INTERNATIONAL STANDARD

ISO/IEC 25062

First edition
2006-04-01

Corrected version
2006-10-01

Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Common Industry Format (CIF) for usability test reports

Ingénierie du logiciel — Exigences de qualité du produit logiciel et évaluation (SQuaRE) — Format commun de l'industrie (CIF) pour les rapports d'essai de rentabilité

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

Contents

Page

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

ISO/IEC 25062 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 7, *Software and system engineering.*

This corrected version of ISO/IEC 25062 incorporates the following correction:

   Figure 1, Appendix C, has been included on page 43.

Introduction

Usability of software is a key factor in predicting successful deployment of that software. Software manufacturers subject software to usability testing at various stages in a product's development; some companies that purchase software also test products for usability before making purchasing decisions. Testing often involves (1) subjects who are representative of the target population of users of the software, (2) representative tasks, and (3) measures of efficiency, effectiveness and subjective satisfaction. When this type of experimental situation exists, the testing is termed summative, i.e., the results can be expressed as statistically meaningful measures of central tendency (e.g. mean or median) and variability (e.g. standard deviation). The Common Industry Format (CIF) for Usability Test Reports is intended for use by usability professionals to report the results of summative usability testing.

The CIF standardizes the types of information that are captured about testing with users. The level of detail allows the same or another organization to replicate the test procedure. The major variables are user demographics, task descriptions, context of the test, including the equipment used, the environment in which the test is conducted, and the protocol by which the subjects and the test administrator(s) interact, as well as the particular metrics chosen to code the findings of the study.

The CIF is intended to replace the proprietary formats employed by companies that perform usability testing, both vendors and purchasers of software. Until now there has been no standard format for reporting usability testing results. Advantages of using a standardized reporting format include (1) a reduction in training time for usability staff since an individual only needs to learn to use one form regardless of how many companies he works for and (2) enhanced potential for increased communication between vendors and purchasing organizations since readers of CIF-compliant reports will share a common language and expectations.

The purpose of this International Standard is to facilitate incorporation of usability as part of the procurement decision-making process for interactive software products so that it is easier to judge whether a product meets usability goals. Examples of decisions include purchasing, upgrading and automating. It provides a common format for human factors engineers and usability professionals in supplier companies to report the methods and results of usability tests to customer organizations.

Audience

The Common Industry Format (CIF) is meant to be used by usability professionals within supplier organizations to generate reports that can be used by customer organizations in the CIF report. The CIF is also meant to be used by customer organizations to verify that a particular report is CIF-compliant. The Usability Test Report itself is intended for two types of readers:

— Usability professionals in customer organizations who are evaluating both the technical merit of usability tests and the usability of the products; and

— Other technical professionals and managers who are using the test results to make business decisions.

The CIF may also be used within a single organization if a formal report of a summative usability test needs to be generated. In this case additional material such as a list of detailed findings may be included.

The report is in two main sections, an Executive Summary and a main body. The main body contains the Methods and Results sections and is aimed at the first audience above. These sections (1) describe the test methodology and results in sufficient technical detail to allow replication by another organization if the test is repeated, and (2) support application of test data to questions about the product's expected costs and benefits. Understanding and interpreting these sections will require technical background in human factors or usability engineering for optimal use. The second audience is directed to the Introduction, which provides summary information for non-usability professionals and managers. The Introduction may also be of general interest to other computing professionals. Decision makers without usability engineering expertise may find the information in the main body to be useful but should rely on expert interpretation when necessary.

Organization

Clause 1 describes the scope of this specification and the conformance criteria. Clause 4 provides definitions

of the terms used throughout the document. Clause 5 is the main description of the specification.

Additional Information

Annex A provides a checklist that can be used to ensure inclusion of required and recommended information. A glossary is provided in Annex B to define terminology used in the report format description. A Word template for report production can be found at: http://www.ncits.org/ref-docs/CIF/CIF_template.dot. A printed version of the template can be found in Annex C. An example is provided in Annex D illustrating how the format is used followed by an informative bibliography.

Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Common Industry Format (CIF) for usability test reports

1 Scope

This International Standard is intended to be used to report the measures obtained from a test of usability as defined in ISO 9241-11: effectiveness, efficiency and satisfaction in a specified context of use.

NOTE   Metrics for other more-detailed usability requirements can be found in ISO/IEC 9126 parts 2 and 3.

This International Standard is intended to be used by:

⎯ usability professionals within supplier organizations to generate reports that can be used by customer organizations;

⎯ customer organizations to verify that a particular report conforms to this International Standard;

⎯ human factors or other usability professionals in customer organizations who are evaluating both the technical merit of usability tests and the usability of the products; and

⎯ other technical professionals and managers in the customer organization who are using the test results to make business decisions about product suitability and purchase.

The Executive Summary and Introduction in 5.2 and 5.3 provide summary information for nonusability professionals and managers.

Subclauses 5.4 and 5.5 describe the test methodology and results in technical detail suitable for replication, and also support application of test data to questions about the product's expected costs and benefits. Understanding and interpreting these sections will require technical background in human factors or usability engineering for optimal use.

The report format assumes sound practice [1, 2] has been followed in the design and execution of the test. Test procedures which produce measures that summarize usability should be used, i.e. the test is summative in nature. Some usability evaluation methods, such as formative tests, are intended to identify problems rather than produce measures; the format is not structured to support the results of such testing methods.

2 Conformance

A usability test report conforms to this International Standard if it complies with all the requirements in this International Standard (stated as "shall"). The recommendations (stated as "should") should be implemented whenever appropriate.

This International Standard specifies the minimum information that should be provided. Additional information may be included. For example, if an organization finds that an additional list of findings is useful, the list may be included even though it is not specified as part of a conformant CIF report.

3 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 9126-1:2001, *Software engineering — Product quality — Part 1: Quality model*

ISO/IEC 9126-2:2001, *Software engineering — Product quality — Part 2: External metrics*

ISO/IEC 9126-3:2001, *Software engineering — Product quality — Part 3: Internal metrics*

ISO/IEC 9126-4:2001, *Software engineering — Product quality — Part 4: Quality in use metrics*

ISO 9241-11:1998, *Ergonomic requirements for office work with visual display terminals (VDTs) — Part 11: Guidance on usability*

ISO 13407:1999, *Human-centred design processes for interactive systems — Annex C*

ISO/IEC 14598-5:1998, *Information technology — Software product evaluation — Part 5: Process for evaluators*

4 Terms and definitions

For the purposes of this document, the following definitions apply. The source for definitions 4.1 to 4.9 is ISO 9241-11:1998.

NOTE   Efficiency, effectiveness, and satisfaction are defined in ISO/IEC 9126 in a similar way to the definitions in ISO/IEC 9241. However, in ISO/IEC 9126 they are defined in terms of the software product where as in ISO/IEC 9241 they are defined in terms of the user's view. The term efficiency in this International Standard is equivalent to the term productivity in ISO 9126, as the term efficiency has been defined with a different meaning in ISO 9126.

4.1
usability
the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use

NOTE   Quality in use is defined in ISO/IEC 9126 in a similar way to the definition of usability in ISO/IEC 9241. Quality in use may be influenced by any of the quality characteristics, and is thus broader than the definition of usability in ISO/IEC 9126 that is in terms of understandability, learnability, operability, attractiveness and compliance.

4.2
effectiveness
the accuracy and completeness with which users achieve specified goals

4.3
efficiency
resources expended in relation to the accuracy and completeness with which users achieve goals

NOTE   Efficiency in the context of usability is related to "productivity" rather than to its meaning in the context of software efficiency.

4.4
satisfaction
freedom from discomfort, and positive attitudes towards the use of the product

4.5
context of use
users, tasks, equipment (hardware, software and materials), and the physical and social environments in which a product is used

4.6
user
person who interacts with the product

4.7
user group
subset of intended users who are differentiated from other intended users by factors such as age, culture or expertise that are likely to influence usability

4.8
goal
intended outcome of user interaction with a product

NOTE   Specific goals relating to user interaction may be referred to as "task goals"

4.9
task
activities required to achieve a goal

NOTE 1   These activities can be physical or cognitive.

NOTE 2   Job responsibilities can determine goals and tasks.

4.10
accessibility
usability of a product, service, environment or facility by people with the widest range of capabilities

NOTE 1   Although "accessibility" typically addresses users who have a disability, the concept is not limited to disability issues.

NOTE 2   The usability-oriented concept of accessibility focuses on achieving levels of effectiveness, efficiency and satisfaction that are as high as possible taking account of the specified context of use, while paying particular attention to the diversity of the capabilities within the user population, and thus aims to minimize the differences in usability experienced by individuals.

4.11
assistive technologies
hardware or software that is added to or incorporated within a system that increases accessibility for an individual

EXAMPLES   Braille displays, screen readers, screen magnification software and eye tracking devices are assistive technologies.

4.12
assist
tester intervention in the form of direct procedural help provided by the test administrator to the test participant in order to allow the test to continue when the participant could not complete the tasks on their own

5 Report Format

5.1 Title page

The following information shall be provided:

– Identify report as: ISO/IEC 25062 Common Industry Format for Usability Test Reports and contact information.

– Name the product and version that was tested.

– Who led the test.

– When the test was conducted.

– Date the report was prepared.

– Who prepared the report.

– Contact name(s) for questions and/or clarifications.

– Supplier phone number.

– Supplier e-mail address.

– Supplier mailing or postal address.

The following information should be provided:

– Customer Company Name.

– Customer Company contact person.

5.2 Executive summary

This section provides a high level overview of the test. The intent of this section is to provide information for procurement decision-makers in customer organizations. These people might not read the technical body of this document.

This section shall begin on a new page and end with a page break to facilitate its use as a stand-alone summary.

A high level overview of the test shall be provided that includes:

– Name and description of the product.

– Summary of method(s) including number(s) and type(s) of participants and tasks.

– Performance and satisfaction results expressed as mean scores or other suitable measure of central tendency.

The following information should be provided:

– Reason for and nature of the test.

– Tabular summary of performance results.

    – If differences between values or products are claimed, the probability that the difference did not occur by chance.

5.3 Introduction

This clause provides a description of the product and the test objectives.

5.3.1 Full product description

The following information shall be provided:

    – Formal product name and release or version.

    – The parts of the product that were evaluated.

    – The user population for which the product is intended.

The following information should be provided:

    – Assistive technologies that are supported by the product.

    – Brief description of the physical and social environment(s) in which the product is intended to be used.

    – The type of user work that is supported by the product.

5.3.2 Test objectives

The following information shall be provided:

    – The objectives for the test including any areas of specific interest.

