



中华人民共和国国家标准

GB/T 29836.3—2013

系统与软件易用性 第3部分：测评方法

Usability of system and software—
Part 3: Method of evaluating and testing

2013-11-12 发布

2014-02-01 实施

中华人民共和国国家质量监督检验检疫总局
中国国家标准化管理委员会 发布

目 次

前言	Ⅲ
引言	Ⅳ
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 概述	1
5 易理解性	1
5.1 明显的功能	1
5.2 描述的完整性	2
5.3 演示能力	2
5.4 演示的有效性	3
5.5 输入的有效性检查	3
6 易学习性	3
6.1 帮助文档的有效性	3
6.2 帮助机制的有效性	4
7 易操作性	4
7.1 使用中默认值的可用性	4
7.2 完成指定任务的步骤	5
7.3 操作的复杂性	5
7.4 完成指定任务过程中误操作的次数	6
7.5 错误的纠正	6
7.6 发生错误的影响力	6
7.7 可还原性	7
7.8 运行差错的易恢复性	7
7.9 使用中的消息的可理解性	7
7.10 运行状态的易监控性	8
7.11 界面元素的易定制性	8
7.12 界面布局的易定制性	9
7.13 快捷方式的易定制性	9
7.14 操作规程的易定制性	9
7.15 特殊辅助功能	10
7.16 无障碍程度	10
7.17 一致性	11
8 吸引力	11
8.1 界面色彩对视觉的吸引力	11
8.2 界面元素形状的舒适度	12
8.3 界面元素尺寸的合理性	12

8.4 布局的合理性	12
8.5 用户的感受度	13
9 测评方法描述	13
9.1 方法的分类	13
9.2 方法实施过程的描述结构	14
9.3 用户模型法	15
9.4 用户测试法	16
9.5 技术测试法	17
9.6 评审法	18
9.7 调查法	20
10 测评中用户的选择	22
10.1 选择原则	22
10.2 计算机技能	23
10.3 业务经验	23
10.4 认知能力	24
11 易用性测试环境	25
11.1 预期效果	25
11.2 应具备的条件	25
11.3 可供选择的设备	25
附录 A (资料性附录) 网上评卷系统易用性测评示例	26
参考文献	35

前 言

GB/T 29836 在《系统与软件易用性》总标题下,目前分为如下三部分:

- 第1部分:指标体系;
- 第2部分:度量方法;
- 第3部分:测评方法。

本部分为 GB/T 29836 的第3部分。

本部分按照 GB/T 1.1—2009 给出的规则起草。

本部分由全国信息技术标准化技术委员会(SAC/TC 28)提出并归口。

本部分起草单位:珠海南方软件网络测评中心、中国电子技术标准化研究院、北京邮电大学、南宁市平方软件新技术有限责任公司、国家应用软件产品质量监督检验中心、教育部考试中心、上海计算机软件技术开发中心、上海浦东软件平台有限公司、上海宝信软件股份有限公司、上海鲁齐信息科技有限公司。

本部分主要起草人:袁玉宇、侯建华、张畅畅、郭新伟、李英华、刘连芳、朱洁、蔡立志、胡宇、黄家裕、杨金翠、韩强、左家平、丁志刚、李家宏、张露莹、崔岩。

引 言

GB/T 29836 的本部分提出了易用性测评方法与测评过程。本部分基本涵盖了目前主流的易用性测评方法,对这些测评方法的过程进行了描述,给出了具体执行步骤,可以便于使用者操作。本部分所列的方法并非一个完备集,不包括尚在探索中的新方法。进行易用性测评的人员可以从本部分中选择合适的单一方法或者将多个方法进行组合,用来测量被测系统的易用性质量情况;亦可以修改或使用本部分未包括的其他方法。本部分适用于各种具有人机交互的计算机软件产品及相关系统,但并非每种方法适用于各种计算机软件产品及相关系统。

本标准预期的主要使用者包括:

- a) 需方(从供方获得或采购系统、软件产品或软件服务的个体或组织);
- b) 独立评价者(与软件无利益关系、受委托实施独立评价的个体或组织);
- c) 开发者(执行开发活动的个体或组织);
- d) 维护者(执行维护活动的个体或组织);
- e) 供方(按所签合同向需方提供系统、软件产品或软件服务的个体或组织),其在合格性测试中确认软件质量时使用;
- f) 用户(使用软件产品执行具体功能的个体或组织),其在验收测试中评价软件产品质量时使用;
- g) 质量管理者(执行软件产品或软件服务的系统性检查的个体或组织),作为质量保证和质量控制的组成部分评价软件质量时使用。

GB/T 29836.1—2013《系统与软件易用性 第1部分:指标体系》给出了易用性指标体系,GB/T 29836.2—2013《系统与软件易用性 第2部分:度量方法》给出了如何获得易用性指标测量值的度量方法。本部分旨在与 GB/T 29836.1—2013 和 GB/T 29836.2—2013 联合使用。

系统与软件易用性

第3部分:测评方法

1 范围

GB/T 29836 的本部分规定了如何获得易用性指标测量值的测评方法。

本部分适用于具有人机交互的各类系统与软件,本部分不适用于无人交互功能或者在使用期间无易用性需求的系统与软件。

注:本部分中所指的系统主要是软件系统。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件,仅注日期的版本适用于本文件。凡是不注日期的引用文件,其最新版本(包括所有的修改单)适用于本文件。

GB/T 11457 软件工程术语

GB/T 16260.1 软件工程 产品质量 第1部分:质量模型(GB/T 16260.1—2006,ISO/IEC 9126-1:2001,IDT)

GB/T 29836.1—2013 系统与软件易用性 第1部分:指标体系

GB/T 29836.2—2013 系统与软件易用性 第2部分:度量方法

3 术语和定义

GB/T 29836.1—2013、GB/T 16260.1 和 GB/T 11457 中界定的术语和定义适用于本文件。

4 概述

系统与软件易用性属于软件外部属性,通常需要用户参与测试。常见的易用性测试方法一般可分为数学模型、主观测评和客观测评三大类。主要的易用性方法的具体说明见第9章。

本部分中测试与测评所指意思一致,不再做明显区别。

附录A给出了网上评卷系统易用性测评示例。

5 易理解性

5.1 明显的功能

明显的功能测试要求如下:

- 前置条件
被测软件的需求已确定,并具备运行条件;
- 测试输入
需求规格说明书、用户文档、软件安装包;
- 测量类型

计数/计数；

——测试方法

用户测试法；

注：本部分提到的测试方法并不唯一，在实践过程中可使用本部分第9章所提到的测评方法或其他适用的方法。

——实施要点：

- 1) 宜选择普通用户进行测试；
- 2) 根据需求规格说明书和用户文档，对要求用户在培训前能够根据界面提示进行操作的功能数进行确定，计为 B ；
- 3) 采用用户测试法对这些功能数进行评价，找出能够在培训前被用户通过界面提示完成操作的功能数，计为 A ；
- 4) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$ ，计算出测量结果。

5.2 描述的完整性

描述的完整性测试要求如下：

——前置条件

被测软件的需求已确定，并具备运行条件；

——测试输入

需求规格说明书、用户文档、软件安装包、其他辅助类文档；

——测量类型

计数/计数；

——测试方法

评审法启发式评估；

——实施要点：

- 1) 宜选择行业专家用户进行测试；
- 2) 查看需求规格说明书和用户文档，确定软件的功能个数，计为 B ；
- 3) 采用评审法中的启发式评估，对用户文档进行阅读，确定通过阅读而理解的功能个数，计为 A ；
- 4) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$ ，计算出测量结果。

5.3 演示能力

演示能力测试要求如下：

——前置条件

被测软件的需求已确定，并具备运行条件；

——测试输入

需求规格说明书、用户文档、软件安装包；

——测量类型

计数/计数；

——测试方法

用户测试法；

——实施要点：

- 1) 宜选择普通用户进行测试；
- 2) 查看需求规格说明书，确定需要演示的功能数，计为 B ；
- 3) 采用用户测试法，对需要演示的功能进行操作，确定这些功能能够演示的个数，计为 A ；

- 4) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$, 计算出测量结果。

5.4 演示的有效性

演示的有效性测试要求如下:

- 前置条件
被测软件的需求已确定, 并具备运行条件;
- 测试输入
需求规格说明书、用户文档、软件安装包;
- 测量类型
计数/计数;
- 测试方法
评审法启发式评估、用户测试法;
- 实施要点:
 - 1) 宜选择行业专家、软件领域专家或普通用户进行测试;
 - 2) 查看需求规格说明书和用户手册, 确定能够演示的功能数, 计为 B ;
 - 3) 采用评审法启发式评估或用户测试法, 对能够演示的功能进行操作, 确定通过演示能够完成的操作的功能, 计为 A ;
 - 4) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$, 计算出测量结果。

5.5 输入的有效性检查

输入的有效性检查测试要求如下:

- 前置条件
被测软件的需求已确定, 并具备运行条件;
- 测试输入
需求规格说明书、用户文档、软件安装包;
- 测量类型
计数/计数;
- 测试方法
技术测试法;
- 实施要点:
 - 1) 宜由专业测试人员进行测试;
 - 2) 查看需求规格说明书和用户文档, 确定要求对输入数据进行有效性检查的功能数, 计为 B ;
 - 3) 采用技术测试法, 找出能够对输入数据进行有效性检查的功能数, 计为 A ;
 - 4) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$, 计算出测量结果。

6 易学习性

6.1 帮助文档的有效性

帮助文档的有效性测试要求如下:

- 前置条件
被测软件的需求已确定, 并具备运行条件;
- 测试输入

需求规格说明书、用户文档、软件安装包；

——测量类型

计数/计数；

——测试方法

评审法启发式评估；

——实施要点：

- 1) 宜选择行业专家、软件领域专家、易用性和人因工程学领域专家进行测试；
- 2) 查看需求规格说明书和用户文档，确定软件的功能个数，计为 B ；
- 3) 采用评审法启发式评估，对软件所附的帮助进行阅读，确定通过阅读帮助而理解的功能个数，计为 A ；
- 4) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$ ，计算出测量结果。

