



普通高等教育“十一五”国家级规划教材

人机界面设计 (实践篇)

李乐山 著



科学出版社
www.sciencep.com

普通高等教育“十一五”国家级规划教材

人机界面设计

(实践篇)

李乐山 著

科学出版社

北京

内 容 介 绍

《人机界面设计(实践篇)》是李乐山于2004年出版《人机界面设计》(科学出版社)之后在该领域的第二部著作。本书主要建立了一套“以用户为本”的软件或有关产品设计过程,该设计方法主要包括三部分内容:第一,调查用户需要和建立用户模型;第二,按照用户模型建立设计指南,然后由软件人员编写代码;第三,按照用户模型建立可用性测试标准和测试方法。作者通过多年实践和研究,主要建立了第一部分和第三部分的设计方法,这些方法在国外也是近十几年才建立起来的。本书主要分析了如何进行用户调查和建立用户模型,如何建立可用性标准化测试方法,并且分析了有关国际标准和国外各种可用性测试方法的优缺点。本书没有分析如何建立设计指南,因为这里面没有复杂的理论原理,主要应该依据各个企业的惯例和要求建立设计指南。

本书适合用作有关专业的本科生或研究生教材,也可作为企业设计人员和有关研究人员的参考书。

图书在版编目(CIP)数据

人机界面设计(实践篇)/李乐山著. —北京:科学出版社,2009

普通高等教育“十一五”国家级规划教材

ISBN 978-7-03-024494-9

I. 人… II. 李… III. 人-机系统-系统设计-高等学校-教材
IV. TB18

中国版本图书馆CIP数据核字(2009)第063275号

责任编辑:巴建芬 潘继敏/责任校对:张 琪

责任印制:张克忠/封面设计:耕者设计工作室

科学出版社 出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

源海印刷有限责任公司印刷

科学出版社发行 各地新华书店经销

*

2009年8月第 一 版 开本:787×1092 1/16

2009年8月第一次印刷 印张:20

印数:1—4 000 字数:440 000

定价:36.00元

(如有印装质量问题,我社负责调换)

前 言

用户界面设计是一个新领域，其主要目的是把人机关系设计从“以技术为本”转向“以用户为本”。这引起了一系列变化，例如，人机关系的描述从数学模型转向心理学模型，在传统功能和结构设计基础上增加了可用性设计，设计过程中增加了用户界面设计指南，在传统质量测试标准中增加了可用性标准和测试方法。这种转变也引起信息技术领域设计思想的巨大变化。计算机应用软件设计包括三个部分：第一，调查用户需要和建立用户模型；第二，建立设计指南和编写代码；第三，完成设计后的质量测试，尤其是可用性测试。这三部分都与用户研究和心理学有关。

国外大约从 20 世纪 80 年代后期开始大力发展用户界面设计领域。1999 年，西安交通大学工业设计系开设了国内第一个人机界面设计（用户界面设计）专业方向。那时，国内企业几乎没有听说过这个专业，也没有这种工作岗位。2004 年，我校工业设计系学生到上海去寻找工作或实习单位时非常困难。幸亏他们在本科学习期间每年暑假都要实习 1 个月，已经锻炼了如何求职。有的学生打电话联系 40 多个单位，他们向一个一个企业解释人机界面设计是什么、他们能干什么、对企业有什么作用，就这样使我国许多企业知道了这个新专业，他们也把作者建立的人机界面设计方法传播到各个企业，由此国内企业也出现了新职业：用户体验工程师、用户界面设计师、交互设计师和可用性工程师等。4 年过去了，情况发生根本变化。该系每年本科毕业设计都必须到企业去完成实际课题，每年都有 80%~90% 的毕业设计项目被企业直接采用，每年上海、北京、深圳、西安都有一些单位向我系要毕业生或实习生，他们普遍反映我们学生人文素质比较高，能力比较强。若干国外大企业也反映我国用户界面设计水准比较高，甚至从美国、韩国等来同我们洽谈合作。

作者从 1989 年开始研究以心理学为基础的人机界面认识方法和设计方法，并建立了以动机心理学和认知心理学为基础的用户调查方法和用户界面设计方法，部分内容已经写入《工业设计心理学》（高等教育出版社）和《人机界面设计》（科学出版社）两本书中。

为什么写这本书？

本书第一个目的是讲述作者自己建立的用户研究方法和用户调查过程，这种方法针对当前国内外主要方法的不足之处进行了改进。例如，系统如何依据动机心理学与认知心理学的框架进行专家访谈，如何设计问卷，如何分析调查数据，如何建立可用性测试标准等，分析用户研究如何与软件设计 3 个阶段融合起来。这些内容对企业是很实用的。企业开展用户研究工作，第一步可以是建立可用性测试方法。请读者注意，问卷设计是一个高难度的工作。有人把别人问卷拿来修改一下，这种问卷无法保证结构效度，也没有测试过信度，因此是不可靠的。

本书第二个目的是分析国外用户界面研究存在的不足。20 世纪 90 年代以后国外进行了大量的用户界面研究，但是迄今仍然存在以下几个缺陷。第一，西方科学认识论经常用简化论（还原论）去认识事物，把复杂的大系统简化为简单因素和关系，然而整体

