DRAFT FOR DEVELOPMENT

DD ISO/PAS
20282-3:2007

Ease of operation of everyday products —

Part 3: Test method for consumer products

ICS 13.180

BSI
British Standards

National foreword

This Draft for Development is the UK implementation of ISO/PAS 20282-3:2007.

This publication is not to be regarded as a British Standard.

It is being issued in the Draft for Development series of publications and is of a provisional nature. It should be applied on this provisional basis, so that information and experience of its practical application can be obtained.

A PAS is a Technical Specification not fulfilling the requirements for a standard, but made available to the public and established in an organization operating under a given procedure.

A review of this Draft for Development will be carried out not later than three years after its publication.

Notification of the start of the review period, with a request for the submission of comments from users of this Draft for Development, will be made in an announcement in the appropriate issue of *Update Standards*. According to the replies received, the responsible BSI Committee will judge whether the validity of the PAS should be extended for a further three years or what other action should be taken and pass their comments on to the relevant international committee.

Observations which it is felt should receive attention before the official call for comments will be welcomed. These should be sent to the Secretary of the responsible BSI Technical Committee at British Standards House, 389 Chiswick High Road, London W4 4AL.

The UK participation in its preparation was entrusted to Technical Committee PH/9, Applied ergonomics.

A list of organizations represented on this committee can be obtained on request to its secretary.

This publication does not purport to include all the necessary provisions of a contract. Users are responsible for its correct application.

This Draft for Development was published under the authority of the Standards Policy and Strategy Committee on 31 January 2008

ISBN 978 0 580 57049 0

Amendments/corrigenda issued since publication

Date	Comments

DD ISO/PAS 20282-3:2007

PUBLICLY AVAILABLE SPECIFICATION

ISO/PAS 20282-3

First edition
2007-12-15

Ease of operation of everyday products —

Part 3: Test method for consumer products

Facilité d'emploi des produits quotidiens —

Partie 3: Méthode d'essai pour produits de consommation courante

Reference number
ISO/PAS 20282-3:2007(E)

=:

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

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

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

In other circumstances, particularly when there is an urgent market requirement for such documents, a technical committee may decide to publish other types of document:

— an ISO Publicly Available Specification (ISO/PAS) represents an agreement between technical experts in an ISO working group and is accepted for publication if it is approved by more than 50 % of the members of the parent committee casting a vote;

— an ISO Technical Specification (ISO/TS) represents an agreement between the members of a technical committee and is accepted for publication if it is approved by 2/3 of the members of the committee casting a vote.

An ISO/PAS or ISO/TS is reviewed after three years in order to decide whether it will be confirmed for a further three years, revised to become an International Standard, or withdrawn. If the ISO/PAS or ISO/TS is confirmed, it is reviewed again after a further three years, at which time it must either be transformed into an International Standard or be withdrawn.

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

ISO/PAS 20282-3 was prepared by Technical Committee ISO/TC 159, *Ergonomics*, Subcommittee SC 1, *Ergonomic guiding principles*.

ISO 20282 consists of the following parts, under the general title *Ease of operation of everyday products*:

— *Part 1: Design requirements for context of use and user characteristics*

— *Part 2: Test method for walk-up-and-use products* [Technical Specification]

— *Part 3: Test method for consumer products* [Publicly Available Specification]

— *Part 4: Test method for the installation of consumer products* [Publicly Available Specification]

Introduction

Many people find everyday products such as ticket machines, washing machines and videorecorders difficult to use, particularly when using them for the first time or infrequently. If the functions provided by a product to support the users' main goals are not easy to operate, many users will find the product difficult, if not impossible, to use. This is clearly not desirable, either for the suppliers of such products or for the users. Information about the ease of operation of a product would therefore be of great value to both suppliers as part of their development process, and to potential purchasers making purchase decisions or comparing alternative products. This would provide an incentive for the production of products that are easier to install and use, and enable potential purchasers to pay specific attention to ease of operation when selecting a product to buy.

This part of ISO 20282 specifies a test method that can be used to provide an operational evaluation of the ease of operation of consumer products. The test method is a summative method that gives performance-based measures that can be used for assessment against predetermined criteria or as the basis for comparisons between different products. Thus the test method is an example of a "performance-related" usability method (see ISO/TR 16982) that can be used to measure ease of operation and establish whether quantitative usability requirements for ease of operation have been achieved.

Manufacturers can use the test method presented here to test whether requirements for ease of operation have been met, or to compare their products with previous versions, or with competitor products. They could communicate the test results to potential purchasers in product descriptions or advertising. Corporate purchasers can use the test method to determine whether products meet their needs, while testing organizations could use it as a basis for providing information to potential users and organizations representing users.

Usability (see ISO 9241-11) is 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. Ease of operation provides a measure of the usability of an everyday product when used by the actual or intended users to achieve their main goals when using the product. It is assumed that users can achieve their intended main goals when the product is operated correctly. The method is intended to be used with products where there is a clear criterion for successfully achieving goals, and not where there is variable quality in the outcome. Ease of operation is primarily concerned with the user interface and the interaction that it supports, rather than the quality of the features that are present within a product or their suitability for the product's intended use.

When the main goals of using a consumer product involve task performance that is fast and of limited complexity, the most important measure of ease of operation is effectiveness. Effectiveness of operation is measured as the percentage of users who can successfully achieve the main goals of use of the product. Efficiency of operation may be important, for example, if a product is to be used by large numbers of users in rapid succession. In addition, for some products to be identified as easy to operate, it is important that users are satisfied with their experience of operation, for example, where users have discretion over whether to use the specified product or not and they can readily choose some alternative means of achieving their goals.

ISO 13407 provides guidance on the human-centred design principles and design activities to be applied throughout development in order to produce usable products. It stresses that usability requirements should be specified prior to development, and that evaluation is an iterative process during development. The summative test method in this part of ISO 20282 can be used to evaluate pre-specified requirements. Other types of usability methods for formative evaluation are more appropriate when the main objective is to provide feedback in the course of design, although this summative method can also provide information on usability problems that need to be corrected. ISO 13407 also emphasizes the importance of identifying the context of use. ISO 20282-1 describes in more detail sources of variance in user characteristics that form part of the context of use that needs to be taken into account when designing for ease of operation. This information is also needed to identify the requirements for testing in this part of ISO 20282.

This part of ISO 20282 has been published as a Publicly Available Specification so that information and experience of its use in practice can be gathered (see Annex G). Versions of the test method for walk-up-and-use products and for the installation of consumer products are given in ISO/TS 20282-2 and ISO/PAS 20282-4.

Ease of operation of everyday products —

Part 3:
Test method for consumer products

1 Scope

This part of ISO 20282 specifies a test method for measuring the ease of operation of consumer products.

The purpose of the test is to provide a basis for predicting the ease of operation of a consumer product, including measures of its effectiveness and efficiency of operation, and the satisfaction of the intended user population in the intended context of its use.

The intended users of this part of ISO 20282 are people with human factors expertise in the design and management of appropriate tests, including manufacturers, suppliers, purchasing organizations and third parties such as consumer organizations.

2 Conformity

A test method reporting values for ease of operation of a consumer product conforms to this part of ISO 20282 if the method used conforms to the applicable requirements in Clauses 7, 8 and 9 and Annexes C, D and E. To provide evidence of conformance, a full report of the results shall be produced using the format specified in Annex F.

