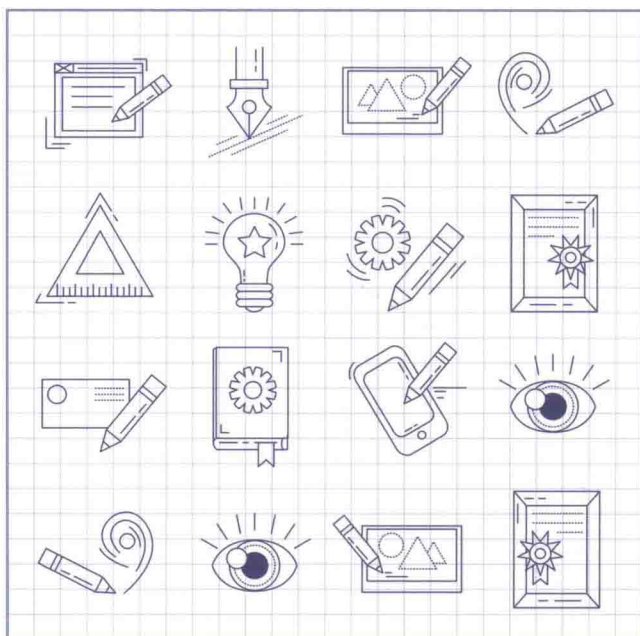




你的产品是否足够友好？

通过阅读和学习本书，可以找出答案

Handbook Of Usability Testing 可用性测试手册

第2版

[美] Jeffrey Rubin Dana Chisnell 著

王超 邹烨 译



中国工信出版集团



人民邮电出版社
POSTS & TELECOM PRESS

第1章

什么造就了可用性

是什么让产品或服务可用？

可用性是很多产品都拥有的品质，但更多的产品缺乏可用性。这是很多原因造成的，如历史、文化、组织架构、金融以及其他原因，这不是本书的讨论范围。好在，仍然有一些通用并可靠的方法可以用来评估哪些设计是可用的，哪些不是，并且评判怎样优化设计可以让一个产品具有良好的可用性，让产品足以在市场立足甚至蓬勃发展。

看上去似乎很难知道是什么造就了可用性，因为除非你有一个突破性的可用性范例，可以切实地推动销售（很容易想起苹果公司 iPod 的案例），否则缺乏或丧失可用性仅仅只是一个问题。想象一下顾客想要在你公司的电子商务网站上买东西。他们与网站的对话可能是这样的：我找不到想要的东西啊。好吧，我找到了想要的东西，但是我不知道卖多少钱。还有现货吗？能运送到我想要的地方吗？我花费这么多可以免运费吗？几乎每一个尝试在网上购物的人都遭遇过类似的问题。

网站的可用性问题非常容易指摘（毕竟网站那么多），但是人们每天会遭遇无数其他的难以使用的产品或服务。你知道怎么用你的闹钟的所有功能吗？手机呢？数码摄影机呢？当你使用自助贩卖机时，根据他们的语音提示做选择容易吗？

1.1 “可用”究竟是什么

大部分时候，产品使用时没有挫败感就具备可用性了。正如本书中所述可用性测试的操作过程和方法，我们需要依托于可用性的定义，即当产品或服务真正可用时，用户可以用他们自己觉得可行的方式去做他们想做的事情，没有阻碍、犹豫或疑问。

但在我们开始定义和探讨可用性测试之前，先谈一谈可用性的概念及其特质。一个产品或服务想要具有可用性，它应该是有用的、高效的、有效的、令人满意的、易学的以及可达的。

有用（**usefulness**）描述的是产品促成用户达成目标的程度，以及用户想使用产品的意愿评价。没有意愿，其他衡量标准都没有意义，因为产品都是摆在货架上任人挑选的。如果一个产品容易使用，容易学习，甚至用起来也令人满意，但是没办法让一个用户达成具体的目标，就算它是免费的也没有用。有趣的是，有用这件事在实验室的实验和研究过程中是最容易被忽略的因素。

