International Standard

ISO/IEC 4944

Information technology — User interfaces — Evaluating usability of natural user interfaces

Technologie de l'information — Interfaces utilisateur — Cadre pour l'évaluation de l'utilisabilité d'une interface utilisateur naturelle

First edition
2024-09

COPYRIGHT PROTECTED DOCUMENT

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.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives or www.iec.ch/members_experts/refdocs).

ISO and IEC draw attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO and IEC take no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO and IEC had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents and https://patents.iec.ch. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 35, *User interfaces*.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html and www.iec.ch/national-committees.

Introduction

As smart devices such as smart phones and smart speakers have become popular, new forms of user interfaces have been introduced. One aspect of the new interfaces is mimicking natural ways of interacting with the devices and utilizing speech, gestures, body language, eye gaze and other natural features of human communication through natural user interfaces (NUIs), which allow for faster, smoother, intuitive, and smarter interaction. However, while the concept of NUI is well defined, practical applications of NUI still need to be developed. This is strongly related to the lack of usability of NUI in smart devices.

When specified users of systems, products or services achieve their specified goals with effectiveness, efficiency and satisfaction in a specified context of use (as defined in ISO/IEC 9241-11), usability can be evaluated. Effectiveness is considered as the ability of users to complete tasks with the system (e.g. the NUI). Efficiency measures resources taken by users to complete specified tasks. Satisfaction measures comfort levels of user experience along with acceptance.

It is important that the usability evaluation of NUI concentrates on recognizing the user's intention and focuses on multiple modalities of information flows between the users and the devices.

This document describes a framework for evaluating the usability of NUI. Although the framework can be applied in most general contexts, it is specifically focused on optimizing user experiences in terms of NUI.

Information technology — User interfaces — Evaluating usability of natural user interfaces

1 Scope

This document provides a framework, requirements and recommendations for evaluating the usability of a natural user interface (NUI) for systems, products or services. Measuring and reporting the usability of the NUI are also described. The usability evaluation focuses on the efficiency, effectiveness and satisfaction of the NUI. This document can also be applied to the usability evaluation of other user interfaces (UIs) for emerging technologies.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements 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 9241-11, *Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 9241-11 and the following apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

3.1 Terms related to natural user

3.1.1
user interface
UI
set of all the components of an interactive system that provide information and controls for the user to accomplish specific tasks with the interactive system

[SOURCE: ISO 9241-110:2020, 3.10]

3.1.2
natural user interface
NUI
user interface where the user operates the interface through a manner related to everyday human behaviours in a real world

Note 1 to entry: Natural refers to the user's expectations, behaviour and feelings that are involved in interacting with a NUI.

Note 2 to entry: Everyday human behaviours refer to behaviours that the user can be expected to have already learned.

Note 3 to entry: Interactions with NUIs typically model interactions performed in the natural world, including (but not limited to) touch, gestures or vocal interactions, or a combination thereof.

3.1.3
natural user interface interaction
NUI interaction
user-system interaction making use of a natural user interface

3.1.4
naturalness
correspondence with everyday interactions between humans and the real-world objects with which they interact

Note 1 to entry: Naturalness relates to everyday behaviours that the user can be expected to have already learned.

EXAMPLE   To perform a "hello" gesture, the user refers to an object by pointing; to order something the user issues a vocal command; or to agree the user nods in the affirmative or voices a vocal onomatopoeia for agreement.

3.2   Terms related to usability

3.2.1
interaction
user-system interaction
exchange of information between a user and an interactive system via the user interface to complete the intended task

Note 1 to entry: Adapted from ISO 9241-110:2020, 3.11 definition of user-system interaction.

3.2.2
diverse users
individuals with differing abilities and characteristics or accessibility needs

[SOURCE: ISO/IEC Guide 71:2014, 2.3]

3.2.3
diverse contexts
differing contexts of use and differing economic, cultural and organizational conditions

[SOURCE: ISO/IEC Guide 71:2014, 2.8]

4   Natural user interfaces (NUI)

Natural user interfaces (NUIs) are a special case of user interfaces that focus on the naturalness of interactions that they support.

