users to recover from mistakes. Messages and help functions are defective and do not communicate well in terms of language, depth of explanation and technical knowledge required or else relevant questions are not treated in an appropriate way for users. It is also possible that novel users do not find many of the help elements that are very interesting for expert users.

Control refers to the amount of transparency as perceived by the end user. Low Control means that users feel they can get stuck with the software, it can do unpredictable things, and they feel safer if they use only those parts of it they know will work. In operating the software the user feels a certain lack of control over the functions. Feedback may be poor, or commands may be too complex, badly named labelled or organised in a way that does not correspond with how users typically understand them.

Efficiency refers to the support for the user to enable the user to get their work done. High Efficiency means that the software supports and helps users in their work, it works at the user's pace in a predictable, consistent manner.

Affect refers to the likeability and the stress-free usage of the product. High Affect means that users enjoy their sessions with this software, they find it mentally stimulating to use, it is satisfying and attractive.

Concerning the LCI and UCI values, it can be seen that the intervals are small, so data is web-collected and the users come from a homogenous group. Concerning the individual user results, there are two users that fall out of the range defined by UF and LF. It would be interesting to interview them to identify the reasons of their point of view.

Concerning the ITEM analysis it can be seen that there are many users who feel undecided, so results differ from the expected results of standard software.

5.3.2 Evaluation in Dortmund

Profile Analysis

Scale	UF	Ucl	Medn	Lcl	LF
Global	61	59	52	45	40
Efficiency	81	54	45	36	9
Affect	74	66	59	52	38
Helpfulness 70	57	51	45	30	
Control	59	55	49	43	33
Learnability 74	56	48	39	29	

Note:

The Median is the middle score when the scores are arranged in numerical order. It is the indicative sample statistic for each usability scale.

The Ucl and Lcl are the Upper and Lower Confidence Limits. They represent the limits within which the theoretical true score lies 95% of the time for this sample of users.

The UF and LF are the Upper and Lower Fences. They represent values beyond which it may be plausibly suspected that a user is not responding with the rest of the group: the user may be responding with an outlier.

Individual User Scores

User	Global	Effic	Affec	Helpf	Contr	Learn	
1	52	59	63	57	50	44	Schafer
2	54	57	58	56	54	65	Rainer koch
3	54	45	60	60	49	58	Ingo Wofaknocht
4	52	54	59	44	43	59	Spetw, Hauke
5	48	40	62	51	49	44	Bongers, Jens
6	25	26	35	36	29	28	Malthias Birbhahn (GACL)
7	47	33	50	44	41	48	Manfred Ostermann

Any scores outside the interval formed by the Upper and Lower Fences are potential outliers. The user who produced an outlier is indicated in the right hand column. The initial letter of the scales in which outliers are found are indicated in parentheses.

Item Consensual Analysis

Detailed results of Item Consensual Analysis are collected in Annex1.

Interpretation

The evaluation in Dortmund offers slightly lower results in all factors, with a large difference in the Efficiency index. The average values obtained are the following:

Affect	59
Helpfulness	51
Control	49
Learnability	48
Efficiency	45

The Affect dimension continues to be the best developed in this software, obtaining good marks, while the rest of dimensions are clearly improvable.

Efficiency, which was good for the French, must be necessarily improved according to Germans. Low Efficiency means that users don't always know what to do next, the software is slow and sometimes it works in a strange, inconsistent way. In this aspect, some preliminary question are: what is the level of training received by the users; are the users being given tasks outside their range of competence? Is the product properly installed on the correct machine?

It can be observed that low Efficiency correlates with low Control results, so in this case the basic functionality is poor in some aspects of the program.

Concerning the values of Ucl and Lcl, the intervals are larger than in the French case (differences between Ucl and Lcl oscillate between 12 and 18), so the sample user group is not homogeneous.

Concerning individual results, there is a user who goes out of the range defined by UF and LF in two aspects of usability and in the global dimension, so it would be interesting to ask him about his point of view.

The ITEM analysis reveals that there are ITEMS in which expected results significantly differ from what is expected in a standard software package. This is because a large group of users feel undecided, an answer hardly expected in the standard case.

5.3.3 Evaluation in Stockholm

Profile Analysis

Scale	UF	Ucl	Medn	Lcl	LF
Global	59	46	43	39	26
Efficiency	63	47	41	35	14
Affect	63	54	50	45	40
Helpfulness	53	44	41	38	26
Control	51	43	39	35	24
Learnability	59	52	46	40	28

Note:

The Median is the middle score when the scores are arranged in numerical order. It is the indicative sample statistic for each usability scale.

The Ucl and Lcl are the Upper and Lower Confidence Limits. They represent the limits within which the theoretical true score lies 95% of the time for this sample of users.

The UF and LF are the Upper and Lower Fences. They represent values beyond which it may be plausibly suspected that a user is not responding with the rest of the group: the user may be responding with an outlier.

Individual User Scores

User	Globa	Effic	Affec	Helpf	Contr	Learn	
1		49	46	51	45	39	50 T. Snickars
2		44	49	59	41	40	48 Liisa Honkaranta
3		36	41	38	35	33	47 Stefan Gustafson (A)
4		37	30	48	32	30	24 Christer Lundell (L)
5		48	46	55	44	39	46 UH Lundstom
6		43	38	50	41	42	38 Anders Carlsson
7		42	28	48	44	48	44 Christer Lihdemar

Any scores outside the interval formed by the Upper and Lower Fences are potential outliers. The user who produced an outlier is indicated in the right hand column. The initial letter of the scales in which outliers are found are indicated in parentheses.

Item Consensual Analysis

Detailed results of Item Consensual Analysis are collected in Annex1.

Interpretation

In the evaluation session in Stockholm, the results obtained are even lower than the results from Dortmund. The usability indexes are as follows:

Affect	50
Learnability	46
Efficiency	41
Helpfulness	41
Control	39

Affect is again the most positively assessed aspect, although according to the Swedish, it is improvable. The rest of aspects must be necessarily reviewed.

The reason for these low values could be that these end users were not properly trained for the evaluation, since all of them felt undecided in many of the SUMI questions.