        NOTE Possible objectives include testing user performance of work tasks and subjective satisfaction in using the product to assess whether a product meets specific success criteria.

    – Functions and components with which the user directly and indirectly interacted.

The following information should be provided:

    – Reason for focusing on a product subset, if the whole product was not tested.

5.4 Method

Sufficient information shall be provided to allow an independent tester to replicate the procedure used in testing.

5.4.1 Participants

The following information shall be provided:

    – The total number of participants tested.

        NOTE In order to generate valid summative statistical analyses, it is necessary to test sufficient numbers of subjects. Eight or more subjects/cell (segment) are recommended for this purpose [9].

    – Segmentation of user groups tested, if more than one.

    – Key characteristics and capabilities of user group.

– How participants were selected; whether they had the essential characteristics.

– Differences between the participant sample and the user population.

EXAMPLE Actual users might attend a training course whereas test subjects were untrained.

– Table of participants (row) by characteristics (columns), including demographics, professional experience, computing experience and user needs for assistive technology.

NOTE Subjects with needs for assistive technology might be individuals with physical or cognitive disabilities.

EXAMPLE The table below is an example; the characteristics that are shown are typical but might not necessarily cover every type of testing situation.

	Gender	Age	Education	Occupation / role	Professional Experience	Computer Experience	Product Experience
P1	●	●	●	●	●	●	●
P2	●	●	●	●	●	●	●
Pn	●	●	●	●	●	●	●

For 'Gender', the table entries are male or female.

For 'Age', the table entries state the chronological age of the participant, or indicate membership in an age range (e.g., 25-45) or age category (e.g., under 18, over 65) if the exact age is not known.

For 'Education', the table entries state the number of years of completed formal education (e.g., in the US, a high-school graduate would have 12 years of education and a college graduate 16 years). In some instances, level of education might be stated as 'highest academic degree'.

For 'Occupation/role', the table entries describe the user's job role when using the product. Use the Role title if known.

For 'Professional experience', the table entries give the amount of time the user has been performing in the role.

For 'Computer experience', the table entries describe relevant background such as how much experience the user has with the platform or operating system, and/or the product domain. This might be more extensive than one column.

For 'Product experience', the table entries indicate the type and duration of any prior experience with the product or with similar products.

The characteristics and capabilities shall be complete enough so that an essentially similar group of participants can be recruited.

Characteristics and capabilities should be chosen to be relevant to the product's usability; they should allow a customer to determine how similar the participants were to the customers' user population.

The following information should be provided:

– Description of any users with needs for assistive technology.

Participants should not be from the same organization as the testing or supplier organization.

5.4.2   Context of product use in the test
The following information shall be provided:

– Any important differences between the evaluated context and the expected context of use.

5.4.2.1 Tasks

The following information shall be provided:

– The tasks for testing.

– The task scenarios for testing.

– Why these tasks were selected.

EXAMPLES The most frequent tasks, the most troublesome tasks.

– The source of these tasks.

EXAMPLES Observation of customers using similar products, product marketing specifications, discussion with users or design team.

– Any task data given to the participants.

EXAMPLE Source materials for data entry.

– Completion or performance criteria established for each task.

5.4.2.2 Test facility

The following information should be provided:

– The setting and type of space in which the evaluation was conducted.

EXAMPLES Usability lab, cubicle office, meeting room, home office, home family room, manufacturing floor.

– Any relevant features or circumstances that could affect the results.

EXAMPLES Video and audio recording equipment, one-way mirrors, or automatic data collection equipment.

5.4.2.3 Participant's computing environment

The following information shall provide enough information to replicate and validate the test, including:

– Computer configuration, including model, OS version, required libraries or settings.

– If used, browser name and version; relevant plug-in names and versions.

5.4.2.3.1 Display devices

The following information shall be provided:

– If screen-based, screen size, resolution, and colour setting.

– If print-based, the media size and print resolution.

– If visual interface elements (such as fonts) can vary in size, specify the size(s) used in the test.

5.4.2.3.2 Audio devices

The following information should be provided:

− If used, the relevant settings or values for the audio bits (determining sound quality), volume, etc.

5.4.2.3.3 Input devices

The following information should be provided:

− If used, the make and model of devices used in the test.

5.4.2.4 Test administrator tools

The following information shall be provided:

− If a standard questionnaire (Section 5.4.4.3 contains a partial list) was used, describe or specify it here.

NOTE Customized questionnaires are included in an appendix.

The following information should be provided:

− Any hardware or software used to control the test or to record data.

5.4.3 Experimental design

The following information shall be provided:

− The experimental design of the test; the plan for assigning experimental conditions to participants and the statistical analysis associated with the plan independent variables and control variables.

− The independent variables and control variables.

− The measures for which data were recorded for each set of conditions.

5.4.3.1 Procedure

The following information shall be provided:

− Operational definitions of measures.

− Descriptions of independent variables or control variables.

− Time limits on tasks.

− Policies and procedures for interaction between tester(s) and subjects.

The following information should be provided:

− Sequence of events from greeting the participants to dismissing them.

− Details of nondisclosure agreements, form completion, warm-ups, pretask training, and debriefing.

− Verification that the participants knew and understood their rights as human subjects [1].

− Steps followed to execute the test sessions and record data.

− Number and roles of people who interacted with the participants during the test session.

− Whether other individuals were present in the test environment.

–   Whether participants were paid or otherwise compensated.

5.4.3.2   Participant general instructions

The following information shall be provided:

–   Instructions given to the participants (here or in an Appendix).

–   Instructions on how participants were to interact with any other persons present, including how they were to ask for assistance and interact with other participants, if applicable.

5.4.3.3   Participant task instructions

The following information shall be provided:

–   Task instruction summary.

5.4.4   Usability metrics

As defined in 4.1, usability is measured by three types of metrics: effectiveness, efficiency, and satisfaction. The choice of measures depends on the goals of a particular study, characteristics of the users, the specific tasks, and context-dependent features; e.g., see Dumas and Redish [4].

The following information shall be provided:

–   Metrics for effectiveness.

–   Metrics for efficiency.

–   Metrics for satisfaction.

Effectiveness and efficiency results shall be reported, even when they are difficult to interpret within the specified context of use. In this case, the report shall specify why the supplier does not consider the metrics meaningful.

EXAMPLE   Suppose that the context of use for the product includes real time, open-ended interaction between close associates. In this case, Time-On-Task may not be meaningfully interpreted as a measure of efficiency, because for many users, time spent on this task is "time well spent".

If it is necessary to provide participants with assists, efficiency and effectiveness metrics shall be provided for both unassisted and assisted conditions, and the number and type of assists shall be included as part of the test results.

5.4.4.1   Effectiveness

Effectiveness relates the goals of using the product to the accuracy and completeness with which these goals can be achieved. Common measures of effectiveness include percent task completion, frequency of errors, frequency of assists to the participant from the testers, and frequency of accesses to help or documentation by the participants during the tasks. It does not take account of how the task goals were achieved, only the extent to which they were achieved. Efficiency relates the level of effectiveness achieved to the quantity of resources expended.

5.4.4.1.1   Completion rate

The completion rate is the percentage of participants who completely and correctly achieve each task goal. If goals can be partially achieved (e.g., by incomplete or suboptimum results), then it might also be useful to report the average goal achievement, scored on a scale of 0 to 100% based on specified criteria related to the value of a partial result. For example, a spell-checking task might involve identifying and correcting 10 spelling errors and the completion rate might be

calculated based on the percent of errors corrected. Another method for calculating completion rate is weighting; e.g., spelling errors in the title page of the document are judged to be twice as important as errors in the main body of text. The rationale for choosing a particular method of partial goal analysis should be stated, if such results are included in the report.

The following information shall be provided:

– The percentage of participants who completely and correctly achieve each task goal.

5.4.4.1.2   Errors

Errors are instances where test participants did not complete the task successfully, or had to attempt portions of the task more than once. Scoring of data should include classifying errors according to some taxonomy, such as in [10].

5.4.4.1.3   Assists

When participants cannot proceed on a task, the test administrator sometimes gives direct procedural help in order to allow the test to proceed. Although 'assist' measures can be related to both learnability and effectiveness, assists have been integrated with effectiveness in this specification to comply with the three elements of usability defined in ISO 9241-11.

The following information shall be provided:

– The unassisted completion rate (i.e., the rate achieved without intervention from the testers).

– The assisted rate (i.e., the rate achieved with tester intervention) where these two metrics differ.

EXAMPLE If a participant received an assist on Task A, that participant should not be included among those successfully completing the task when calculating the unassisted completion rate for that task. However, if the participant went on to successfully complete the task following the assist, he could be included in the assisted Task A completion rate.

When assists are allowed or provided, the number and type of assists shall be included as part of the test results. When assists are allowed by the study design, the need for presenting the assisted completion rate is not meaningfully distinguishable from the unassisted rate. In such cases, only a single rate needs to be provided but the number and types of assists shall be included.

EXAMPLE In an educational setting, students might be encouraged to seek assistance.

In some usability tests, participants are instructed to use support tools such as online help or documentation, which are part of the product, when they cannot complete tasks on their own. Accesses to product features that provide information and help are *not* considered assists for the purposes of this report. It may, however, be desirable to report the frequency of accesses to different product support features, especially if they factor into participants' ability to use products independently.

5.4.4.2   Efficiency

Efficiency 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 might also relate to other resources (e.g., total cost of usage). A common measure of efficiency is time on task.

The following information shall be provided:

– The mean time taken to complete each task, together with the range and standard deviation of times across participants.

5.4.4.2.1   Completion rate/mean time-on-task

Completion Rate/Mean Time-On-Task is another measure of efficiency [10]. The relationship of success rate to time allows customers to compare fast error-prone interfaces (e.g., command lines with wildcards to delete files) to slow easy interfaces (e.g., using a mouse and keyboard to drag each file to the trash).

5.4.4.3 Satisfaction

The following information shall be provided:

– One or more measures of user satisfaction.

Satisfaction describes a user's subjective response when using the product. User satisfaction might be an important correlate of motivation to use a product and might affect performance in some cases. Questionnaires to measure satisfaction and associated attitudes are commonly built using Likert and semantic differential scales.

A variety of instruments are available for measuring user satisfaction of software interactive products, and many companies create their own. Whether an external, standardized instrument is used or a customized instrument is created, subjective rating dimensions such as Satisfaction, Usefulness, and Ease of Use should be considered for inclusion, as these will be of general interest to customer organizations.