在产品研发的早期，通常是由市场团队来明确哪些产品或系统功能是用户想要和需要的，影响可用性的其他因素甚至完全没有考虑。正是由于缺乏这样的考虑，开发团队想要站在用户视角是非常困难的，他们可能只是简单的猜测，更有甚者把自己作为用户。这是系统导向的设计经常发生的事情。

高效（**efficiency**）是指准确完整地完成任务的敏捷性，通常是一种时间上的衡量。举个例子，你可能会设置一个可用性测试的基准是“95%的用户可以在10分钟内下载完软件。”

有效性（**effectiveness**）是指一个产品在多大程度上能以用户预期的那样运行，以及用户使用它来完成想要的事情的难易程度。这个通常以定量的出错率来衡量。正如测量效率一样，有效性的可用性测试也需要以一定百分比的用户量来做依据。以高效的案例作延伸，测试基准的描述可能是“95%的用户在第一次尝试后就可以正确下载软件”。

易学性（**learnability**）是有效性的一部分，和用户操作系统的能力有关，这种能力水平的定义要经过一定量以及一段时间的训练（也可能根本没有时间）。易学性通常也指经过一段时间以后，不活跃的用户重新学习使用系统的能力。

满意度（**satisfaction**）是指用户对产品的知觉、感受和想法，通常由书面及口头的问卷来收集。当产品很好地满足用户需求，并提供更佳的满意度时，用户使用产品的表现可能会更好。一般情况下，要求用户对试用产品进行评价和排名，进而暴露问题的本质。

可用性的目标和宗旨就是将上述特性以可测量的方式来定义。然而要注意的是，让产品足够可用，绝不是单纯地生成一串使用和满意度的数字。当数字告诉我们一个产品是否“起作用”时，同样也有一个独特的定性的因素衡量产品可用性如何，这是难以用数据量化来做出判断的。行为数据可以告诉你为什么会有这样的问题，定性因素则影响如何解释行为数据才能够解决问题。任何医生都能够测量病人的生命体征，如血压和脉搏。但解释这些数据并给每一位特定的病人推荐合适的疗程，才是医生真正的价值所在。为了设计有效的疗程，

判断可能引发问题的多个不同原因，并找出哪些有可能是特殊情况，这通常意味着要看得更远，远超单个个体数据。没有经过训练的双眼可能会错过这些细微的差别。

可达性（accessibility）和可用性相伴而生。就最广义的理解来说，人们会觉得可达性说的是可以通过产品来完成目标。但本书中谈及可达性时，指的是对于身有残疾的人来说，产品如何更加可用。为身有残疾的用户改善产品可用性，或为了某种处于特殊情况下的用户改善，抑或是两者都考虑，这种做法通常也会让没有残疾的用户获益。为了残障人士考虑可达性可以使设计清晰简洁，对于那些遇到临时的限制（如受伤）或情景的人也有益处（如注意力分散、糟糕的环境条件、光线太亮或是不够亮）。有很多的工具以及指南可以帮你进行可达性设计（与本书同步的网站上有许多可达性资源链接（访问 www.wiley.com/go/usabilitytesting 可以获得更多信息）。你应该更加熟悉可达性的最佳范例，以便在组织进行以用户为中心的设计时，可以与可用性测试以及其他方法共同运用。

可用性和可达性是以用户为中心设计（UCD）的众多准则中的一部分，以用户为中心的设计包含很多方法以及技巧，在本章后面的部分我们会谈到。反过来，以用户为中心的设计来源于更大更全面的概念，叫作体验设计。顾客可能会在你的网站上完成购买过程，但是如何与之后的事情紧密配合，如商品的运输、保养、售后服务，抑或是退货？你的组织如何支持研究并决定购买的过程？所有这些都是体验设计的重要环节。

这让我们继续回到可用性这个话题。

真正的可用性是看不见的。如果事情运行顺利，你根本不会注意到它。就像如果房间里面温度适宜，那么没有人会抱怨。但产品的可用性通常是一个连续的过程。你的产品可用性如何？就算用户可以完成目标，产品是否还能提升可用性？值得提升吗？