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 20282-1:2006, *Ease of operation of everyday products — Part 1: Design requirements for context of use and user characteristics*

4 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

4.1
actual users
group(s) of people who directly interact with a product

NOTE Before a product is released this is the intended user group, and after release this is based on what is known about the actual user group.

[ISO 20282-1:2006, definition 3.1]

4.2
consumer product
product that is intended to be acquired and used by an individual for personal rather than professional use

[ISO 20282-1:2006, definition 3.2]

4.3
context of evaluation
users, tasks, equipment (hardware, software and materials), and physical and social environments in which a product is evaluated

[ISO/TS 20282-2:2006, definition 4.3]

4.4
context of use
users, tasks, equipment (hardware, software and materials), and physical and social environments in which a product is used

[ISO 9241-11:1998, definition 3.5]

4.5
ease of operation
usability of the user interface of an everyday product when used by the intended users to achieve the main goal(s) supported by the product

NOTE 1 Ease of operation is a specific aspect of usability as defined in ISO 9241-11, which in this case is applied to the operation of everyday products. Ease of operation assumes that the functionality of the product other than the user interface operates correctly.

NOTE 2 Ease of operation is measured as effectiveness of operation, optionally including efficiency of operation and satisfaction with operation.

[ISO 20282-1:2006, definition 3.4]

4.6
effectiveness
accuracy and completeness with which users achieve specified goals

[ISO 9241-11:1998, definition 3.2]

NOTE For the purpose of this test method, effectiveness of operation is measured by the percentage of users who achieve the main goal(s) of use of a product accurately and completely. Measures of effectiveness of operation are based on the end result independently of whether the goal is achieved in the most efficient way.

4.7
effectiveness of operation
percentage of users who achieve the main goal(s) of use of a product accurately and completely

NOTE Measures of effectiveness of operation are based on success in achieving the end result independently of whether the goal is achieved in the most efficient way.

[ISO 20282-1:2006, definition 3.6]

4.8
efficiency
resources expended in relation to the accuracy and completeness with which users achieve goals

[ISO 9241-11:1998, definition 3.3]

NOTE For the purpose of this test method, efficiency of operation is measured as the time taken to achieve the main goal(s).

4.9
efficiency of operation
time taken to achieve the main goal(s)

NOTE This identifies a specific resource for efficiency as defined in 4.8.

[ISO 20282-1:2006, definition 3.8]

4.10
everyday product
consumer product or walk-up-and-use product designed for use by members of the general public

NOTE 1 Some products are designed for use by the general public as well as for professional use, but this definition only applies to non-professional use of the product.

NOTE 2 “Everyday” does not imply that the product must be used every day by the user, rather that it is found in everyday life.

[ISO 20282-1:2006, definition 3.9]

4.11
goal
intended outcome

[ISO 9241-11:1998, definition 3.8]

NOTE A goal is stated independently of the functionality used to achieve it.

4.12
intended users
group(s) of people for whom a product is designed

NOTE 1 Adapted from ISO 9241-9:2000, definition 3.4.6.

NOTE 2 In many cases the actual user population is different from that originally intended by the manufacturer. The intended user group is based on realistic estimations of who the actual users of the product will be.

[ISO 20282-1:2006, definition 3.12]

4.13
interaction
bi-directional information exchange between users and equipment

[IEC/TR 61997:2001, definition 3.4]

NOTE 1 Equipment includes both hardware and software.

NOTE 2 Information exchange may include physical actions, resulting in sensory feedback.

4.14
main goal(s)
most frequent or important outcome(s) that all or a large majority of users want to achieve when using a product

[ISO 20282-1:2006, definition 3.14]

NOTE Examples of main goals are given in Annex A.

4.15
satisfaction
freedom from discomfort, and positive attitudes towards the use of the product

[ISO 9241-11:1998, definition 3.4]

NOTE For the purposes of this test method, satisfaction with operation is measured by the attitude towards the operation of the product.

4.16
satisfaction with operation
measures of attitude towards the operation of the product user interface

[ISO 20282-1:2006, definition 3.16]

4.17
task
activities required to achieve a goal

NOTE These activities can be physical and/or cognitive.

[ISO 9241-11:1998, definition 3.9]

4.18
usability
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

[ISO 9241-11:1998, definition 3.1]

4.19
user
person who interacts with the product

[ISO 9241-11:1998, definition 3.7]

4.20
user characteristics
attributes of a user that can influence usability

[ISO 20282-1:2006, definition 3.20]

4.21
user interface
elements of a product used to control it and receive information about its status, and the interaction that enables the user to use it for its intended purpose

EXAMPLE The user interface of a shower tap is the water control lever, where the movement of the lever controls the temperature of the water and the position of the lever communicates the temperature to the user.

NOTE A list of operating instructions permanently displayed on the product is part of the user interface.

[ISO 20282-1:2006, definition 3.21]

4.22
user test group
group of persons selected to participate in a test of usability, sampled according to specific requirements

[ISO/TS 20282-2:2006, definition 4.22]

4.23
walk-up-and-use product
everyday product that provides a service to the general public

NOTE This includes products intended for use by the general public in commercial premises such as in a shop or hotel.

[ISO 20282-1:2006, definition 3.22]

5 Purpose of testing ease of operation

The purpose of the test method is to evaluate whether ease of operation objectives have been met, or to compare products with previous versions or competitor products.

NOTE 1 This method is complementary to test methods for formative evaluation that use a smaller number of participants (see Annex B).

The test method may be used by

— a manufacturer to test whether usability requirements have been met and/or to provide evidence of the ease of operation of a product for marketing purposes, or

— a potential corporate purchaser or test organization to compare alternative products.

EXAMPLE A manufacturer wants to demonstrate that a videorecorder is very easy to program. A representative sample of 50 people who have purchased a videorecorder with the intention of programming it to record programmes are recruited. They are selected to be representative in age and education of owners of different brands of videorecorder (as these are expected to be the major factors that influence ease of operation). Each session is expected to take a total of 15 min to 20 min, and two videorecorders are set up with televisions for programming. Each person is asked to program the videorecorder to record a specified programme in a TV magazine. It is expected that most people will use the published code for the programme, but direct programming of the channel and time is also acceptable. The test result is given as the measured success rate together with a confidence interval. To be 95 % confident that 80 % of the user population can program the videorecorder, 45 of the 50 people tested will have to be successful.

NOTE 2 ISO/TS 20282-2 specifies a similar test method for walk-up-and-use products, and ISO/PAS 20282-4 specifies a similar method for the ease of installation of consumer products.

6 Test procedure

To measure ease of operation, the following steps shall be followed.

a) Identify the product to be tested, see 7.1.

b) Identify the expected context of use (users, tasks and environment), see 7.2.

c) Check that the product is compatible with intended user characteristics, see 7.3.

d) Decide whether to test one or more groups, see 7.4.1.

e) Identify which measures are required, whether there are required values for the measures, or whether two results are being compared, see 7.4.2.

f) Select a group of users that represents the intended user group of the product, see 7.4.4.

g) Establish a test procedure that includes a representative group of users using the products to achieve the main goals of use, see 7.4.7.

h) Measure success rate and, optionally, task time and satisfaction (using a questionnaire), see 7.5.

i) Calculate effectiveness of operation (percentage success rate) and, optionally, efficiency of operation (median task time) and satisfaction with operation (mean questionnaire scores), see Clause 8.

j) Prepare a full report and/or a short summary, see Clause 9.