A number of questionnaires are available that are widely used. They include: ASQ [6], CUSI [7], PSSUQ [7], QUIS [3], SUMI [5], and SUS [2]). While each offers unique perspectives on subjective measures of product usability, most include measurements of Satisfaction, Usefulness, and Ease of Use. General information on questionnaire construction are addressed in Dumas and Redish [4].

Suppliers may choose to use validated published satisfaction measures or may submit satisfaction metrics they have developed themselves.

5.5 Results

5.5.1 Data analysis

The following information shall be provided in sufficient detail to allow replication of the data analysis methods by another organization if the test is repeated:

– Data collection – the differences between the data that was planned to be collected and the data that was actually collected.

   EXAMPLE How missing data was treated. How data was treated with respect to exclusion of outliers.

– Data scoring – the mapping between the data values that were collected and the values used in further analysis.

   EXAMPLE How errors were categorized. How actual ages map to age ranges. How assisted errors are mapped to a set of values.

– Data reduction – the method used to compute the measure of central tendency and to characterize the variation in the data.

   EXAMPLE Which measure of central tendency was used (e.g. mean or mode). How variation was measured (e.g., standard deviation or range).

– Statistical analysis – the statistical procedures used to analyze the data.

   EXAMPLE How groups were compared (e.g., t-test. F-test …).

Data that are reported as means shall include the standard deviation and optionally the standard error of the mean.

5.5.2 Presentation of the results

The following information shall be provided:

– The measures of effectiveness and efficiency that characterize the performance results per task or task group in tabular form.

EXAMPLE The following table would be replicated for each task in a task group.

Task A

User #	Unassisted Task Effectiveness [(%)Complete]	Assisted Task Effectiveness [(%)Complete]	Task Time (min)	….	Errors	Assists	Efficiency
1							
2							
N							
Mean							
Standard Deviation							
Min							
Max							

Various graphical formats are effective in describing usability data at a glance. Bar graphs are useful for describing subjective data such as that gleaned from Likert scales. A variety of plots can be used effectively to show comparisons of expert benchmark times for a product vs. the mean participant performance time. The data might be accompanied by a brief explanation of the results, but a detailed interpretation is discouraged in the body of the report. If necessary, a detailed interpretation may be included in an appendix.

The following information may be provided:

– Graphs of the measures of effectiveness and efficiency that characterize the performance results.

5.5.2.1 Performance results

The following information should be provided:

– Summary Table(s) of the measures of effectiveness and efficiency that characterize the performance results across all tasks.

– Graphical Presentation of the measures of effectiveness and efficiency that characterize the performance results.

– Additional tables of metrics if they are relevant to the product's design and a particular application area.

EXAMPLE The following is a summary table of performance results across all tasks.

Summary

User #	Total Unassisted Task Effectiveness [(%)Complete]	Total Assisted Task Effectiveness [(%)Complete]	Total Task Time (min)	…	Total Errors	Total Assists	Efficiency
1							
2							
N							
Mean							
Standard Deviation							
Min							
Max							

A table of results may be presented for groups of related tasks (e.g., all program creation tasks in one group, all debugging tasks in another group) where this is more efficient and makes sense. If a unit task has subtasks, then the subtasks may be reported in summary form for the unit task.

EXAMPLE If a unit task is to identify all the misspelled words on a page, then the results might be summarized as a percent of misspellings found.

5.5.2.2 Satisfaction results

The following information shall be provided:

– Summary table(s) of the measures that characterize the satisfaction results.

EXAMPLE The following table is a summary table of partial satisfaction results.

Satisfaction

Participant Number	Ease of Use	Usefulness	Appearance	Clarity	…
1					
2					
N					
Median					
Min					
Max					

The following information should be provided:

–   Graphical presentation of the measures that characterize the satisfaction results.

5.6   Appendices

The following information shall be provided:

–   Custom questionnaires (Sections 5.4.2.4 and 5.4.4.3), if used.

–   Participant general instructions (if not in the body of the report); see Section 5.4.3.2.

–   Participant task instructions (Section 5.4.3.3).

The following information may be provided:

–   Release Notes explaining or updating the test results.

Annex A
(informative)

Checklist

Use the following checklist to assure that required elements (❖) are addressed in your CIF report. Recommended items are denoted with a '♦'.
Title Page
♦ Company Logo or Name
❖ Identify report as: ISO/IEC 25062 Common Industry Format for Usability Test Reports
❖ Name the product and version that was tested
❖ Who led the test
❖ When the test was conducted
❖ Date the report was prepared
❖ Who prepared the report
♦ Customer Company Name
♦Customer Company contact person
❖ Contact name(s) for questions and/or clarifications
❖ Supplier phone number
❖ Supplier e-mail address
❖ Supplier mailing or postal address
Executive Summary
❖ Provide a high level overview of the test
❖ Name and describe the product
❖ Summary of method(s) including number and type of participants and tasks
❖ Results expressed as mean scores or other suitable measure of central tendency
♦ Reason for and nature of the test
♦ Tabular summary of performance results
♦ If differences are claimed, the associated statistical probability
❖ Start on new page; end with page break
Introduction
Full Product Description
❖ Formal product name and release or version
❖ Describe what parts of the product were evaluated
❖ The user population for which the product is intended
♦ Assistive technologies that are supported by the product
♦ Brief description of the environment in which it should be used
♦ The type of user work that is supported by the product

Test Objectives
❖ State the objectives for the test and any areas of specific interest
❖ Functions and components with which the user directly and indirectly interacted
◆ Reason for focusing on a product subset
Method
Participants
❖ The total number of participants tested
❖ Segmentation of user groups tested, if more than one
❖ Key characteristics and capabilities of user groups
❖ How participants were selected; whether they had the essential characteristics
❖ Differences between the participant sample and the user population
❖ Table of participant (row) characteristics (columns)
◆ Description of users with needs for assistive technologies
Context of Product Use in the Test
❖Any important differences between the evaluated context and the expected context of use
Tasks
❖ Describe the tasks for testing
❖ Describe the task scenarios for testing
❖ Explain why these tasks were selected
❖ Describe the source of these tasks
❖ Include any task data given to the participants
❖ Completion or performance criteria established for each task
Test Facility
◆ Describe the setting, and type of space in which the evaluation was conducted
◆ Detail any relevant features or circumstances which could affect the results
Participant's Computing Environment
❖ Computer configuration, including model, OS version, required libraries or settings
❖ If used, browser name and version; relevant plug-in names and versions
Display Devices
❖ If screen-based, screen size, resolution, and color setting
❖ If print-based, the media size and print resolution
❖ If visual interface elements can vary in size, specify the size(s) used in the test
Audio Devices
◆ If used, specify relevant settings or values for the audio bits, volume, etc
Manual Input Devices
◆ If used, specify the make and model of devices used in the test
Test Administrator Tools
❖ If a questionnaire was used, describe or specify it here
◆ Describe any hardware or software used to control the test or to record data

Experimental Design
❖ Describe the experimental design of the test
❖ Define independent variables and control variables
❖ Describe the measures for which data were recorded
Procedure
❖ Operational definitions of measures
❖ Operational definitions of independent variables or control variables
❖ Time limits on tasks
❖ Policies and procedures for interaction between tester(s) and subjects
◆ Sequence of events from greeting the participants to dismissing them
◆ Non-disclosure agreements, form completion, warm-ups, pre-task training, and debriefing
◆ Verify that the participants knew and understood their rights as human subjects
◆ Specify steps followed to execute the test sessions and record data
◆ Number and roles of people who interacted with the participants during the test session
◆ Specify if other individuals were present in the test environment
◆ State whether participants were paid
Participant General Instructions
❖ Instructions given to the participants (here or in Appendix)
❖ Instructions on how participants were to interact with any other persons present
Participant Task Instructions
❖ Task instruction summary
Usability Metrics
❖ Metrics for effectiveness
❖ Metrics for efficiency
❖ Metrics for satisfaction
Results
Data Analysis
❖ Data Collection
❖ Data Scoring
❖ Data Reduction
❖ Statistical Analysis
Presentation of the Results
❖ The measures of effectiveness and efficiency that characterize the performance results per task or task group in tabular form
Performance Results
◆ Summary Table(s) of Performance Results across all tasks
◆ Graphical Presentation of Performance Results
◆ Additional tables of metrics if relevant

Satisfaction Results	
❖ Summary Table(s) of measures that characterize the satisfaction results	
❖ Graphical Presentation of the measures that characterize the satisfaction results	
Appendices	
❖ Custom Questionnaires, if used	
❖ Participant General Instructions	
❖ Participant Task Instructions	
◆ Release Notes	

Annex B
(informative)

Glossary

A.1
automatic data collection equipment
hardware and/or software for automatic collection of usability test data. Examples of such data collection include screen capture and video, audio, and keystroke recording, often with time stamping.

A.2
configuration
The specific collection of hardware and software, including model and version numbers, that constitutes the computing environment on which the product was tested. The configuration includes the platform.

A.3
customer
organization considering the purchase of the product

A.4
debriefing
post-test interview with a participant

A.5
experimental design
plan for assigning experimental conditions to participants and the statistical analysis associated with the plan Typically, this includes a specification of the independent variables, dependent variables, number of participants and sampling strategy, procedure for assigning participants to experimental conditions, and order in which test tasks are given.

A.6
formative evaluation
a method of testing as part of an iterative design process Often involves use of a thinking aloud procedure. (Contrast with Summative test.)

A.7
independent variable
variable manipulated by the tester to set up the test conditions (Contrast with the dependent variable, which is a variable that is being measured by the study.)

EXAMPLE If two systems are being compared with the same tasks, the independent variable is the system, and the dependent variable is the level of task performance or user satisfaction.

A.8
likert scale
rating scale designed to measure user attitudes or reactions by quantifying subjective information It typically contains a set of favorable and unfavorable statements about the product, to which a respondent indicates his or her level of agreement or disagreement. A Likert questionnaire may have an odd or even number of points, is often numerically numbered, and typically has at least two anchors describing the extremes of the scale dimension.

EXAMPLE

This product was easy to use.

Strongly disagree	Somewhat disagree	Slightly disagree	Neutral	Slightly agree	Somewhat agree	Strongly agree

I felt frustrated while using this product.

Strongly disagree	Somewhat disagree	Slightly disagree	Neutral	Slightly agree	Somewhat agree	Strongly agree

A.9
mean
arithmetic 'average' of a set of quantities Specifically, the sum of the quantities divided by the number of quantities in the set.