因素和关系是无法简化的。西方科学得益于它，偏见也来自于它，受害也出自于它。其中一个表现是国外某心理学教授提出的用户脑力模型（mental model）太简单，国外按照这个用户模型设计调查问卷和可用性测试问卷也过于简单，低估了用户操作行动和认知特性的复杂性，不足以解决用户的可用性问题，因此他们设计的用户界面仍然存在许多可用性问题。正因为此，本书提供的调查问卷要比国外的复杂得多。西方简化论只适合认识简单事物，不适合研究人，因为人太复杂了，把人类迄今全部知识综合在一起也不足以把人的心理搞清楚。其实，自己动脑子思考一下就能够搞明白的。第二，西方文化自古受机械论影响很深，工业革命以来把人变成了机器，连醉汉喝酒时也记得定量描述：“我要 0.33 升。”他们认为已经实现“以人为本”的用户界面，我们可能认为仍然是“以机器为本”的设计。例如，E-mail 的用户界面，采用的就是按键式，这是录音机按键的概念，我们日常写信不用这些操作概念。当新手用户第一次遇到这些概念时，仍然不知道如何操作电子邮件，他们日常写信发信的行为方式无法在这里得到延伸，其问题出在他们往往只调查经验用户或熟练用户，没有调查新手用户。经验用户和熟练用户实际上已经完成了基本学习阶段，适应机器了，已经成为“机器人”了。本书作者强调，调查不同用户具有不同作用，不能以专家用户和经验用户代替新手用户的需要。

本书第三个目的是对国外用户界面可用性测试标准的目的和缺陷进行深入分析。1985 年出现了可用性测试的国际标准 ISO9241，它把可用性定义为：有效性、效率和满意度。实质上那是一个初级阶段的标准，因为 20 世纪 80 年代欧美计算机界才开始从“以机器为本”转向“以人为本”，作者在这里再次提醒，不要盲目跟随国外。虽然国外主要软件都经过了可用性测试，可是大家都会感到国外有些软件并不一定好用，国外甚至有人说 90% 网站的可用性测试是无用的。问题在哪里？问题之一是国外的可用性标准和测试观念存在问题。国外的可用性标准可以被分为“以技术为本”、“以设计人员为本”和“以人机学专家为本”的，恰恰缺少“以用户为本”的可用性测试标准和测试方法。

最后，针对国外用一个可用性标准测试各种产品，本书作者提出一个新观点，不应该用一个可用性标准衡量各种应用软件。由于各种产品的可用性含义不一样，都应该建立有针对性的可用性测试标准和测试方法，本书具体讲述了如何建立以用户为本的可用性测试方法。这是作者提出的新观点、新方法，可能是对西方可用性标准和测试观念的一个挑战，也是本书的第四个目的。作者提出的标准和方法是“以用户为本”的，应该按照用户模型提出测试依据，它不同于国外“以技术为本”、“以设计人员为本”和“以人机学专家为本”的测试标准和方法。另外，实验结果表明，国外的抽样方法也存在明显缺陷，本书也提出了可用性测试抽样的改进方法。本书总结了许多实践经验，提供了实际调查问卷案例，供读者参考。

感谢张煜参与编写了本书第二章第十二节和第十三节；此外，作者还要感谢以下人员：张若思、李江、雷淑芬、李见为、丁嫣、宗威、周熙、白明、刘静照、关鑫、万波、李海龙和王靓，他们为本书的写作做出了贡献。

李乐山

西安交通大学工业设计系主任

lilshan2@sina.com

2009 年 1 月 10 日

目 录

前言

第一章 绪论：用户界面设计过程的要点	1
第二章 用户调查	11
第一节 如何进行访谈	11
第二节 如何调查用户的行动需要	27
第三节 用户任务模型	41
第四节 如何调查用户的认知需要	43
第五节 用户认知模型与出错模型	58
第六节 如何设计问卷	60
第七节 设计调查中经常遇到的问题	71
第八节 调查效度分析	76
第九节 如何改善调查效度	92
第十节 调查信度分析	95
第十一节 如何改善调查信度	103
第十二节 可用性测试的抽样问题	105
第十三节 数据统计分析	119
第三章 当前国外的可用性测试	140
第一节 依据 ISO9241 第 11 部分的可用性测试问卷	140
第二节 依据 ISO9241 第 10 部分的可用性固定问卷	147
第三节 其他各种可用性结构框架	157
第四节 固定问卷可用性测试方法	164
第五节 多媒体用户界面人机学标准	174
第四章 对国外可用性测试方法的分析和改进	182
第一节 对国外可用性测试现状的分析	182
第二节 对满意度测试的分析与改进	196
第三节 当前国外可用性测试方法及存在的主要问题	202
第四节 改进的可用性测试方法	209
第五节 供设计人员的可用性分析方法	213
第六节 供设计人员的可用性检查清单	220
第七节 专题研究	228
第五章 工作量评价方法	237
第一节 工作量评价概述	237
第二节 主任务测试法	241
第三节 次要任务测试法	242
第四节 主观测试法	245
第五节 美国宇航局任务负荷指数	249