7 Test method

7.1 Identify product

Identify the specific product and version to be tested.

The product should be installed in its normal configuration.

Consumer products may be tested with separate instructions, if these would typically be available.

7.2 Specify expected context of use

7.2.1 Identify main goals of use of the product

The main goal(s) of use of the product shall be identified based on information provided by the manufacturer. This shall include the most frequent and/or important user goals that the product is intended to support. It is assumed that it is possible for a user to achieve these goals when the product is operated correctly. The goals shall be expressed in terms of the intended outcome of the task activity expressed independently of the means by which it is achieved.

Annex A contains a list of typical consumer products, with the main goals of use of each product. These are accompanied by examples of the task activity that are typically associated with using the product to achieve the main goal. If the product being tested is listed in Annex A, a test shall be carried out including the main goal listed for that product. If other main goals are identified for the product being tested the results should be reported separately, see 8.1.

NOTE For many everyday products there will be one self-evident main goal, such as using a telephone to make a phone call.

The criteria for complete and accurate goal achievement shall be specified.

EXAMPLE For a camera, acceptable exposure and focus of a picture.

7.2.2 Identify the tasks

The task activities that need to be carried out in order to use the product to achieve the main goal(s) shall be identified. This could be based on information provided by the manufacturer and/or a representative sample of potential users.

This information is needed to plan the test and analyze the results, but is not included in the task instructions (see 7.4.7).

EXAMPLE 1 Camera: insert film [storage media], switch on, take aim, take picture, take steps to review picture.

EXAMPLE 2 Washing machine: insert clothes, insert detergent, select program, start machine, remove clothes after machine stops, switch off machine.

EXAMPLE 3 Public telephone: activate the phone (lift the speaker), provide payment, activate a number (dial).

7.2.3 Specify intended or actual user groups

The actual users of the product, if it is in use and known, or otherwise the intended user groups of the product, shall be identified, for example, based on information provided by the manufacturer. Some products are aimed at specific segments of the population (e.g. related to income or age group), or could be in a family of related consumer products for different market sectors. The range of each user characteristic listed in ISO 20282-1:2006, Clause 7, expected to have a significant effect on ease of operation, shall be specified.

EXAMPLE The user population of a mobile phone is expected to be 80 % English-speaking, 10 % French-speaking and 10 % speakers of other languages. It is assumed that all the users will have prior experience of using mobile phones to make phone calls.

For those user groups to be included in the testing, the following information is needed.

— Which user characteristics could affect the ease of operation of the product?

— What range of characteristics exists in the intended or actual user group?

— What is the expected distribution of each relevant user characteristic?

a) Physical characteristics

Identify whether the product requires particular physical characteristics (such as body dimensions, strength and biomechanical abilities, visual abilities, auditory abilities, or handedness). Either prior tests of users near the expected limits or the use of existing sources of data on the range of human characteristics may be used to establish the range of people who can use the product.

EXAMPLE Existing data could be used to determine the size of text required for legibility by particular segments of the population.

b) Psychological and social characteristics

It is important to establish what cognitive characteristics or capabilities (such as knowledge and experience, cultural differences, literacy and language) are likely to have a major effect on ease of operation. This may be done by expert assessment, collection of actual usage data, or empirical studies to evaluate the effect of different user characteristics.

c) Special needs

Identify to what extent the product is intended to support people (such as older users) whose physical or psychological characteristics (body dimensions, strength, biomechanical abilities, visual abilities, auditory abilities, handedness, knowledge, experience, culture, literacy or language) are towards the end of the range. Where possible, use existing data.

EXAMPLE 1 The product is intended also to be used by elderly users. The required text size for the display is identified from existing data.

If there is particular interest in whether subgroups or individuals at the end of the ranges of variation in physical or cognitive abilities can use the product, additional evaluation may be carried out for these groups.

EXAMPLE 2 Additional tests are carried out to establish whether the product can be used by people with poor sight, or wheelchair users.

7.2.4 Specify environmental characteristics

The main environments in which the product is (or is intended to be) used shall be identified. Those attributes of the physical or social environment that are likely to have significant effects on the effectiveness with which goals can be achieved shall be specified.

EXAMPLE The camera could be used at night where there might be no other illumination. The environmental temperature range is −5 °C to 40 °C.

7.3 Check that the product is compatible with intended user characteristics

Where inspection of the product indicates that successful attainment of the main goal(s) identified in 7.2.1 will not be possible for some members of the intended user population, it may be appropriate to use existing data to predict how many people will be affected. If the number of people who cannot use the product represents an unacceptable proportion of the intended users, no user-based testing can be carried out.

7.4 Design the test

7.4.1 Decide whether to test one or more groups

The widest possible range of users within the expected user population should be tested, but if any differences in the context of use would have a major impact on the ease of operation, each user group or contextual condition should be tested or analysed separately (see Annex C).

Decide whether to test separately any special groups who have characteristics that might make it difficult to use the product, for example, with a particular disability.

7.4.2 Identify measures

a) Identify which measures are required.

- Measure effectiveness of operation: when the main goals of using an everyday product involve task performance that is fast and of low complexity, the most important measure of ease of operation is effectiveness.
- Establish whether measures of efficiency of operation and/or satisfaction with operation are needed.

b) Identify whether there are required values for the measures, or whether two results are being compared. If so, identify the required statistical confidence (confidence that the result has not occurred by chance).

NOTE The higher the required success rate and the higher the confidence, the more users will be required. In scientific work, it is usual to require at least 95 % confidence in order to be able to draw a conclusion, but for some business purposes, a lower degree of confidence (such as 80 %) might be sufficient.

EXAMPLE 1 The objective is to achieve a success rate of 80 % with 95 % confidence (see the example in Clause 5).

EXAMPLE 2 The objective is to compare the ease of operation of two alternative products, and establish with 80 % confidence whether there are any differences.

7.4.3 Specify test scenarios and conditions

A test scenario shall be specified for each main goal that has been identified (see 7.2.1).

The test scenario shall include a precise statement of the goal to be attained.

EXAMPLE Make a phone call to 01509 123456 using this telephone.

For each test goal, the circumstances(s) associated with successful goal attainment shall be identified, together with a way of monitoring when it has been achieved. Since successful completion is the criterion, alternative methods of achieving the goal(s) (including errors that are successfully corrected) shall be included as positive results, unless the means used have undesirable consequences, such as posing a risk to the well-being of the individual, or causing damage to the product.

To enable task timing, the criteria for task completion shall be specified (typically, a specific action or statement by the user).

A maximum amount of time allowed for successful goal achievement shall be established. This period should be sufficient to allow test persons to try different ways of achieving the goal before they give up. The expected time to complete the task should be estimated from pretests, and the maximum time allowed to users before they are categorized as having failed should be at least three times the expected time.