6.2 帮助机制的有效性

帮助机制的有效性测试要求如下：

——前置条件

被测软件的需求已确定，并具备运行条件；

——测试输入

需求规格说明书、用户文档、软件安装包；

——测量类型

计数/计数；

——测试方法

评审法启发式评估；

——实施要点：

- 1) 宜选择行业专家、软件领域专家、易用性和人因工程学领域专家进行测试；
- 2) 查看需求规格说明书和用户文档，确定软件的需要使用帮助功能的个数，计为 B ；
- 3) 采用评审法启发式评估，需要使用帮助功能的功能按照提供的帮助进行操作，能顺利完成操作的功能个数，计为 A ；
- 4) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$ ，计算出测量结果。

7 易操作性

7.1 使用中默认值的可用性

使用中默认值的可用性测试要求如下：

——前置条件

被测软件的需求已确定，并具备运行条件；

——测试输入

需求规格说明书、用户文档、软件安装包；

——测量类型

计数/计数；

——测试方法

评审法启发式评估；

——实施要点：

- 1) 宜选择行业专家、软件领域专家、易用性和人因工程学领域专家进行测试；

- 2) 根据需求规格说明书和用户文档,确定要求可供用户自己选择参数的功能个数,计为 B ;
- 3) 采用评审法启发式评估对这些功能进行评价,找出符合要求,用户可自己选择参数的功能个数,计为 A ;
- 4) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$,计算出测量结果。

7.2 完成指定任务的步骤

完成指定任务的步骤测试要求如下:

- 前置条件
被测软件的需求已确定,并具备运行条件;
- 测试输入
需求规格说明书、用户文档、软件安装包;
- 测量类型
计数/计数;
- 测试方法
用户模型法、专家评审走查法;
- 实施要点:
 - 1) 宜选择行业专家进行测试;
 - 2) 根据需求规格说明书和用户文档,通过用户模型法计算出完成每个任务所需要的步骤数目,并由专家确认这些数目,所有的任务个数,计为 B ;
 - 3) 采用专家走查法对这些任务进行实际操作,找出完成同样任务所需要的步骤数目;
 - 4) 过程 2) 中任务所需步骤数和过程 1) 任务所需的步骤数相一致的任务,统计出个数,计为 A ;
 - 5) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$,计算出测量结果。

7.3 操作的复杂性

操作的复杂性测试要求如下:

- 前置条件
被测软件的需求已确定,并具备运行条件;
- 测试输入
需求规格说明书、用户文档、软件安装包;
- 测量类型
计数/计数;
- 测试方法
用户模型法、评审法走查法;
- 实施要点:
 - 1) 宜选择行业专家、软件领域专家、易用性和人因工程学领域专家进行测试;
 - 2) 根据需求规格说明书和用户文档,通过用户模型法计算出完成指定某个任务所需要的动作(操作)次数,并由专家确认这些次数,所有的任务个数,计为 B ;
 - 3) 采用专家走查法对这些任务进行实际操作,找出完成同样任务所需要的动作(操作)次数;
 - 4) 过程 2) 中任务所需动作(操作)次数和过程 1) 任务所需的动作(操作)次数相一致的任务,统计出个数,计为 A ;
 - 5) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$,计算出测量结果。

7.4 完成指定任务过程中误操作的次数

完成指定任务过程中误操作的次数测试要求如下：

- 前置条件
被测软件的需求已确定,并具备运行条件;
- 测试输入
需求规格说明书、用户文档、软件安装包;
- 测量类型
计数/计数;
- 测试方法
用户模型法、用户测试法;
- 实施要点:
 - 1) 宜选择普通用户进行测试;
 - 2) 首先使用用户模型法确定需要测试任务的完成步骤数,计为 B ;
 - 3) 采用用户测试法,观察用户在完成该任务时误操作的步骤数,计为 A ;
 - 4) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$,计算出测量结果。

7.5 错误的纠正

错误的纠正测试要求如下：

- 前置条件
被测软件的需求已确定,并具备运行条件;
- 测试输入
需求规格说明书、用户文档、软件安装包;
- 测量类型
计数/计数;
- 测试方法
用户测试法;
- 实施要点:
 - 1) 宜选择普通用户进行测试;
 - 2) 首先使用用户测试法确定用户完成某项任务时出现错误的次数,计为 B ;
 - 3) 观察用户在完成该任务时纠正错误的次数,计为 A ;
 - 4) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$,计算出测量结果。

7.6 发生错误的影响力

发生错误的影响力测试要求如下。

- 前置条件
被测软件的需求已确定,并具备运行条件;
- 测试输入
需求规格说明书、用户文档、软件安装包;
- 测量类型
计时/计时;
- 测试方法
用户模型法、用户测试法;

——实施要点：

- 1) 宜选择普通用户进行测试；
- 2) 用户发生错误之后,对完成任务的影响程度,这个影响常常在实际中转化为可测得对时间的影响程度；
- 3) 测试该指标时,首先用用户模型法估算一下待测任务的完成时间,这个时间,计为 B ；
- 4) 采用用户测试法测试用户在完成任务过程中出现误操作后,完成任务所需要的时间,计为 A ；
- 5) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$,计算出测量结果。

7.7 可还原性

可还原性测试要求如下。

——前置条件

被测软件的需求已确定,并具备运行条件；

——测试输入

需求规格说明书、用户文档、软件安装包；

——测量类型

计数/计数；

——测试方法

技术测试法；

——实施要点：

- 1) 宜由专业测试人员进行测试；
- 2) 查看需求规格说明书和用户文档,确定界面状态能恢复到原状态的功能个数,计为 B ；
- 3) 采用技术测试法对这些界面进行评价,找出可恢复到原状态的功能个数,计为 A ；
- 4) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$,计算出测量结果。

7.8 运行差错的易恢复性

运行差错的易恢复性测试要求如下：

——前置条件

被测软件的需求已确定,并具备运行条件；

——测试输入

需求规格说明书、用户文档、软件安装包；

——测量类型

计数/计数；

——测试方法

评审法启发式评估；

——实施要点：

- 1) 宜选择行业专家、软件领域专家进行测试；
- 2) 查看需求规格说明书和用户文档,确定软件的功能个数,计为 B ；
- 3) 采用评审法启发式评估对这些界面进行评价,找出能够容忍误操作并能帮助用户恢复的功能个数,计为 A ；
- 4) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$,计算出测量结果。

7.9 使用中的消息的可理解性

使用中的消息的可理解性测试要求如下：

- 前置条件
被测软件的需求已确定,并具备运行条件;
- 测试输入
需求规格说明书、用户文档、软件安装包;
- 测量类型
计数/计数;
- 测试方法
用户调查问卷法;
- 实施要点:
 - 1) 宜选择普通用户进行测试;
 - 2) 查看需求规格说明书和用户文档,设计调查问卷;
 - 3) 发放问卷对使用过系统的人进行调查,调查人数,计为 B ;
 - 4) 问卷中,对系统消息可以理解的人数,计为 A ;
 - 5) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$,计算出测量结果。

7.10 运行状态的易监控性

运行状态的易监控性测试要求如下:

- 前置条件
被测软件的需求已确定,并具备运行条件;
- 测试输入
需求规格说明书、用户文档、软件安装包;
- 测量类型
计数/计数;
- 测试方法
技术测试法;
- 实施要点:
 - 1) 宜由专业测试人员进行测试;
 - 2) 查看需求规格说明书和用户文档,确定要求向用户提供运行状态查看的功能个数,计为 B ;
 - 3) 找出能够向用户提供运行状态查看的功能个数,计为 A ;
 - 4) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$,计算出测量结果。

7.11 界面元素的易定制性

界面元素的易定制性测试要求如下:

- 前置条件
被测软件的需求已确定,并具备运行条件;
- 测试输入
需求规格说明书、用户文档、软件安装包;
- 测量类型
计数/计数;
- 测试方法
评审法启发式评估;
- 实施要点:

- 1) 宜选择行业专家、软件领域专家、易用性和人因工程学领域专家进行测试；
- 2) 查看需求规格说明书和用户文档，确定要求用户可以改变的界面元素数量，计为 B ；
- 3) 采用评审法启发式评估对这些功能进行评价，找出用户能够改变的界面元素数量，计为 A ；
- 4) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$ ，计算出测量结果。

7.12 界面布局的易定制性

界面布局的易定制性测试要求如下：

- 前置条件
被测软件的需求已确定，并具备运行条件；
- 测试输入
需求规格说明书、用户文档、软件安装包；
- 测量类型
计数/计数；
- 测试方法
评审法启发式评估；
- 实施要点：
 - 1) 宜选择行业专家、软件领域专家、易用性和人因工程学领域专家进行测试；
 - 2) 查看需求规格说明书和用户文档，确定要求用户可以改变的界面布局的数量，计为 B ；
 - 3) 采用评审法启发式评估对这些功能进行评价，找出用户能够改变的界面布局的数量，计为 A ；
 - 4) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$ ，计算出测量结果。

7.13 快捷方式的易定制性

快捷方式的易定制性测试要求如下：

- 前置条件
被测软件的需求已确定，并具备运行条件；
- 测试输入
需求规格说明书、用户文档、软件安装包；
- 测量类型
计数/计数；
- 测试方法
评审法启发式评估；
- 实施要点：
 - 1) 宜选择行业专家、软件领域专家、易用性和人因工程学领域专家进行测试；
 - 2) 查看需求规格说明书和用户文档，确定要求用户可以自定义快捷键的功能个数，计为 B ；
 - 3) 采用评审法启发式评估对这些功能进行评价，找出用户能够自定义快捷键的功能个数，计为 A ；
 - 4) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$ ，计算出测量结果。