第六节 主观工作量评价法	257
第六章 科学论	262
第一节 本体论	262
第二节 认识论	275
第三节 方法论	294
第四节 计算机发展的可能前景	301
参考文献	306

第一章 绪论：用户界面设计过程的要点

本章要点

本章主要面对初学者，介绍用户界面的大致设计过程，分析设计过程中的常见问题。各个企业的设计过程是按照企业实际情况建立起来的，彼此的设计过程不完全一致。本章只是给出了用户界面设计过程的一般工作要点，不可把本章介绍的方法当作死板的教条。

一、用户界面设计过程要点概述

一个产品或应用软件设计过程大致包括以下几个要点。

第一，确立设计项目。发现问题，分析设计任务，确立可行性，提出设计项目。

第二，设计调查。包括市场调查、设计人员调查、用户调查，建立用户模型。

第三，制定设计指南。调查软件人员设计规范需求，按照软件人员期待编写设计指南，软件人员根据设计指南规划应用软件功能和结构，并编写代码，当他们发现设计指南不恰当时，要与用户界面设计师沟通修改。

第四，产品测试。建立产品测试标准，进行各种测试，其中可用性测试是综合性的。

具体说，设计过程的要点大致如图 1-0-1 所示。

1. 下达设计项目

由以下几方面因素确立企业的设计项目。

第一，我国许多企业的设计项目主要是由总经理或少数几个负责人决策的。

第二，设计人员根据专业知识（主要是人机学、心理学等）、市场反馈信息、专家建议等，提出对现有产品的改进建议，作为未来新设计项目。这些信息也会被送到下面用户模型中。

第三，根据长期积累的设计数据库的信息，例如，长期积累的各种用户人群喜好的颜色、喜欢的造型、追求的生活方式等，提出对新产品的基本要求。这些信息也会被送入用户模型中。

2. 任务分析

分析该设计项目的定位和可行性。主要包括以下几方面。

第一，市场策略分析。它主要指从整体上考虑用什么方法在市场上立足，是否能够持续发展，确立企业在市场上的立足点、用户人群、地域及销售服务方式。

第二，企业策略分析。有些企业要求“做强做大”；有些企业认为“今日争强，明日必衰”，因此把企业策略立足于如何能够长期稳守一方；有些企业依赖某个企业生存；有些企业只搞设计，让其他企业进行生产；有些企业不但要搞设计还要搞生产；有些企业也许还采取其他生存方式。

第三，生产策略分析。这关系到企业如何经营生产，如何从技术和管理上扬长避

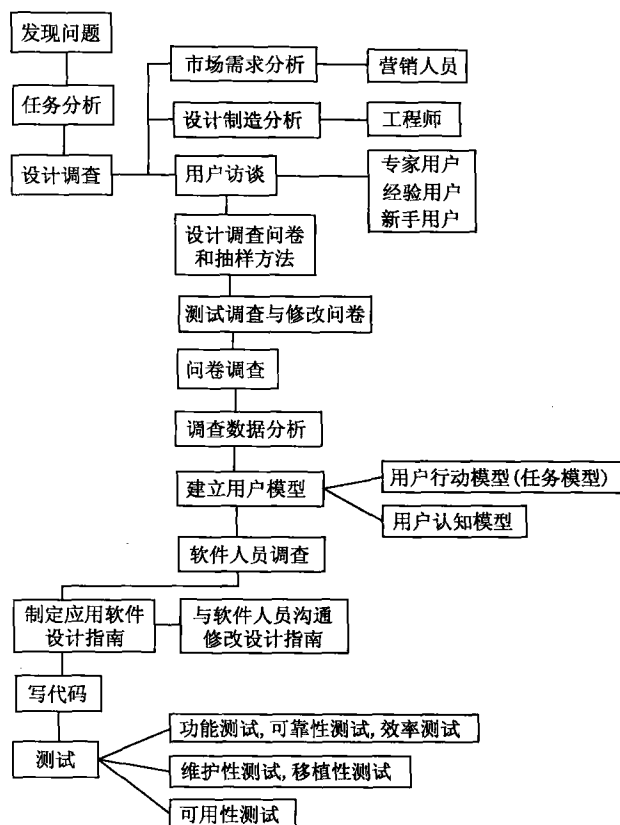


图 1-0-1 用户界面设计过程示意图

短。有些企业把提高效率摆在第一位；有些企业考虑长远目标而把内部和谐放在第一位；有些企业采取劳动密集型生产；有些企业采取高科技高附加值产品；也许要引进技术专家。

第四，产品设计策略分析。一个企业规划了若干产品，其中有主导产品，还有辅助产品，彼此能够相互弥补市场兴衰。产品策略就是考虑如何搞这些产品才能维持企业生存和发展。

第五，新概念产品分析。新概念产品是对社会价值、追求的生活方式、用户人群的审美以及各种新技术、新思想的研究结果。它有若干含义。它可能指当前市场上不存在的新产品，也许会引起新的行为方式，例如手机、电动自行车等产品改变了人的行为方式。它也可能是综合了以往的若干产品，例如把手机、MP3、数码照相机综合在一起。新概念产品也许通过采用新技术提高产品功能，例如智能手机。在汽车领域中，新概念汽车只不过是新设计的样车。有些新概念也许是分析社会核心价值体系的结果，也许是设计人员偶然发现的，也许是某个新科技引起的，也许是新材料导致的。总之，这些新概念产品仅仅是设计人员提出的新想法。