The evaluation of the usability of an NUI includes:

a)   evaluation of the usability of the user interface to accomplish tasks;

   NOTE 1   This is similar to the evaluation of any other user interface.

b)   evaluation of the usability of the naturalness of the user interface.

   NOTE 2   This involves an additional, more specialized evaluation that is described in this document.

The naturalness of the user interface is provided by the NUI interactions, which are recognized and used by the NUI.

NUI interactions attempt to maximize their achievement of the interaction principles identified in ISO 9241-110 as:

—   Suitability for the user's tasks: an interactive system is suitable for the user's tasks when it supports the users in the completion of their tasks, i.e. when the operating functions and the user-system interactions are based on the task characteristics.

— Self-descriptiveness: the interactive system presents appropriate information, where needed by the user, to make its capabilities and use immediately obvious to the user without unnecessary user-system interactions.

— Conformity with user expectations: the interactive system's behaviour is predictable based on the context of use and commonly accepted conventions in this context.

— Learnability: the interactive system supports discovery of its capabilities and how to use them, allows exploration of the interactive system, minimizes the need for learning and provides support when learning is needed.

— Controllability: the interactive system allows the user to maintain control of the user interface and the interactions, including the speed and sequence and individualization of the user-system interaction.

— Use error robustness: the interactive system assists the user in avoiding use errors and, in case of identifiable use errors, treats them tolerantly and assists the user when recovering from the errors.

— User engagement: the interactive system presents functions and information in an inviting and motivating manner supporting continued interaction with the system.

NUI maximizes the naturalness of its interactions by supporting the following:

— Suitability for the user's tasks: the users of the interactive system with NUI perform tasks in a reasonable manner to how they would perform in the real world with everyday interactions (if the computer was not involved).

— Self-descriptiveness: the interactive system with NUI presents appropriate information to the user in a manner that is related to how the information exists in the real world (rather than in a manner that is optimized for the computer).

— Conformity with user expectations: the interactive system with NUI recognises and processes predictable user interactions in a predictable manner that is consistent with the real world (rather than designed based on typical human-computer interaction styles).

— Learnability: the interactive system with NUI supports the users in the discovery and exploration of its capabilities and information (without the need for formal training regarding the user interface and supported interactions before they can use the system).

— Controllability: the interactive system with NUI allows the user to interact with the NUI in the manner and sequence that is most natural to their preferred manner (rather than in a single prescribed manner).

— Use error robustness: the interactive system with NUI allows the users to try different interactions while avoiding and minimizing any resulting use errors and by providing suitable error recovery where needed.

— User engagement: the interactive system with NUI invites and motivates the users to interact with the NUI in a similar manner to everyday human behaviour in the real world.

5 Evaluating usability of a natural user interface (NUI)

5.1 Framework

Usability as defined in ISO 9241-11 is a complex concept.

NOTE   Usability is one of the outcomes of use of a system, product, or service. There are other outcomes of use, e.g. accessibility, user experience, avoidance from harm from use. This document only focuses on usability.

Different users will experience different levels of usability of the same system, product or service in achieving different goals in different contexts of use. Therefore, evaluation of the usability involves a number of different more specific evaluations that are described in the following subclauses.

Evaluating usability of NUI involves applying the definition of usability, as illustrated in Figure 1, by considering the following three issues:

— the system, product or service being evaluated (see 5.2), which in this case is the NUI;

— the users, goals, tasks and context of use (see 5.3, 5.4 and 5.5) which determine how NUI will be evaluated and used;

— the components of usability evaluation (see 5.6), e.g. effectiveness, efficiency and satisfaction.