If there is more than one main goal to be tested, it will be necessary to decide the order in which the tasks are to be undertaken. In general, if there is a normal sequence in which tasks would be undertaken, this should be the order for testing. If no normal sequence exists it will be necessary to avoid testing in one fixed sequence, in order to avoid arbitrary order effects. In this case, the order may be systematically varied, or assigned randomly for each person tested.

7.4.4 Recruit a representative sample of users

A sample representing the intended or actual user groups shall be recruited to take part in the test. A sample of users shall be selected that models the distribution of relevant user characteristics within the user groups (see Annex C). Nobody who takes part in the test shall fall outside the ranges specified for the variation in user characteristics identified in 7.2.3.

Decide whether the characteristics of the intended or actual user groups identified in 7.2.3 divide users into distinct categories that will have a major impact on ease of operation. For example, a global product would probably need separate groups of each nationality.

It is important to decide whether to test first-time users, or people with previous experience. Even with a new product, people could have experience with similar products or a similar previous model, so the type of previous experience may need to be defined. Any relevant background experience of use of similar or related products shall be recorded for each user.

NOTE For products used occasionally, it is appropriate to test first-time use. But for a product primarily used regularly, it may be better to test users after initial practice sessions.

Systematically select a sample of users with the same distribution of values or attributes of the relevant user characteristics as the intended or actual user groups identified in 7.2.3, and record the characteristics of each user.

Many general physical and psychological characteristics (such as strength, visual acuity and cognitive capabilities) can be adequately sampled through demographic variables such as age and education. The main exceptions are relevant previous experience, specialized knowledge or skills and culture.

EXAMPLE An alarm clock requires manual dexterity to operate the controls and good visual acuity to read the instructions. These relevant user characteristics are sampled by recruiting users with a representative distribution of age and educational attainment. Owing to the fact that 5 % of purchasers are expected to have previous experience of this type of alarm clock, and as the task of setting the alarm is expected to be much more difficult for those who do not have this experience, only people without previous experience will be tested.

The number of users selected and the accuracy of sampling will determine the extent to which the results accurately estimate the values for the population sampled (see Annex C).

Test persons shall not be from the testing or manufacturing organization.

7.4.5 Product and equipment

The version of the product that is tested should normally be a standard production model. If testing is being carried out before development has been completed, the version used shall closely resemble the completed final product, and any areas that are different shall be noted. Evaluation shall be conducted with the product being used in conjunction with any other equipment that normally influences its use.

When the test commences, the product should use the default settings or be set in a typical configuration.

Any instructions provided with the product may be made available to the users (see 7.4.7).

7.4.6 Test environment

a) Testing shall be carried out in an environment that resembles as closely as possible the environment in which the product would be used. If a usability laboratory is used, the normal conditions of use shall be simulated. If tested in the field, a controlled set of users and tasks shall be analysed.

   If goal attainment while using the product could be adversely affected by particular environmental conditions, it may be appropriate to carry out the testing in those conditions in order to be assured of high levels of ease of operation at all times.

   EXAMPLE Normally used in cold weather by people wearing gloves.

b) If the product is normally used in a stressful environment, this should be reproduced.

   EXAMPLE A noisy environment or time pressure to take a photo.

7.4.7 Establish test procedure

a) A written script shall be prepared for each task goal, containing instructions that are presented to each user. The script shall include a description of the scenario within which testing is taking place (e.g. taking a photo at a family gathering) and shall state the particular goal and the conditions which apply to it. No hints shall be included on how to achieve the goal or which features to use.

b) Users shall not be given any hints or assistance from the personnel conducting the test, but depending on the purpose of the test, the instructions that come with the product may be provided.

c) Those conducting the test shall note the time taken to achieve each goal. If the user reaches the maximum amount of time allowed without attaining the goal, they shall be asked to stop and, if appropriate, move on to the next goal.

7.5 Measures

7.5.1 Effectiveness of operation

All tests shall measure effectiveness of operation. The effectiveness of operation shall be measured by the percentage of users who achieve the main goal(s) of use of a product accurately and completely. The precise criteria for successful achievement of each goal identified in 7.2.1 shall be used to determine the number of users successfully achieving each goal.

7.5.2 Efficiency of operation

If time to achieve the goal is important, this shall be measured as the time taken to achieve the main goal(s). The time taken to achieve the goal(s) shall be measured from the time the user is asked to start to the time the goal has been achieved.

7.5.3 Satisfaction with operation

If satisfaction with operation is being measured, it shall be measured by the attitude towards the operation of the product user interface with the scale specified in Annex E. If there is more than one main goal, satisfaction with operation should be measured for each main goal and for overall usage.

7.5.4 Other data

The nature of any errors or other difficulties encountered by users may be recorded in order to identify ease of operation problems that can be provided as feedback to design. This could include partial goal achievement.

8 Results

8.1 Main goals

If there is more than one main goal, results shall be given separately for each main goal, together with a result for all the goals.

8.2 Effectiveness of operation

The result(s) calculated shall be the percentage of users successfully achieving the main goal (success rate), together with the confidence intervals, as specified in Annex D. If there is more than one goal, the result shall be reported for the percentage of users successfully achieving each goal.

EXAMPLE 45 out of 50 users achieved the main task goals of use of the videorecorder. 90 % of users were successful and it is estimated with 95 % confidence that at least 80,1 % of the intended users will be successful.

8.3 Efficiency of operation

The result(s) reported shall be the mean or median time taken by users who successfully achieved the specified goal(s), together with the 95 % confidence intervals.

If the values of task time are not approximately normally distributed (which is the case when some times are particularly long), a logarithmic transformation may be used before making a statistical analysis, and the median task time reported.

EXAMPLE The median time taken by those users who successfully bought the correct ticket was 20 s (with a 95 % probability that the true value is between 15 s and 25 s).

8.4 Satisfaction with operation

The result reported shall be the mean satisfaction scores on the satisfaction scale for all users (see Annex E).

9 Report

A full report of the results shall be prepared using the format specified in Annex F.

If a short version of the report is provided, it should include the following information:

a) which user group(s) and goal(s) have been tested;

b) overall effectiveness of operation for (all) the goal(s);

c) if more than one goal, success rate for each goal individually;

d) whether instructions included with the product were provided to the user, and whether they were used;

e) if measured,

- mean or median time taken for each goal, and
- overall satisfaction.

Annex A
(informative)

Product goals

Table A.1 gives the main goals for examples of consumer products.

NOTE It might not be sufficient to only test these main goals if there are other important main goals.