如何估计当前用户界面的普遍设计水准？设计观念受“技术为本”的机械论影响很大，把人当作机器的一部分，强调用户学习，就是让人去适应机器。用户界面设计思想

从 20 世纪 90 年代才从“以技术为本”转到“以人为本”的设计观念。计算机似乎像一个畸形产品，已经按照“以机器为本”设计定型，然后按照“以用户为本”观念进行修改，这很难。不得已，人们试图通过用户界面设计来弥补这个不足。

3. 用户调查

用户调查的基本目的是了解用户需要，建立用户界面设计指南和可用性标准。按照调查形式来分，用户调查主要包括用户访谈、观察用户行动和用户问卷调查等。用户调查主要包含以下内容：①调查新概念产品与人们行动方式（生活方式、工作方式、休闲方式等）的融合程度，最终确定是否设计这个新概念产品。例如，某医院想把 CT 图和核磁共振扫描图等用计算机统一进行管理，通过对医务人员的调查，搞清楚他们有关职业的工作过程，最终用计算机去实现这些过程。②了解潜在用户的有关认知过程，构建潜在用户界面的认知的基本结构。例如，大夫或护士如何识别命令图标，如何存储图片、调取图片，如何解读图片，如何输入诊断信息等。③对现有产品进行调查，发现问题，改进设计。④向软件专家了解软件设计指南格式要求，按照他们的需要和要求，制定用户界面设计指南。它可能包括以下几个部分文件：软件系统的功能描述，软件系统的结构描述，用户界面结构描述，用户界面初始状态描述，运行过程的各个状态描述，用户界面风格描述以及设计依据的各种标准。然后软件人员编写代码。与此同时，用户界面设计人员（交互设计师等）开始设计用户界面。

4. 建立产品测试标准，进行各种测试

全部设计工作完成后，开始进行测试。按照用户观点，主要应该进行 6 方面测试：功能性、可靠性、可用性、效率、维护性、移植性。对用户界面设计师来说，最重要的是可用性测试。测试前必须先建立测试标准。

二、用户调查的基本方法

遇到问题怎么办？①翻书。②求助权威。这两种方法都不适合开发产品。在现实中，往往采用另外两种方法：①调查研究。②实验尝试。调查可以包含以下几种形式：参观，用户访谈，问卷调查，专题讨论（焦点访谈），现场观察和访谈，跟踪调查，查询资料等。这些方法有不同作用，这些调查方法的目的和针对问题不同，用来解决不同问题。下面分别进行介绍。

1. 参观

参观展览会是一种难得的学习机会。它有如下优点：①可以在一天内大致了解到国内外有关情况，在平日想参观其中任何一个企业都是十分困难的。②很快就能比较各种同类产品的特点。③可以请他们进行示范操作，拆卸产品。④可以拍照，可能得到有关资料或样品。⑤能认识有关人员，建立联系，询问实习与合作的可能性，也可以询问求职的可能性。⑥可以提出各种问题，不用担心自己是外行。所以要珍惜参观展览会的机会，因为实际上你一生很少有这样的机会和时间。

2. 访谈

访谈是一种学习方法。遇到陌生任务不会干时，最简捷的方法是访谈，详细请教那些了解情况的人。访谈谁？专家，你在设计中遇到的各种问题都可以去请教专家用户，访谈主要适合开放性的问题（问答题）、多重选择性的问题和排序性问题，而在问卷调

查中主要搞清楚量表型的问题和统计性的问题。

经常访谈的四类专家是软件专家、专家用户、新手用户和市场营销专家。所谓专家，一般具有 10 年专业经验，他们可以帮助解决选题方面的问题。

访谈专家用户起什么作用？①搞清楚设计目的。这并不是说，专家用户能够告诉你设计目的，而是说，他们也许一句话就能够给你提供线索，能够给你一些启发。②搞清楚行业情况，这种背景信息对决策不可缺少。③建立一个调查因素结构框架，也就是应该调查哪些因素。④通过专家用户分析和验证调查的真实性和全面性，这种分析被称为效度分析。如果不考虑这些效度，意味着调查和设计很可能不真实，不全面，很可能你不知道已经出现的严重偏差。如果一份调查报告没有效度分析，其设计数据是不可靠的。效度的分析和验证缺乏系统的数学方法，主要依靠专家的水平。

对专家访谈前，要准备一个调查提纲，或者调查框架性问题。主要从整体角度列出调查问题。访谈过程中，最好先简要讲清楚调查目的和基本调查问题，然后把话语权交给专家用户，跟随他们的谈话。遇到不清楚的问题时，可以记下来，也可以在适当时候打断对方。最好能够用录音机把交谈过程记录下来，以便事后整理资料。还要带上照相机，遇到实际演示操作情景时，能够及时拍照。

3. 新手用户调查

新手用户是那些缺乏基本操作使用经验的人，甚至没有见过该产品。对新手用户的调查可以采用访谈或问卷调查，或者观察他们的操作过程。访谈新手用户起什么作用？主要调查哪些问题？