A.10
nondisclosure agreement
An agreement between the supplier and the customer, in which they recognize that their cooperation requires them to divulge some of their proprietary information to each other, and in which each agrees not to disclose any of the other's proprietary information to anyone else.

A.11
one-way mirror
piece of glass that appears as a mirror from one side and a window from the other, depending on the lighting conditions This device allows testers to observe a usability test without themselves being observed by or distracting the participants.

A.12
participant
person who participates in a usability test as a representative of the product's target user population A participant performs the test tasks and follows other instructions given by the tester.

A.13
platform
the kind(s) of computer (operating system and hardware) used for the testing Examples of common platforms include Windows on COMPAQ, MacOS on Power Macintosh, UNIX on Sparc or HP/UX, Linux on Dell, and OS/2 on IBM. Novel devices such as PDAs can also be considered platforms.

A.14
population
the set of individuals from which a statistical sample is taken

A.15
procurement decision maker
a person responsible for making decisions about which products to acquire for the customer organization

A.16
replication
the collection of two observations under identical test conditions In this standard, it is used to mean that the customer organization should be able to repeat the test described in the report with its own sample of participants.

A.17
scenario
a description of the sequence of events from the user's perspective to perform a task in a specified context

A.18
semantic differential
a rating scale designed to measure user attitudes or reactions by quantifying subjective information It typically contains a set of bipolar adjective pairs (e.g., good/bad, easy/hard) that describe some characteristic of the product. A semantic differential questionnaire typically has a seven- to ten-point scale.

EXAMPLE

This product was

easy to use	1	2	3	4	5	6	7	hard to use
frustrating	1	2	3	4	5	6	7	satisfying

A.19
standard deviation
the square root of the variance. Provides a measure of the variability of the distribution of measurements.

A.20
standard error of the mean
square root of (variance divided by sample size). Estimates the likely closeness of the sample mean to the population mean.

A.21
statistically significant
unlikely to have occurred by chance (i.e., likely to reflect real differences)

A.22
summative test
a usability test of a completed product to determine how well it meets its usability objectives

A.23
supplier
the producer of a product being considered for purchase. Organization that provides a product to the customer [ISO/IEC TR 15504-9:1998].

A.24
task scenario
the presentation and description of a task as part of a realistic, very short story

A.25
taxonomy
the systematic classification of the objects in a field of study

A.26
test protocol
a list of the steps to be followed in the test

A.27
tester / test administrator
a person who designs and/or conducts a usability test. (The tester is not the person who interacts with the product to exercise task scenarios; that person is a test participant.)

A.28
usability professional/specialist/engineer
person who engages in the practice of usability engineering/testing Such practitioners know how to design and execute summative performance tests with human subjects and how to analyze the resulting data. Their training is usually an advanced degree in human factors, behavioral science, industrial engineering, human-computer interaction, industrial design, computer science, or a related field.

A.29
validation
a determination of the validity of a measure such as a questionnaire or a usability test

A.30
variance
the sum of squared differences of a set of measurements from their <u>mean</u>, divided by the sample size. (See <u>Standard Deviation</u>)

Annex C
(informative)

Report Template

1) ISO/IEC 25062 Common Industry Format for Usability Test Reports

[❖ Name the product and version that was tested]

[❖ Who led the test]

[❖ When the test was conducted]

[❖ Date the report was prepared]

[❖ Who prepared the report]

For:

[♦ Customer Company Name]

[♦ Customer Company contact person]

Address inquiries to: [❖ Contact name(s) for questions and/or clarifications]

Phone: [❖ Enter supplier's phone number]

Email: [❖ Enter supplier's email address]

Address: [❖ Enter supplier's mailing or postal address]

---

1) Legend for use: Items prefixed with a ❖ are Required; items prefixed with a ♦ are recommended.

Executive Summary

[❖ Provide a high level overview of the test]

[❖ Name and describe the product]

[❖ Summary of method(s) including number and type of participants and tasks]

[❖ Results expressed as mean scores or other suitable measure of central tendency]

[✦ Reason for and nature of the test]

[✦ Tabular summary of performance results]

[✦ If differences are claimed, the associated statistical probability.]

Introduction
Full Product Description

[❖ Formal product name and release or version]

[❖ Describe what parts of the product were evaluated]

[❖ The user population for which the product is intended]

[♦ Assistive technology supported by the product]

[♦ Brief description of the environment in which it should be used]

[♦ The type of user work that is supported by the product]

Test Objectives

[❖ State the objectives for the test and any areas of specific interest]

[❖ Functions and components with which the user directly and indirectly interacted]

[♦ Reason for focusing on a product subset]

Method
Participants

[❖ The total number of participants tested]

[❖ Segmentation of user groups tested, if more than one]

[❖ Key characteristics and capabilities of user group]

[❖ How participants were selected; whether they had the essential characteristics]

[❖ Differences between the participant sample and the user population]

[❖ Table of participants (row) and characteristics (columns)]

	Characteristic 1	Characteristic 2	…	Characteristic N
Participant 1				
Participant 2				
…				
Participant N				

[♦ Description of any users with needs for assistive technology]

Context of Product Use in the Test

[❖Any known differences between the evaluated context and the expected context of use]

Tasks

[❖ Describe the tasks for testing]

[❖ Describe the task scenarios for testing]

[❖ Explain why these tasks were selected]

[❖ Describe the source of these tasks]

[❖ Include any task data given to the participants]

[❖ Completion or performance criteria established for each task]

Test Facility

[⬧ Describe the setting, and type of space in which the evaluation was conducted]

[⬧ Detail any relevant features or circumstances which could affect the results]

Participant's Computing Environment

[❖ Computer configuration, including model, OS version, required libraries or settings]

[❖ If used, browser name and version; relevant plug-in names and versions]

Display Devices

[❖ If screen-based, screen size, resolution, and color setting]

[❖ If print-based, the media size and print resolution]

[❖ If visual interface elements can vary in size, specify the size(s) used in the test]

Audio Devices

[⬧ If used, specify relevant settings or values for the audio bits, volume, etc]

Manual Input Devices

[⬧ If used, specify the make and model of devices used in the test]

Test Administrator Tools

[❖ If a questionnaire was used, describe or specify it here]

[⬧ Describe any hardware or software used to control the test or to record data]

Experimental Design

[❖ Describe the experimental design of the test]

[❖ Define independent variables and control variables].

[❖ Describe the measures for which data were recorded]

Procedure

[❖Operational definitions of measures]

[❖ Operational definitions of independent variables or control variables]

[❖Time limits on tasks]

[❖Policies and procedures for interaction between tester(s) and subjects]

[⬧ Sequence of events from greeting the participants to dismissing them]

[⬧ Non-disclosure agreements, form completion, warm-ups, pre-task training, and debriefing]

[⬧ Verify that the participants knew and understood their rights as human subjects]

[⬧ Specify steps followed to execute the test sessions and record data]

[⬧ Number and roles of people who interacted with the participants during the test session]

[⬧ Specify if other individuals were present in the test environment]

[⬧ State whether participants were paid]

Participant General Instructions

[❖ Instructions given to the participants (here or in Appendix)]

[❖ Instructions on how participants were to interact with any other persons present]

Participant Task Instructions

[❖ Task instruction summary]

Usability Metrics

[❖ Metrics for effectiveness]

[❖ Metrics for efficiency]

[❖ Metrics for satisfaction]

Results

Data Analysis

[❖ Data Collection]

[❖ Data Scoring]

[❖ Data Reduction]

[❖ Statistical Analysis]

Presentation of the Results

[❖Measures of effectiveness and efficiency that characterize the performance results per task or task group in tabular form]

User #	Measure 1	Measure 2	…	Measure N
Participant 1				
Participant 2				
…				
Participant N				
Mean				
Standard Dev.				
Min				
Max				

Performance Results

[⬥ Summary Tables of the measures of effectiveness and efficiency that characterize the performance results across all tasks]

User #	Total Measure 1	Total Measure 2	…	Total Measure N
Participant 1				
Participant 2				
…				
Participant N				
Mean				
Standard Dev.				
Min				
Max				

[⬥Graphical Presentation of Performance Results]

[⬥ Additional tables of metrics as relevant]

Satisfaction Results

[❖ Summary table(s) of the measures that characterize the satisfaction results]

User #	Scale 1	Scale 2	…	Scale N
Participant 1				
Participant 2				
…				
Participant N				
Mean				
Standard Dev.				
Min				
Max				

[⬥Graphical Presentation of Satisfaction Results]

Appendices

[❖ Custom Questionnaires, if used]

[❖ Participant General Instructions]

[❖ Participant Task Instructions]

[⬥ Release Notes]

Annex D
(informative)

Example

ISO/IEC 25062 Common Industry Format for Usability Test Reports

Oracle Universal Installer NT for Oracle 8.1.3

Tested by: Kathy Baxter

October, 1998

Date Prepared: February 1, 2000

Prepared by: Mark Stimson

For:

The Boeing Company

Keith Butler

Address inquiries to: Anna Wichansky

Phone: 650-506-8007

Email: awichans@us.oracle.com

Address: 500 Oracle Parkway MS2op10, Redwood Shores, CA. 94065

Executive Summary

Oracle Universal Installer is the current revised GUI provided since 1998 to guide users through the installation process for Oracle products such as the database client and server software. A usability test was conducted on the task of using the Oracle Universal Installer to install the Oracle8 Enterprise Edition for NT.

Four Oracle database administrators (DBAs) and 4 non-Oracle database NT system administrators (NT SysAdmins) participated in this study. Each participant performed 3 installation tasks using the Universal Installer: the typical install; custom install; and DBA client install. There were 2 task orders in an attempt to reduce learning effects. Thus, the experimental design was a 2 x 2 x 3 mixed factors design, with user and order factored between groups, and the installation task factored within groups. There were 5 dependent variables: unassisted task completion rate; time to task completion; number of errors; number of assists; and on-line help accesses. A subjective questionnaire was also administered at the end to obtain qualitative exploratory data. The mean task completion time for all 8 participants was XX minutes for all tasks.   Both user groups had the same error rate of X.X and the same unassisted task completion rate of XX.

Summary Performance Results By User Type

| User Type	Assisted Task Completion Rate	Unassisted Task Completion Rate	Mean Total Task Completion Time	Errors	Assists	Efficiency
DBAs	XXXXXX	XXXXXX	XXXXXX	XXX	XXX	XXXXXX
NT SysAdmin	XXXXXX	XXXXXX	XXXXXX	XXX	XXX	XXXXXX
Mean	XXXXXX	XXXXXX	XXXXXX	XXX	XXX	XXXXXX

Introduction

Full Product Description

Oracle8 is the object oriented relational database management system that is designed to support all of a customer's users and data, providing a high-speed and cost-effective system for running business applications. Oracle 8.1.3 installs via the Oracle Universal Installer which allows users to install the various files contained within Oracle8, as well as to perform custom installs and uninstalls for user specified files. All major parts of the Universal Installer that are needed for installing Oracle 8.1.3 Enterprise Edition were evaluated. This study tested the post-beta code version.