Regarding the Lcl and Ucl values, the intervals are small. This means that data is web-collected and the users come from a homogenous group.

Concerning the individual user results, there are two users who fall out of the interval defined by UF and LF, one in Affect and the other in Learnability, so it would be interesting to interview them to identify the reasons of their point of view.

Concerning the ITEM analysis, it can be seen that there is a high level of undecidedness in all users, so the results obtained differ from those of the standard software.

5.4 Focus Group

To better analyze the acceptance of the Virtual Fires system and to learn more about the user expectations for a product like this, we organized a focus group session embedded in each end user evaluation session. The evaluation coordinator helped by a local representative in each evaluation site acted as moderator.

Topic: General discussion of the Virtual Fires System

Guide for Questions:

- What are your expectations of such a product?
- What other products are the competitors of this one from your point of view?
- What are the advantages of this new technology compared to traditional techniques?
- How much would you be ready to pay for a product like this per year?

The results of this focus group were the following:

Lyon

- The Virtual Fires system was deemed too complex and too generic. A system tailored to a certain underground facility with more limited configuration possibilities would be more useful.
- Realism of the simulation was considered a critical point: modelling of fire and nozzles, types of nozzles allowed, etc.
- The system was considered a good tool for designers and risk analyzers, also for real time training of transport operators.
- A 1D representation of smoke progress was considered a necessary feature.
- Systems faster than RT for decision making were considered of interest.
- The Virtual Fires system could be used to produce films for training purposes. One out of fifteen drivers could be instructed in this way to improve common people response to fire emergencies. This would help to illustrate some unknown facts about fire, i.e. that fire progress can be faster than a running person.

- 18000 euros a year was quite an affordable price for transport operators if training was improved significantly.
- Models of people escaping should be included in the system to improve its usefulness.
- Loss of visibility caused by smoke opacity was considered an important feature.
- Lattice-Boltzmann algorithms look suspicious in the eyes of some designers, they are considered not to be mature enough and to show some problems such as excess viscosity or not well tested turbulence behaviour.
- Fluctuating boundary conditions were considered a necessary feature.
- The competitor software developed by NIST was considered very good but not yet applicable to tunnels because it has only been validated for room fires.
- In order to compete with existing CFD packages, the Virtual Fires system has certain advantages: it is highly user friendly and platform independent, cases are easier to define and graphical output is better.
- Validation is very important. Some appropriate testcases have been defined by Ercoftac and include channels and backsteps.
- Five to ten companies in the tunnel design world at least can be envisioned as potential buyers of the system.
- Another competitor for firefighting training is VectorCommand, which is based on playbacks rather than interactive simulation.
- Simulating one second of fire in one minute on a PC with ICE was considered very good performance, but more or less average if the number of cells was considered.
- The influence of traffic on fire spread should be considered.
- The question arises whether transport operators could use a system like this without external support.
- The time consumption of setting up a new scenario is asked (3d for Catia translation, 2d for CFD geometry setup and 1d for adjustments). Modifying the scenario would take a few hours for the geometry and half a day for the CFD domain. Comparatively, a study in Fluent can take four weeks and mastering the software another month.
- The validation must be much more convincing if the system is going to be used for engineering than if it is going to be used for training.
- Transport operators could find such a product useful to define the needs of equipment inside a tunnel.
- Designers consider that the product just now is good for research but not mature enough for industry.
- In order to use the system to validate tunnels, the tool itself must be validated

- A lot of studies have to be performed before constructing a tunnel (fire, toxicity, etc.). But you need to provide a lot of detail to make this tool useful for this purpose.
- Apparently, only a scientist can set up this tool
- Constructors rely on experts who use a variety of calculation tools: hand-made, zone, 1D, CFD, etc.
- The tool would be useful for tunnel constructors if it allowed them to get rid of some of the external consultancy. A tool that would allow them to interpret the results themselves

Dortmund

- The simulated case was too nice in that the smoke went directly to the ventilation.
- Since the high temperature area does not coincide with the smoke area, it is important to have a graphical superimposed representation. Temperature isosurfaces are very useful.
- Representing the air currents can help to understand the behaviour of ventilation.
- Smoke progression representation is considered important.
- Simple controls such as buttons and scrolls should be provided to facilitate interaction.
- A combined sensor/forecasting system could be useful for decision making.
- Current technology for simulating fire includes zone models.
- The system has good visualization capabilities but it is probably too complex to use.
- Validity of the results is a key factor.
- The system is limited for simulating complex networks of tunnels and rooms.
- A clear idea of the work required to use the system is necessary before making a purchase decision. Another relevant question is how many testcases per station would be necessary.
- A larger catalogue of testcases would be necessary.
- It would be interesting to collect real probe data from stations on fire and representing them on Virtual Fires.
- CGNS format is an asset of the system in that it allows to use data from different CFD packages.
- Without knowing in detail benefits and costs, it is difficult to speak about prices, but a useful system would be worth paying 18000 euros per year.
- If there were CAD models of all stations, it would be easier to use the system.
- Real time simulation would be very useful for strategy planning.
- Fast feedback is important for a group session.
- The system is useful for what if studies and best strategy finding.

- Time consumption for station definition is a key issue.
- Common firefighters are just comfortable with smoke and fire representations; they do not have the capability of interpreting representations such as temperature isosurfaces.
- This would be a useful tool for external consultants at the time of designing. For training it is not so good
- Collaboration in a fire emergency, which was not considered here, is a very important feature.
- Having a graphical tool like Virtual Fires would be useful to compare results of different CFD packages.
- The system is a good prototype, but probably not mature enough for commercialization.
- For groups larger than about 25 people practical training is nearly impossible because you need such a big area. Simulation could be useful!
- Importance of moving trains: piston effect that can be simulated with a 1D solver.
- It is more important to measure tendencies rather than extremely precise calculations.
- It is necessary to define summer and winter atmospheric conditions and cover 95% of typical situations. After that, you can define the optimal mission.
- A CFD simulation can cost about 3000 euro and two to 3 weeks to enter the data.
- Simulation can help to define the right safety measures for a station instead of installing everything possible just in case.
- 18000 euros would be a very good price.
- When would it be available commercially? With Viricity in the distribution it could take six to twelve months.
- It would be interesting to create videos to illustrate emergency procedures.