7.14 操作规程的易定制性

操作规程的易定制性测试要求如下：

- 前置条件

被测软件的需求已确定,并具备运行条件;

——测试输入

需求规格说明书、用户文档、软件安装包;

——测量类型

计数/计数;

——测试方法

评审法启发式评估;

——实施要点:

- 1) 宜选择行业专家、软件领域专家、易用性和人因工程学领域专家进行测试;
- 2) 查看需求规格说明书和用户文档,确定要求用户可以自定义操作规程的功能个数,计为 B ;
- 3) 采用评审法启发式评估对这些功能进行评价,找出用户能够自定义操作规程的功能个数,计为 A ;
- 4) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$,计算出测量结果。

7.15 特殊辅助功能

特殊辅助功能测试要求如下:

——前置条件

被测软件的需求已确定,并具备运行条件;

——测试输入

需求规格说明书、用户文档、软件安装包;

——测量类型

计数/计数;

——测试方法

用户测试法;

——实施要点:

- 1) 宜选择对特殊辅助功能有需要的普通用户;
- 2) 查看需求规格说明书和用户文档,确定要求提供辅助功能的种类数,计为 B ;
- 3) 经用户测试,能够提供辅助功能的种类数,计为 A ;
- 4) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$,计算出测量结果。

7.16 无障碍程度

无障碍程度测试要求如下:

——前置条件

被测软件的需求已确定,并具备运行条件;

——测试输入

需求规格说明书、用户文档、软件安装包;

——测量类型

计数/计数;

——测试方法

用户测试法;

——实施要点:

- 1) 宜选择对无障碍功能有需要的普通用户;

- 2) 查看需求规格说明书和用户文档,确定功能数,计为 B ;
- 3) 经用户测试,能够让有身体障碍的用户操作的功能数,计为 A ;
- 4) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$,计算出测量结果。

7.17 一致性

一致性测试要求如下。

——前置条件

被测软件的需求已确定,并具备运行条件;

——测试输入

需求规格说明书、用户文档、软件安装包;

——测量类型

计数/计数;

——测试方法

评审法启发式评估;

——实施要点:

- 1) 宜选择行业专家、软件领域专家、易用性和人因工程学领域专家进行测试;
- 2) 根据需求规格说明书和用户文档,确定系统内、外各部分需符合一致性的界面元素个数,计为 B ;
- 3) 采用评审法启发式评估对这些界面元素进行评价,找出符合界面一致性要求的元素个数,计为 A ;
- 4) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$,计算出测量结果。

8 吸引力

8.1 界面色彩对视觉的吸引力

界面色彩对视觉的吸引力测试要求如下:

——前置条件

被测软件的需求已确定,并具备运行条件;

——测试输入

需求规格说明书、用户文档、软件安装包;

——测量类型

计数/计数;

——测试方法

用户调查问卷法;

——实施要点:

- 1) 宜选择普通用户进行测试;
- 2) 查看需求规格说明书和用户文档,设计用户调查问卷,主要涉及色彩、形状、大小对视觉的舒适度;
- 3) 选定调查范围,发放调查问卷,确定调查人数,计为 B ;
- 4) 回收调查问卷,在确定有效的调查问卷中,统计认为舒适的人数,计为 A ;
- 5) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$,计算出测量结果。

8.2 界面元素形状的舒适度

界面元素形状的舒适度测试要求如下：

- 前置条件
被测软件的需求已确定,并具备运行条件;
- 测试输入
需求规格说明书、用户文档、软件安装包;
- 测量类型
计数/计数;
- 测试方法
用户调查问卷法;
- 实施要点:
 - 1) 宜选择普通用户进行测试;
 - 2) 查看需求规格说明书和用户文档,设计用户调查问卷,主要涉及界面元素形状是否舒适;
 - 3) 选定调查范围,发放调查问卷,确定调查人数,计为 B ;
 - 4) 回收调查问卷,在确定有效的调查问卷中,统计认为舒适的人数,计为 A ;
 - 5) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$,计算出测量结果。

8.3 界面元素尺寸的合理性

界面元素尺寸的合理性测试要求如下：

- 前置条件
被测软件的需求已确定,并具备运行条件;
- 测试输入
需求规格说明书、用户文档、软件安装包;
- 测量类型
计数/计数;
- 测试方法
用户调查问卷法;
- 实施要点:
 - 1) 宜选择普通用户进行测试;
 - 2) 查看需求规格说明书和用户文档,设计用户调查问卷,主要涉及界面尺寸的合理性;
 - 3) 选定调查范围,发放调查问卷,确定调查人数,计为 B ;
 - 4) 回收调查问卷,在确定有效的调查问卷中,统计认为合理的人数,计为 A ;
 - 5) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$,计算出测量结果。

8.4 布局的合理性

布局的合理性测试要求如下：

- 前置条件
被测软件的需求已确定,并具备运行条件;
- 测试输入
需求规格说明书、用户文档、软件安装包;
- 测量类型
计数/计数;

——测试方法

评审法启发式评估；

——实施要点：

- 1) 宜选择行业专家、软件领域专家、易用性和人因工程学领域专家进行测试；
- 2) 查看需求规格说明书和用户文档，确定界面布局和元素应符合的标准；
- 3) 根据需求规格说明书和用户文档，确定需符合标准的界面布局和元素个数，计为 B ；
- 4) 采用评审法启发式评估对这些界面布局和元素进行评价，找出符合标准的个数，计为 A ；
- 5) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$ ，计算出测量结果。

8.5 用户的感受度

用户的感受度测试要求如下：

——前置条件

被测软件的需求已确定，并具备运行条件；

——测试输入

用户文档、软件安装包；

——测量类型

计数/计数；

——测试方法

用户调查法；

——实施要点：

- 1) 宜选择普通用户进行测试；
- 2) 设计用户感受度调查问卷，分为满意和不满意；
- 3) 调查问卷的总发放数，计为 B ；
- 4) 统计调查结果为满意的问卷数，计为 A ；
- 5) 根据 GB/T 29836.2—2013 测量公式 $X=A/B$ ，计算出测量结果。

9 测评方法描述

9.1 方法的分类

易用性测评方法可分为数学模型、客观评价和主观评价三大类，如图 1。

数学模型是用构建数学模型的方式来模拟人机交互的过程。这种方法把人机交互的过程看作是解决问题的过程。此类方法适合于无法进行用户测试的情形。数学模型类方法主要包括用户模型法，GOMS 是目前最普遍使用的数学模型。用户模型法的实施过程参加见“9.3 用户模型法”。

客观评价分为两类，一类是用户测试法，一类是技术测试法。其中用户测试法是以用户行为作为观测对象，由测试人员分析判断并记录测试结果的一类方法。此类方法适合于可以由用户行为决定测试结果的情形。由测试地点的不同，用户测试法可分为实验室测试和现场测试；由于实验设计方法的不同，用户测试法可分为有控制条件的统计试验和非正式的易用性观察测试。用户测试法的实施过程见“9.4 用户测试法”。技术测试法是不需要分析判断用户的行为，而直接由专业测试人员进行测试的一种方法。技术测试法的实施过程见“9.5 技术测试法”。

主观评价是以系统的使用者（专家或普通用户）对被测系统的主观判断为最终测试结果的一类方法。此类方法适合于有用户主观评价决定测试结果的情形。主观评价类方法包括评审法和调查法两种。其中评审法根据具体实施方式的不同可细分启发式评估和走查法，其中走查法可分为认知走查和协作走查。评审法的实施过程见“9.6 评审法”。调查法根据具体实施方式的不同可细分为问卷调查法

和访谈法,调查问卷按照问卷方式可分为纸面问卷、网上问卷和电话问卷,访谈法可分为结构性访谈、开放化访谈和集体访谈。调查法的实施过程见“9.7 调查法”。

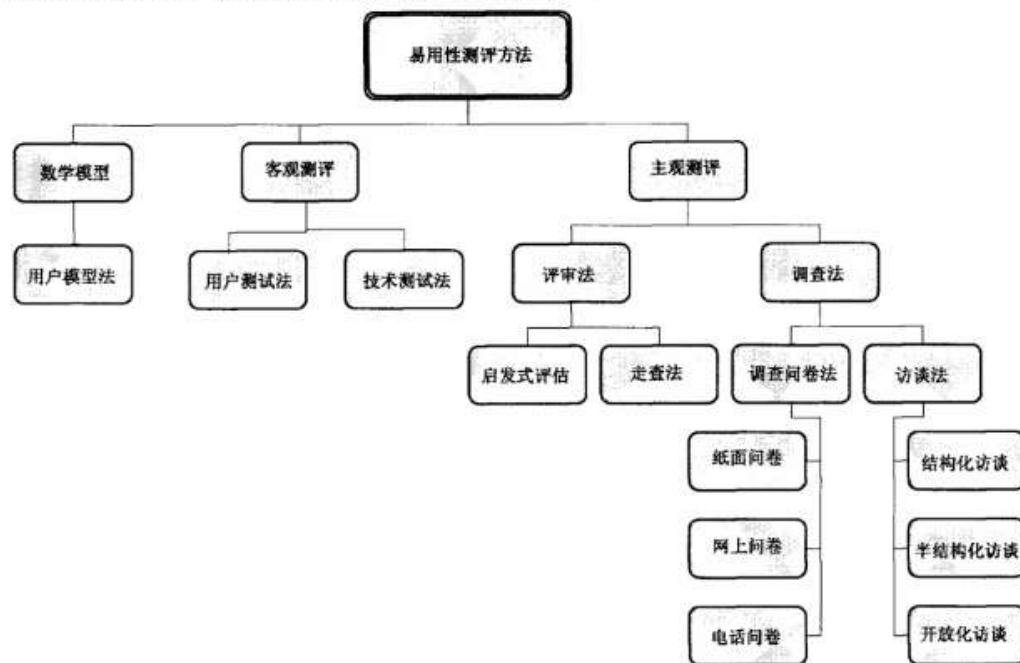


图1 易用性测评方法分类

9.2 方法实施过程的描述结构

方式实施过程的描述结构如下：

- a) 测试目的
使用该方法可以达到的测评目标；
- b) 测试准备
开始正式测评活动之前需要做的准备工作；
- c) 相关人员
参与测评活动的人员。相关人员包括：
 - 1) 专家用户：具备与被测系统相关领域的专业知识和经验的人员，在测评活动中可以提供专业评价，其具备的专业知识是测评活动的关键因素；
 - 2) 普通用户：参与测评活动，使用被测系统的普通人员，此类人员能够较为真实地体现最终用户的使用情况；
 - 3) 专业测试人员：实施整个测评活动的人员，并在测评活动中起主导作用。专业测试人员可以包括但不限于测试设计人员、测试实施人员、软件测试领域专家等；
- d) 测试活动
测评活动的具体实施步骤；
- e) 测试总结
通过测评的实施，可以获得的测评结果及相关情况总结。