大部分可用性专业人员花费大量的时间来研究如何排除设计问题，让用户的挫败感最小化。这真的是一个值得称颂的目标！但是你要知道，要让使用产品的每一位用户都达到这一点是非常困难的，对于用户达成目标来说，这是用户体验中影响非常小的一部分，并且就算有定量的方法可以测试产品的可用性，有一些东西的可用性也是没有办法衡量的。只能衡量产品有多么的不好用：用户使用一件东西的时候会遇到多少问题，这些问题都是什么，以及造成问题的原因。

在设计的迭代过程中，只有通过综合评估的方法，如可用性测试，才有可能让产品和服务更加有用、好用，甚至令人欣喜。

1.2 导致可用性低下的原因

究竟是什么导致许多高科技产品如此难用？

接下来，我们将研究这个主题，讨论这个现象存在的原因，并全面研究问题的对策。书中很多案例不仅仅涉及消费类硬件、软件、网站，还包括用户使用手册和内嵌的电子帮助文档以及错误提示信息。解决方法一并适用于乐器、移动手机、游戏操控台，等等，甚至类似超声波影像操作平台或数码相机的用户手册都是本书涉及的范围。

导致产品可用性低下的五大原因

对于当前从事产品研发的工作者，如工程师、用户界面设计师、技术协调员、培训专员、管理者等而言，导致产品或系统可用性低下的原因是如此惊人地相似。

- 聚焦于设备或系统的研发。
- 目标受众的扩展与适应。
- 可用性产品本身存在设计难点。
- 团队内各专业人员缺乏协同。
- 设计与技术实现的不匹配。

原因一：聚焦于设备或系统的研发

产品的设计和研发过程，只强调和关注设备或系统，而非最终用户。图 1-1 清晰地展示了用户行为的通用模型。

贝利用户行为模型指出任何类型的用户行为都由以下三大因素构成。

- 人群
- 环境
- 活动

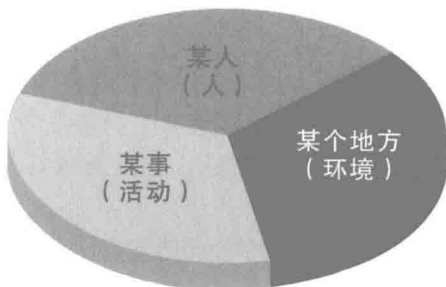


图 1-1 贝利用户行为模型

因为系统或产品的研发的目的是提升人类某个领域的表现，设计师应该在设计过程中纳入对这三个因素的思考。三个因素都将最终影响人们最终的表现。可惜，设计师、工程师、程序员只习惯聚焦在活动这个因素上，而往往不够重视人群和环境的因素。同时也忽略了三个因素相互之间的关系。出现这种不平衡的状态可能有以下原因。

- 一种基本的假设是因为人类与生俱来的灵活性和适应性，所以人们更容易适应设备，而反过来让设备适应人是行不通的。
- 开发工程师历来更适应看上去“非黑即白”绝对化的、科学化的、抽象联系的系统工作。而人类的问题是灰色的、混乱而模糊的。
- 开发工程师被任用和奖励的原因历来不是因为他们通晓人性，而是他们具备解决技术问题的能力。
- 人的需求成为忽略因素最重要的原因是，设计师以往设计的产品面向的最终用户与他们自己非常相近。这自然没有理由去研究一个跟自己相近的人群。让我们接下来看第二个原因。

原因二：目标受众的扩展与适应

随着技术渗透了主流消费市场，目标受众不断壮大并持续演变。开发团队无法快速顺应这种进化。

早期数字化产品的用户（早期的尝鲜者）热衷于计算机专业知识，热爱科技、沉醉于修修补补，为自己能搞定任何问题而骄傲。这些数字化产品的开发者有着相似的特质。本质上，用户和开发者是一体的。正因为这种相似性，开发者只需要为“邻座”设计，其实就是为你身边的人做设计。毫无疑问，这种方法相对成功，用户也几乎不曾抱怨过使用障碍。

这些用户怎么可能抱怨呢？很多用户享受使用产品过程中需要付出的修补和调整。狂热的用户会因为他们能搞定复杂的产品功能而无比自豪。因此，“设备导向”或“系统导向”的方式自然而然地成为了研发标准。