第一，体验新手用户的行动方式和思维方式。学习计算机操作要花费大量时间和精力，强调用户学习就是“以机器为本”。改进用户界面设计的目的，就是减少用户的这种面向机器的学习花费。用户界面设计的基本思想是尽量采用新手用户的日常生活经验，尽量减少或避免计算机专业术语和机器逻辑。一个设计良好的用户界面需要用户的学习时间很短，需要记忆的新概念和操作过程很少，这主要根据新手用户去测试判断。

第二，观察新手学习的学习特性和学习过程。所谓“学习”，主要包含以下过程：①概念、命令和菜单的理解和记忆；②操作过程的理解和记忆；③如何构思一个简单操作过程，例如，开机、关机、打开文件等；④如何构思和完成一个复杂的任务链，例如，用样条曲线命令去绘一个任意曲线，然后修改这个曲线，移动和复制这个曲线；⑤如何规划和完成一个完整的项目。其中有如下两个观察重点：用户如何把一个任务计划转换成计算机操作过程；用户如何把计算机反馈信息解释为与任务有关的提示。

第三，通过新手用户学习中的出错情况，可以发现用户界面的设计问题。

根据不同目的，还可以对新手用户进行跟踪调查和实验观察。对新手的调查可以采用访谈、问卷调查和观察。

4. 一般用户调查

一般用户有时候也被称为熟练用户、普通用户或经验用户，他们的基本特征如下：①了解计算机的基本行为特征；②能够用计算机完成自己的基本任务；③不完全使用快捷操作。这个人群是产品操作使用的主要用户人群，他们对该产品的态度和接受程度往往决定你设计的成败。

5. 偶然用户调查

这种用户指偶然操作，而不是经常操作。偶然用户是针对银行提款机、家用数码相机等设备而提出的一个用户概念。例如，银行提款机的用户特性如下：①他们可能对数字设备没有操作体验，甚至没有计算机概念。②用户按照用户界面和提示，必须能够完全顺利地完成操作。③用户提款时不能出错，不能由于操作出错而造成钱的损失。经常可能出现的错误是操作结束后取回了提款卡忘记了取钱，就匆匆忙忙走了。他们也可能在操作半途中出错，如果出现这种错误，能够查询识别出来，提醒用户纠正。④要假设他们必须独立操作，哪怕是第一次操作，他们也必须是独立操作，不能通过询问别人去学会操作。⑤完成操作后，他们很可能忘记了操作过程，下一次操作时，他们应该能够依靠用户界面和提示完成任务。⑥应该适应文化程度比较低的用户，必须适应可能的行为不便的人。⑦用户操作中可能出现各种稀奇古怪的行为，操作必须安全可靠，这种界面必须经过反复的可靠性实验。

6. 观察用户操作行动

访谈或问卷调查都属于用户主观调查方法，因此必须还要观察用户的操作行动。各种明显操作行动都适合通过观察去了解，例如，用户的学习过程、操作过程、出错情况以及非正常情景中的操作等。观察用户操作行动应该注意以下几点：①传统实验室方法。它主要有三个特点：首先它隔离了外界干扰，而实际操作环境存在各种干扰；其次实验室的设计往往是观察单一因素对行动后果的影响，实际上用户是在各个因素作用下操作计算机；最后，被测试人在实验室里进行实验时，心理状态与平时正常心理状态不一样，也许会超常发挥，也许会过度紧张而失常，这样测试的结果不同于他们的正常行为。此外还需要在真实情景中进行实验。②在真实情景中观察用户的操作行动，你的观察最好不要引起用户的注意。为了消除“实验影响”，可以让他先了解实验情况，或者放松一下。③用眼动仪观察用户的感知过程。当用户操作计算机时，我们无法观察他的认知过程，采用眼动仪可以弥补这个缺陷。国外眼动仪价格过高，我国计算机专业硕士研究生完全能够完成这个仪器的设计，大约需要 10 万元。④实验前要做好准备。例如，设计好用户操作任务，建立实验表格，其中包括实验过程、如何观察、如何记录数据等。⑤要对实验进行效度和信度分析。效度分析确认是否能够真实达到实验目的，信度分析确认数据是否可靠。没有效度和信度分析的数据无法判断其真实性和可靠性。

7. 跟踪调查

目的是了解短期调查无法了解到的信息。例如，用户在日常真实生活情景中对产品的学习过程，使用产品后操作行动和认知的变化情况，一个产品对一代人或两代人的生活方式和行动方式的影响，一个产品在长期使用中所出现的问题，一个产品对生活方式的影响或对儿童行为方式的影响等。从事跟踪调查必须事先设计一个调查表格，列出背景信息（姓名、年龄、性别、职业、联系方式等），调查内容，观察内容，访谈问题或问卷，这样才能把调查信息汇集起来，不至于口头上问一问，因缺乏系统记录而无法汇总备案。这种长期跟踪的调查信息需要花费大量资源，因此这种信息价值很高，其他方式的用户调查无法替代。我国非常缺乏这样的调查信息。