9.3 用户模型法

9.3.1 测试目的

通过调查用户操作心理,分析用户行为,建立用户模型,将系统测试出的值与该模型计算出的值进行比对,据此判断所评价系统是否是高效且符合用户习惯的。

9.3.2 测试准备

测试准备应包括:

- a) 对用户操作心理的调查研究;
- b) 分析用户行为;
- c) 选择用户模型;
- d) 建立用户模型。

9.3.3 相关人员

人类因素学领域的专家用户。

9.3.4 测试活动

对用户操作心理的调查研究,最终体现在建立的用户模型中,建立用户模型时全面考虑人、机器、环境三者的相互作用,即非理性用户模型。建立用户模型需要考虑以下几个方面。

9.3.4.1 用户感知模型

它是从用户的基本感知特性总结出来的。计算机的输出信息以视觉和听觉等方式被眼睛耳朵等感知系统接收后,传输到感知处理器中进行理解。人机界面的布局、图标设计、选单设计都应该按照用户的感知特性进行设计,目的是为了减少用户的识别、记忆和理解困难。例如,图形用户界面的设计应当尽量减少用户不必要的眼球移动;设计易于浏览的格式和布局。

9.3.4.2 用户的认知模型

人类认知过程是由思维处理器与短时记忆和长时记忆器的协调工作完成的。首先,被人感知的视觉、听觉等信息被感知处理器处理后会有选择地被传送到短时记忆器中。短时记忆器中的部分信息也会有选择地被传送到长时记忆器中。用户的认知特性为界面设计提供了一些设计准则,例如,为了不超过短时记忆器的能力范围,在设计中应当尽量将大批的信息按照其相互关系分类组织起来,同时,人机界面的设计应当简单明了,避免在用户面前显示与任务无关的信息以分散注意力。

9.3.4.3 任务模型

用户的任务模型主要内容是用户操作使用计算机完成各种任务的行动过程。“任务”在心理学中被称为“行动”。按照动机心理学,一个行动包含4个基本过程:意向、计划、实施、评价。这一理论是“理性用户模型”。当前除了一直研究的生理和自然因素外,也开始更多关注情感、个性和动机等非认知因素。非理性用户任务模型的核心思想是不存在普遍适用的用户任务模型,针对每一个具体设计项目,都应进行用户行动过程的具体调查,系统了解他们的行动特性,建立具体的用户任务模型,按照用户的行动特性设计操作过程。

9.3.4.4 任务出错模型

用户在操作过程中会出错。这往往是由两个原因造成的:首先,出错是人心理固有的特性,人天生

就会出错,例如,不小心误操作,思维不连续,忘了预定操作,多个任务导致分心等,疲劳导致注意分散,属于人心理固有的容易导致出错的因素;人机界面设计应该调查用户本身容易出错的特性,设法减少用户出错的机会。其次,如果人机界面设计不符合用户的任务习惯,不符合用户思维方式等,很容易引起用户出错。

9.3.4.5 用户学习模型

设计研究用户的学习过程,就是要设法改变人机界面设计,以减少学习困难。例如,人们学习和解决问题是经常会联想到类似问题和解决方法,所以使用比拟或暗喻等方法设计学习材料和界面元素有利于用户对信息的分析和掌握。

评价者依照以上模型来判断系统是否存在易用性问题。

9.3.5 测试总结

测试总结的内容包括所建立的用户模型以及根据用户模型所计算出数值、测试系统所得出的值,并对以上两个值进行对比和分析。

9.4 用户测试法

9.4.1 测试目的

将软件完全处于用户使用状态下进行测试,其目的是为了能够得到用户使用的真实感受,并能够发现易用性需求的问题。

9.4.2 测试准备

在准备阶段,实验室人员应该确保测试房间已经准备好,计算机系统已经处于测试计划的初始状态,所有的测试材料、指南和问卷都已经准备就绪。

为使用户的不安和紧张减至最低,所有的准备工作都要在用户到来之前完成。

要关闭所有的屏幕保护程序和其他会干扰实验的系统组件,如病毒软件的扫描提示等。

9.4.3 相关人员

普通用户、专业测试人员。

实验员的人数一般应等于或者大于用户人数。

9.4.4 测试活动

9.4.4.1 介绍

介绍时,实验室人员向测试用户表示欢迎,并对测试目的做简短解释。如果用户对计算机配置不熟悉,实验室人员还需对其进行说明。

接下来介绍实验过程,主要包括以下因素:

- a) 测试目的是对软件进行评价,而非对用户个人进行测试;
- b) 除非专业测试人员就是系统设计人员,否则专业测试人员应该对用户提到他与被测系统没有关系,因此用户可以自由表达意见,而不必担心伤害到专业测试人员的感情。如果专业测试人员就是系统设计人员,不要将这个事实说出来,以免得到相反的结果;
- c) 测试结果将用于改进系统,因此测试时的系统与最终发布的系统可能会有所不同;
- d) 提醒用户对所测试的系统进行保密,不要与他人谈论有关内容;
- e) 说明参与测试是自愿的,用户可以随时停止测试;

- f) 让用户放心,对测试的结果会保密;
- g) 对于使用的任何录音录像设备向用户进行说明和解释;
- h) 向用户说明欢迎提问,但是专业测试人员不会在测试时回答他们的问题;
- i) 关于实验的任何特殊要求,如需要大声说出感受,或者需要保证尽可能迅速不出错地完成既定目标的要求等;
- j) 回答在实验开始之前用户提出的需要澄清的任何问题。

如果使用单向玻璃观察室进行观测,在介绍结束之后需要带领用户参观观察室并作简要说明。

9.4.4.2 测试执行

在测试期间,专业测试人员通常不要与用户进行交流,也不要有任何个人观点或关于用户操作好坏的表示。专业测试人员可以发出一些不代表什么态度的词汇,如“嗯”、“哦”,来表示知道了用户的评论,并且让用户继续下去,应注意声调和语气,不能让用户提出专业测试人员的情绪和判断。

当用户陷入了困境,如果测试的目的是记录用户能否独立完成任务或者完成某个任务的时间,专业测试人员不能给用户提供任何帮助;当用户已经明显停滞,并对当前处境感到不快的时候,由指定的专业测试人员提供简单提醒式的帮助,并且专业测试人员应清楚用户遇到的问题是什么,以及用户随后还会遇到什么样的问题;当用户遇到前面几个用户多次碰到的同样的问题时,对用户提供帮助。应杜绝过早和过多地提供帮助。

在多人同时观察测试的情况下,应预先制定其中一人为正式的专业测试人员,只有他(她)才可以在测试中向用户提供指令和进行交谈。为了不干扰用户,即使其他观察人员有任何不同意见,也要保持缄默。确实需要发表意见,可以以纸条方式或者休息时交流,不影响用户是应遵循的原则。

9.4.4.3 事后交流

在测试之后,要询问用户,并要求用户填写一份主观满意度的问卷。为了避免专业测试人员的评论所带来的偏向,用户应当在有关系统的讨论之前填写问卷。

在交谈中,请用户对系统的使用情况进行评论并提出改进建议。这些建议并不一定要对具体的设计改动起到作用,但是用户建议可以作为重新设计时构思的丰富来源。因此专业测试人员不能对建议进行褒贬评价。

在询问时,专业测试人员也可以询问用户在测试过程中发生的专业测试人员难以理解的事件和举止进行进一步的解释。即使用户可能不记得他们那么做的原因,但是可以通过他们的陈述获得用户的假设和目标。

最后,测试用户离开后专业测试人员尽快检查测试结果,对用户编号、计算机记录、问卷和表格进行标识,并整理专业测试人员自己的笔记。撰写完整的测试报告。

9.4.5 测试总结

包括用户的任务完成情况、问卷情况及计算机记录情况等的内容汇总与分析。

9.5 技术测试法

9.5.1 测试目的

由专业测试人员进行测试,其目的是为了能够获得更为直接、准确的测试结果,并能够发现易用性设计的问题。

9.5.2 测试准备

在准备阶段,专业测试人员应该确保测试房间已经准备好,计算机系统已经处于测试计划的初始状

态,所有的测试材料、指南和问卷都已经准备就绪。

要关闭所有的屏幕保护程序和其他会干扰实验的系统组件,如病毒软件的扫描提示等。

9.5.3 相关人员

专业测试人员。

9.5.4 测试活动

9.5.4.1 用例设计

根据测试需求,确定易用性的功能点。设计测试步骤和测试方法。

9.5.4.2 测试执行

按照用例设计文档,在已准备好的测试环境中由专业测试人员完成测试。

9.5.5 测试总结

记录、汇总并分析测试中出现的易用性问题

9.6 评审法

9.6.1 启发式评估

9.6.1.1 测试目的

评估人员依据已经确立认可的易用性规则,发现界面设计的可用性问题,并将其作为再设计的重点。

9.6.1.2 测试准备

测试准备应包括:

- a) 确认被测系统可以及时有效的提供其被使用过程中的状态;
- b) 确认系统符合用户的认知加工习惯;
- c) 用户应可控制和自由使用系统;
- d) 必要的帮助提示与说明文档。

9.6.1.3 相关人员

专家用户。采取“角色扮演”的方法,模拟典型用户使用产品的情形,从中找出潜在的问题,不需要普通用户参与。

参与者一般为不同领域的专家用户,如:软件领域专家、行业专家、易用性和人因工程学等领域的专家。在人数上建议采用3~5人。

9.6.1.4 测试活动

9.6.1.4.1 评价界面

由评价者单独对界面进行评价。评价者根据启发式浏览界面至少两次,一次集中于过程,另一次集中于个别的对话成分。评价者也可以考虑其他易用性原则或根据自己的经验、习惯对界面进行评价。一般来说,完成对一个界面的启发式评估需要1 h~2 h。在评价完成之后,评价者还需完成一份问卷,用来收集等级评定信息以帮助确定所发现的易用性问题的优先级。

9.6.1.4.2 结果的记录和整理

对评价结果进行记录和整理有两种方式:每一个评价者都单独完成一份书面报告,或评价者在评价过程中进行出声报告,由现场观察员现场记录。两种方式各有利弊。书面报告的优点是提供了正式的评价记录,但需要评价者付出额外的努力,最后还需要人员对报告进行阅读并将信息汇总。使用观察的方法增加了每一个评价单元的开支,但是减轻了评价者的工作负担,提高效率。

9.6.1.4.3 讨论阶段

最后,在个人评价之后,可进行一次小型的讨论。讨论阶段以头脑风暴的模式进行,主要集中于重新设计的问题,报告主要的易用性问题以及设计方面存在的一般问题,并提出一些设计和改进的建议。

9.6.1.4.4 评价的成果

启发式评估不是提供一种系统的方法来对产生的易用性问题进行修正,也不是对重新设计的质量进行评估的方式,而是利用启发式解释每一个观察到的易用性问题,当易用性问题被确认时,它们也同时得到相当明显的修正。

9.6.1.5 测试总结

使用启发式评估方法的测试总结是易用性问题的列表,每一个问题后面还注明了评价者认为它所违反的易用性原则。

9.6.2 走查法

9.6.2.1 测试目的

是通过分析用户的心理加工过程逐步检查使用系统执行的过程,从中找出易用性问题。

9.6.2.2 测试准备

测试准备应包括:

- a) 定义用户群;
- b) 选择样本任务;
- c) 确定任务操作的正确序列;
- d) 确定每个操作前后的界面状态;
- e) 样本任务的确定;
- f) 正确操作步骤的制定;
- g) 按照序列确定界面的状态。

9.6.2.3 相关人员

参与者一般为三个领域的专家用户:软件程序领域或易用性、行业领域和人的因素领域的专家。在人数上建议采用3~5人。

9.6.2.4 测试活动

9.6.2.4.1 构建故事

分析阶段中,分析者按照正确的序列操作,比较每一操作前后的状态,以确定用户按照序列操作的可能性有多大。在考虑到用户固有知识背景的基础上构建一个“可信的故事”解释为什么设想的用户会

选择正确的操作步骤完成任务,或者构建一个“失败的故事”解释为什么用户没有按照正确的操作步骤进行下去。分析者构建故事的根据是走查过程中对以下四个问题的回答。

a) 用户会努力获得正确的结果吗?

这一点所关注的问题是用户的意图如何,思路是否清晰,如果用户不试图获求正确操作所要求的步骤,那么他们不可能选择正确的操作。比如,用户的任务是打印一篇文档,但是系统要求在打印之前要选择好打印机,此时分析者就要回答用户是否知道需要选择打印机。

b) 用户会注意到正确的操作步骤或方式吗?

对于直接的显而易见的操作如选单和按钮,一般来说用户使用起来不会有什么问题,但是像双击这样的操作,用户就不容易觉察到。分析者在走查的过程中对这些不明显的操作应予以特别的关注。

c) 用户会把正确的操作步骤与努力获得的结果联系起来吗?

即使用户努力获求正确的结果,也能看到可用的正确操作,但是如果用户不知道通过其所意识到的操作步骤可以获得想要的结果,依然会使任务的完成受到影响。比如:用户看到显示有“select printers”的选单条目,就能很清楚的知道通过此选单的操作可以获得其想要的打印机;但是如果选单显示为“SysP”,即使用户知道在打印文档之前应该选择打印机,而且知道下一操作为“SysP”选单,由于用户很难把二者联系起来,同样会影响任务的完成。

d) 在正确的操作之后,用户是否能得到合适的反馈?

这一问题关注的是界面能否提供正确的及时的反馈,由于没有得到反馈,用户可能会取消正确的操作而去进行他们所认为的更为“合适的”操作。所以分析者在走查的过程中要检测是否由于界面反馈的问题而影响了任务的完成。

9.6.2.4.2 重要信息的记录

分析者在走查过程中除了记录准备阶段和实施阶段必要记录的信息之外,有两类信息应予以特别的关注:对于一个基于用户拥有特定知识的假设上被认为是成功的操作,如果怀疑设想的用户群是否具有这样的知识对用户有这样的知识,这样的假设应记录下来;对于影响任务完成的操作,分析者应记录所有问题及认为有问题的原因,不能仅仅关注对CW的4个问题的回答,比如用户没有把正确的操作和它的结果联系起来可能不是因为实现该操作的按钮不好,而是有其他按钮看起来与其一样好,所以记录应尽可能的全面。

9.6.2.4.3 后续工作

如果构建的故事是“可信的”,那么这部分的工作就是简单的整理了。但实际上,走查不可能没有发现任何界面设计的问题,也就是说总会或多或少的存在“失败的故事”,所以评价者在这一阶段应该给设计者整理出一个指导性的建议,为他们前期界面的修改和后期界面的再设计提供依据。

9.6.2.5 测试总结

记录所有易用性问题记录及认为有问题的原因。

9.7 调查法

9.7.1 调查问卷法

9.7.1.1 测试目的

目的在于收集用户意见并对其结果进行统计分析。

9.7.1.2 测试准备

测试准备应包括：

- a) 确定需要调查研究的问题和目的；
- b) 了解整个测试过程；
- c) 设计调查问卷，设计时既要考虑与整个测试流程相呼应，又要充分考虑参与测试的用户的特点，使其能够顺利准确的回答问卷上提出的问题；
- d) 当参与测试用户人数较多时，需要培养对问卷调查目的和内容都很熟悉的工作人员；
- e) 帮助测试人员学习调查问卷设计目的及填写方法。

9.7.1.3 相关人员

普通用户。

专业测试人员，至少应包括实验观察人员、调查问卷的设计人员。

9.7.1.4 测试活动

测试活动应包括：

- a) 调查问卷一般要求用户在测试中或测试后填写，故在测试工作开始前向用户说明调查问卷的目的，并向用户讲解问卷的填写方法及问卷的提交时间和方式。当测试用户地域分布较广或不便于面对面交流的时，也可将调查目的及填写方法注在问卷的开始部分；
- b) 在填写问卷期间，当测试用户集中在一个试验室内，工作人员应主动观察用户是否有困难填写问卷，并积极帮助用户填写问卷，但在帮助过程中不应引导用户向某一方向填写。当测试用户地域分布较广或不便于面对面交流时，工作人员应在测试用户提交问卷时，主动询问用户在填写过程中是否有困难或者是某些题目不易填写；
- c) 问卷回收后，保留有效问卷，并应根据测试报告的要求将有效问卷分类汇总，得出测试报告。用户在体验被测试的软件产品后，被邀请进行问卷调查，可根据问卷形式的不同，测试活动也会在以上活动步骤上有所调整；
- d) 纸面调查：
将调查题目呈现在纸介质上，要求用户在体验被测试的产品后填写相关调查题目。这种调查方式实现简单，调查者也可以较好的控制调查过程，可以得到比较准确的用户反馈信息；
- e) 电话问卷：
电话问卷方式在商业领域的售后服务中应用比较广泛。在信息用户调查中也有一定的应用。这种调查方式具有口头交谈方式调查的直接性，但需要投入较大的调查人力、物力，因此，此种方法应用范围比较有限；
- f) 网上问卷：
网上问卷方式是在传统的问卷调查方法的基础上，通过计算机网络的电子交互平台，以网页的形式发布调查问题，用户通过表单的形式提交自己的答案。这种用户调查方式可以不受时间的限制，用户可以 24 h 随时获得问卷并参与调查，并且可以在用户填写调查问卷前，先通过表单的形式采集用户的基本个人信息，并利用数据库对用户基本信息和反馈的问卷信息进行管理，这样在调查后期可以通过系统对信息进行自动统计分析处理。此外，通过计算机网络的使用，可以大大提高调查信息的传递效率和安全性，并可以为用户提供定制的信息。

9.7.1.5 测试总结

将试卷回收后所做的分类汇总结果，根据测试实施者的测试目的进行统计分析，并将结果汇总为测

试总结,以客观、清晰的方式向测试实施者汇报本次问卷调查的结果。

9.7.2 访谈法

9.7.2.1 测试目的

目的在于同用户交流软件使用感受,收集用户意见并对其结果进行统计分析。

9.7.2.2 测试准备

测试准备应包括:

- a) 明确访谈目的,并列出访谈的问题;
- b) 选择用户专家,并请其体验被测试的软件产品;
- c) 营造与用户面对面交流的轻松环境,使用户在心情舒畅的环境下客观地与访谈者进行交流。