The evaluation of the NUI usability requires evaluation metrics as follows:

a) how to evaluate usability of NUI (see 5.7), which includes:

    1) evaluating the usability of the user interface of the NUI for accomplishing goals;

    2) evaluating the usability of the naturalness of the interactions supported by the NUI;

b) how to measure the components of usability of NUI (see Clause 6);

c) how to report the results of the usability evaluation (see Clause 7).
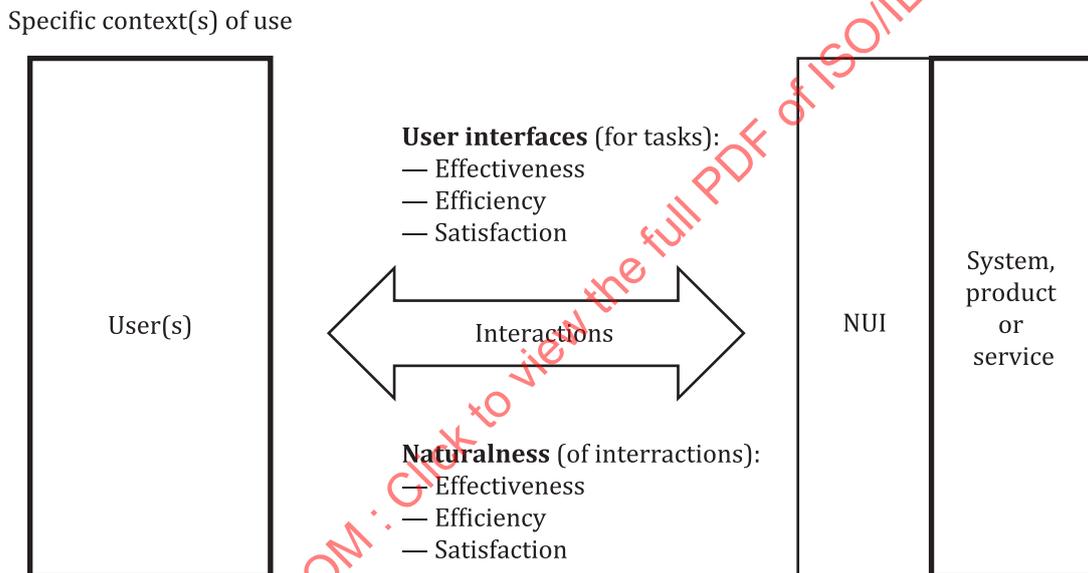
Figure 1 — Framework for evaluating usability of NUI

5.2 NUI as a system, product or service being evaluated

While usability can be applied to many different things, it is important that usability evaluation focuses on a well-defined object of the evaluation.

While NUI is a part of a system that exists within a context of use, it is important that usability evaluation of NUI is limited to the context of use of the interface and does not extend to the whole system.

NUI can involve multiple media and modalities. It is important to clarify whether the evaluation is focusing on individual components of the NUI or the entire NUI.

5.3 Evaluating usability for diverse NUI users

In order for NUI to be natural, it is important that it be usable by a wide range of users.

When establishing the specified users for evaluating the usability of NUI, it is important to consider the usability of the NUI for diverse users.

Distinct user groups that together represent the set of diverse users should be identified as the specified users for the purposes of the usability evaluation of NUI.

5.4   Evaluating usability for goals, tasks and interactions of NUI

Traditional usability focuses on goals and the various tasks that can be used to achieve these goals. This focus limits the evaluation of usability to specific outcomes.

While evaluating usability, NUI still needs to be limited to consideration of specific outcomes. It needs to be open to more diverse approaches to achieving these intended outcomes than the typical set of activities that are used to describe a single task.

In NUI, users are likely to try the set of interactions that they find to be most natural to use. While the system will recognize and act upon some of these interactions, it can miss others.

Interactions each have an identifiable purpose that is part of accomplishing a task. Interactions are each made up of one or more sets of:

— a user action (that is intended to be recognized by the NUI);

— a system response (that is presented via the NUI).

The system response to a user action can be:

— nothing (where the NUI failed to recognize the user action);

— an indication that the user action has been recognized and acted upon;

  — resulting in the system doing what was intended;

  — resulting in the system doing something other than what was intended;

  — resulting in the system not being able to determine what to do.

Evaluating the usability of NUI shall evaluate both:

a)   the usability of the user interface;

   NOTE 1   This identifies the usability of the NUI for users to achieve their goal (either directly or after multiple attempts).