三、问卷设计

通过各种访谈只能了解个别人的观点和情况，要了解和提取用户人群的整体特性，必须进行一定数量的问卷调查。设计问卷会涉及以下几方面问题。

1) 注重调查实践和积累数据。西方国家的许多计算机公司非常注重用户调查，主要依赖大量的调查实践积累了丰富的用户调查经验、调查问卷、标准格式和信息。美国公司能够在用户调查方面投入大量资金，进行系统的调查，这是作为开拓型或领头型企业必须具备的品质。他们不能依靠其他企业经验，不能从任何渠道得到廉价信息，只能依靠自己调查。如果我国企业想成为国际一流企业，就应该具有这样的眼界。积累的数据多少是衡量设计水准的主要标志之一，这些数据是无法用专家取代的。

2) 用户调查的基本内容包括以下四个方面：①用户价值、工作方式、生活方式、需要等，这些内容属于社会心理学新的研究范围。②用户的行动需要和操作过程，这些属于动机心理学新的研究内容。③用户大脑活动情况，这些属于认知心理学新的研究内容。④参照社会学（或者社会心理学）的抽样方法，可以重新研究和确定用户调查的抽样方法。

3) 用户调查的结构框架决定了调查水准，这是用户调查的第一个关键点，其核心问题是如何提高用户调查的全面性、真实性和有用性，这里特别强调“全面”、“真实”和“有用”，这三点决定了调查的有效性，否则会给后面设计的造成难以弥补的负面影响。

4) 是否存在可以通用的标准调查问卷？不存在。不可能用一份问卷作为“普遍规律”。任何调查都有一定目的，都依据一定假设和结构框架进行了信度计算，修改别人的问卷，就失去了它的信度了。

5) 设计问卷的过程需要考虑几个问题：①调查的问题能够对设计提供有用信息。②调查信息的效度分析。效度分析指调查信息的真实性和完整性的分析，调查效度可以被分为结构效度、预测效度、内容效度、交流效度和分析效度。效度分析主要依靠一定数量的有经验的专家。③对调查数据进行信度分析。主要考虑以下三个方面：用户心理稳定性、内部一致性及评价人的稳定性和一致性信度。用户心理稳定性指对待同一类问题的回答基本一致。例如，他说：“我喜欢使用计算机”，但是他后来又说：“我不喜欢使用计算机”。要剔除那些矛盾信息的答卷。其次，内部一致性指对各个相关问题的回答也是相关的。例如，他说：“我喜欢使用计算机”，那么对“使用次数？”他却回答：“每年使用1次”。这两个答案显然不一致。一份调查问卷的各个问题彼此是相关的，或者说，一份问卷的问题可以被分为彼此相关的几组问题。信度分析对这些问题的彼此相关分析被称为心理一致性或同质信度分析，主要通过计算 Cronbach alpha 来表现内部一致的程度。如果一名评价人对一个调查结果或测试结果的前后两次评价明显不同，或者几名评价人对一个测试结果的评价明显不一致，那么这种评价就不可靠。评价信度是通过下列系数来表示的：Gamma 系数、Spearman ρ 等级相关系数或 Kendall 和谐系数。近几年，有些人采用 Krippendorff alpha，它的优点是适合多种格式的调查问题，而 Cronbach alpha 只适合量表型的数据格式，限制了调查问题。

四、试调查

一般来说,在进行正式问卷调查之前,要进行试调查,其目的是修改问卷,改进抽样方法。试调查抽样采用如下方法:①按照小样本抽样,也就是随机抽样 30 人进行问卷调查。②按照总抽样的 10% 进行抽样,一般小于 300 人。抽样调查之后,要分析试调查数据的基本目的,试分析其效度和信度。

五、正式抽样调查

根据以上分析,修改调查问卷,然后进行正式问卷抽样调查。在正式抽样中考虑的主要问题是,如何使抽样调查人数能够比较真实全面反映该用户人群的整体情况。当前采样三类抽样方法:概率抽样方法、非概率抽样方法和 Nielsen 提出的方法。概率抽样和非概率抽样提供了一些有关的数学方法,但是这些数学方法往往太简单抽象,实际上抽样遇到的问题很复杂。Nielsen 提出可用性测试的抽样方法,这种方法也遇到问题。要想使得用户人群的抽样能够比较真实全面,一定要进行实验积累经验,知道如何实现这些数学条件,如何从实际出发考虑抽样的全面真实。

六、如何建立用户模型

1. 用户模型

用户模型是用户特性的综合描述,主要包括设计所需要的用户特性方面的信息,其目的是为用户界面提供设计信息和可用性测试标准。应该从哪些方面建立用户模型?国外有人只提出了一个智力模型(mental model)。多年设计实践表明,这一用户模型对用户估计得过分简单,不能满足“以用户为本”的功能设计和用户界面设计的基本要求。作者对这个问题思考了 10 年,提出了一个用户调查的系统结构框架,它包含用户行动模型、用户认知模型、用户学习模型和用户出错模型。为什么需要这四个模型?因为人的复杂程度远超过其中任何一个模型所概括抽象的特性,其中任何一个模型都不能全面真实反映用户的使用特性。这四个方面是从四个不同角度去描述用户特性的,从中提取有用信息建立用户界面设计指南,才能够比较真实、全面地反映用户需要和用户操作特性。用户行动模型描述了用户的需要和各种任务的操作过程。用户认知模型描述了他们的感知、注意、思维、记忆、理解、表达、交流、语言、学习、选择、决断、发现问题与解决问题等特性。用户学习模型描述了用户的学习过程及其需要。用户出错模型描述了操作过程中他们出现的各种错误。最终要建立一系列标准化表格。