Table A.1 — Examples of different products and their main goal(s)

Consumer product	Main goal(s)	Example tasks	Installation and other notes
Alarm clock	Get woken at a certain time Re-establish silence	Adjust time, set alarm time, switch alarm on, switch alarm off. Switch alarm off.	Insertion of batteries is an installation and servicing task.
Blood pressure measuring device	Measurement of blood pressure	Put on, start measurement, take reading, take off.	Insertion of batteries is an installation and servicing task.
Camcorder	Make and play a selected moving picture	Insert cassette, take movie, stop, restart, play movie.	
Cassette recorder	Record and play auditory information	Insert new cassette, control levels, take recording, stop recording, rewind tape, play tape, adjust loudness, remove cassette.	Setup with stereo is an installation task.
CD player	Play a CD	Insert CD, play CD, select track, stop CD, eject CD.	Scan, shuttle, etc. are arbitrary.
Dishwasher	Wash a set of dishes	Insert dishes, insert detergent(s), select program and start machine, stop machine and remove dishes.	The correctness of the selected program is arbitrary, but the intended selection should be correct; the insertion of salt is a servicing task.
Dryer	Dry clothes	Insert clothes, select program, switch off machine, get out cloth, clean filter.	
Electric heater	Heat an area	Switch on, select heat, switch off.	
Electric mixer	Mix food	Unravel cord, plug in, insert food and mixer tool, switch on, mix, switch off, retain food, clean, pack up.	
Electric oven	Heat food	Switch on a selected plate with a selected heat.	Plate and heat are arbitrary, but the intended selections should be correct.
Fax machine	Send page(s) to a specified phone number	Insert page, dial and send fax to a number, get back paper.	Connection to phone line and switching on is an installation task.
Fire extinguisher	To extinguish a fire	Take extinguisher from place, activate extinguisher.	
Gas hob	Heat food	Switch on a selected plate, ignite gas, adjust heat.	Plate and heat are arbitrary, but the intended selections should be correct.

Table A.1 (*continued*)

Consumer product	Main goal(s)	Example tasks	Installation and other notes
Light switch	Get light in a specific location	Switch on selected room's light, switch off a selected room's light.	With multiple switches, ease of operation can target a single switch as well as an arrangement.
Microwave	Heat food to an appropriate temperature	Insert food, select power and time, start, get out food.	The values are arbitrary, but the intended selection should come out correct.
Mobile telephone	1) Take a call 2) Make a call	Activate the phone, activate a number (dial).	Numbers are arbitrary, but the intended numbers should be correctly dialled.
Photo camera	Take pictures for later review	Insert film [storage media], switch on, take aim, take picture, take steps to review picture.	Insertion of batteries is an installation and servicing task. Normal daylight conditions should be given. Quality of picture is technical.
Radio	Listen to a selected radio programme	Switch on, select station, adjust loudness, switch off.	The selection of the station is arbitrary, but the intended selection should be correct.
Refrigerator	Keep food at an appropriate temperature	Insert food, switch to minus, select appropriate temperature, get out food.	The temperature is arbitrary, but the intended selection should be correct.
Room telephone	1) Take a call 2) Make a call	Activate the phone (lift the speaker), activate a number (dial).	Numbers are arbitrary, but the intended numbers should be correctly dialled.
Satellite set-top box	Watch selected TV channels	Switch on, select station, switch off.	Setting up is an installation task.
Telephone answering machine	Access and play recorded calls	Playing received calls, playing a specific received call, deleting all calls.	Connecting to a telephone, adjusting reaction to number of calls and recording messages belong to setup.
Television with remote control	Watch selected TV channels	Switch on TV, select programme and loudness with remote control, switch off TV.	The setup of channels is an installation task.
Toaster	Toast bread to desired brownness	Insert bread, select brownness, switch on, remove bread.	The temperature selected is arbitrary, but the intended selection should be correct.
Vacuum cleaner	Clean an area	Unravel cord, plug in, switch on, select strength, move brush, switch off, unplug, ravel cord.	The exchange of dust bins is a service task. The quality of the result is outside the scope.
Videorecorder	Record and play videos	Switch on, insert cassette, take recording from running TV, stop recording, rewind tape, play tape, eject tape, switch off.	Setup with TV is an installation task.
Washing machine	Clean clothes using an appropriate program	Insert clothes, insert detergent, select program, start machine, get out clothes after machine has stopped, switch off machine.	The correctness of the selected program is arbitrary, but the intended selection should be realized.
Water tap	Obtain water of required temperature	Switching to cold and off.	The first movement shall be recorded, as a wrong movement could result in burns.

Annex B
(informative)

Background to the test method

B.1 Difference between formative and summative methods

This part of ISO 20282 contains a summative test method that may be used, for example, to estimate the percentage of people that can successfully achieve their goals in actual use. Other formative test methods are intended to be used primarily to identify usability problems. Formative methods typically employ fewer users and involve the user thinking aloud, so do not give accurate estimates of effectiveness, efficiency or satisfaction.

B.2 Difference between usability and ease of operation

Ease of operation is a specific instantiation of usability for everyday products. The differences are illustrated in ISO 20282-1:2006, Annex B, as shown in Table B.1.

Table B.1 — Difference between usability and ease of operation

Measure	Usability according to ISO 9241-11	Ease of operation according to ISO 20282-1
Users	Appropriate to the situation	General public
Usage	Appropriate to the situation	Primarily first time
Environment	Appropriate to the situation	Personal or public (not a part of professional work)
Product	Any hardware, software and materials	Interface of everyday products
Effectiveness measure	Appropriate to the situation	Success rate
Efficiency measure	Appropriate to the situation	Task time
Satisfaction measure	Appropriate to the situation	Smiley scale
Goals	Appropriate to the situation	Main goals

Ease of operation is an appropriate measure for the simple products that are the main focus of this part of ISO 20282, but for more complex tasks and products other usability measures (for example, measures of partial task success, mental workload or satisfaction with particular features) may also be relevant.

B.3 Choice of measures

To obtain reliable measures of ease of operation, a summative evaluation method consistent with ISO 9241-11 is needed. This contrasts with typical formative evaluation methods that use a small number of participants to identify usability problems.

Effectiveness is defined in ISO 9241-11 in terms of accuracy and completeness, but for the simple tasks within the scope of this part of ISO 20282, only success rate is used.

NOTE Accuracy is quite difficult to measure: see, for example, ISO/IEC 25062.

Efficiency is defined in ISO 9241-11 in terms of the resources expended in relation to the accuracy and completeness. For practical purposes, time is the most important resource (and is more easily interpreted on its own rather than as a ratio).

Satisfaction is defined in ISO 9241-11 as freedom from discomfort, and positive attitudes towards the use of the product. For ease of operation the most important component of satisfaction is a positive attitude.

B.4 Number of users

While small numbers of users are sometimes sufficient to identify a large proportion of usability problems, larger numbers are required to confidently estimate the population mean for success rate. Larger numbers also make it easier to obtain a representative sample of users.

Annex C
(normative)

Recruiting a representative sample of users

C.1 General

A sample representing the intended or actual user groups should be recruited to take part in the test. A sample of users should be selected that models the distribution of relevant user characteristics within the intended or actual user groups.

NOTE The user group for testing may be more specific or more general than the intended user group for designing.

C.2 Decide whether to divide users into groups

Decide whether the characteristics of the intended or actual user groups identified in 7.2.3 divide users into distinct categories that will have a major impact on ease of operation.

If there is more than one distinct user group, decide whether to test all or only some user groups, depending on the purpose of the test.

If information is required on the ease of operation for different user groups, include sufficient representatives of each user group to enable statistical conclusions to be made.

EXAMPLE 1 If previous knowledge or experience or language spoken will have a major impact on ease of operation, each group can be tested separately.

If information is only required about ease of operation for the whole population, use one representative sample drawn from the population (see C.4). To maximise statistical validity, use as large a sample size as possible. However, it will not be possible to draw any statistical inferences about the ease of operation for any constituent groups that represent only a small percentage of the population.