9.7.2.3 相关人员

访谈者为软件易用性专家或测试实施者,被访谈者为专家用户或普通用户。

9.7.2.4 测试活动

用户在体验被测试的软件产品后,被访谈者邀请进行面对面的交流,由于各种访谈类型的有所不同,相应的测试活动也会不同:

- a) 结构化访谈又名标准化访谈,是一种定量研究方法,这种方法的目的在于确保对每一个被访谈者精确地呈现以同样的顺序出现的同样的问题,确保答案总体上可靠,并确信不同样本群之间或不同测量周期之间具有可比性。即同样的问题按照同样的顺序呈现给被访谈者,并尽可能用同样的言辞;
- b) 半结构化访谈是将开放式访谈与结构化访谈相结合的一种方式。访谈者可在被访谈者回答事先预定好的题目后,和被访谈者在进行交流,这时可不必采用标准化访谈严格的形式而采用开放式访谈灵活的交谈方式;
- c) 开放化访谈又称为非标准化访谈,深度访谈,自由访谈,它是一种无控制或半控制的访谈。在进行此类访谈时,要求访谈者引导被访谈者对其所关心的问题畅所欲言,访谈者尽量不要打断被访谈者,充分发挥访谈双方的主动性,积极性,灵活性和创造性。访谈者要对被访谈者的发言进行记录。

9.7.2.5 测试总结

对于结构化访谈和半结构化访谈中的标准化访谈内容,需进行分类汇总并对其结果进行统计分析,生成测试总结,以客观、清晰的方式向测试实施者汇报本次访谈的结果。

对于开放式访谈和半结构化访谈中的自由发言内容,需进行缩减,提炼出测试实施者所需的内容,可将内容直接整理为测试总结,也可再次对内容进行分类汇总并对其结果进行统计分析,生成测试总结。

10 测评中用户的选择

10.1 选择原则

通过对参与测试的用户进行评价和分类,确定参与测试的用户自身的能力和本能极限,以降低用户的差异性对度量结果的负面影响,增加度量的准确性。包括:

- a) 计算机技能:主要考察用户具备的计算机使用能力,此类考察可以用于评价用户是否适合参加常规系统的使用,或者是否适合参加针对初学者的计算机系统与软件培训类软件;
- b) 业务经验:主要考察用户是否具备相关的业务知识,是否能够正确理解业务逻辑,此类考察可以用于评价用户是否适宜参加与某些行业应用关联度较大的系统,例如财会系统软件,或者是否适合参加多产品易用性比较分析类型的测试;
- c) 认知能力:主要考察用户认识能力和实践能力所达到的水平。这些能力是用户完成系统操作所应具备的一般能力,某种活动才需要的特殊能力如没有特殊需求,不在测评之列。

10.2 计算机技能

10.2.1 基本知识

基本知识可以考察被测用户对计算机系统相关基本知识的熟悉程度,包括三项可选内容:

- 计算机硬件了解程度,是指被测用户对计算机常用硬件了解的程度,例如对 CPU、内存、显示器、鼠标、键盘、网卡、打印机、扫描仪等硬件的了解;
- 计算机软件了解程度,是指被测用户对计算机常用软件的了解程度,例如对系统软件、办公软件、杀毒软件、绘图软件、媒体播放软件等软件的了解;
- 网络知识了解程度,是指被测用户对网络知识了解程度,例如对浏览器、FTP、BBS、传输速度、搜索引擎等网络知识的了解。

10.2.2 单机操作

单机操作可以考察用户对于计算机单机系统使用的熟悉程度,包括五项可选内容:

- 键盘熟练程度,是指被测用户使用键盘的熟练程度,例如对键盘布局的熟悉程度、操作速度、快捷键和组合键的运用能力等;
- 鼠标熟练程度,是指被测用户使用鼠标的熟练程度,例如对鼠标的单击、双击、鼠标滑轮的使用、右键的使用等;
- 其他必要外设使用熟练程度:是指被测用户对于特定系统要求的外设的使用熟练程度,例如打印机、扫描仪、外接存储设备、音视频视频外设等;
- 系统软件使用熟练程度:对操作系统使用的熟练程度,例如登陆、设置密码、拷贝/粘贴/删除文件、硬件驱动安装、安装/卸载软件、系统参数查询与设置等;
- 普通应用软件使用熟练程度:对常用的应用软件使用的熟练程度,例如使用系统软件、办公软件、杀毒软件、绘图软件、媒体播放软件等。

10.2.3 网络应用

网络应用可以考察用户对于计算机网络系统使用的熟悉程度,包括两项可选内容:

- 局域网络应用能力,是指被测用户在局域网络中连通服务器、其他终端机的能力,例如网上邻居查找、网络打印机使用、文件共享等;
- Internet 网络应用能力:是指被测用户连通 Internet 网络,并使用 Internet 各种应用的能力,例如浏览网页、收发电子邮件、文件传输、信息搜索、电子交易等。

10.3 业务经验

10.3.1 业务熟悉程度

业务熟悉程度可以考察用户对于被测计算机系统相关的行业和业务的熟悉程度,包括两项可选内容:

- 行业背景,是指用户与被测计算机系统涉及的行业的相关程度,例如,用户所从事的工作,所处的环境、经历等是否与被测软件涉及的行业相关;
- 业务流程了解程度,是指对被测计算机系统涉及的业务流程的熟悉程度,例如,用户从事被测软件涉及的业务的时间,是否通过业务考试、具有职称的级别等。

10.3.2 产品使用经验

产品使用经验可以考察用户对于与被测计算机系统具有相同功能和操作逻辑处理流程的系统的使用经验,包括两项可选内容:

- 同类软件产品使用经验,是指用户是否使用过同类软件产品,使用过多少种同类软件产品,例如描述与被测软件涉及业务相同业务、相关业务的软件产品;
- 本产品使用经验,用户是否使用过本产品,例如使用、学习或者试用过本产品,或使用过本产品的其他版本。

10.4 认知能力

10.4.1 逻辑判断力

逻辑判断力可以考察用户借助于概念、判断、推理反映现实的能力,测评用户是否具备一种确定的,而不是模棱两可的;前后一贯的,而不是自相矛盾的;有条理、有根据的思维,包括可选内容:

- 分析,是指测试用户能够对系统功能理解、并且分析、进一步使用的能力;
- 判断,是指测试用户是否有能力判断该问题是有好的影响还是有不好的影响;
- 综合,是指测试用户有将测试中遇到的情况、问题、以及学习的内容加以综合,举一反三;
- 归纳,是指测试用户能够根据测试中所遇到的情况和发现的问题,总结出系统所存在的问题。

10.4.2 记忆能力

记忆能力可以考察用户识记、保持、再认识和重现客观事物所反映的内容和经验的能力,包括三项可选内容:

- 程序记忆,是指用户对习得行为和技能的记忆,包括基本条件反射和各种习得的动作,如打字、鼠标操作、快捷键使用时的动作等,这种记忆通过动作来表达;
- 语义记忆,是指用户关于基本事实性知识的记忆,包括所掌握的有关字词或其他语言符号、其意义和指代物、它们之间的联系,以及有关规则、公式和操纵这些符号、概念和关系的算法的有组织知识;
- 情景记忆,是指测试用户可将测试时所发生错误,还原出来,供测试方进行观察。

10.4.3 理解能力

理解能力指标可以用于测评用户在考查的知识范围内,在再认或再现的基础上对有关知识的转换能力或解释、推理能力,包括推五项可选内容:

- 推断,是指被测用户界定问题的能力,例如,在测试过程中能在运行结果中区分正确或错误推断的程度;
- 假设识别,是指用户识别在某个陈述中未表达的假设或构想,即用户能够理解系统所描述的业务及场景;
- 推论,是指决定在某个陈述中是否必然地得出某种结论,例如用户能够根据系统的界面设计、消息提示等来进行下一步的操作,并能够完成该项任务;
- 解释,是指用户在测试过程中能在运行结果中区分正确或错误推断的程度;

- 论据评价,是指测试用户可以向测试方指出存在的问题,并给出存在问题的证据。

11 易用性测试环境

11.1 预期效果

搭建易用性测试环境的预期效果应满足:

- a) 硬件投入除计算机外,其他设备节俭开支;
- b) 参与人机器的性能几乎没有任何影响;
- c) 实现一次性参与人机器屏幕、参与人面部表情、实验室场景(参与人动作)同步记录,不需要后期合成,而且视频上还有时间记录显示。

在满足以上3点的基础之上,再考虑可以随意将这个实验室携带到参与人环境。

11.2 应具备的条件

易用性测评环境应至少包括:

- a) 安排一个安静的空间;
- b) 具有“正在测试,请勿打扰”的标签;
- c) 切断电话线以及传真和手机;
- d) 保证充足的光线。

11.3 可供选择的设备

根据不同需求,实验室建设还可能需要有:

- a) 测试用计算机,显示器;
- b) 摄像机及三脚架,摄像机专用话筒;
- c) 话筒数个,由于时常需要将话筒的位置调到嘴的旁边,所以一个单独的话筒是不够的;
- d) 单向镜子,台灯或者摄像灯,用来观看摄像机图像的监视器;
- e) 录像放映机、录像软件,电线和电源插座;
- f) 用于隔离测试室和观察室的单向玻璃;
- g) 喇叭;
- h) 动作捕捉器和人眼信息捕捉器;
- i) 其他。

附录 A (资料性附录)

网上评卷系统易用性测评示例

A.1 引言

网上评卷系统易用性测评体系作为本部分在实际应用中的案例,该体系分为指标体系、测评方法、评价方法三部分。网上评卷系统指标体系已在 GB/T 29836.1—2013 附录 B 建立,评价方法已在 GB/T 29836.2—2013 附录 A 进行详细分析与介绍。本附录将给出如何通过测评获取部分指标值的示例。