2. 用户分类

用户分类是根据设计目的确定的,不存在一个死板的用户分类方法。可以按照角色把用户进行分类,他们各自具有不同的目的和任务。也可以按照对应用软件 and 用户界面的熟练程度,把用户分为以下几种:专家用户、经验用户(熟练用户、一般用户)、偶然用户和新手用户。他们分别具有不同的行动特性、认知特性、学习特性和出错特性。这些用户分类是有明确定义的(李乐山, 2004a; 李乐山, 2004b)。

七、建立用户界面设计指南

用户模型是对用户调查后获取的关于用户特性的抽象综合，其目的是作为用户界面及有关功能和结构的设计依据。可能存在以下几种类型的设计指南。

1) 从用户行动和认知角度看，设计指南主要依据如下内容。

第一，用户行动描述。主要包括以下几方面：①功能描述，把用户行动目的（用户任务）转换成机器功能。②操作过程说明，把用户行动计划（任务计划）转换成机器操作过程或操作步骤。③功能和操作特性说明，例如操作顺序，操作灵活性、连续性、透明性、简单性、单一性，任务链之间转换的连续性等。④操作条件说明，每一步操作应该具备的硬件和软件条件，按照人机学方法设计各种操作器件。⑤操作反馈描述，每一步操作后应该提供适当反馈信息显示操作结果及有关提示，表现用户每步操作后引起的机器反应和状态变化，此信息是供用户评价操作结果的。⑥提出评价标准，用户评价根据反馈信息评价行动结果，用户的评价标准来自最初用户的行动目的。⑦非正常情景中的操作特性描述，非正常情景包括嘈杂、黑暗、晃动以及各种突发事件等，这时用户的感知、认知、行为特性不同于正常情况，对功能和用户界面也提出特殊要求。以往设计人员在功能和界面上往往忽视了这种非正常情景，因此用户遇到非正常情景时，更容易出现操作事故。

第二，用户认知的描述。例如，用户视觉特性和视觉需要，这是屏幕显示信息的依据，如按照用户认知意图提供显示内容、符号种类、布局等，显示信息量为 7 ± 2 个信息块，信息符合用户的搜索、发现、识别等特性，颜色要符合环境对比下用户的感知特性等。信息符号符合用户对含义的理解，而不是为了好看。信息处理符合用户认知处理过程。例如，写实的图标易于表达实物，而难以表达动作。计算机命令适合采取用户生活经验进行比喻。人机互动方式要符合用户的交流期待、交流方式、人机各自角色、交流进程等。采用多种信息通道，减少视觉负担、触觉形式和机器声音等非视觉信息，符合用户评价对信息的需要。

第三，用户学习特性描述。要新手用户、经验用户、专家用户的学习特性各有特性，首先要考虑新手用户的学习特性。

2) 从程序员角度看，设计指南包括以下几方面。

第一，图形界面标准。它可以大大减少程序员工作量，明显提高设计效率，然而也容易限制了发展。制定标准需要积累大量设计经验，要兼顾未来发展。图形界面标准规定了窗口的外观和结构，常用控件的标准图形，控件功能、尺寸和位置（如 edit 控件需要控制它的最大长度以及输入限制等）等。如果缺乏图形界面标准，在制定用户界面设计指南时，要逐步把各个图形趋向固定功能、固定外观、固定位置，逐步朝标准化方向发展。

第二，启动后初始界面的描述。要通过用户调查搞清楚这些描述是否符合用户期待。

第三，软件系统功能描述。用户界面主要通过几种功能表达用户的行动：键盘和鼠标、窗口控制、光标控制、对话控制、命令和控件。通过这些功能的相互配合使用来完成用户的各种行动。有时无法直接完成用户行动，只好用间接方法，要注意收集这类问

题并逐步改进它们。机器的功能包含了用户期待的任务功能，也包含了面向机器行为控制的功能。后者局限了用户的行动，例如，用键盘或鼠标难以描绘渐变的颜色和线条，不符合用户的行动期待；用键盘不符合绘画方式。由于人机行为方式的差异，用户不得不打破自己正常的行动计划，然后用计算机具有的功能去拼凑自己的行动意图。这正是要通过用户调查去搞清楚的问题，也是长远设计创新的目标。

第四，软件系统结构描述，用户界面结构、位置及尺寸的描述。结构指组成因素（元素）和各个因素之间的关系。用户界面上主要有几种结构描述。①图形结构指屏幕显示各个窗口和对话框的结构，例如，主要互动窗口、对话窗口、信息窗口等。②操作结构指用什么输入输出设备（硬件），如何实施操作行动，例如，如何用键盘实现全部行动任务的操作，如何用鼠标实现全部行动任务的操作。③命令结构指每个命令的组成，例如，绘制直线的命令结构是：直线图标（宾语）+动词（命令）。④用户每个任务的目的和作用，它对应各种菜单、命令和控件的结构，这些图形界面的形式和内容。⑤用户的各个行动（如翻书页）对应机器功能操作（用键盘操作翻书页）对应的转换的结构。⑥机器状态的变化。