   NOTE 2   This is typically evaluated by evaluating the accomplishment of a selection of user's tasks that achieve the specified goals.

b)   the usability of the "naturalness" of the interactions of the NUI.

   NOTE 3   This recognizes the extent to which a NUI is actually "natural".

   NOTE 4   This can be evaluated by evaluating the ratio of success of user attempts to interact in with the NUI, without prior training in what actions the NUI supports.

5.5   Evaluating usability for diverse contexts of use of NUI

In order for NUI to be natural, it is important that it be usable by a wide range of contexts of use.

When establishing the specified contexts of use for evaluating the usability of NUI, it is important to consider diverse contexts.

In order to consider actual users, it is important to specify distinct natural contexts of use that together represent the set of diverse contexts.

Distinct contexts of use that together represent the set of diverse contexts should be identified as the "specified context of use" for the purposes of NUI usability evaluation.

5.6 Components of usability

The definition of usability recognises that it involves three components (effectiveness, efficiency, and satisfaction) that, while they may interact with one another, are intended to be considered separately. ISO 9241-11 further divides the three components into the following subcomponents:

a) Effectiveness is composed of:

1) accuracy, which is the extent to which an actual outcome matches an intended outcome;

2) completeness, which is the extent to which users of the system, product, or service are able to achieve all intended outcomes.

b) Efficiency is composed of:

1) time used, which is the time expended in attempting to achieve a goal;

2) human effort expended, which is the mental and physical effort expended to complete specified tasks;

3) financial resources expended, which include the costs of using the system, product, or service, such as paying wages, or the cost of energy or connectivity;

4) materials expended, which include physical items (e.g. raw materials, water, paper) used as input to the task (including maintenance tasks) and processed by the system, product, or service.

c) Satisfaction is composed of:

1) physical responses, which include feelings of physical comfort or discomfort;

2) cognitive responses, which include attitudes, preferences and perceptions;

3) emotional responses, which result from experience while using the object of interest.

Fully evaluating the usability of NUI involves evaluating each of these nine subcomponents for each combination of the conditions of evaluation.

The evaluation of usability of NUI shall involve separate evaluations of the effectiveness, efficiency and satisfaction of the NUI under the same context(s) of use.

NOTE 1   An evaluation involving less than all three of effectiveness, efficiency and satisfaction is only an evaluation of those components and not an evaluation of usability.

The evaluation of usability of NUI should consider the various subcomponents of each of effectiveness, efficiency, and satisfaction and should include evaluation of the most appropriate subcomponents.

NOTE 2   Ideally it will involve evaluation of all of the subcomponents. However, at a minimum it will involve at least one subcomponent of each of the three main components as long as it clearly identifies the subcomponents that have been evaluated to produce the component evaluation.

5.7 Evaluation based on use

ISO/IEC 25066 identifies the following types of evaluation approaches:

a) Inspection to identify usability defects and potential usability problems including:

1) deviations of the object of evaluation from specified criteria such as user requirements, principles, design guidelines or established conventions;

2) potential usability problems when attempting to complete one or more tasks with the object of evaluation;

b) Observation of users including:

1) observing user behaviour to identify actual usability findings;

2) measuring user performance and response (e.g. time taken to perform a task, number of use errors, skin conductance or eye pupil dilation);

NOTE 1   The observation of users can be carried out either as an explicit usability test or conducted in a "real life" setting, or both.

NOTE 2   The usability problems are either identified during the observation or are identified from subsequent analysis.

c) User surveys including:

1) eliciting problems, opinions and impressions from users (qualitative user surveys);

2) measuring level of user satisfaction or perception, e.g. rating scale values for satisfaction or for subjectively perceived effectiveness or efficiency (quantitative user surveys);

3) other user reported data (e.g. data collected from an individual in conjunction with observation data).

NOTE 3   Collection of information about participants such as demographic data does not constitute a user survey.

Regardless of the approach or approaches taken, usability is defined and described in ISO 9241-11 as an outcome of use.