A.2 网上评卷系统易用性的测评步骤

软件易用性的测试和测评最具权威的结果应该来自用户,并且在用户的实际工作任务和操作环境下产生的。根据这一特点,借助软件测试中的用户测试方法来帮助完成对网上评卷软件易用性的测评。具体步骤如下:

- 建立网上评卷系统易用性测评指标体系,这实际上是提炼归纳测评内容的过程。这是非常关键的一步,也是整个测评方法的重中之重。该步骤已在第四章中进行了讨论;
- 选择测评各测评指标的测试方法;
- 收集测评数据,进行综合分析,得出测评结果。

步骤如图 A.1 所示。



图 A.1 网上评卷系统易用性的测评步骤

A.3 网上评卷系统易用性测评指标值的测评思路

真实系统中得到的值与预期达到的值比例,预期达到的值不是最低值,而是理想值,测试得到的值越接近该值,说明该指标越好。

理想值由评审法得到,与传统方法不同,是根据专家的专业经验和使用系统的经验确立这些值。

专家一般由评卷教师、评卷工作管理者及有网上评卷工作开发经验的人员组成。

A.4 测试人员的筛选

根据网上评卷工作特点及测评工作需要,我们将参与测试的人员分为评卷的管理者(测评方)、软件开发方(被测评方)和评卷教师(用户)组成,他们分别有各自的工作内容和关注点,如表 A.1 所示。

表 A.1 网上评卷系统易用性测评相关人员

人员	工作内容	关注点
评卷的管理者 (测评方)	建立测评指标、制定测评计划,并对采集回来的测评数据进行整理和分析	1. 测评过程中用户的表现; 2. 测评结果
软件开放方 (被测评方)	提供被评价的软件产品,搭建测试环境	1. 测评过程中系统的表现; 2. 测评结果
评卷教师 (测试用户)	按照评价方制定的测评计划测试相关功能,并按要求记录测试结果	系统是否能帮助其完成相关工作,并且使用感觉舒服、顺畅

为使测评结果更加准确且符合实际,需要将评卷教师(测评用户)根据职责分成不同的小组,完成不同的任务。具体分组情况如表 A.2 所示。

表 A.2 网上评卷教师的分组依据

小组名称	所在人数比率	工作特点	小组成员特点
填空题组	40%	答案比较简单,容易判断	<ul style="list-style-type: none"> • 年龄在(25~30)岁之间; • 评卷经验较少,甚至有部分从未评过卷; • 有较强的计算机使用能力和丰富的使用经验
作文组	50%	答案内容较长,需要较长的判断时间	<ul style="list-style-type: none"> • 年龄在(30~45)岁之间; • 都有评卷经验; • 计算机使用能力和经验普遍较逊于填空题组评卷教师
专家组	10%	对所有评卷员的工作质量进行监控,对异常情况要立即处理。该工作需要非常丰富的教学和评卷经验	<ul style="list-style-type: none"> • 年龄在(50~60)岁之间; • 计算机使用能力和经验较差

A.5 产品指标的获取方法

A.5.1 外观指标

A.5.1.1 布局

布局的合理性

该指标主要测评网上评卷系统界面的布局及元素是否符合评卷工作要求,是否便于评卷教师操作。

在测评该指标时,首先需要确定那些界面布局及元素要符合何种标准,再用评审法中的启发式评估法,来确认这些布局或元素是否符合标准,即确认易用性原则,符合标准的布局 and 元素数量占要求的数量的比例越高,该指标值越好。故测量公式为:

$$X=A/B$$

A 是符合评卷工作要求及操作方便性要求的界面布局及元素数;

B 是要求应符合评卷工作要求及操作方便性要求的界面布局及元素数。

$$0 \leq X \leq 1$$

X 越接近 1, 表明该评卷系统界面布局越符合评卷工作要求且操作方便性好。

界面布局及元素需要符合的易用性原则, 以及通过评审法启发式评估对这些布局及元素的评价结果, 如表 A.3 所示。

表 A.3 界面布局及元素易用性原则

序号	界面布局及元素	易用性原则	符合原则的布局及元素
1	答题显示区	应作为评卷界面的主体, 清晰, 醒目	符合
2	给分区	在设计时应考虑与答案显示区相邻, 并摆放位置符合评卷教师自上而下看卷的试卷习惯。即不要让评卷教师看试卷后还需要寻找给分区, 这会给评卷教师造成视觉疲劳	符合
3	回评、设置雷同卷、统计功能按钮	应便于评卷教师查找, 但不得对评卷教师造成视觉干扰	符合
4	评卷教师工作信息, 工作时间、评卷份数、速度	该信息评卷教师隔一段时间就会查看, 应便于评卷教师查找, 但不得对评卷教师造成视觉干扰, 该项目的显示明显性应次于回评、设置雷同卷、统计功能按钮	符合
5	评卷教师个人信息	应便于评卷教师查找, 但不得对评卷教师造成视觉干扰。该项目的显示明显性应次于回评教师工作信息	符合
6	培训	1. 专家应能比较轻松地本别出哪些评卷教师状态, 例如通过培训、正在培训中、未通过培训等。 2. 专家在查看用户状态后, 应可以方便地对其进行操作, 例如使某用户通过培训、终止某用户的评卷工作等	符合
7	查看工作进度	1. 专家应能首先能够查看到工作的整体进度。 2. 进而可以查各大题的工作进度。 3. 之后又可查看到每小道的进度。 设计原则应符合由粗到细的原则, 在查看某道小题时应也能识别出其属于哪个大题	符合
8	查看工作情况	1. 专家首先能够查看到各工作组的工作情况。 2. 进而可查到每位评卷员的工作情况。 设计原则应符合由粗到细的原则, 在查看某评卷员时应该也能够识别出其属于哪组	符合

根据表 A.3 所示, 可以得出测量公式中的 $A=8, B=8$ 。

A.5.1.2 一致性

一致性

该指标主要测评网上评卷系统的界面元素样式是否符合一致性要求。

测评方法同“布局的合理性”指标, 故测量公式为:

$$X = A/B$$

A 是符合一致性要求的界面元素数量;

B 是要求必须符合一致性要求的界面元素的数量。

$$0 \leq X \leq 1$$

X 越接近 1, 表明该评卷系统界面元素一致性越好。

界面元素需要符合一致性要求, 以及通过评审法启发式评估对这些界面元素的评价结果, 如表 A.4 所示。

表 A.4 界面元素一致性要求

序号	界面元素	一致性要求	是否符合
1	答题情况显示区	各题目的答题情况显示区布局应一致	是
2	评分界面	各题目的评分界面布局应一致	是
3	统计界面	各统计界面风格应一致	是
4	查询界面	各查询界面风格应一致	否
5	各界面的风格	各界面中菜单、按钮等界面元素的字体、配色方案等应保持一致	是

根据表 A.4 所示, 可以得出测量公式中的 $A=4, B=5$ 。

A.5.1.3 可辨别的功能

明显的功能

该指标主要测评评卷教师在培训之前, 能够识别的并可以根据界面提示进行操作的功能。该项指标主要反应了网上评卷系统的情景涉及是否合理, 描述是否准确。

测评方法同“布局的合理性”指标, 故测量公式为:

$$X = A/B$$

A 是在培训前, 用户能够并可以根据界面提示进行操作的功能数;

B 是要求在培训前, 用户能够并可以根据界面提示进行操作的功能数。

$$0 \leq X \leq 1$$

X 越接近 1, 表明该项指标越好。

要求在培训前, 用户能够并可以根据界面提示进行操作的功能为本附录在 GB/T 29836.1—2013 附录 B 中分析出的核心功能模块中的功能。在评卷模块中的功能为: 评分、分数提交、回评试卷、问题卷的处理、雷同卷的处理、统计工作量, 因此在测评评卷模块时, $B=6$; 在质量监控模块中的功能为: 查看培训情况、查看工作进度、查看工作情况、强制考核, 因此在测评质量监控模块时, $B=4$ 。

通过选择普通用户进行用户测试法测试, 得到的结果是: 评卷模块中的功能, 用户均能够并可以根据界面提示进行操作, $A=6$; 在质量监控模块中, 用户也均能够并可以根据界面提示进行操作, $A=4$ 。

A.5.2 操作指标

A.5.2.1 操作复杂度

使用中默认值的可用性

该指标是测评评卷教师是否能方便的对便于自己操作的参数进行选择, 例如字体、背景颜色, 显示比例等。该指标主要反应软件对不同用户不同要求的适应性。

测评方法同“布局的合理性”指标, 故测量公式为:

$$X = A/B$$

A 是评卷教师可自己选择的参数的功能数;

B 是要求可供评卷教师自己选择参数的功能数。

$$0 \leq X \leq 1$$

X 越接近 1, 表明该项指标越好。

可进行修改默认值的参数, 以及通过评审法启发式评估对这些默认值的参数是否可进行修改的结果, 如表 A.5 所示。

表 A.5 可修改的默认值参数

序 号	需要修改默认值	是否可修改
1	答题情况显示区中图片的显示比例	否
2	答题情况显示区中背景的颜色	否
3	整体界面的配色方案	是
4	统计功能中的时间、范围的默认值的修改	是
5	查询功能中查询条件的选择的默认值的修改	是

根据表 A.5 所示, 可以得出测量公式中的 $A=3, B=5$ 。

完成指定任务的步骤

该指标是测评在评卷系统中所设计的评卷、设置雷同卷等常用功能所需要操作的步骤是否符合业务及用户要求。该指标主要反应系统完成任务过程是否符合业务要求及用户习惯。

测试该值时, 首先用用户模型法计算出完成某任务所需要的步骤数目, 并由专家确认该数目; 之后由专家操作该系统, 用走查法检测用系统完成同样任务所需的步骤数目, 最后由专家对两个数目进行对比, 并判断系统设计是否符合要求。测评结果应为步骤数目符合要求的任务数量与要求进行测评的任务数量的比例, 测量公式为:

$$X = A/B$$

A 是符合完成所用的步骤数目的任务数;