The architectural redesign of the original Oracle Installer was undertaken to address today's challenges for software packaging, installation and distribution. The new Oracle Universal Installer is based on a powerful Java engine that provides an extensible environment able to handle more complex internal and customer requirements. The Universal Installer allows for a common installation process on any platform.

The user population for the Universal Installer is very broad. It includes high end users such as DBAs and System Administrators, as well as the occasional end user.
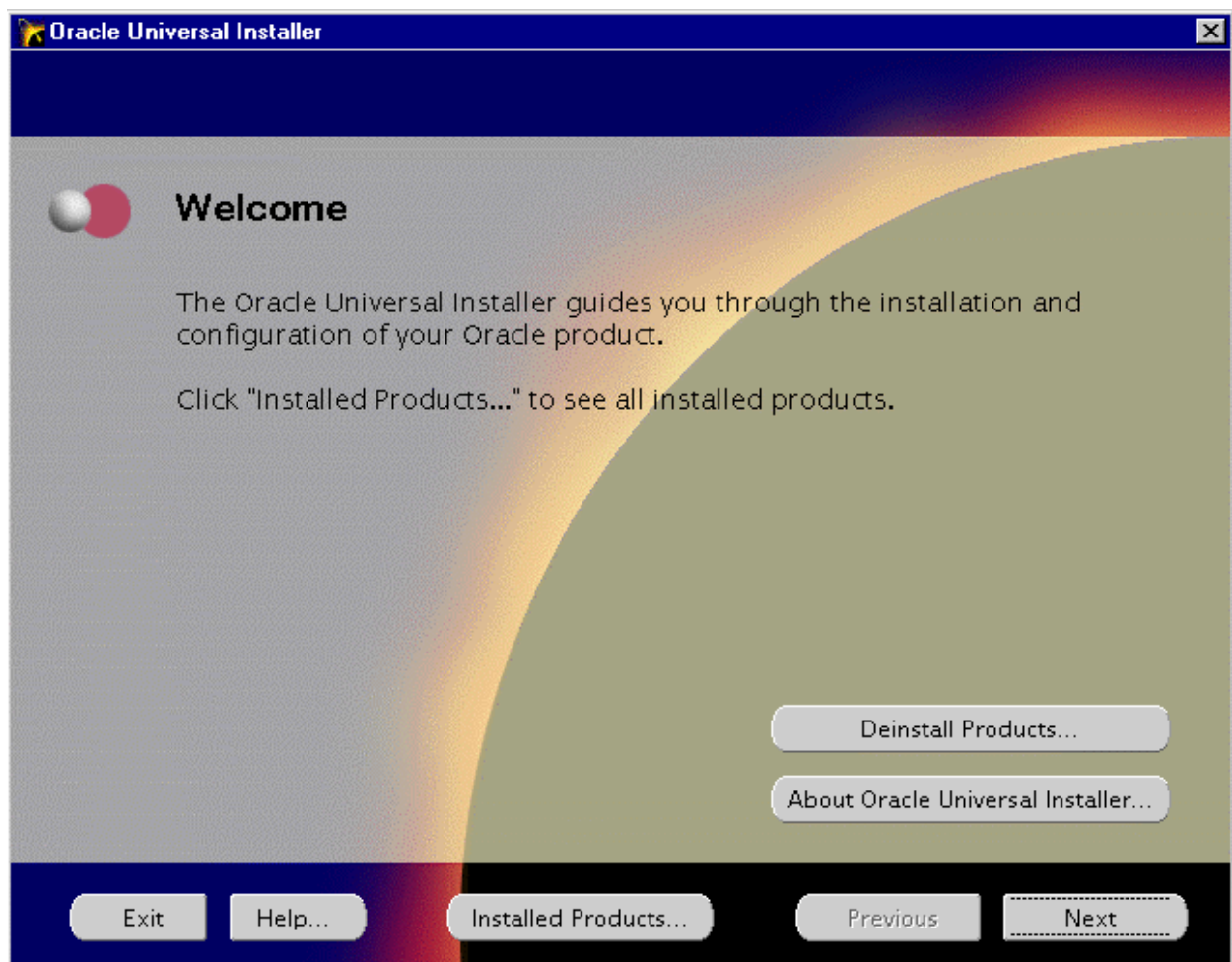
Test Objectives

The test was performed to evaluate the Universal Installer's: usefulness (do the intended users of the installer find it to be useful); ease of use (does the Oracle8 Installation provide an interface that is easy to use for both Oracle and non-Oracle users); and functionality (do the installation tools contain all the features/functions the intended users require). The test specifically focused on the custom and typical enterprise server installation, as well as the installation of the DBA client.

Method
Participants

Four Oracle database administrators (DBAs) and 4 NT system administrators (NT SysAdmins) without Oracle experience participated in this study. All participants were required to have the following characteristics to participate:

a)   speaks English fluently;

b)   willing to sign a non-disclosure agreement;

c)   did not work for a company that produces software;

d)   were either Oracle DBAs or NT SysAdmins.

Participants were selected from an in-house subject database developed from subscription lists of database and computing journals. There is no reason to believe that there were any significant differences between the participant sample and the user population. These were not the same participants as those used in the first study on Oracle Installer for Oracle 8.0.3.

Participant Profiles

Participant	Job Function	Company	Years Experience
1	DBA	A	X
2	DBA	B	X
3	DBA	C	X
4	DBA	D	X
5	NT System Admin	E	X
6	NT System Admin	F	X
7	NT System Admin	G	X
8	NT System Admin	H	X

Participant Self Ratings on Level of Experience (on 1-low to 7-high scale)

Participant	Oracle DBA Tasks	Oracle8	Windows NT
1	X	X	X
2	X	X	X
3	X	X	X
4	X	X	X
5	X	X	X
6	X	X	X
7	X	X	X
8	X	X	X

Context of Product Use in the Test
Tasks

Description of task scenarios: An explanatory scenario was given verbally. The participants were simply given a list of tasks. Each participant completed three tasks:

Typical installation of the database server.

Custom installation of the database server.

Database administration client.

Why tasks were selected:

- The typical install task was selected to test the ease of use of the simplest installation. This task was selected because when users evaluate the Oracle database, they usually do a standard install on a clean system. If this process is not very easy, users may evaluate the overall usability of the database as low.
- The custom install task was selected to test the ease of use of the custom installation. The custom installation covered the case where there were pre-existing Oracle products on the server that had to be uninstalled prior to the new database installation.
- The client install task was selected to test the ease of use of installing just the database administration client. Customers frequently administer databases from remote workstations which would require the Client to be installed. We wanted to ensure that the installation of this database client was easy.

Task source: The source of the tasks was collaboration between product development, product management and usability. The key consideration was the development of realistic tasks that represented customer use of the product.

Task data given to participants: The task data were given to participants for all sub-tasks on an instruction sheet which can be seen in Appendix B. Briefly, they were given…

- the name of the new Oracle home (**NewOracleHome**) and the path (**D:\Oracle\Ora81new**)
- a list of the software they needed to install (**Oracle8 Server, Oracle Net8 Products, and Oracle Configuration Assistants**).
- the name of another new Oracle home (**NewOracleHome**) and the path (C**:\Oracle\Ora81new**) for the typical installation.

Task performance criteria: The tasks were considered successful when the user received the "Database has been installed successfully" message at the end of the installation.

Test Facility

Intended context of use: Any workplace where computers need to have a database installed. These settings include office, manufacturing, or other workplace environments where standard desktop computer workstations may be used.

Context used for the test: The usability test was conducted in the Oracle Headquarters' Usability Lab. The user room was configured to represent a closed computer room with a desktop computer. Participants worked alone and were asked to think aloud while doing the tasks. Participants were observed through a one way mirror, and by video cameras. A usability engineer tracked the participants' actions and comments using data logging software.

Participant's Computing Environment

Intended context of use: Oracle Universal Installer NT version 8.1.3 is intended for use on any PC running Windows NT.

Context used for the test: The test was run using 2 PCs. The "typical" install task was performed on one PC and the "custom" install and DBA client install tasks were performed on the second PC. Both PCs ran the Windows NT 4.0 operating system. Both PCs had 128 MB RAM and 400 MHz processors. One CD ROM disk that contained the installation software was given to each participant. The display units were 17" ViewSonic monitors set at 1024

by 768 pixels, true color, and a 75 Hertz refresh rate. It should be noted that the hardware systems were upgraded since the first study on Oracle Installer for Oracle 8.0.3.

Test Administrator Tools

Tasks were timed using the Oracle Usability Logger, designed especially for Oracle's usability tests. Sessions were videotaped (a combined picture of the screen and a view of the participant), although information derived from the videotapes is not part of this report. At the end of the sessions, participants completed a subjective questionnaire that was specifically designed to address issues with the Oracle Universal Installer NT version 8.1.3. Subjects rated the questionnaire on a 7-point Likert scale and made a verbal comment for the record.

Experimental Design

The experimental design was a 2 x 2 x 3 mixed factors design, with user and order factored between groups, and the installation task factored within groups. No control variables (other than user type and order) were accounted for either experimentally or statistically. The tasks were counterbalanced so that there were two orders. Complete counterbalancing was not done because two of the tasks needed to occur in chronological order. Thus, order was Typical-Custom-Client; and Custom-Client-Typical (Appendix B shows the former order only). Several additional control or individual difference variables were recorded, such as users' experience with operating systems and database environments; however differences between groups were not analyzed.

There were 5 dependent variables: unassisted task completion rate; time to task completion; number of errors; number of assists; and on-line help accesses. A subjective questionnaire was also administered at the end to obtain qualitative satisfaction data.

Procedure

Upon arrival, participants were reminded that they had been asked to participate in a usability evaluation of Oracle Universal Installer NT version 8.1.3, and that by participating in this evaluation, they would help us make our software easier to learn and use. They were told this evaluation was designed to test the software, and was not intended to test their individual performance in any way. They were told that they would work on some typical database applications tasks, and that while they worked their interactions be videotaped and recorded. Confidentiality was assured and they were allowed to take breaks if needed at any time. Next, participants were asked to sign a video tape release form and a confidential non-disclosure agreement. They were then asked to confirm the information they had provided about themselves before participating: Job function, years experience, company name, database experience (7-point Likert scale), and operating system experience (7-point Likert scale).