Stockholm

- Key factors are the validity of the results and the length of tunnel that can be evaluated.
- There is interest in simulating a complex network system like the central station in Stockholm.
- Some fire phenomena that create high risk for firefighters such as backlayering (smoke that goes down behind the firefighters) should be simulated.
- Correct modelling of different extinguishing agents is very important.
- It appears that a flat screen system would be better to show a simulation to other people.
- The end user should not go to the CAVE system site. Portable CAVES should be used instead.
- Some possible use in the future for a system like this is to communicate with decision makers or people who do not understand fire:

- Transport operators, police
- Architects
- Authorities who must give permissions or funding
- The software should be combined with an evacuation programme
- Different actions of the firefighters must be contemplated as well as different arrival routes for the firefighters.
- The system is not usable yet although it is exciting to see it.
- It can be a good educational tool.
- It should be more realistic. Not only one user should have a good visual representation of the system (because of the tracker). The image oscillates a bit.
- Good for designers and people who have to understand the dynamics of fire. It is a good tool for city planning
- Sometimes it seems that complicated techniques like 3D is more important than the functionality. It is possible to get the same functionality in 1D
- In a complex system like Tcentral, if possible to simulate, the system would be largely useful for judging and planning and evacuating operations
- This is a technique which I am convinced will develop our way to work and make decisions.
- It is impressive to really be on site
- Iso surfaces with temperature was the most useful tool.
- This could help to convince the safety agency that a tunnel or station could be considered safe.
- Most of the simulations are done by external consultants at SL (Stockholm Local Transport) so it's hard to say how much we would pay for such a system.
- It would be valuable to integrate an evacuation model in the system.

5.5 Difficulties encountered

Transport operators attended the evaluation sessions in very low numbers (one or two people on each evaluation site). Therefore, they have only been considered in the focus group section.

Most end users reacted negatively towards taking part in hands-on test. Apparently the reason was that people who are not familiar with PCs are afraid of looking stupid when using the system. In other contexts with more resources for the evaluation process, a stand-alone package is created with the software application and online task instructions, support and questionnaire filling in facilities.

The evaluation in Dortmund had to be postponed two weeks because of a hardware failure.

The first evaluation session in Stockholm was affected by network problems that made it impossible to test real time interaction with the system.

6 Annex 1. Results of Item Consensual Analysis.

In the following tables, the numbers in the row labelled 'Profile' are the observed responses of the actual users to each item. The numbers in the row labelled 'Expected' are the number of responses expected on the basis of the standardisation database.

The Goodness of Fit between the observed and expected values is summarised using Chi Square, and these statistics are presented on the line below the expected values.

The number at the end of the Goodness of Fit line is the total Chi Square which applies to that item. The greater the value of the total Chi Square, the more likely it is that the obtained values differ from what is expected from the standardisation database.

Each total Chi Square marked with

*** is at least 99.99% certain to be different

** is at least 99% certain to be different

- is at least 95% certain to be different

Total Chi Square values without asterisks are not likely to differ much from the standardisation database. In this output, the SUMI items which differ most from the standardisation are presented first.

6.1 Evaluation in Lyon

The instructions and prompts are helpful.

Item 3		Agree	Undecided	Disagree
Profile		1	5	0
Expected	3,74	1,28	0,99	
Chi Sq		2,0	10,84	0,99
				13,83***

I have to look for assistance most times when I use this software.

Item 50	Agree	Undecided	Disagree		
Profile		0	4	2	
Expected	0,79	0,88	4,33		
Chi Sq		0,79	11,12	1,25	13,17**

There have been times in using this software when I have felt quite tense.

Item 32	Agree	Undecided	Disagree		
Profile		0	4	2	
Expected	2,27	0,91	2,82		
Chi Sq		2,27	10,46	0,24	12,97**

The software has at some time stopped unexpectedly.

Item 4	Agree	Undecided	Disagree		
Profile		0	3	3	
Expected	2,81	0,63	2,55		
Chi Sq		2,81	8,83	0,08	11,72**

Learning how to use new functions is difficult.

Item 35	Agree	Undecided	Disagree		
Profile		0	5	1	
Expected	0,92	1,47	3,61		
Chi Sq		0,92	8,46	1,88	11,27**

The software documentation is very informative.

Item 15	Agree	Undecided	Disagree		
Profile		0	6	0	
Expected	2,09	2,99	0,92		
Chi Sq		2,09	3,02	0,92	6,03*

The software allows the user to be economic of keystrokes.

Item 34	Agree	Undecided	Disagree		
Profile	2	4	0		
Expected	3,69	1,5	0,81		
Chi Sq	0,77	4,16	0,81	5,74	

The software hasn't always done what I was expecting.

Item 41	Agree	Undecided	Disagree	
Profile	0	2	4	
Expected	2,79	1,36	1,85	
Chi Sq	2,79	0,3	2,49	5,58

It is easy to forget how to do things with this software.

Item 45	Agree	Undecided	Disagree	
Profile	2	3	1	
Expected	1,45	1,07	3,49	
Chi Sq	0,21	3,51	1,77	5,5

Getting data files in and out of the system is not easy.

Item 49	Agree	Undecided	Disagree	
Profile	1	5	0	
Expected	0,83	2,46	2,71	
Chi Sq	0,04	2,62	2,71	5,36

If this software stops it is not easy to restart it.

Item 9	Agree	Undecided	Disagree	
Profile	0	5	1	
Expected	0,97	2,31	2,72	
Chi Sq	0,97	3,12	1,08	5,17

I would recommend this software to my colleagues.

Item 2	Agree	Undecided	Disagree
Profile	2	4	0

Expected	3,59	1,61	0,81	
Chi Sq	0,7	3,57	0,81	5,08

This software is awkward when I want to do something which is not standard.

Item 24	Agree	Undecided	Disagree	
Profile	0	5	1	
Expected	1,84	2,45	1,71	
Chi Sq	1,84	2,65	0,29	4,78

Error prevention messages are not adequate.

Item 38	Agree	Undecided	Disagree	
Profile	0	5	1	
Expected	1,49	2,43	2,08	
Chi Sq	1,49	2,73	0,56	4,78

It is obvious that user needs have been fully taken into consideration.