B 是要求必须进行测评的任务数目。

$$0 \leq X \leq 1$$

X 越接近 1, 表明该项指标越好。

要求必须进行测评的任务为在 GB/T 29836.1—2013 附录 B 中分析的测评核心任务, 故 $B=6$ 。

下面将确定各任务实际的完成步骤数, 如表 A.6 所示, $A=6$ 。

表 A.6 操作复杂度测评步骤

任务名称	操作步骤	步骤数目
评卷	选择分数→确认分数→提交分数	3
回评	选择需要回评的试卷→选择分数→确认分数→提交分数	4
问题卷的提交	提交→选择问题类型→确认提交	3
设置雷同卷	选择需要比对的试卷→设置雷同→确认雷同	3
查询培训情况	选择题目→选择组→选择评卷员→设置培训通过	4
强制考核	输入查选条件→选择评卷员→选择培训题目→设置→确认将其设置为考核状态	5

A.5.2.2 误操作的容错能力

发生错误后,完成任务的时间

用户在使用评卷系统完成某项任务过程中出现错误后,系统应能在较短的时间内恢复,并继续使用户完成任务。

测试该指标时,首先用用户模型法估算一下待测任务的完成时间,再使用用户测试法测试用户在完成任务过程中出现错误时所需要的完成时间,两值越接近越好。测量公式为:

$$X = A/B$$

A 是用户在使用评卷系统完成某项任务出现错误后,整个任务完成的时间;

B 是用户在正常情况下,完成某项任务所用时间。

$$0 \leq X \leq 1$$

X 越接近 1,表明该项指标越好。

要求应进行测评的任务为在 GB/T 29836.1—2013 附录 B 中分析的测评核心任务,下面将确定各任务的完成时间。

易发生误操作测评时间要求如表 A.7 所示。

其中 A 值为正常完成时间,B 值为出错后完成时间。

表 A.7 易发生误操作测评时间要求

任务名称	正常完成时间	出错后完成时间
评卷	小题:10 份/min 作文题:2 份/10min	小题:11 份/min 作文题:2 份/11min
回评	小题:10 s 作文:5 min	小题:14 s 作文:5 min
问题卷的提交	5 s	8 s
设置雷同卷	小题:1 min 作文:8 min	小题:2 min 作文:8 min
查询培训情况	1 min	2 min
强制考核	3 min	4 min

可还原性

用户在使用评卷系统完成某项功能时,界面状态是否能回到原状态。

测评方法同“布局的合理性”指标,测量公式为:

$$X = A/B$$

A 是可以恢复到原状态的功能数;

B 是要求可以恢复到原状态的功能数。

$$0 \leq X \leq 1$$

X 越接近 1,表明该项指标越好。

要求可以恢复到原状态的功能有回评试卷、问题卷的提交、设置雷同卷、统计工作量、查看培训状态和查看工作状态。故 B=6。

由专业测试人员通过技术测试法测试后,可以恢复到原状态的功能有回评试卷、问题卷的提交、设置雷同卷、统计工作量、查看培训状态和查看工作状态。故 A=6。

A.6 交互指标的获取方法

A.6.1 消息交互

A.6.1.1 输入信息

输入的有效性检查

该指标是反应系统是否能够对分数、时间等关键输入信息进行有效性检查,以保证进入到系统数据的有效性。

测评方法同“布局的合理性”指标,故测量公式为:

$$X=A/B$$

A 是能够对输入数据进行有效性检查的功能数;

B 是要求对输入数据进行有效性检查的功能数。

$$0 \leq X \leq 1$$

X 越接近 1,表明该项指标越好。

要求对输入数据进行有效性检查的功能,以及通过技术测试法实际检查的情况,如表 A.8 所示。

表 A.8 输入数据有效性检查

序 号	功能	检查内容	检查情况
1	分数提交	检查分数是否有效	有效
2	统计工作量	对时间、范围等信息是否有效	有效
3	统计工作进度		有效
4	查看培训情况	检查输入的查询条件是否有效	有效
5	查看工作情况	检查输入的查询条件是否有效	有效
6	强制考核	检查考核是否有效	有效

根据表 A.8 所示, $A=6$, $B=6$ 。

A.6.1.2 运行状态

运行状态的易监控性

该指标主要反应评卷系统是否有向用户提供试卷分发、统计等操作过程状态的信息能力。在数据量较大的情况下,时间分发、统计等功能有可能会比较慢,系统应向用户有提示信息,表明系统正在运行,而且运行到何种程度。

测评方法同“布局的合理性”指标,故测量公式为:

$$X=A/B$$

A 是能够向用户提供运行状态查看的功能数目;

B 是要求向用户提供运行状态查看的功能数目。

$$0 \leq X \leq 1$$

X 越接近 1,表明该项指标越好。

需向用户提供运行状态的功能有:试卷分发、统计工作量、查看培训情况、查看工作进度、查看工作情况, $B=5$ 。

由专业测试人员通过技术测试法测试后,能够向用户提供运行状态查看的功能有:试卷分发、统计

工作量、查看培训情况、查看工作进度、查看工作情况， $A=5$ 。

A.6.2 可定制指标

A.6.2.1 界面内容

界面元素易定制性

用户是否在使用评卷系统的某些功能时可以根据自己的喜好改变界面的颜色、字体及颜色等项目
测评方法同“布局的合理性”指标，故测量公式为：

$$X=A/B$$

A 是用户可以改变的界面元素数量；

B 是要求用户可以改变的界面元素数量。

$$0 \leq X \leq 1$$

X 越接近 1，表明该项指标越好。

用户可以改变以下界面元素，以及通过评审法启发式评估对这些界面元素进行评价的结果，如表 A.9 所示。

表 A.9 用户可以改变的界面元素

序 号	界面元素	是否能够改变
1	答题情况显示区中图片的显示比例	否
2	答题情况显示区中背景的颜色	否
3	整体界面的配色方案	是
4	界面字体的改变	是

根据表 A.9 所示， $A=2, B=4$ 。

界面布局的易定制性

用户是否在使用评卷系统的某些功能时可以根据自己的喜好改变界面布局风格，例如改变选单摆放位置、隐藏选单栏等。

测评方法同“布局的合理性”指标，故测量公式为：

$$X=A/B$$

A 是用户可以改变的界面布局元素数量；

B 是要求用户可以改变的界面元素数量。

$$0 \leq X \leq 1$$

X 越接近 1，表明该项指标越好。

用户可以定制的界面布局，以及通过评审法启发式评估对界面布局进行评价的结果，如表 A.10 所示：

表 A.10 用户可定制的界面布局

序 号	界面元素	是否能够改变
1	图片显示区、给分区、选单栏的摆放位置	是
2	改变显示区、给分区、选单栏三个区域的大小比例	是
3	为了扩大显示区，而隐藏选单栏	是

根据表 A.10 所示， $A=3, B=3$ 。

A.6.2.2 操作方式

快捷方式的易定制性

评卷教师可自定义增加分数、减少分数、提交分数等功能的快捷键盘。

测评方法同“布局的合理性”指标,故测量公式为:

$$X=A/B$$

A 是用户可以自定义快捷键的功能数;

B 是要求用户可以自定义快捷键的功能数。

$$0 \leq X \leq 1$$

X 越接近 1,表明该项指标越好。

用户可以自定义的快捷方式,以及通过评审法启发式评估对快捷方式进行评价的结果,如表 A.11 所示。

表 A.11 用户可以自定义的快捷方式

序 号	界面元素	是否能够自定义
1	增加分数	是
2	减少分数	是
3	提交分数	是
4	上下移动	是
5	左右移动	是

根据表 A.11 所示, $A=5, B=5$ 。

参 考 文 献

- [1] GB/T 5271.1 信息技术 词汇 第1部分:基本术语(GB/T 5271.1 2000,eqv ISO/IEC 2382-1:1993)
- [2] GB/T 5271.20 信息技术 词汇 20部分:系统开发(GB/T 5271.20 1994,eqv ISO/IEC 2382-20:1990)
- [3] GB/T 8566—2007 信息技术 软件生存周期过程(ISO/IEC 12207:1995,IDT)
- [4] GB/T 16260.2 软件工程 产品质量 第2部分:外部度量(GB/T 16260.2—2006,ISO/IEC TR 9126-2:2003,IDT)
- [5] GB/T 16260.3 软件工程 产品质量 第3部分:内部度量(GB/T 16260.3—2006,ISO/IEC TR 9126-3:2003,IDT)
- [6] GB/T 16260.4 软件工程 产品质量 第4部分:使用质量度量(GB/T 16260.4—2006,ISO/IEC TR 9126-4:2004,IDT)
- [7] GB/T 18905.2—2002 软件工程 产品评价 第2部分:策划和管理(ISO/IEC 14598-2:2000,IDT)
- [8] GB/T 18905.3—2002 软件工程 产品评价 第3部分:开发者用的过程(ISO/IEC 14598-3:2000,IDT)
- [9] GB/T 18905.4—2002 软件工程 产品评价 第4部分:需方用的过程(ISO/IEC 14598-4:1999,IDT)
- [10] GB/T 18905.5—2002 软件工程 产品评价 第5部分:评价者用的过程(ISO/IEC 14598-5:1998,IDT)
- [11] GB/T 18905.6—2002 软件工程 产品评价 第6部分:评价模块的文档编制(ISO/IEC 14598-6:2001,IDT)
- [12] ISO 9241-10:1996 Ergonomic requirements for office work with visual display terminals (VDTs) — Part 10: Dialogue principles
- [13] ISO 9241-11:1998 Ergonomic requirements for office work with visual display terminals (VDTs) — Part 11: Guidance on usability
- [14] ISO 13407:1999 Human-centred design processes for interactive systems