Since usability is an outcome of use, the evaluation of usability of NUI shall be performed either for a completed NUI or during the development of NUI.

NOTE 4   If the system, product, or service containing the NUI is completed this can involve observing user behaviour and measuring user performance and responses.

NOTE 5   If the system, product, or service containing the NUI is under development this can involve user testing of a prototype of the NUI.

NOTE 6   Usability evaluation of NUI can also include obtaining information from users after they have used the NUI or its prototype.

NOTE 7   It is not appropriate to base a usability evaluation of NUI solely on an inspection-based evaluation of the design of the interface.

6   Measuring usability of a NUI

6.1   Purposes of and measures for evaluation

There are various purposes for the evaluation of NUI, including:

— evaluations conducted during the development or redevelopment of a system, product or service;

— evaluation of a completed system, product or service;

— comparison of various systems, products or services.

There are various types of measures or evaluation results that can be produced by the usability evaluation:

a) Quantitative measures can identify whether or not there are problems and the severity of problems, but do not identify the source of the problems.

NOTE 1   Quantitative measures are especially useful in the evaluation of a completed system, product or service, and in the comparison of various systems, products or services.

b) Qualitative measures can identify what particular problems are occurring.

NOTE 2   Qualitative measures are especially useful development or redevelopment or in understanding whether problems experienced by some users are likely to be of concern to other users.

While there are various possible measures for evaluating usability, the measures in this clause are especially useful for measuring the usability of NUI.

The choice of measures and the level of details of each measure is dependent on the objectives of the parties involved in the measurement.

6.2 Measures for evaluating the user interface of NUI

6.2.1 Evaluating user interface of NUI for accomplishing tasks

An evaluation of user interfaces of NUI shall include the evaluation of the completion of at least one task per specified goal of the system, product or service.

NOTE 1   Evaluating the completion of more than one task per goal can improve the accuracy of the results.

NOTE 2   The evaluation of user interface of NUI is comparable to the evaluation of other types of systems, products and services. The evaluation of naturalness of interactions with NUI (see 6.3) goes beyond this more typical evaluation to focus on what makes NUI natural.

6.2.2 Effectiveness measures

6.2.2.1 Accuracy measures

Accuracy is the extent to which an actual outcome matches an intended outcome. With some goals (such as balancing financial accounts) absolute correctness is necessary. With other goals (such as dressing for the weather outside) there are a range of outcomes that can all be considered to be accurate, but also a range of outcomes that will be considered inaccurate.

NOTE 1   It is possible to partially achieve a goal (e.g. dressing for outside), without fully achieving accurately.

NOTE 2   It is possible that individual outcomes can be accurate without the outcomes being fully complete.

The accuracy of achieving the goals of NUI can be measured by:

a)   how fully users achieve their goals via supported tasks using any available NUI features (e.g. facial expressions, gestures, voice);

b)   the percentage of tasks successfully completed by users on their first attempt;

c)   the percentage of users who can successfully complete their tasks;

d)   the absence or presence of use errors and unnecessary outputs that interfere with the user completing the task.

6.2.2.2 Completeness measures

Completeness is the extent to which users of the system, product or service are able to achieve all intended outcomes.

NOTE 1   Completeness deals with the number of goals achieved, to a level sufficient to be recognized as having been achieved.

NOTE 2   It is possible that outcomes can be complete without all the outcomes being fully accurate. The completeness of achieving the goals of a NUI can be measured by:

a)   the percentage of the goals that the users have achieved, via supported tasks using the NUI;

b)   the percentage of the goals that the users are not to be able to achieve, while interacting through the NUI.

6.2.2.3 Additional qualitative measures

It is useful to report:

a) the failures to achieve a goal that have been identified;

b) the use errors that have been made in attempting to achieve goals;

c) unnecessary or misleading information that was presented.

6.2.3 Efficiency measures

6.2.3.1 Time used measures

The time used efficiency of user interface of NUI can be measured by:

a) the time taken to complete all of the specified goals;

NOTE 1   While the time to complete all specified goals measures efficiency of the entire system, it is highly dependent on the efficiency of the NUI.

b) the ratio of (total time spent minus time spent dealing with use errors) to the (total time spent) completing all of the specified goals.