第五，人机互动过程描述，又被称为交互设计。机器状态在用户界面上体现为某些元素的功能、结构、尺寸、亮度、速度、声音和颜色的变化。用户界面设计指南在描述人机互动时，要把用户的行动计划分解转换成机器的操作过程，同时要描述机器在实施每个步骤后达到什么状态。设计指南要描述各个功能、菜单、结构的操作过程（方式）以及它们在操作前、操作中、操作后的状态。前一个动作的结束状态应该符合后续动作的初始条件，或者说，后续动作的初始条件应该符合前一个动作的结束状态，这样，才能使前一种操作的操作逻辑和后续的操作逻辑保持一致。即使如此，操作中用户不得不把行动计划（如习惯的用笔画图的过程）分解，转换成机器的操作过程（用各种命令去画图的过程）。机器的操作过程试图按照用户行动计划而设计，同时又受机器行为方式的限制。这些问题是在用户调查中应该发现和改进的问题。具体说，它主要描述以下三方面。①机器初始状态和用户界面初始状态。②用户的操作过程与机器的反馈信息，用户与机器相互配合及其用户界面状态变化的描述。例如，用户每个操作行动（或任务）的目的和过程，引起屏幕状态、菜单和控件变化；再例如，图标正常颜色，鼠标覆盖的图标颜色，鼠标操作时图标的颜色等。③用户每个操作之后，机器状态的变化。这一部分的描述是程序员编程的主要依据。

第六，容错要求。为了减少用户出错，设计指南提出容错要求和容错方式。当前容错设计主要表现为：设计各种出错提示，禁止操作出错，例如，用灰色屏蔽菜单项。

第七，用户界面设计风格。例如，国外当前流行的主要风格是金属色，按键式。这个用户界面图形标准是 OSF（Open Software Foundation）1995 年公布的 OSF/Motif，它是一个美国图形用户界面工业标准，IBM、微软、惠普、西门子、飞利浦、日立公司参与该组织，该图形标准属于 IEEE1295，已经被 200 多个硬件和软件平台所采用。这个标准认为自己的主要目的是实现“以用户为本”，然而，按键本身就来自机器概念，仍然体现了“以机器为本”的设计思想。这对某些应用软件并不能实现“以用户为本”的设计思想。

八、可用性测试方法

进行可用性测试前要确定三个问题：①建立测试目的和测试任务，选择用户去完成任务或实验。②要确定这些实验或任务中要观察或记录什么内容，要确认是否存在某些问题。例如，观察用户在完成各种任务过程中，他的感知和认知存在什么问题，出现哪些操作错误。③把用户测试得到的数据进行分析，包括统计分析、相关分析、聚类分析、效度和信度分析。这些分析分别可以满足不同目的。

一般可用性测试采用以下三种方法：探索法（heuristics）、有声思维（think aloud）和认知预演法（cognitive walkthrough）。

国外从事可用性测试主要有三类人员：用户、设计人员和人机学专家。让他们完成一定操作，然后填写调查问卷。当前国内外主要采用以下几种可用性测试方法。

1) 国际标准。最重要的是 1985 年建立的 ISO9241，其中在第 11 部分规定了可用性包含三个因素：有效性、效率和用户满意度。然而 ISO9241 并没有规定标准测试方法，它只在其附件列举了一些例子，说明如何测试这三个因素。实施该标准以来，人们发现它存在若干问题，例如，效率是机械论概念，带有“以机器为本”的痕迹。再例如，“用户满意度”或“舒适度”是一个模糊概念，缺乏一致的心理学术概念，由此测试的条款有 90 多个不同概念。如今人们很少采用这三个参数，而采用这个标准中有关的“以用户为本”的思想。

2) 国际标准 ISO9241 第 10 部分的七条标准测试可用性，包括适合任务、能够自我描述、用户可控制、符合用户期待、容错、适合个性化和适合学习。这些标准是以设计人员为本的。

3) 采用固定人机学调查问卷，20 世纪 80 年代以来西方若干国家的大学或研究机构建立了可用性调查问卷。例如，美国施奈德于 1987 年在他的书中公布了 QUIS（questionnaire for user interaction satisfaction）固定问卷；20 世纪 90 年代爱尔兰的 University College Cork 的人因素研究组（HFRG）建立的 SUMI（software usability measurement inventory，软件可用性测试量表）。请注意，这些问卷都受知识产权保护，必须在得到书面允许后才能使用，为此一般要交一些费用。这些问卷都比较简单，主要用途是把若干产品都用同一测试方法进行测试，其结果可以进行横向比较，很难再有更多用途。

4) 本书作者提出“以用户模型为基础”的可用性测试方法。针对以上方法的缺点提出改进措施。

第一，不再按照设计人员或人机学家单独提出测试任务，而是与用户协商后确定。

第二，不再由设计人员、人机学家单独观测测试，而是由他们合作。

第三，不再由单个用户继续操作测试，而是由两三个人合作完成。

第四，不再按照以往那种问卷，而是按照具体操作过程记录问题。

第五，不再采取死板简单的可用性标准，而由用户确定可用性标准。