Item 31	Agree	Undecided	Disagree	
Profile	5	1	0	
Expected	2,43	2,17	1,4	
Chi Sq	2,71	0,63	1,4	4,74

The software has a very attractive presentation.

Item 42	Agree	Undecided	Disagree	
Profile	6	0	0	
Expected	3,38	1,63	0,99	
Chi Sq	2,03	1,63	0,99	4,65

I feel safer if I use only a few familiar commands or operations.

Item 14	Agree	Undecided	Disagree	
Profile	2	3	1	
Expected	2,42	1,12	2,45	
Chi Sq	0,07	3,13	0,86	4,06

I think this software has made me have a headache on occasions.

Item 37	Agree	Undecided	Disagree	
Profile	0	3	3	
Expected	1,45	1,25	3,29	
Chi Sq	1,45	2,43	0,03	3,91

Using this software is frustrating.

Item 27	Agree	Undecided	Disagree	
Profile	0	0	6	
Expected	1,03	1,28	3,69	
Chi Sq	1,03	1,28	1,44	3,75

I prefer to stick to the facilities that I know best.

Item 20	Agree	Undecided	Disagree	
Profile	1	3	2	
Expected	2,53	1,25	2,22	
Chi Sq	0,92	2,46	0,02	3,4

I keep having to go back to look at the guides.

Item 30	Agree	Undecided	Disagree	
Profile	0	3	3	
Expected	1,19	1,34	3,47	
Chi Sq	1,19	2,07	0,06	3,32

This software responds too slowly to inputs.

Item 1	Agree	Undecided	Disagree
Profile	0	0	6
Expected	1,14	0,92	3,94
Chi Sq	1,14	0,92	1,07 3,13

This software seems to disrupt the way I normally like to arrange my work.

Item 16	Agree	Undecided	Disagree
Profile	1	3	2
Expected	0,58	1,43	3,98
Chi Sq	0,3	1,72	0,99 3,0

This software occasionally behaves in a way which can't be understood.

Item 46	Agree	Undecided	Disagree
Profile	0	2	4
Expected	1,93	1,51	2,56
Chi Sq	1,93	0,16	0,81 2,91

It is relatively easy to move from one part of a task to another.

Item 44	Agree	Undecided	Disagree
Profile	6	0	0
Expected	4,14	0,93	0,93
Chi Sq	0,83	0,93	0,93 2,69

Tasks can be performed in a straight forward manner using this software.

Item 26	Agree	Undecided	Disagree
Profile	6	0	0

Expected	4,17	1,03	0,8	
Chi Sq	0,8	1,03	0,8	2,63

Either the amount or quality of the help information varies across the system.

Item 43	Agree	Undecided	Disagree	
Profile	1	5	0	
Expected	1,78	3,13	1,09	
Chi Sq	0,34	1,12	1,09	2,56

It takes too long to learn the software commands.

Item 10	Agree	Undecided	Disagree	
Profile	0	0	6	
Expected	0,64	1,05	4,31	
Chi Sq	0,64	1,05	0,66	2,35

It is easy to make the software do exactly what you want.

Item 39	Agree	Undecided	Disagree	
Profile	4	2	0	
Expected	2,5	2,07	1,44	
Chi Sq	0,9	0,0	1,44	2,34

I feel in command of this software when I am using it.

Item 19	Agree	Undecided	Disagree	
Profile	3	3	0	
Expected	3,41	1,59	1,01	
Chi Sq	0,05	1,26	1,01	2,32

The software has helped me overcome any problems I have had in using it.

Item 28	Agree	Undecided	Disagree	
Profile	0	4	2	
Expected	1,65	3,0	1,35	
Chi Sq	1,65	0,34	0,32	2,31

Working with this software is satisfying.

Item 12	Agree	Undecided	Disagree	
Profile	5	1	0	
Expected	3,23	1,95	0,82	
Chi Sq	0,98	0,46	0,82	2,26

It is easy to see at a glance what the options are at each stage.

Item 48	Agree	Undecided	Disagree	
Profile	4	0	2	
Expected	3,35	1,43	1,22	
Chi Sq	0,13	1,43	0,49	2,05

I find that the help information given by this software is not very useful.

Item 8	Agree	Undecided	Disagree	
Profile	0	3	3	
Expected	1,33	1,93	2,74	
Chi Sq	1,33	0,6	0,02	1,95

Working with this software is mentally stimulating.

Item 17	Agree	Undecided	Disagree	
Profile	4	1	1	
Expected	2,43	2,08	1,49	
Chi Sq	1,01	0,56	0,16	1,73

There are too many steps required to get something to work.

Item 36	Agree	Undecided	Disagree	
Profile	2	0	4	
Expected	1,24	1,19	3,56	
Chi Sq	0,46	1,19	0,05	1,71

I enjoy my sessions with this software.

Item 7	Agree	Undecided	Disagree	
Profile	2	3	1	
Expected	3,48	1,7	0,82	
Chi Sq	0,63	1,0	0,04	1,67

There is never enough information on the screen when it's needed.

Item 18	Agree	Undecided	Disagree	
Profile	0	1	5	
Expected	1,03	1,34	3,63	
Chi Sq	1,03	0,09	0,52	1,64

The way that system information is presented is clear and understandable.

Item 13	Agree	Undecided	Disagree	
Profile	5	1	0	
Expected	3,78	1,32	0,9	
Chi Sq	0,39	0,08	0,9	1,37

There is too much to read before you can use the software.

Item 25	Agree	Undecided	Disagree
Profile	0	2	4

Expected	0,97	1,33	3,71	
Chi Sq	0,97	0,34	0,02	1,33

I sometimes don't know what to do next with this software.

Item 6	Agree	Undecided	Disagree	
Profile	2	2	2	
Expected	1,92	1,04	3,04	
Chi Sq	0,0	0,9	0,36	1,26

I sometimes wonder if I am using the right command.

Item 11	Agree	Undecided	Disagree	
Profile	1	1	4	
Expected	2,13	0,95	2,92	
Chi Sq	0,6	0,0	0,4	1,0

I think this software is inconsistent.