NOTE 2   Attempts to speed-up users to achieve higher time-based efficiency, where they do not involve reducing the errors that delay users, can lead to decreases in effectiveness and satisfaction.

6.2.3.2 Human effort measures

The human effort efficiency of user interface of NUI can be measured by:

a) measures of the mental and physical workload involved in completing the goals;

b) measures of the peak workload required to complete the most difficult tasks;

c) indirect measures of the fatigue experienced by users completing the goals.

NOTE 1   The number of steps that are required of a user is not necessarily a good measure of human effort required, since some users can find it easier to work in smaller steps while other users can find it easier to work in larger steps.

NOTE 2   The amount of human effort expended is separate from any salary paid to the users, since it measures workload that is needed (regardless of any compensation for this workload) to accomplish the goals.

6.2.3.3 Financial resources expended

Financial resources expended include the costs of using the system, product or service, such as paying wages, or the cost of energy or connectivity. Financial costs can also include the costs of disposal of used equipment or waste. This can include some proportion of the financial costs for reusable resources acquired for this set of goals (such as specialized hardware and software). See ISO 9241-11.

The financial efficiency of user interface of NUI can be measured by:

a) the amount of money that will have to be spent to achieve the goals for a given time period (e.g. per week, per month, or per year);

b) The amount of money that will have to be spent to achieve a particular task once.

NOTE   This is intended to measure the amount of financial expenditures that will actually be incurred in order to use the NUI to achieve the goals. It is not intended to include amounts of money already expended (such as for materials already purchased).

6.2.3.4 Materials expended

Materials expended focuses on physical items that are consumed or transformed in the achievement of goals.

The material expended efficiency of user interface of NUI can be measured by:

a) the amount of materials that will have to be expended to achieve the goals for a given time period (e.g. per week, per month, or per year);

b) the amount of materials that will be expended to achieve a particular task.

NOTE   This measure is intended to be separate from any measure of the costs of the materials, since it measures how efficient (or wasteful) the system is in using material resources.

6.2.4   Satisfaction measures

The three subcomponents of satisfaction, identified in ISO 9241-11, are often combined, rather than being evaluated separately, when evaluating usability.

Satisfaction of user interface of NUI is more likely to be measured qualitatively rather than quantitatively, due to the subjective nature of the concept of satisfaction.

Some common measures of satisfaction of user interface of NUI include:

a) user reported satisfaction (in response to satisfaction surveys or to a general request for comments about the level of satisfaction that users have with using the user interface);

   NOTE   Several formalized questionnaires exist such as SUS (System Usability Scale) and QUIS (Questionnaire for User Interaction Satisfaction).

b) observations (by the evaluators) of discomfort and frustration experienced by users of the NUI;

c) user complaints and suggestions for improvements;

d) user willingness to use or recommend the system after initial use.

6.3   Measures for evaluating naturalness of interactions with NUI

6.3.1   Evaluating naturalness of interactions with NUI for accomplishing tasks

While traditional usability focuses on achieving application-directed goals, the naturalness presents its own unique goal, in addition to these other goals. It is typically associated with individual interactions that can be used in various tasks, rather than with specific tasks that achieve specified goals.

Evaluating the naturalness of interactions, therefore, applies to the overall use of the system. This evaluation can proceed along with the evaluation of using the user interface to achieve the goals, but involves its own measures, which are accumulated across all NUI interactions.

Where the purpose of the evaluation is to evaluate the naturalness of the interactions with one or more available NUI features (e.g. facial expressions, gestures, voice), the full set of tasks used in evaluating the naturalness of the NUI should be evaluated using each of the NUI features of interest.

6.3.2   Effectiveness measures

6.3.2.1   Accuracy measures

Accuracy as it applies to naturalness focuses on the extent to which the outcome of recognizing individual and natural interactions matches their intended outcomes. The accuracy in classifying user commands means closeness of agreement between users' performance of executing commands and the classified commands.