It is important to decide whether to test first-time users and/or people with previous experience. Even with a new product, people could have experience with similar products or a similar previous model, so the type of relevant previous experience may need to be defined.

EXAMPLE 2 Users of a mobile phone are divided into groups determined by their expertise in setting up this type of mobile phone:

a) those who have never set up a mobile phone;

b) those who have set up a mobile phone of a different brand;

c) those who have previously set up a similar mobile phone.

Each group will be tested separately, as they are expected to have different results for ease of operation.

Only one group might be tested, for example, the most frequent users or a critical user group.

EXAMPLE 3 A hotel alarm clock is expected to be difficult to set for the first time, but easy to set on subsequent occasions. Only those who have not used the alarm clock before are tested.

C.3 Choose sample size

Given the variation in user characteristics associated with a consumer product a sample size of at least 50 per group is desirable in order to obtain a sufficiently representative sample of the intended or actual user group, and to provide results with sufficient accuracy. (Smaller sample sizes may be sufficient if most or all users are expected to be successful.)

The larger the sample size, the more accurate the prediction (provided the sample is representative of the user population).

EXAMPLE If the measured success rate is 90 %:

— for 18 successful users out of 20 users, the 95 % confidence interval is −18 %;

— for 45 successful users out of 50 users, the 95 % confidence interval is −10 %;

— for 90 successful users out of 100 users, the 95 % confidence interval is −6 %.

C.4 Recruit a representative sample of users

A sample that is representative of the relevant characteristics of the intended or actual user group should be recruited to take part in the test. To obtain a truly representative sample for an everyday product would need a very large number of users. As this is usually uneconomic, stratified sampling may be used to obtain a sample that represents the major relevant characteristics.

As it is usually impossible to obtain a truly random sample, a sample of users should be selected that models the distribution of relevant user characteristics within the intended or actual user groups.

Many general physical and psychological characteristics (such as strength, visual acuity and cognitive capabilities) that effect ease of operation can be indirectly sampled through appropriate demographic variables such as age, gender and educational attainment. Users should be selected to be representative of these demographic variables and any other significant characteristics that will not be adequately sampled by the demographic variables.

EXAMPLE 1 In a multi-ethnic population, users are selected to have a representative cultural background.

EXAMPLE 2 When testing a consumer product, users are selected to be representative in their experience of different brands.

EXAMPLE 3 An alarm clock requires manual dexterity to operate the controls and good visual acuity to read the instructions. These relevant user characteristics are sampled by recruiting users with a representative distribution of age and educational attainment. As 5 % of purchasers are expected to have previous experience of this type of alarm clock, and as the task of setting the alarm is expected to be much more difficult for those who do not have this experience, only people without previous experience will be tested.

Stratified sampling should be used to distribute users across as many permutations of characteristics as is practical. The preferred combinations of user characteristics can be laid out in a table.

EXAMPLE 4 Age, gender and reading ability are the relevant attributes, and the distribution in the user population is:

Gender	male: 45 %, female: 55 %.
Age	20 to 39: 40 %, 40 to 64: 40 %, 65 or above: 20 %.
Reading ability	low: 10 %, medium: 20 %, high: 70 %.

An example stratified sample of size 60 with interrelated controls is given in Table C.1.

Table C.1 — Example stratified sample of size 60 with interrelated controls

Age	Reading ability						Total
	low		medium		high		
	Gender						
	Male	Female	Male	Female	Male	Female	
20 to 39	1	1	2	3	8	9	24
40 to 64	1	1	2	3	8	9	24
65 or more	1	1	1	1	3	5	12
Total	6		12		42		60

C.5 Minority groups

If it is not possible to include minority groups (such as a specific nationality or specific disability) because they constitute a very small percentage of the intended user population within a representative sample, additional tests may be carried out with these groups to determine whether members of these groups are able to use the product.

NOTE If ease of operation is known to be acceptable for most of the user population, testing could be restricted to specific groups that are expected to have more difficulty.

C.6 Selecting users

Having identified the important user characteristics, potential participants should be screened to select a sample that matches the required profile. For a laboratory test, the screening takes place in advance. For an installed walk-up-and-use product in a public place, there are three possibilities:

a) participants could be screened and invited in advance as for a laboratory test;

b) passing members of the general public could be screened for potential participation;

c) if it is acceptable to monitor normal usage, actual users could be approached after use of the product and screened for a match to the required sample.

Annex D
(normative)

Confidence intervals

D.1 Effectiveness of operation

Effectiveness of operation measures the percentage of users who successfully achieve the main goal(s) of use of a product. The confidence intervals for effectiveness of operation shall be calculated using the exact binomial statistic (see Bibliography). The establishing of a high success rate with a high degree of confidence requires a large sample of users. Table D.1 gives examples of the maximum number of unsuccessful users required to estimate population success rates greater than 75 % and 80 %, with 80 % and 95 % confidence (using single-tailed confidence intervals).

EXAMPLE If 43 out of 50 users (86 %) are successful, there is a 95 % probability that the success rate is at least 75 %, and an 80 % probability that the success rate is at least 80%.

NOTE 1 These statistics are only valid for random samples. It is important that the users are selected randomly and, when stratifying a sample, it is important that the users are selected randomly in each category (see Annex C).

NOTE 2 A 95 % level of confidence is conventional in scientific work and where definitive claims are being made; however, an 80 % level of confidence could be sufficient for business decisions.

Table D.1 shows that the maximum number of unsuccessful users required to estimate a population success rate greater than 80 % with 95 % confidence is

— 0 users out of 14 to 21 users tested,

— 1 user out of 22 to 29,

— 2 users out of 30,

— 3 users out of 40,

— 5 users out of 50, and

— 13 users out of 100.

The maximum number of unsuccessful users required to estimate a population success rate greater than 75 % with 95 % confidence is

— 0 users out of 11 to 16 users tested,

— 1 user out of 17 to 22,

— 2 users out of 23 to 28,

— 3 users out of 30,

— 5 users out of 40,

— 7 users out of 50, and

— 17 users out of 100.

NOTE 3 Examples of other estimates of population success rates are given in Table D.1.

Table D.1 — Examples of confidence intervals

No. of users	Max. no. unsuccessful users required to estimate a population success rate greater than 75 % with 95 % confidence	Estimated minimum population success rate with 95 % confidence	Estimated minimum population success rate with 80 % confidence	Max. no. unsuccessful users required to estimate a population success rate greater than 80 % with 95 % confidence	Estimated minimum population success rate with 95 % confidence	Estimated minimum population success rate with 80 % confidence
10						
11	0	76,2 %	86,4 %			
12	0	77,9 %	87,4 %			
13	0	79,4 %	88,4 %			
14	0	80,7 %	89,1 %	0	80,7 %	89,1 %
15	0	81,9 %	89,8 %	0	81,9 %	89,8 %
16	0	82,9 %	90,4 %	0	82,9 %	90,4 %
17	1	75,0 %	83,4 %	0	83,8 %	91,0 %
18	1	76,2 %	84,3 %	0	84,7 %	91,4 %
19	1	77,4 %	85,0 %	0	85,4 %	91,9 %
20	1	78,4 %	85,8 %	0	86,1 %	92,3 %
21	1	79,3 %	86,4 %	0	86,7 %	92,6 %
22	1	80,2 %	87,0 %	1	80,2 %	87,0 %
23	2	75,1 %	82,3 %	1	81,0 %	87,5 %
24	2	76,0 %	83,0 %	1	81,7 %	88,0 %
25	2	76,9 %	83,7 %	1	82,4 %	88,5 %
26	2	77,7 %	84,3 %	1	83,0 %	88,9 %
27	2	78,5 %	84,8 %	1	83,6 %	89,3 %
28	2	79,2 %	85,3 %	1	84,1 %	89,7 %
29	3	75,4 %	81,8 %	1	84,7 %	90,0 %
30	3	76,1 %	82,4 %	2	80,5 %	86,3 %
40	5	75,5 %	81,0 %	3	81,7 %	86,6 %
50	7	75,3 %	80,2 %	5	80,1 %	84,6 %
100	17	75,6 %	79,0 %	13	80,1 %	83,3 %