Item 21	Agree	Undecided	Disagree	
Profile	0	2	4	
Expected	0,83	1,55	3,62	
Chi Sq	0,83	0,13	0,04	1,0

The organisation of the menus or information lists seems quite logical.

Item 33	Agree	Undecided	Disagree	
Profile	5	1	0	
Expected	4,3	0,97	0,74	
Chi Sq	0,12	0,0	0,74	0,85

I can understand and act on the information provided by this software.

Item 23	Agree	Undecided	Disagree	
Profile	5	1	0	
Expected	4,34	1,1	0,56	
Chi Sq	0,1	0,01	0,56	0,68

This software is really very awkward.

Item 47	Agree	Undecided	Disagree	
Profile	0	1	5	
Expected	0,4	1,06	4,54	
Chi Sq	0,4	0,0	0,05	0,45

I will never learn to use all that is offered in this software.

Item 40	Agree	Undecided	Disagree	
Profile	1	2	3	
Expected	1,7	1,65	2,65	
Chi Sq	0,29	0,07	0,05	0,41

The speed of this software is fast enough.

Item 29	Agree	Undecided	Disagree	
Profile	4	1	1	
Expected	3,36	1,0	1,64	
Chi Sq	0,12	0,0	0,25	0,37

Learning to operate this software initially is full of problems.

Item 5	Agree	Undecided	Disagree	
Profile	1	1	4	
Expected	1,3	1,11	3,59	
Chi Sq	0,07	0,01	0,05	0,13

I would not like to use this software every day.

Item 22	Agree	Undecided	Disagree
Profile	1	1	4
Expected	1,29	1,01	3,7
Chi Sq	0,06	0,0	0,02 0,09

6.2 *Evaluation in Dortmund*

Tasks can be performed in a straight forward manner using this software.

Item 26	Agree	Undecided	Disagree
Profile	0	3	4
Expected	4,87	1,2	0,93
Chi Sq	4,87	2,69	10,12 17,68***

Working with this software is mentally stimulating.

Item 17	Agree	Undecided	Disagree
Profile	7	0	0
Expected	2,84	2,42	1,74
Chi Sq	6,1	2,42	1,74 10,26**

There is too much to read before you can use the software.

Item 25	Agree	Undecided	Disagree
Profile	4	2	1
Expected	1,13	1,55	4,32
Chi Sq	7,32	0,13	2,55 10,01**

I find that the help information given by this software is not very useful.

Item 8	Agree	Undecided	Disagree	
Profile	0	6	1	
Expected	1,55	2,25	3,2	
Chi Sq	1,55	6,26	1,51	9,32**

It is easy to see at a glance what the options are at each stage.

Item 48	Agree	Undecided	Disagree	
Profile	2	5	0	
Expected	3,91	1,66	1,43	
Chi Sq	0,93	6,68	1,43	9,04*

Learning how to use new functions is difficult.

Item 35	Agree	Undecided	Disagree	
Profile	0	5	2	
Expected	1,08	1,72	4,21	
Chi Sq	1,08	6,28	1,16	8,51*

It is obvious that user needs have been fully taken into consideration.

Item 31	Agree	Undecided	Disagree	
Profile	0	6	1	
Expected	2,84	2,53	1,64	
Chi Sq	2,84	4,78	0,25	7,86*

I have to look for assistance most times when I use this software.

Item 50	Agree	Undecided	Disagree	
Profile	3	2	2	
Expected	0,92	1,02	5,05	
Chi Sq	4,66	0,93	1,84	7,43*

Either the amount or quality of the help information varies across the system.

Item 43	Agree	Undecided	Disagree	
Profile	0	7	0	
Expected	2,08	3,65	1,28	
Chi Sq	2,08	3,08	1,28	6,43*

I keep having to go back to look at the guides.

Item 30	Agree	Undecided	Disagree	
Profile	2	4	1	
Expected	1,39	1,56	4,05	
Chi Sq	0,26	3,81	2,29	6,37*

The way that system information is presented is clear and understandable.

Item 13	Agree	Undecided	Disagree	
Profile	3	4	0	
Expected	4,41	1,54	1,05	
Chi Sq	0,45	3,94	1,05	5,45

This software responds too slowly to inputs.

Item 1	Agree	Undecided	Disagree	
Profile	2	3	2	
Expected	1,33	1,07	4,6	
Chi Sq	0,34	3,49	1,47	5,29

The software hasn't always done what I was expecting.

Item 41	Agree	Undecided	Disagree
Profile	2	4	1

Expected	3,25	1,58	2,16	
Chi Sq	0,48	3,69	0,62	4,79

There have been times in using this software when I have felt quite tense.

Item 32	Agree	Undecided	Disagree	
Profile	1	3	3	
Expected	2,65	1,06	3,29	
Chi Sq	1,03	3,53	0,03	4,58

It is easy to make the software do exactly what you want.

Item 39	Agree	Undecided	Disagree	
Profile	1	5	1	
Expected	2,91	2,41	1,68	
Chi Sq	1,26	2,78	0,27	4,31

It is easy to forget how to do things with this software.

Item 45	Agree	Undecided	Disagree	
Profile	0	3	4	
Expected	1,69	1,24	4,07	
Chi Sq	1,69	2,48	0,0	4,17

I sometimes wonder if I am using the right command.

Item 11	Agree	Undecided	Disagree	
Profile	1	3	3	
Expected	2,49	1,11	3,4	
Chi Sq	0,89	3,22	0,05	4,16

It is relatively easy to move from one part of a task to another.

Item 44	Agree	Undecided	Disagree	
Profile	3	3	1	
Expected	4,83	1,08	1,09	
Chi Sq	0,69	3,41	0,01	4,11

The speed of this software is fast enough.

Item 29	Agree	Undecided	Disagree	
Profile	2	3	2	
Expected	3,92	1,17	1,91	
Chi Sq	0,94	2,88	0,0	3,82

I think this software is inconsistent.

Item 21	Agree	Undecided	Disagree	
Profile	1	4	2	
Expected	0,97	1,81	4,22	
Chi Sq	0,0	2,65	1,17	3,82

I prefer to stick to the facilities that I know best.