而今，一切都发生了改变。用户往往对计算机和机械设备知之甚少，对购买的产品缺乏调教的耐心，已经完全与产品设计师的设想脱节。更重要的是，现在的用户远远无法毗及设计师，包括技能、资质、期望或伴随设计过程需要的任何特质。在过去，公司可能需要聘请化学博士操作他们的产品，现在他们会找高中毕业生来完成同样的操作。显然，“邻座”设计方法在面临用户和设计师之间巨大差距时荡然失效。一些忽略趋势的公司继续秉承这一策略，将持续生产出可用性低下的产品。

设计师也不再是爱好驱动的狂热份子。大多数设计师接受过专业的人机交互、工业设计、人因工程学、计算机科学或混合型学科的教育。那些曾经让不谙技术

的人望而却步的电子化或数字化设备，而今却不可避免地被普通用户在工作或生活中接触和使用。那些无论是在工作，还是生活中举足轻重的大众化产品，如手机、数码录像机、网站或精密的测试仪器，他们的用户都不具备技术素养。当今的用户需要的是一个工具而非一种爱好。

原因三：可用性产品本身存在设计难点

尽管很多组织将设计可用的系统视为“常识”，但它的困难却难以预测。

虽然有很多著作提及可用性的方法论，但对于不具有行为学和社会学背景的人而言，却是难以参透的概念。可用性定义以及如何获得可用性，既是艺术问题，也是科学问题。似乎每个人对于可用性都有一个概念和实施方法。直到可用性可被衡量，才跳出了概念的范畴（这需要可执行的定义和精确的测量）。

比起抛弃可用性原则的设计师而言，轻视可用性原则更加危险，因为这会招至更坏的结果。正如 Will Rogers 所说：“产生问题的原因不是你无知的部分，而是你以为自己知道的部分。”很多组织已然将可用性工程漠视为“常识”。

本书在 1994 年首次出版时，没多少系统设计者和开发者了解以用户为中心的设计原则。现在，多数设计师或多或少了解以用户为中心的设计原则，至少是有所耳闻。然而，认知和实践之间仍然存在着鸿沟。可用性准则不够明晰，依然需要教育、辅助设备和系统化方法来共同运用设计过程中这所谓的“常识”。

原因四：团队内专业人员缺乏协同

组织招聘专业化团队完成产品和系统的研发，却忽视了成员间的相互协同。

出于提升效率的目的，很多组织将生产流程拆分成相互独立运作的系统元素。例如，软件开发的组成要素包括用户界面、帮助系统和书面文档。诚然，这些要素由独立的个人或团队研发。专业化分工本身没有问题。问题在于独立要素之间缺乏协同融合，不同研发团队间沟通不足。

通常产品研发在独立拆分的部门之间推进。在外界看来，研发的结果会被感知成图 1-2 所示的结果。



图 1-2 缺乏协同的产品研发

每个研发组像地窖一般独立运作，最终的产品也反映了这个过程。帮助中心不能给到用户界面这一层足够的支持，甚至与用户界面设计背道而驰。用户帮助

文档因为缺乏交叉索引而出奇地冗长。亦或文档呈现的内容已与最新版用户界面脱节。这些都可想而知。

产品发布后问题产生了。最终用户面对新产品时，期望它作为一个整体工作。如图 1-3 所示。用户不会刻意区分这三个部分。用户预期产品的每个部分都应该紧密配合在一起。即便产品有诸多专业化的优势，也会因为无法满足用户预期而失败。

有意思的是，组织在不经意间加剧了割裂的状态。他们将产品的各个要素独立进行可用性测试。用户文档独立于用户界面作测试，用户界面也独立于帮助测试。这个方法最终将无济于事，因为自身可用并不意味着什么。只有产品的各要素配合良好，才能真正满足用户需求。



图 1-3 协同式的产品研发

所幸，近年研发方法体系有了长足的进步，开始不断强调设计迭代和团队跨学科构成。此外，大量以可用性理念指导下产生的无边界的产品和服务案例，在市场上获得了巨大的成功，如 Netflix、eBay、Yahoo!、iPod 及 iPhone、Whirlpool（惠而浦）家用电器最新产品线。它们成功的最大因素是协同研发。