D.2 Efficiency of operation

The confidence intervals for efficiency of operation shall be calculated using the standard error of the mean for task time. If the task time data is not normal, it should be transformed to a normal distribution — for example, using a logarithmic transformation — before calculating the standard error of the mean.

NOTE The size of the confidence interval will depend on the size of the sample and the variance in the population tested. The variance will generally decrease with increased sample size, and will generally be larger for more diverse populations.

EXAMPLE The standard error of the mean is calculated after a logarithmic transformation, and is equivalent to 4,2 s. The median task time is 22,2 s, so the results could be reported as a median of 22,2 s, with a 95 % probability that the true value is between 18,0 s and 26,4 s.

Annex E
(normative)

Satisfaction with operation scale

The circles shall be 10,5 mm in diameter and the scale should be used as is, stretched over 69 mm between the centres of the extremes.

For users with visual disabilities, these sizes may be increased proportionately.

When using the scale, the faces presented in Figure E.1 shall be used as they are, as the convexness and concaveness of every Smiley's mouth has been carefully selected for equidistance.

NOTE The Smiley scale was first brought up by Kunin (1955), has been widely used in psychology to measure emotional aspects of satisfaction, and has also been applied in industrial usability tests. A five-point version of the scale has recently received factor analytical validation and has been positively tested for equidistance by Ruth Jäger (2004).

"Please tick the face that corresponds most closely with your degree of satisfaction with the operation of..."

Figure E.1 — Smiley scale

The Smileys correspond to the values −2, −1, 0, +1, +2. No intermediate measures are allowed. Group values may be given as an arithmetic average.

Annex F
(normative)

Format for test reports

NOTE This format is based on the common industry format (CIF) for usability test reports according to ISO/IEC 25062.

F.1 Title page

The following information shall be provided:

a) identification of report as ISO/PAS 20282-3, Test report for ease of operation of consumer products;

b) name of product and version tested;

c) when the test was conducted;

d) where the test was conducted;

e) date the report was prepared;

f) test organization name;

g) test organization contact person and contact details.

F.2 Executive summary

This section provides a high-level overview of the test. The intent of this section is to provide a stand-alone summary.

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

a) name and description of the product,

b) summary of method(s), including number(s) and type(s) of participants and tasks, and

c) percentage value for ease of operation and central tendency (mean) for efficiency and satisfaction, if measured.

F.3 Product and expected context of use

F.3.1 Full product description

The following information shall be provided:

a) formal product name, model and version;

b) the parts or functions of the product that were evaluated.

F.3.2 Expected context of use

The following information shall be provided:

a) the main goal(s) of use of the product;

b) the intended or actual user groups of the product;

c) for those user groups to be included in the testing,

- the user characteristics that could affect the ease of operation of the product,
- the range of characteristics that exists in the intended or actual user group,
- the expected distribution of each relevant user characteristic, and
- whether the product is suitable for older users or any other groups with special needs;

d) the main environments in which the product is (or is intended to be) used.

F.4 Method

F.4.1 General

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

F.4.2 Participants

The following information shall be provided:

a) any physical or cognitive user characteristics that are expected to affect ease of operation;

b) the total number of participants tested;

c) which user group(s) were tested, and why they were selected;

d) key characteristics and capabilities of each user group;

e) how participants were selected; whether they were representative of the essential characteristics;

f) any differences between the participant sample and the user population;

g) table of participants (row) by characteristics (columns), including relevant demographics, experience and any special needs.

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

Characteristics should be chosen to be relevant to the product's ease of operation; they should allow a reader to determine how similar the participants are to the user population.

EXAMPLE Table F.1 is an example; the characteristics that are shown are typical but might not necessarily cover every type of testing situation.

Table F.1 — Example table for information about participants

Participant	Gender[a]	Age[b]	Linguistic ability[c]	Physical attributes[d]	Task experience[e]	Related products experience[f]
1						
2						
n						

[a] Male/female.

[b] The chronological age of the participant, or indicate membership in an age range (e.g. 25 to 45) or age category (e.g. under 18, over 65) if the exact age is not known.

[c] Any relevant language skills.

[d] Any physical attributes related to ease of operation.

[e] Description of the relevant background such as how much experience the user has in carrying out this type of task.

[f] Type and duration of any prior experience with any similar products.

F.4.3 Context of product use in the test

F.4.3.1 General

Any known differences between the evaluated context and the expected context of use shall be provided.

F.4.3.2 Tasks

The following information shall be provided.

a) The task scenarios for testing.

b) Why these tasks were selected.

EXAMPLE 1 The most frequent tasks, or the most everyday tasks.

c) How these tasks were identified.

EXAMPLE 2 Observation of consumers using similar products, product marketing specifications.

d) Any task data given to the participants.

e) Completion criteria established for each task.

F.4.3.3 Test environment

The following information should be provided.

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

EXAMPLE 1 Usability lab, meeting room, home office, home family room, or public hall.

b) Any relevant features or circumstances that could affect the results.

EXAMPLE 2 Video and audio recording equipment, one-way mirrors, or automatic data collection equipment.

F.4.3.4 Technical environment

F.4.3.4.1 General

If the product needs surrounding equipment or service to function, this product or service shall be specified, e.g. cell phone service provider, TV controlled by a remote control.

F.4.3.4.2 Participant's computing environment

This is only applicable to a software product. The following information shall provide enough information to replicate and validate the test, including

a) computer configuration, including model, OS version, required libraries or settings, and

b) if used, browser name and version, relevant plug-in names and versions.

F.4.3.4.3 Display devices

The following information shall be provided:

a) if screen-based, screen size, resolution, and colour setting;

b) if print-based, the media size and print resolution;

c) if visual interface elements (such as fonts) can vary in size, the size(s) used in the test.

F.4.3.4.4 Audio devices

If used, the relevant settings or values for the audio bits, volume, etc. should be provided.

F.4.3.4.5 Manual input devices

If used, the make and model of devices used in the test should be provided.

F.5 Experimental design

F.5.1 General

The measures for which data were recorded shall be provided.

F.5.2 Procedure

F.5.2.1 General

The following information shall be provided:

a) operational definitions of measures;

b) time limits on tasks;

c) policies and procedures for interaction between tester(s) and subjects.