Item 20	Agree	Undecided	Disagree	
Profile	1	1	5	
Expected	2,95	1,46	2,6	
Chi Sq	1,29	0,14	2,23	3,66

There are too many steps required to get something to work.

Item 36	Agree	Undecided	Disagree	
Profile	0	3	4	
Expected	1,45	1,39	4,16	
Chi Sq	1,45	1,85	0,01	3,31

Getting data files in and out of the system is not easy.

Item 49	Agree	Undecided	Disagree	
Profile	2	4	1	
Expected	0,97	2,87	3,16	
Chi Sq	1,11	0,44	1,48	3,03

The instructions and prompts are helpful.

Item 3	Agree	Undecided	Disagree	
Profile	4	3	0	
Expected	4,36	1,49	1,15	
Chi Sq	0,03	1,53	1,15	2,71

This software is awkward when I want to do something which is not standard.

Item 24	Agree	Undecided	Disagree	
Profile	1	5	1	
Expected	2,14	2,86	1,99	
Chi Sq	0,61	1,6	0,5	2,71

I feel safer if I use only a few familiar commands or operations.

Item 14	Agree	Undecided	Disagree	
Profile	2	3	2	
Expected	2,83	1,31	2,86	
Chi Sq	0,24	2,17	0,26	2,67

The software has at some time stopped unexpectedly.

Item 4	Agree	Undecided	Disagree
Profile	2	2	3

Expected	3,28	0,74	2,98	
Chi Sq	0,5	2,15	0,0	2,65

There is never enough information on the screen when it's needed.

Item 18	Agree	Undecided	Disagree	
Profile	0	3	4	
Expected	1,2	1,57	4,23	
Chi Sq	1,2	1,31	0,01	2,53

Using this software is frustrating.

Item 27	Agree	Undecided	Disagree	
Profile	1	0	6	
Expected	1,2	1,5	4,31	
Chi Sq	0,03	1,5	0,67	2,2

I will never learn to use all that is offered in this software.

Item 40	Agree	Undecided	Disagree	
Profile	1	1	5	
Expected	1,98	1,93	3,09	
Chi Sq	0,49	0,45	1,19	2,12

Learning to operate this software initially is full of problems.

Item 5	Agree	Undecided	Disagree	
Profile	0	2	5	
Expected	1,51	1,29	4,19	
Chi Sq	1,51	0,38	0,16	2,05

The software allows the user to be economic of keystrokes.

Item 34	Agree	Undecided	Disagree	
Profile	4	3	0	
Expected	4,3	1,75	0,95	
Chi Sq	0,02	0,89	0,95	1,86

The software has helped me overcome any problems I have had in using it.

Item 28	Agree	Undecided	Disagree	
Profile	1	3	3	
Expected	1,93	3,5	1,57	
Chi Sq	0,45	0,07	1,3	1,82

If this software stops it is not easy to restart it.

Item 9	Agree	Undecided	Disagree	
Profile	0	4	3	
Expected	1,13	2,7	3,17	
Chi Sq	1,13	0,63	0,01	1,77

The organisation of the menus or information lists seems quite logical.

Item 33	Agree	Undecided	Disagree	
Profile	4	1	2	
Expected	5,01	1,13	0,86	
Chi Sq	0,2	0,01	1,51	1,72

This software occasionally behaves in a way which can't be understood.

Item 46	Agree	Undecided	Disagree	
Profile	1	3	3	
Expected	2,25	1,76	2,98	
Chi Sq	0,7	0,87	0,0	1,57

I would recommend this software to my colleagues.

Item 2	Agree	Undecided	Disagree	
Profile	3	2	2	
Expected	4,18	1,87	0,94	
Chi Sq	0,34	0,01	1,19	1,53

I enjoy my sessions with this software.

Item 7	Agree	Undecided	Disagree	
Profile	4	3	0	
Expected	4,06	1,98	0,95	
Chi Sq	0,0	0,52	0,95	1,48

Working with this software is satisfying.

Item 12	Agree	Undecided	Disagree	
Profile	4	3	0	
Expected	3,76	2,28	0,96	
Chi Sq	0,01	0,23	0,96	1,2

I would not like to use this software every day.

Item 22	Agree	Undecided	Disagree	
Profile	2	2	3	
Expected	1,5	1,18	4,32	
Chi Sq	0,16	0,57	0,4	1,14

It takes too long to learn the software commands.

Item 10	Agree	Undecided	Disagree	
Profile	0	1	6	

Expected	0,75	1,23	5,03	
Chi Sq	0,75	0,04	0,19	0,97

The software documentation is very informative.

Item 15	Agree	Undecided	Disagree	
Profile	2	3	2	
Expected	2,44	3,49	1,07	
Chi Sq	0,08	0,07	0,81	0,96

Error prevention messages are not adequate.

Item 38	Agree	Undecided	Disagree	
Profile	1	4	2	
Expected	1,74	2,83	2,43	
Chi Sq	0,31	0,48	0,08	0,87

The software has a very attractive presentation.

Item 42	Agree	Undecided	Disagree	
Profile	3	2	2	
Expected	3,94	1,9	1,16	
Chi Sq	0,23	0,01	0,62	0,85

I feel in command of this software when I am using it.

Item 19	Agree	Undecided	Disagree	
Profile	3	2	2	
Expected	3,98	1,85	1,17	
Chi Sq	0,24	0,01	0,58	0,83

I think this software has made me have a headache on occasions.

Item 37	Agree	Undecided	Disagree	
Profile	1	1	5	
Expected	1,69	1,46	3,84	
Chi Sq	0,28	0,15	0,35	0,78

This software is really very awkward.

Item 47	Agree	Undecided	Disagree	
Profile	1	1	5	
Expected	0,47	1,23	5,3	
Chi Sq	0,6	0,04	0,02	0,66

This software seems to disrupt the way I normally like to arrange my work.

Item 16	Agree	Undecided	Disagree	
Profile	1	2	4	
Expected	0,68	1,67	4,65	
Chi Sq	0,15	0,06	0,09	0,3

I can understand and act on the information provided by this software.