原因五：团队内专业人员缺乏协同

用户页面的设计和技术实现是两种不同的活动，需要截然不同的技能。现今人们更强调和需要设计技能，而技术实现需要另一种思维模式和技术模式为主导。

设计关注产品沟通方式，技术实现则关注产品运作方式。早期设计和开发的分歧还罕为人知。早期，工程师和设计师被雇佣都是因为他们的技术专长（编程和机器导向的思维）而非设计专长（信息沟通和用户导向的思维）。这是因为早期利用计算机编程语言完成产品设计是项很大的挑战，所以情有可原。当然，优雅沟通可以带来更好的效果，但它不是当时的最高指导原则。

随着新一代编程语言的发展以及自动化编程工具的产生，技术实现的挑战骤然减小。而为了满足日益增长的用户规模，适应他们经验不足、对可用性期待较高的特质，设计的挑战陡然增加。就拿计算机做类比，工作的焦点已经从设备内部（机器如何运行）转移至外围的最终用户（如何与产品交互）。

工作焦点的转移也趋使设计师技能的变化。设计将逐渐从技术实现中脱离出来。或许有一天，在设计用户界面时将完全不需要编写代码。

上述5个原因只不过粗略地阐释了缺乏可用性的产品和系统不断出现的原因。更重要的是，这些问题和误区背后有着共同的主题：过于强调产品本身而忽视了产品本身应该实现的预期效果。尤其当开发周期变得愈发短促而忙乱，对用户的关注和理解就更愈发少了。

设计师容易对产品设计的本质失察，同时更容易忘记他们正在创建一种产品与人的关系。此外，为了设计这种关系，设计师必须允许用户专注于手头的任务，帮助他们达成目标而不是让用户关注完成任务的方法。设计师同时需要设计各个产品相互配合的关系。这意味着设计的实体之间可以无间地沟通，同时需要扩大体验维度，卷入产品的整个使用周期或工作场景。曾经的工作模式已经无法适应当今的用户和技术了。

我们需要新的方法和技术，帮助设计师转变视角和方法。这种由外及内的，适应最终用户的需求和能力的工作方法，就是以用户为中心的设计方法。基于以用户为中心的前提，可用性测试才体现价值并得以发展。后面会更详尽地探讨关于以用户为中心的理念。

1.3 产品可用性的成因

以用户为中心的设计（UCD）在过去数十年中被冠以不同的定义描述，如人因工程学、人机工程学和可用性工程（其中人因工程学和人机工程学几乎可以相互取代，与其区分两者在方法和实现方式上的差异，不如理解为地域因素导致了两者的差异。在美国，人因工程学传播广泛，而在其他国家尤其在欧洲，人机工程学则更为普及）。以用户为中心的设计意味着设计可用的产品和系统的各种技术、过程控制、方法和步骤。不过同样重要的是它是过程中将用户置于核心地位的哲学理念。

尽管设计团队必须首先考虑产品技术（如何实现我们的设想）以及功能特性（产品是否完成特定功能），团队必须同时考虑用户在使用产品时的体验问题。以用户为中心的设计，应以聚焦用户为起点，同时顾及底层技术的能力和边界，以及公司设想中希望提供的功能特性。

在设计过程中，UCD 会遵循目标用户实际使用方式，而不是强势地改变用户

使用产品的方式。国际标准组织（ISO）第 13407 条规范这样定义 UCD：“卷入用户并清晰理解用户和任务目标；恰当地配置用户因素和技术因素于产品功能中；迭代式的设计解决方案；多学科融合的设计活动。”

超越以用户为中心设计一个产品，应该关注用户对一个产品的拥有权的整个闭环体验。在理想情况下，整个过程包括与潜在消费者的交互、从最初的销售和营销触点贯穿到被用户购买其他产品或换代当前产品的全过程，都是以用户为中心应涉及的问题。在这样的情景下，公司应该进一步关注包括所有售前售后与用户的接触点和交互内容。不过，让我们先迈进一步，聚焦设计过程的探讨。