The test administrator erased the hard drive and rebuilt the operating system prior to each individual test. Participants were tested one at a time. The test administrator stayed in the control/observer room and interacted remotely with the participant through microphones and speakers.   There were no time limits on the tasks. After the last task, participants were asked to complete a subjective questionnaire. Finally, they were given $XXX for their participation.

Participant Task Instructions

Please see Appendix B for the full list of instructions.

Summary of Typical Install Task: Complete a typical installation of the database server.

Summary of Custom Install Task: Complete a custom installation of the database server.

Summary of Client Install Task: Complete a database administration client installation.

Usability Metrics
Effectiveness

Completion Rate: Unassisted completion rate was defined as the percentage of participants who completed each sub-task correctly without any assistance from the test administrator. Assisted completion rate was defined as the percentage of participants who completed each sub-task correctly with test administrator intervention.

Errors: An error was defined as an action taken that could not lead to the completion of the task, including omissions and commissions.

Assists: An assist was defined as verbal help given by the test administrator to guide the participant to the next step in completing the task. Users were given three tries to complete the task before an assist was given. There was no set limit on assists once they were given, unless the user timed out of the task.

Efficiency

Task time: The amount of time to complete the task.

Completion rate efficiency: Mean assisted completion rate/mean task time.

Number of references to the manual: number of separate references made to the online manual. Hard copy documentation was not available for this test.

Satisfaction

At the end of the sessions, participants completed a subjective questionnaire that was specifically designed to address issues with the Oracle Universal Installer NT version 8.1.3. Subjects rated the questionnaire on a 7-point Likert scale and made a verbal comment for the record. Scores were given for each participant's perception of: usefulness; appearance; ease of use; and clarity and understandability.

Results
Data Analysis

Data Scoring: Participant behavior was categorized into groups of successes, errors, verbal assists and accessing on-line help. Behaviors were marked as errors when a participant made an action that could not lead to the completion of the task. Likewise, for each time the test administrator felt that the participant needed verbal assistance in order to complete the task, and when that assistance was given, the participant received a score for verbal assistance. Each time the participant accessed on-line help for one of the sub-tasks they were scored for accessing on-line help.

Data Reduction: Data for each task was analyzed separately and summarized together. Data were also separated into 2 groups based on user type.

Statistical Analyses: Descriptive statistics used included: means, standard deviations, standard errors of the means, minimum values, and maximum values. There were no inferential statistical analyses performed.

Presentation of the Results
Performance Results

Of the 8 participants, N6 were able to complete all the tasks without assistance from the test administrator. NTwo users made errors trying to figure out how to start the Universal Installer. Thus, the range of unassisted task completion rate was XX% to XX%. The mean total time to completion was approximately XX minutes. There were/were no differences in performance measures between the different user types. DBAs and NT SysAdmins as a group made the same/different number of errors, needed the same/different number of assists, and took the same/different amounts of time to complete the tasks.

It should be noted that the typical install in this study is similar to the first sub-task in the first study of the Oracle installer, Oracle Installer for Oracle 8.0.3. The mean task completion time for the typical install was XX minutes.

Typical Install Performance Results by Participant and User Type

Participant	Assisted Task Completion Rate	Unassisted Task Completion Rate	Total Task Time (h:m:s)	Errors	Assists	Help Access	Efficiency
1	XXXX	XXXX	XXXX	X	X	XXX	XXXX
2	XXXX	XXXX	XXXX	X	X	XXX	XXXX
3	XXXX	XXXX	XXXX	X	X	XXX	XXXX
4	XXXX	XXXX	XXXX	X	X	XXX	XXXX
5	XXXX	XXXX	XXXX	X	X	XXX	XXXX
6	XXXX	XXXX	XXXX	X	X	XXX	XXXX
7	XXXX	XXXX	XXXX	X	X	XXX	XXXX
8	XXXX	XXXX	XXXX	X	X	XXX	XXXX
DBAs	XXXX	XXXX	XXXX	X	X	XXX	XXXX
NT SysAdmin	XXXX	XXXX	XXXX	X	X	XXX	XXXX
Mean	XXXX	XXXX	XXXX	X	X	XXX	XXXX
Standard Dev	XXXX	XXXX	XXXX	X	X	XXX	XXXX
Stand Error	XXXX	XXXX	XXXX	X	X	XXX	XXXX
Min	XXXX	XXXX	XXXX	X	X	XXX	XXXX
Max	XXXX	XXXX	XXXX	X	X	XXX	XXXX

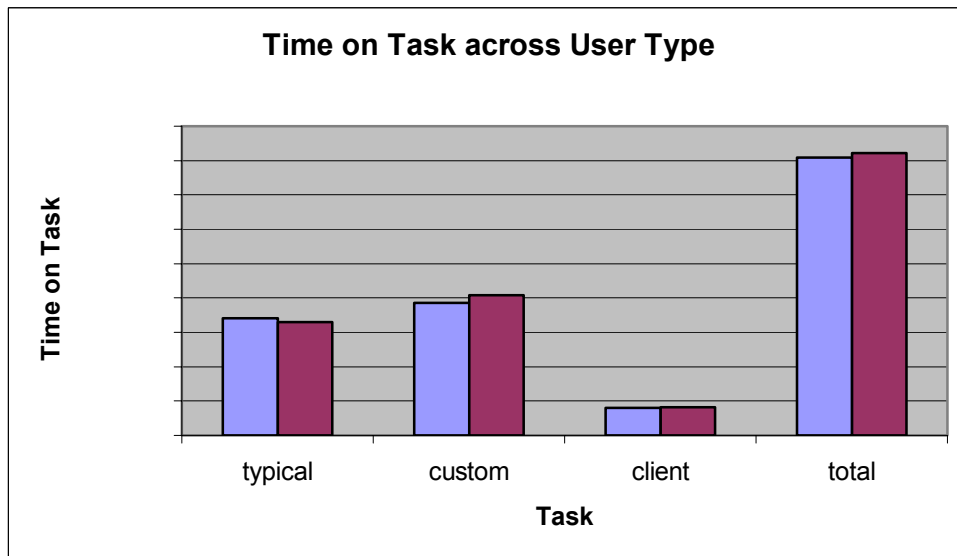### Custom Install Performance Results by Participant and User Type

Participant	Assisted Task Completion Rate	Unassisted Task Completion Rate	Total Task Time (h:m:s)	Errors	Assists	Help Access	Efficiency
1	XXXX	XXXX	XXXXXXX	X	X	XXX	XXXX
2	XXXX	XXXX	XXXXXXX	X	X	XXX	XXXX
3	XXXX	XXXX	XXXXXXX	X	X	XXX	XXXX
4	XXXX	XXXX	XXXXXXX	X	X	XXX	XXXX
5	XXXX	XXXX	XXXXXXX	X	X	XXX	XXXX
6	XXXX	XXXX	XXXXXXX	X	X	XXX	XXXX
7	XXXX	XXXX	XXXXXXX	X	X	XXX	XXXX
8	XXXX	XXXX	XXXXXXX	X	X	XXX	XXXX
DBAs	XXXX	XXXX	XXXXXXX	X	X	XXX	XXXX
NT SysAdmin	XXXX	XXXX	XXXXXXX	X	X	XXX	XXXX
Mean	XXXX	XXXX	XXXXXXX	X	X	XXX	XXXX
Standard Dev	XXXX	XXXX	XXXXXXX	X	X	XXX	XXXX
Stand Error	XXXX	XXXX	XXXXXXX	X	X	XXX	XXXX
Min	XXXX	XXXX	XXXXXXX	X	X	XXX	XXXX
Max	XXXX	XXXX	XXXXXXX	X	X	XXX	XXXX

Client Install Performance Results by Participant and User Type

Participant	Assisted Task Completion Rate	Unassisted Task Completion Rate	Total Task Time (h:m:s)	Errors	Assists	Help Access	Efficiency
1	XXXX	XXXX	XXXXX	X	X	XXX	XXXXXX
2	XXXX	XXXX	XXXXX	X	X	XXX	XXXXXX
3	XXXX	XXXX	XXXXX	X	X	XXX	XXXXXX
4	XXXX	XXXX	XXXXX	X	X	XXX	XXXXXX
5	XXXX	XXXX	XXXXX	X	X	XXX	XXXXXX
6	XXXX	XXXX	XXXXX	X	X	XXX	XXXXXX
7	XXXX	XXXX	XXXXX	X	X	XXX	XXXXXX
8	XXXX	XXXX	XXXXX	X	X	XXX	XXXXXX
DBAs	XXXX	XXXX	XXXXX	X	X	XXX	XXXXXX
NT SysAdmin	XXXX	XXXX	XXXXX	X	X	XXX	XXXXXX
Mean	XXXX	XXXX	XXXXX	X	X	XXX	XXXXXX
Standard Dev	XXXX	XXXX	XXXXX	X	X	XXX	XXXXXX
Stand Error	XXXX	XXXX	XXXXX	X	X	XXX	XXXXXX
Min	XXXX	XXXX	XXXXX	X	X	XXX	XXXXXX
Max	*XXXX*	*XXXX*	*XXXXX*	*X*	*X*	*XXX*	*XXXXX*

Summary Performance Results by Participant and User Type

| Participant	Assisted Task Completion Rate	Unassisted Task Completion Rate	Total Task Time (h:m:s)	Errors	Assists	Help Access	Efficiency
1	XXXX	XXXX	XXXXX	X	X	X	XXXXX
2	XXXX	XXXX	XXXXX	X	X	X	XXXXX
3	XXXX	XXXX	XXXXX	X	X	X	XXXXX
4	XXXX	XXXX	XXXXX	X	X	X	XXXXX
5	XXXX	XXXX	XXXXX	X	X	X	XXXXX
6	XXXX	XXXX	XXXXX	X	X	X	XXXXX
7	XXXX	XXXX	XXXXX	X	X	X	XXXXX
8	XXXX	XXXX	XXXXX	X	X	X	XXXXX
DBAs	XXXX	XXXX	XXXXX	X	X	X	XXXXX
NT SysAdmin	XXXX	XXXX	XXXXX	X	X	X	XXXXX
Mean	XXXX	XXXX	XXXXX	X	X	X	XXXXX
Standard Dev	XXXX	XXXX	XXXXX	X	X	X	XXXXX
Standard Error	XXXX	XXXX	XXXXX	X	X	X	XXXXX
Min	XXXX	XXXX	XXXXX	X	X	X	XXXXX
Max	XXXX	XXXX	XXXXX	X	X	X	XXXXX

Time on Task across User Type

Satisfaction Results

Test participants were asked to rate the Universal Installer on 4 metrics by assigning a score between 1 and 7 to each metric, with 7 being high. The data was collected from 7 of the 8 test participants. The eighth participant's data was not included due to a data collection error.