Item 23	Agree	Undecided	Disagree	
Profile	5	1	1	
Expected	5,06	1,28	0,66	
Chi Sq	0,0	0,06	0,18	0,24

I sometimes don't know what to do next with this software.

Item 6	Agree	Undecided	Disagree	
Profile	2	1	4	
Expected	2,24	1,21	3,55	
Chi Sq	0,03	0,04	0,06	0,12

6.3 Evaluation in Stockholm

It is relatively easy to move from one part of a task to another.

Item 44	Agree	Undecided	Disagree	
Profile	0	7	0	
Expected	4,83	1,08	1,09	
Chi Sq	4,83	32,41	1,09	38,33***

This software responds too slowly to inputs.

Item 1	Agree	Undecided	Disagree	
Profile	0	6	1	
Expected	1,33	1,07	4,6	
Chi Sq	1,33	22,73	2,82	26,88***

It takes too long to learn the software commands.

Item 10	Agree	Undecided	Disagree	
Profile	1	6	0	
Expected	0,75	1,23	5,03	
Chi Sq	0,09	18,6	5,03	23,72***

I have to look for assistance most times when I use this software.

Item 50	Agree	Undecided	Disagree	
Profile	2	5	0	
Expected	0,92	1,02	5,05	
Chi Sq	1,25	15,46	5,05	21,76***

I think this software is inconsistent.

Item 21	Agree	Undecided	Disagree
---------	-------	-----------	----------

Profile	0	7	0	
Expected	0,97	1,81	4,22	
Chi Sq	0,97	14,89	4,22	20,08***

I think this software has made me have a headache on occasions.

Item 37	Agree	Undecided	Disagree	
Profile	0	6	1	
Expected	1,69	1,46	3,84	
Chi Sq	1,69	14,08	2,1	17,88***

The way that system information is presented is clear and understandable.

Item 13	Agree	Undecided	Disagree	
Profile	0	6	1	
Expected	4,41	1,54	1,05	
Chi Sq	4,41	12,95	0,0	17,36***

I sometimes wonder if I am using the right command.

Item 11	Agree	Undecided	Disagree	
Profile	2	5	0	
Expected	2,49	1,11	3,4	
Chi Sq	0,09	13,64	3,4	17,14***

There is too much to read before you can use the software.

Item 25	Agree	Undecided	Disagree	
Profile	1	6	0	
Expected	1,13	1,55	4,32	
Chi Sq	0,01	12,79	4,32	17,13***

I keep having to go back to look at the guides.

Item 30	Agree	Undecided	Disagree
Profile	1	6	0
Expected	1,39	1,56	4,05
Chi Sq	0,11	12,63	4,05 16,78***

The software hasn't always done what I was expecting.

Item 41	Agree	Undecided	Disagree
Profile	1	6	0
Expected	3,25	1,58	2,16
Chi Sq	1,56	12,31	2,16 16,04***

I find that the help information given by this software is not very useful.

Item 8	Agree	Undecided	Disagree
Profile	0	7	0
Expected	1,55	2,25	3,2
Chi Sq	1,55	10,04	3,2 14,79***

It is easy to forget how to do things with this software.

Item 45	Agree	Undecided	Disagree
Profile	1	5	1
Expected	1,69	1,24	4,07
Chi Sq	0,28	11,36	2,32 13,95***

I feel in command of this software when I am using it.

Item 19	Agree	Undecided	Disagree
Profile	0	6	1
Expected	3,98	1,85	1,17

Chi Sq	3,98	9,31	0,03	13,31**
--------	------	------	------	---------

I can understand and act on the information provided by this software.

Item 23	Agree	Undecided	Disagree
---------	-------	-----------	----------

Profile	2	5	0
---------	---	---	---

Expected	5,06	1,28	0,66
----------	------	------	------

Chi Sq	1,85	10,77	0,66	13,27**
--------	------	-------	------	---------

The instructions and prompts are helpful.

Item 3	Agree	Undecided	Disagree
--------	-------	-----------	----------

Profile	0	5	2
---------	---	---	---

Expected	4,36	1,49	1,15
----------	------	------	------

Chi Sq	4,36	8,25	0,63	13,24**
--------	------	------	------	---------

There are too many steps required to get something to work.

Item 36	Agree	Undecided	Disagree
---------	-------	-----------	----------

Profile	1	5	1
---------	---	---	---

Expected	1,45	1,39	4,16
----------	------	------	------

Chi Sq	0,14	9,34	2,4	11,88**
--------	------	------	-----	---------

I sometimes don't know what to do next with this software.

Item 6	Agree	Undecided	Disagree
--------	-------	-----------	----------

Profile	3	4	0
---------	---	---	---

Expected	2,24	1,21	3,55
----------	------	------	------

Chi Sq	0,26	6,45	3,55	10,26**
--------	------	------	------	---------

Getting data files in and out of the system is not easy.

Item 49	Agree	Undecided	Disagree
---------	-------	-----------	----------

Profile	0	7	0	
Expected	0,97	2,87	3,16	
Chi Sq	0,97	5,93	3,16	10,06**

The software has at some time stopped unexpectedly.

Item 4	Agree	Undecided	Disagree	
Profile	4	3	0	
Expected	3,28	0,74	2,98	
Chi Sq	0,16	6,9	2,98	10,04**

There is never enough information on the screen when it's needed.

Item 18	Agree	Undecided	Disagree	
Profile	1	5	1	
Expected	1,2	1,57	4,23	
Chi Sq	0,03	7,53	2,47	10,03**

There have been times in using this software when I have felt quite tense.

Item 32	Agree	Undecided	Disagree	
Profile	2	4	1	
Expected	2,65	1,06	3,29	
Chi Sq	0,16	8,11	1,59	9,86**

The organisation of the menus or information lists seems quite logical.

Item 33	Agree	Undecided	Disagree	
Profile	2	4	1	
Expected	5,01	1,13	0,86	
Chi Sq	1,81	7,32	0,02	9,15*

Working with this software is satisfying.

Item 12	Agree	Undecided	Disagree	
Profile	1	6	0	
Expected	3,76	2,28	0,96	
Chi Sq	2,03	6,08	0,96	9,07*

It is easy to see at a glance what the options are at each stage.