现在有很多文章或书籍关于以用户为中心的设计（UCD）（在随书的网站 www.wiley.com/go/usabilitytesting 上列举了一些我们喜欢的文章和书单）。然而重要的是，读者应该理解 UCD 的基本原则，以便理解可用性测试操作实况。可用性测试并不是 UCD 本身，它仅仅是帮助实现一个优秀的以用户为中心的设计的诸多方法中的一种。

我们希望能着重强调以用户为中心的设计有以下几点基本原则。

- 及早聚焦用户和用户任务。
- 评估和测量产品使用方法。
- 迭代式设计。

及早聚焦用户和用户任务

仅仅辨识和分类用户还不足够，我们建议贯穿整个研发周期，设计团队和用户直接接触。当然，你的团队需要接受训练和指导，以便更好地管理接触用户的过程。你自己也有必要承担起更多的学习和练习。

用户接触过程需要一个目标来指导，要警惕只是简单地在用户行为满意度表格上打满勾勾。真正需要的是系统化和结构化地收集用户信息。设计师需要专业采访以及组织数据搜集等工作技能的训练。此外，结果应该确保不产生误导。

评估和测量产品使用方法

这里，我非常强调在设计过程早期就应该针对真实用户，通过开发和原型测试，开展易学性和易用性的行为测量。

迭代式设计和测试

我们一再重申迭代式设计的重要性。不过，在整个开发闭环中并没有真正执行到位。真正的迭代式设计鼓励通过测试早期概念原型和设计创意彻底检查和反思设计。如果设计师不准备投入这么重要的步骤，就将导致迭代式设计影响极少，流于表面。本质上，真正的迭代式设计鼓励产品形成的过程经历设计、测试、再设计和再测试的循环。

实践 UCD 的组织特性

以用户为中心的设计要求大部分公司反思自身的商业举措、产品研发和目标消费人群。尽管，目前没有现成的保证成功的公式，不过能实践 UCD 的公司具备一些共性特质，如下所示。

- 强调卷入用户信息。
- 多学科背景团队。
- 参与式的开明的管理。
- “随时随地习得”的洞察力。
- 明确的可用性目的和目标。

1. 强调卷入用户信息

与传统开发方法论中强调内容不同的是，一个以用户为中心的方法是在进入下一研发阶段前，在本阶段回收用户反馈或输入信息。这会卷入各种技术和可用性测试手段。

如今，最主流的以技术驱动的产品或系统开发的公司，会在产品研发周期内卷入若干可用性工程或人因学工作过程。在过程中，会产生一些问题。问题和一些针对性的改进建议请见图 1-4。



图 1-4 问题和对应的解决方法

每个阶段的可用性工程事项是不同的。注意，尽管特定的研发周期的人因学活动是出自专业人员的判断，但仍然有多处工作需要跨团队协作。所以下文就是我要道出的 UCD 团队特性。

2. 多学科背景的团队

设计不再是一个人甚至一个专业人员能独立掌控的。虽然一个设计师可以主导产品设计的多重职责，但依然没法了解所有的环节。为非技术背景的最终用户设计一个非常复杂的产品需要考虑太多的因素。以用户为中心的设计要求多样化的技能、知识，以及最重要的目标用户特征和使用方式。如今，团队由多领域的专家构成逐渐成为一种标准配备，如工程学、市场、培训、用户界面设计、人因学和多媒体等领域。多个领域的专家转而在互补的领域得到培训，因此交叉原则的工作较以往更容易展开，也更有活力。

3. 参与式的开明的管理

将可用性纳入运作体系，通常与公司管理层是否顺应公司阶段性发展需要，利用条例保证设计团队获得充分话语权有紧密关联。管理者知道可用性将带来经济价值并赢得市场份额。

4. “随时随地习得”的洞察力

UCD 是一个伴随产品最终成形的进化过程。它要求设计师具备一种观念，最佳设计是在试错、发现和优化的过程中产生的。关于推进过程的假设，如果不能被实物化并经历最终用户的评估，终究还只是一个假设。最终用户的表现和偏好将决定设计决策。