The following information should be provided:

d) sequence of events from greeting the participants to dismissing them;

e) details of non-disclosure agreements, form completion, warm-ups, pre-task training, and debriefing;

f) verification that the participants knew and understood their rights as human subjects;

g) steps followed to execute the test sessions and record data;

h) number and roles of people who interacted with the participants during the test session;

i) whether other individuals were present in the test environment;

j) whether participants were paid or otherwise compensated.

F.5.2.2 Participant general instructions

The following information shall be provided:

a) instructions given to the participants (here or in an appendix);

b) 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.

F.5.2.3 Participant task instructions

A task instruction summary shall be provided.

F.5.3 Ease of operation metrics

Metrics for effectiveness of operation shall be provided as specified in 7.5.

The following information should be provided as specified in 7.5:

a) metrics for efficiency of operation;

b) metrics for satisfaction with operation.

F.6 Results

F.6.1 General

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

a) Data collection and scoring.

b) Data reduction.

EXAMPLE 1 How data were collapsed across tasks or user groups.

c) Statistical analysis.

EXAMPLE 2 Explaining any statistical transformation and reduction. (Consistent with earlier clauses.)

Tabular performance results per task or user group shall be provided.

Various graphical formats are effective in describing ease of operation data at a glance. A variety of plots can be used effectively to show comparisons of expert benchmark times for a product versus the mean participant performance time.

F.6.2 Performance results

A table of results may be presented for groups of related tasks where this is more efficient and makes sense.

The following information should be provided:

a) summary table(s) of performance results across all tasks;

b) graphical presentation of performance results.

Table F.2 is an example of such a table.

Table F.2 — Example of summary of performance of user group

User Group A			
User no.	**Task success**	**Task time** min	**Satisfaction score**
1			
2			
n			
Success			
Median time			
Mean score			
Confidence interval: min. max.			

F.7 Appendices

The following information shall be provided:

a) participant general instructions (if not in the body of the report);

b) participant task instructions.

Release notes explaining or updating the test results may also be provided.

Annex G
(informative)

Feedback on this part of ISO 20282

G.1 Address for feedback

This part of ISO 20282 is being published initially as a Publicly Available Specification so that information on and experience of its use in practice can be gathered. Comments on the content can be sent to ISO/TC 159/SC 1 at:

iso_tc159@din.de

G.2 Test methods for consumer products

This test method for consumer products is being published in this form because testing consumer products involves issues that imply that the results could be less reliable or more difficult to interpret than for a walk-up-and-use product. In some cases, the same product can be used in a similar way to either a walk-up-and-use product or a consumer product (e.g. a washing machine).

In other cases, use of a consumer product introduces extra complexity that can make it difficult to design an appropriate test of ease of operation. There may be a greater variety in the context of use, including different market segments, differences in the purposes for which the product will be used and different usage environments. This could make it difficult to organize a test with sufficient numbers of users to represent all user segments.

EXAMPLE Different user groups could have different main goals. For a mobile phone, these can be to make a call, send a text message or take a photo.

Feedback is invited (see G.3) on whether the results obtained when testing consumer products are reliable and easy to interpret.

The intended users and usage environment of a walk-up-and-use product are usually well defined and easy to identify. Many consumer products are designed for market segments that are not so well defined, and can be used in several environments.

The test method is intended for main goals, and does not take into account ease of operation of additional tasks that can be performed with more complex products. At the same time, it is intended for relatively simple goal achievement where there is a clear definition of success, and neither does it take into account the quality of the outcome.

G.3 Feedback invited

Feedback on this part of ISO 20282 is particularly invited on the following issues and questions.

— Experiences of usage, case studies and any published results.

— Types of products, area of application, numbers of groups, numbers of users, incorporation of disabilities.

— Any problems with the method?

— Suggestions for improvements?

— How easy is it to identify the main goals?

— How easy is it to recruit a representative sample of users?

— If two organizations independently test the same product, do they obtain essentially the same results?

— Are the results obtained reliable and easy to interpret?

— Is the method suitable for testing other types of products?

Bibliography

[1] ISO 8317:2003, *Child-resistant packaging — Requirements and testing procedures for reclosable packages*

[2] ISO/IEC TR 9126-4:2004, *Software engineering — Product quality — Part 4: Quality in use metrics*

[3] ISO 9241-9:2000, *Ergonomic requirements for office work with visual display terminals (VDTs) — Part 9: Requirements for non-keyboard input devices*

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

[5] ISO 13407:1999, *Human-centred design processes for interactive systems*

[6] ISO/TR 16982, *Ergonomics of human-system interaction — Usability methods supporting human-centred design*

[7] ISO/TS 20282-2:2006, *Ease of operation of everyday products — Part 2: Test method for walk-up-and-use products*

[8] ISO/PAS 20282-4, *Ease of operation of everyday products — Part 4: Test method for the installation of consumer products*

[9] ISO/IEC 25062:2006, *Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Common Industry Format (CIF) for usability test reports*

[10] IEC/TR 61997:2001, *Guidelines for the user interface in multimedia equipment for general purpose use*

[11] JÄGER, R. *Konstruktion einer Ratingskala mit Smilies*. *Diagnostica*, 2004, 50, 1, 31-38

[12] KUNIN, TH.: The construction of a new type of attitude measure. *Personnel Psychology*, 1955, 8, 65-77

[13] Sample Size Calculator <http://www.dssresearch.com/toolkit/sscalc/size_p1.asp>

BSI — British Standards Institution

BSI is the independent national body responsible for preparing British Standards. It presents the UK view on standards in Europe and at the international level. It is incorporated by Royal Charter.

Revisions

British Standards are updated by amendment or revision. Users of British Standards should make sure that they possess the latest amendments or editions.

It is the constant aim of BSI to improve the quality of our products and services. We would be grateful if anyone finding an inaccuracy or ambiguity while using this British Standard would inform the Secretary of the technical committee responsible, the identity of which can be found on the inside front cover. Tel: +44 (0)20 8996 9000. Fax: +44 (0)20 8996 7400.

BSI offers members an individual updating service called PLUS which ensures that subscribers automatically receive the latest editions of standards.

Buying standards

Orders for all BSI, international and foreign standards publications should be addressed to Customer Services. Tel: +44 (0)20 8996 9001. Fax: +44 (0)20 8996 7001. Email: orders@bsi-global.com. Standards are also available from the BSI website at http://www.bsi-global.com.

In response to orders for international standards, it is BSI policy to supply the BSI implementation of those that have been published as British Standards, unless otherwise requested.

Information on standards

BSI provides a wide range of information on national, European and international standards through its Library and its Technical Help to Exporters Service. Various BSI electronic information services are also available which give details on all its products and services. Contact the Information Centre. Tel: +44 (0)20 8996 7111. Fax: +44 (0)20 8996 7048. Email: info@bsi-global.com.

Subscribing members of BSI are kept up to date with standards developments and receive substantial discounts on the purchase price of standards. For details of these and other benefits contact Membership Administration. Tel: +44 (0)20 8996 7002. Fax: +44 (0)20 8996 7001. Email: membership@bsi-global.com.

Information regarding online access to British Standards via British Standards Online can be found at http://www.bsi-global.com/bsonline.

Further information about BSI is available on the BSI website at http://www.bsi-global.com.

Copyright

BSI
389 Chiswick High Road
London
W4 4AL