Summary Satisfaction Results by Participant

Participant Number	Ease of Use	Usefulness	Appearance	Clarity and Understandability
1	X	X	X	X
2	X	X	X	X
3	X	X	X	X
4	X	X	X	X
5	X	X	X	X
6	X	X	X	X
7	X	X	X	X
Mean	XXX	XXX	XXX	XXX
Standard Dev	XXX	XXX	XXX	XXX
Standard Error	XXXXXX	XXXXXX	XXXXXX	XXXXXX
Min	X	X	X	X
Max	x	x	x	x

Oracle Universal Installer Subjective Usability Metrics

Composite Usability Metrics: Easy, Useful, Attractive, Clear and Understandable
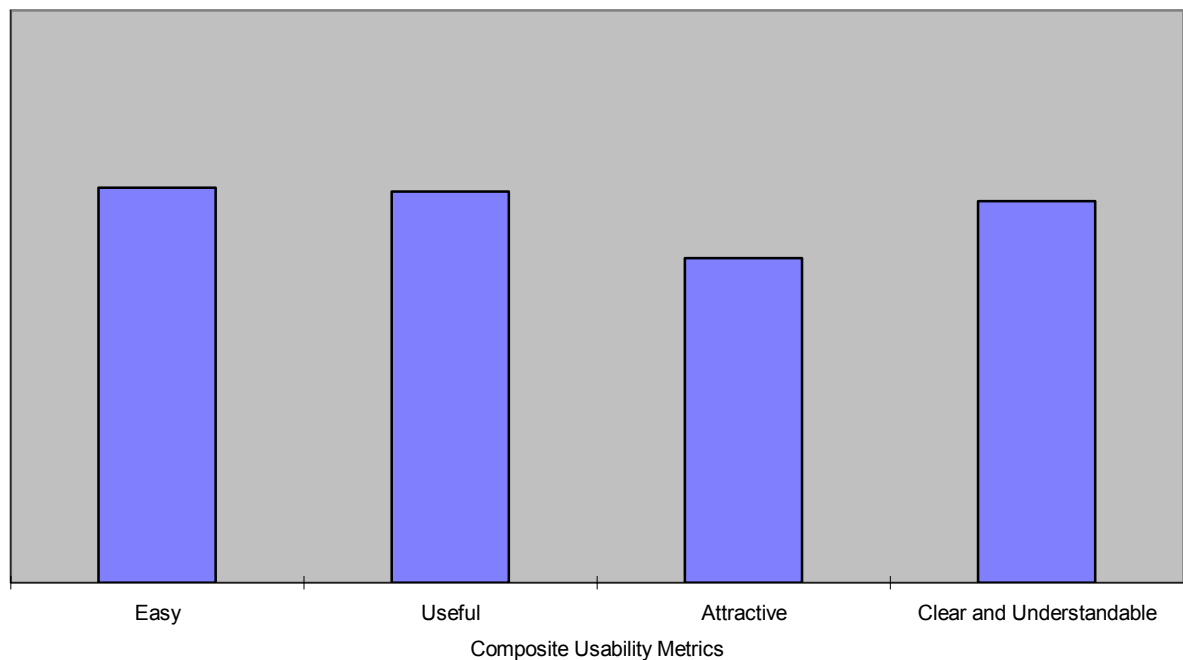
Appendix A
Universal Installer Post-Test Questionnaire

General

What was your general impression of the new Oracle 8.1.3 installation?

Visual Design

What did you think of the overall look of the screen? _____ liked _____did not like.

Would you say this installation product looked ____better or ____ worse or ____ the same as other installation products you have used?

Was the text readable on the screen? ____yes ____no

Interaction Design

Did you understand immediately what was on the screen and how to use it? _____yes _____no

Was the installation product well organized? ____yes ____no (where not?)

Did you get enough feedback on the status of the actions you had taken? ____yes ____no

Was the response time satisfactory? ____yes ___no

How easy were these tasks to do the first time?:

    A. De-installing a product   Very difficult 1 2 3 4 5 6 7 Very easy

    B. Typical server installation   Very difficult 1 2 3 4 5 6 7 Very easy

    C. Custom server installation   Very difficult 1 2 3 4 5 6 7 Very easy

    D. DBA client installation   Very difficult 1 2 3 4 5 6 7 Very easy

    E. Database Configuration Assistant   Very difficult 1 2 3 4 5 6 7 Very easy

What did you think about the control behavior when picking items to install or de-install? Did it make sense?

If the top level was checked what did you assume about the items listed under it?

Terminology

What terminology in this installation product was new to you? Was it clear at first how those terms were being used? Would there have been better terms to use?

What did you think of the wording with in the interface? Was is clear, specific enough, etc.?

Were the error dialogs informative?

Ease of Use

What part of the installation did you find easy? Why?

What part of the installation did you find difficult? Why?

Features

Which features of the installation do you think are the most useful? Why?

What features or functions did you find confusing or difficult to understand?

What would you do to make them easier to understand?

Do you feel this installation product would meet your needs _____yes _____no ___don't know?

What extra features you think this installation product should have?

Overall Appeal
      A.   How useful did you find this product?   Not at all 1 2 3 4 5 6 7 Very useful

      B.   How attractive is its on-screen appearance?   Not at all 1 2 3 4 5 6 7 Very attractive

      C.   How easy is it to use?   Very difficult 1 2 3 4 5 6 7 Very easy

      D.   How clear and understandable is it?   Not at all clear 1 2 3 4 5 6 7 Very clear

      E.   How easy was it for you to accomplish your tasks? Very difficult 1 2 3 4 5 6 7 Very easy

Would you want to use this software in your work? ____ yes _____ no. If no, why not?

What did you think of the disk space allocation interface? Did you notice it?

During a "typical" install do you think the documentation should be automatically installed as a default. Why or why not? Do you want to have control over where the documentation is put on the disk?

Are you familiar with the product "Install Shield"? How does this Oracle installation product compare to the Install Shield product?

Have you used any other installation products? How does the Oracle installation product compare to those?

Additional Comments:

Appendix B

Test Tasks

Oracle Universal Installer NT installation

Typical) You will install Oracle 8.1 onto machine uiclass1-pc. This machine runs a Windows NT 4.0 operating system.

You will create a new Oracle Home named: NewOracleHome

 at path: C:\Oracle\Ora81new

Install the Server using the "typical" installation option.

You will be creating a new database.

Custom) You will install Oracle 8.1 onto the machine uiclass2-pc. This machine runs a Windows NT 4.0 server operating system.
Before you begin installing the new software, check if there are any Oracle products currently on the machine. If there are, you will need to de-install them before installing the new software.

You will create a new Oracle Home named: NewOracleHome

 at path: D:\Oracle\Ora81new

Install the Server using the option that lets you customize the install by giving you full control over every aspect of the installation.

Install **only** the following:

> Oracle8 Server,
>
> Oracle Net8 Products,
>
> Oracle Configuration Assistants.

You will create a new database. Do not migrate any databases.

You will be using TCP/IP protocol.

When creating the database choose the "typical" database and "copy existing files from CD" options.

Client) Install the Database Administrator client onto machine uiclass2-pc. This client will include products needed to administer a database from this machine. Use the same Oracle Home as you did in the last task.

Appendix C

Release Notes

The 8.1.3 usability test revealed design issues with the custom install checkbox… selection mechanism shown below (Figure 1). The current version of this mechanism places too much information in the navigator tree to be easily interpretable. It requires the use of italics and grayed out checkboxes to show states such as availability and installation status. …
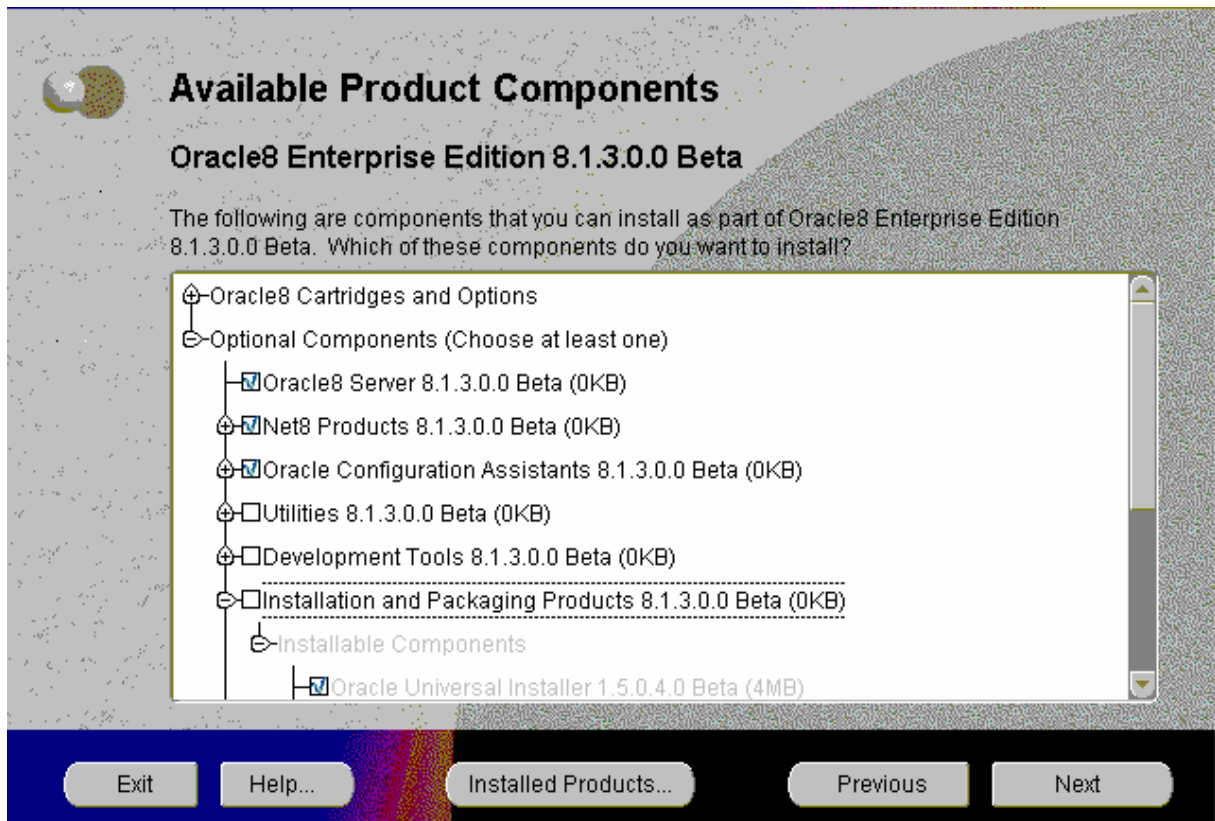