5. 明确的可用性的目的和目标

设计一个可用的产品必须基于一个结构化和系统化的过程，同时在启动初期设定了高标准的目的和具体的目标。模糊和错误的构想无法实现包括可用性在内的各种目标。你的团队甚至需要明确定义“可用性”本身。让你的产品具备可用性的可操作定义可以包括以下几种。

- 有用
- 高效
- 有效性
- 满意度
- 可达性

这几项从本质上全面定义了如何生产可用的产品。我们现在来回顾下一名可用性专家在完成以用户为中心的设计时常用的技术和方法。

1.4 实现可用性的技术

UCD 是在产品开发周期的不同节点中应用不同的技术、方法和实践工作。回顾常用方法可以提供可用性测试一些场景。可用性测试本身就是诸多方法中的一种。请注意下文描述的技术被提及的顺序从一定程度上代表了它们应用到一个产品开发周期中的顺序。

1.4.1 人机工程学研究

人机工程学研究借鉴于人类学学科体系。它在实际使用场所中观察用户（如工作场所、家、咖啡吧等），收集目标用户是谁，他们需要完成的与产品相关的任务和目标是什么，以及他们要完成目标的场景。通过这一定性研究，就可以构建用户轮廓、人物画像（典型用户）、情景和任务描述。基于此，你和设计团队可以在整个开发周期中获得设计决策的依据。

1.4.2 参与式设计

UCD 与其说是一门技术，不如说是具象化的过程。参与式设计是邀请一位或多位用户代表参与到设计团队。这个方法通常用于家用系统的研发项目。在项目一开始就卷入最终用户参与设计的核心环节，以便标记用户的知识、技能，甚至对设计的情绪反应。用户代表过于融入设计团队是一种潜在的风险。他们不再以自己原本的方式反应和思考，相反他们更倾向于避免训诫团队工作人员，隐瞒重要的顾虑或评论。

参与式设计可以有一些技术调整，安排短暂的独立式的工作坊，让用户、设计师和开发就一个具体的设计问题工作在一起。例如，用户、设计师和工程师用一个可工作的原型，一起完成产品最佳尺寸和外形的推敲打磨。

1.4.3 焦点小组研究

用户焦点小组研究是在项目非常早期让用户代表参与评估初步概念的过程。焦点小组研究可以理解为概念可行性的检验。某些时候，可以同时区分和确认代表用户的类型。所有的焦点小组研究会把多个被试者安排在一起，这是该技术有别于其他技术的重要因素。

小组共同参与评估的概念可以用非常雏形方式呈现，比如手绘铅笔稿、故事版、屏幕上详细描绘的原型或塑料手板模型。目标是探明概念的可接受程度，用户在哪些地方无法接受或感到不满意，哪些改进可以使设计更被接受和更有用。焦点小组研究的美妙之处在于可以深度研究小部分人的评判和感受，同时学习到

最终用户是如何思考和感受的。这样，焦点小组对比可用性测试，是截然不同且无可替代的。焦点小组适合收集大体的定性的信息，不适合研究表现和真实行为。记住，焦点小组里的人表达的是主观感受，往往会跟他们的真实行为相左。可用性测试适合观察行为和测量表现，顺便结合部分定性信息的收集。

1.4.4 调研

在执行调研时，可以从较大基数范围理解用户对现有产品或潜在产品的偏好。尽管调研无法覆盖焦点小组研究纵深地研究反馈和设计依据，但可以从全量人群中收集大量的样本。比如，尼尔森量表，现行最流行的调研量表，曾基于1500人的偏好研究全国数百万美金的商业决策。调研可以用于产品周期的任意时段，其中大多数时候用于早期研究潜在用户。调研至关重要的一点是必须保证语言描述清晰，所有阅读者都不会产生歧义。调研如果没有多轮测试和充足的准备时间，就不能保证好的效果。此外，再重申一遍，询问用户会怎么做或做了什么永远无法替代在可用性测试中直接观察他们的行为。