Item 48	Agree	Undecided	Disagree	
Profile	1	5	1	
Expected	3,91	1,66	1,43	
Chi Sq	2,16	6,68	0,13	8,97*

This software occasionally behaves in a way which can't be understood.

Item 46	Agree	Undecided	Disagree	
Profile	2	5	0	
Expected	2,25	1,76	2,98	
Chi Sq	0,03	5,94	2,98	8,95*

The software allows the user to be economic of keystrokes.

Item 34	Agree	Undecided	Disagree	
Profile	1	5	1	
Expected	4,3	1,75	0,95	
Chi Sq	2,53	6,03	0,0	8,56*

Tasks can be performed in a straight forward manner using this software.

Item 26	Agree	Undecided	Disagree
Profile	2	4	1
Expected	4,87	1,2	0,93

Chi Sq	1,69	6,51	0,01	8,2*
--------	------	------	------	------

I will never learn to use all that is offered in this software.

Item 40	Agree	Undecided	Disagree
---------	-------	-----------	----------

Profile	1	5	1
---------	---	---	---

Expected	1,98	1,93	3,09
----------	------	------	------

Chi Sq	0,49	4,88	1,41	6,78*
--------	------	------	------	-------

I enjoy my sessions with this software.

Item 7	Agree	Undecided	Disagree
--------	-------	-----------	----------

Profile	2	5	0
---------	---	---	---

Expected	4,06	1,98	0,95
----------	------	------	------

Chi Sq	1,05	4,59	0,95	6,59*
--------	------	------	------	-------

I feel safer if I use only a few familiar commands or operations.

Item 14	Agree	Undecided	Disagree
---------	-------	-----------	----------

Profile	6	1	0
---------	---	---	---

Expected	2,83	1,31	2,86
----------	------	------	------

Chi Sq	3,56	0,07	2,86	6,5*
--------	------	------	------	------

Error prevention messages are not adequate.

Item 38	Agree	Undecided	Disagree
---------	-------	-----------	----------

Profile	0	6	1
---------	---	---	---

Expected	1,74	2,83	2,43
----------	------	------	------

Chi Sq	1,74	3,54	0,84	6,12*
--------	------	------	------	-------

This software is awkward when I want to do something which is not standard.

Item 24	Agree	Undecided	Disagree
---------	-------	-----------	----------

Profile	0	6	1	
Expected	2,14	2,86	1,99	
Chi Sq	2,14	3,44	0,5	6,08*

Using this software is frustrating.

Item 27	Agree	Undecided	Disagree	
Profile	0	4	3	
Expected	1,2	1,5	4,31	
Chi Sq	1,2	4,18	0,4	5,78

It is easy to make the software do exactly what you want.

Item 39	Agree	Undecided	Disagree	
Profile	0	5	2	
Expected	2,91	2,41	1,68	
Chi Sq	2,91	2,78	0,06	5,76

It is obvious that user needs have been fully taken into consideration.

Item 31	Agree	Undecided	Disagree	
Profile	1	5	1	
Expected	2,84	2,53	1,64	
Chi Sq	1,19	2,42	0,25	3,86

Either the amount or quality of the help information varies across the system.

Item 43	Agree	Undecided	Disagree	
Profile	1	6	0	
Expected	2,08	3,65	1,28	
Chi Sq	0,56	1,52	1,28	3,35

The software has a very attractive presentation.

Item 42	Agree	Undecided	Disagree	
Profile	2	4	1	
Expected	3,94	1,9	1,16	
Chi Sq	0,96	2,32	0,02	3,29

I prefer to stick to the facilities that I know best.

Item 20	Agree	Undecided	Disagree	
Profile	1	3	3	
Expected	2,95	1,46	2,6	
Chi Sq	1,29	1,64	0,06	2,99

The software has helped me overcome any problems I have had in using it.

Item 28	Agree	Undecided	Disagree	
Profile	0	5	2	
Expected	1,93	3,5	1,57	
Chi Sq	1,93	0,65	0,12	2,69

Working with this software is mentally stimulating.

Item 17	Agree	Undecided	Disagree	
Profile	4	3	0	
Expected	2,84	2,42	1,74	
Chi Sq	0,48	0,14	1,74	2,35

Learning how to use new functions is difficult.

Item 35	Agree	Undecided	Disagree
Profile	1	3	3
Expected	1,08	1,72	4,21

Chi Sq	0,01	0,96	0,35	1,31
--------	------	------	------	------

If this software stops it is not easy to restart it.

Item 9	Agree	Undecided	Disagree	
Profile	1	4	2	
Expected	1,13	2,7	3,17	
Chi Sq	0,02	0,63	0,43	1,07

The speed of this software is fast enough.

Item 29	Agree	Undecided	Disagree	
Profile	4	2	1	
Expected	3,92	1,17	1,91	
Chi Sq	0,0	0,59	0,43	1,03

I would recommend this software to my colleagues.

Item 2	Agree	Undecided	Disagree	
Profile	3	3	1	
Expected	4,18	1,87	0,94	
Chi Sq	0,34	0,68	0,0	1,02

This software is really very awkward.

Item 47	Agree	Undecided	Disagree	
Profile	0	2	5	
Expected	0,47	1,23	5,3	
Chi Sq	0,47	0,48	0,02	0,97

This software seems to disrupt the way I normally like to arrange my work.

Item 16	Agree	Undecided	Disagree
---------	-------	-----------	----------

Profile	0	2	5	
Expected	0,68	1,67	4,65	
Chi Sq	0,68	0,06	0,03	0,77

Learning to operate this software initially is full of problems.

Item 5	Agree	Undecided	Disagree	
Profile	1	1	5	
Expected	1,51	1,29	4,19	
Chi Sq	0,17	0,07	0,16	0,4

I would not like to use this software every day.

Item 22	Agree	Undecided	Disagree	
Profile	2	1	4	
Expected	1,5	1,18	4,32	
Chi Sq	0,16	0,03	0,02	0,22

The software documentation is very informative.

Item 15	Agree	Undecided	Disagree	
Profile	2	4	1	
Expected	2,44	3,49	1,07	
Chi Sq	0,08	0,07	0,0	0,16