Figure 1: 8.1.3 Custom … Selection Screen

As part of the preparation for the 8.1.6 release, the usability group created and iteratively tested two alternative prototypes for custom install… selection mechanisms. These prototypes included variations of the use of icons …with a legend bar (Figure 2).
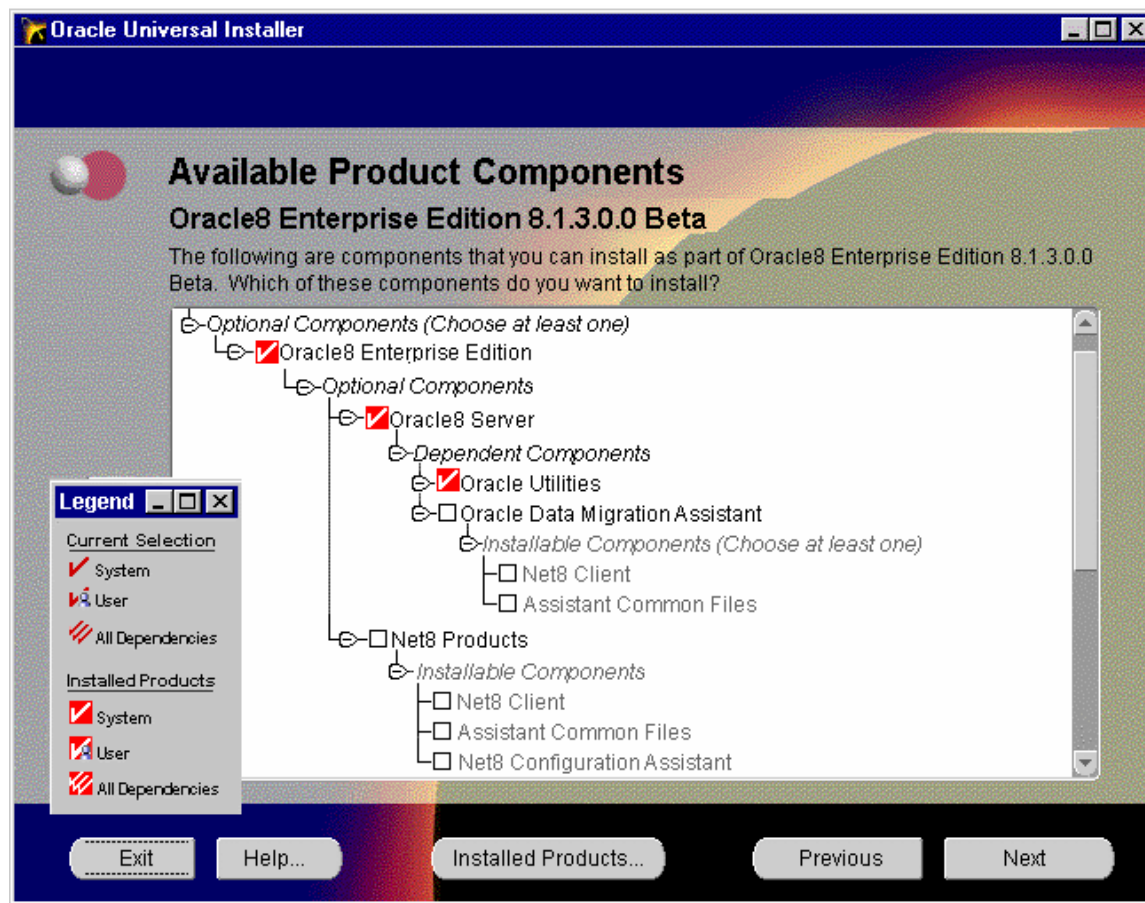
Figure 2: Rejected design alternative for the custom selection screen.

However the usability walkthroughs on these products indicated that these design solutions were ultimately not effective. These designs tried to include too much information in individual icons and were not easily interpretable.... The best solution ended up being the creation of a new control type called the …HGRID control…. The HGRID … control (Figure 3) allows for the presentation of detailed install information in a separate column from the tree control itself. This allows for the preservation of a tree hierarchy while having a second column that can be used for more detailed installation data.….
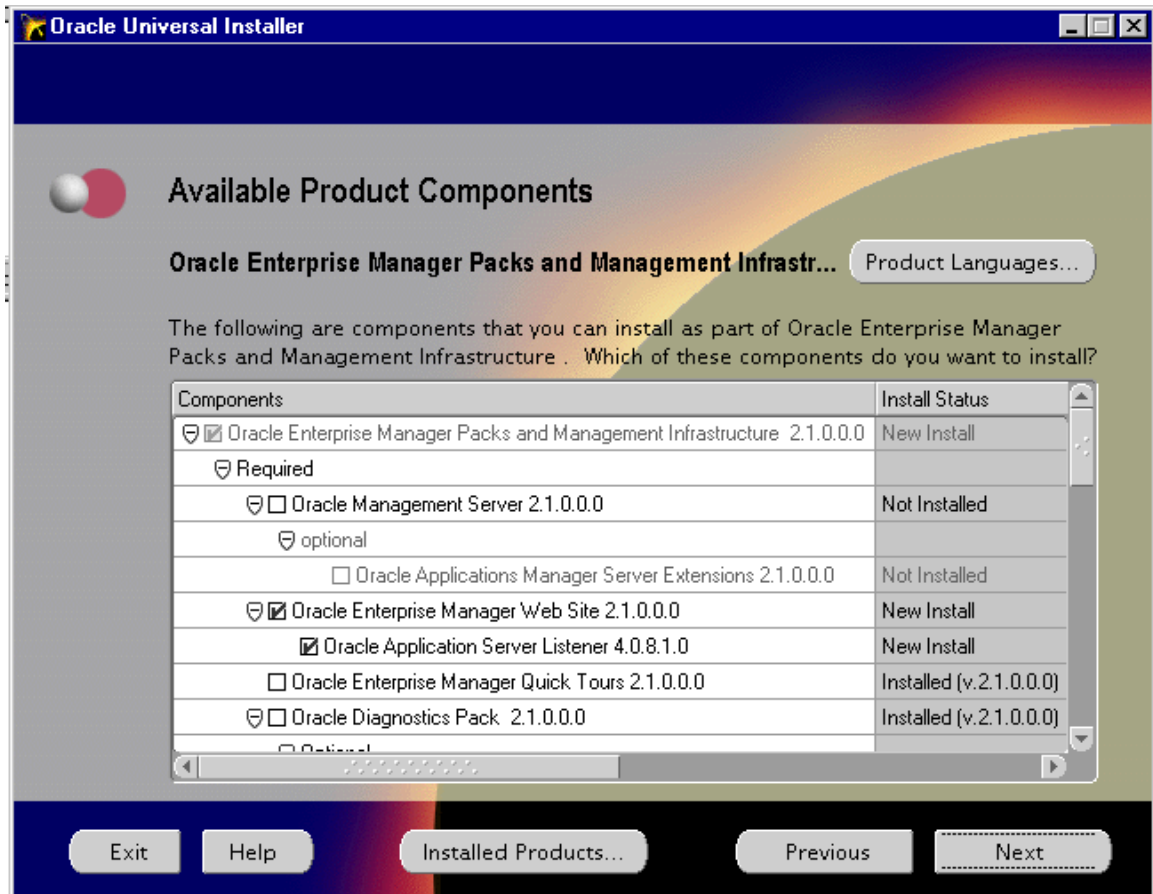
Figure 3: New version of the installer with the HGRID … control to be tested in Oracle 8.1.6 usability test.

Oracle 8.1.6 is shipping with the new version of the custom screen that includes the HGRID … control. There are several steps that the usability team plans for improving the user interface of the Oracle Universal Installer for Oracle 8.1.7 and Oracle 8.2. These include:

1. Review the feedback from Boeing on current issues with the Oracle Universal Installer for Oracle 8.
2. Make a site visit to Boeing and other customers of the Oracle Universal Installer for Oracle 8 to gather detailed user requirements and usability feedback from DBA's.
3. Conduct usability test on the Oracle Universal Installer for Oracle 8.1.6. The tasks for this test will focus on complex custom installation scenarios (e.g. Installing Oracle 8.1.6 on machine with pre-existing Oracle database or other applications). Invite Boeing DBA's to participate in this study.

Bibliography

[1]   American Psychological Association. *Ethical Principles in the Conduct of Research with Human Participants*. 1982.

[2]   Brooke, J. (1996). SUS: A "quick and dirty" usability scale, In Jordan, P. W., Thomas, B. T., and Weerdmeester, B. A. (Eds.), *Usability Evaluation in Industry*. UK: Taylor and Francis, pp. 189–194.

[3]   Chin, J. P., Diehl, V. A., and Norman, K. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. In: *CHI '88. Conference Proceedings on Human Factors in Computing Systems.* New York: ACM, pp. 213–218.

[4]   Dumas, J., and Redish, G. (1999). *A Practical Guide to Usability Testing.* UK: Intellect.

[5]   Kirakowski, J. (1996). The software usability measurement inventory: Background and usage. In Jordan, P., Thomas, B., and Weerdmeester, B. (Eds.), *Usability Evaluation in Industry*. UK: Taylor and Francis, pp. 169–177.

[6]   Lewis, J. R. (1991). Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the ASQ. *SIGCHI Bulletin*, 23(1), 78–81.

[7]   Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, 7, 57–78.

[8]   Macleod, M., Bowden, R., Bevan, N. and Curson, I. (1997). The MUSiC Performance Measurement method. *Behaviour and Information Technology*, 16, 279–293.

[9]   Nielsen, J., and Landauer, T. K. (1993). A mathematical model of the finding of usability problems. In: *CHI '93. Conference Proceedings on Human Factors in Computing Systems*. New York: ACM, pp. 206–213.

[10]   Norman, D.A. (1983). Design rules based on analyses of human error. *Communications of the ACM*, 26(4), 254–258.

[11]   Rubin, J. (1994). Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests. New York: John Wiley and Sons, Inc.

ICS 35.080

Price based on 46 pages