1.4.5 走查

一旦明确了谁是你的目标用户以及任务目的，走查就可以在概念早期或产品原型阶段，以用户视角研究一个用户将如何对待一个产品。通常，设计师需要引导他的同事走查真实用户的任务（有时甚至要扮演用户的），同时另一个组员记录使用障碍或对团队的顾虑。结构化的走查由IBM在代码检查的工作中被首次开创使用。执行被试者设定一个具体的角色（如主持人、记录员），并且在清晰的指引下完成任务，确保操作效用（走查过程不应超过2小时）。与其让设计师扮演用户的角色，不如邀请真实用户特别是领袖型用户进来。

1.4.6 封闭式和开放式卡片分类法

卡片分类法用于研究内容或功能的“可查找性”。这个方法可以有效地获得用户关于内容组织、文案措辞和用户界面标签等方面的反馈。你既可以向被试者展示没有标题的内容或分类，让用户自己命名（开放式卡片分类），也可以向被试者告知初步的或已有的分类并请他们往里面自主分组内容或功能（封闭式卡片分类）。

1.4.7 纸面原型

在纸面原型的技术中，需要在纸面上向用户展示产品的一个部分并询问相关的问题或从另一个角度了解用户的反应。也可以在草稿上用铅笔手绘页面稿、线框稿、几个分屏界面、几个页面或由不同状态的面板组成的框架稿，在逐一展示

给用户的过程中发现用户的预期与预设的匹配程度。例如，在一个电子商务网站关于购物车的原型中，可以展示一个有待购商品的购物车的页面，接下来里面的商品发生了一些变化，接着运费和税费增加了（或者，在进展到下一部分时，直接跟用户再设定出有待购商品的购物车）。

如果需要研究标签对用户预判下一步的作用或预设的分组是否符合用户对任务的思考和表述，可以向用户展示一级导航。用户给出一级导航的选择时，可以展示下一级相应的导航内容。这个过程可以持续到你设计或准备的导航深度。

当然，也可以直接就原型询问用户。问题比较宽泛，可以针对组织形式和布局的具体表现，也可以关于用户对特定选择或信息类型的预期。

纸面原型的价值在于通过快速而便捷的方式获取关键性信息。在没有写一行代码之前，设计师就可以确认哪些功能和特性是符合预期的，哪些不是。此外，技术文档撰写人也可以利用纸面原型的技术在撰写之前先行评估内容列表。这项技术可以以最小资源投入反复应用。

1.4.8 专家启发式评估

专家启发式评估是卷入一个不了解项目的可用性专家或人因工程学专家对产品或系统进行检查。专家们会根据符合研究主体的可用性原则（启发性的）、人因学文献以及已有的专家级经验来完成评估。得出的观点将针对使用产品的目标细分人群。

双重专家既深谙可用性原则或人因工程学，又对某个领域（如健康、金融服务等）或产品运用的专项技术很在行。双重专家可以让启发式评估更加有效。

1.4.9 可用性测试

可用性测试，本书的核心是观察具有代表性的最终用户完成真实任务，并凭借经验采集数据。测试可以简单地分为两大类别。一类测试比较正式，要求执行真实的实验来证明或证伪某些假设。第二类测试则不那么正式但依然严格（本书重点阐述的方法），要求以迭代式的测试来发掘可用性缺陷，并逐渐完善问题产品的形态。

1.4.10 追踪调查

追踪调查应用于正式产品发布之后。目的是通过调研、采访和观察为下一次产品发布收集数据。结构化的追踪调查或许是最可信和最准确的可用性评估技术了。因为能保证真实的用户、产品和环境作用在一起。不幸的是追踪调查的方法很少采用，因为设计师可能需要耗费2年的时间研究产品才有可观的收获。销售数据的好坏固然有一定价值，但也无法帮我们认识到产品的优势和劣势。

上述这些无论如何都不是不可更改的方法集合。这里也不太可能为读者评估出执行 UCD 所需的各种技术资金投入程度和复杂程度。很少有组织会实施所有的技术，也不太会以原始单纯的形式实施技术。通常情况下，这些技术会根据具体需求和项目限制作出调整和融合。大家可以在本书配套的网站 www.wiley.com/go/usabilitytesting 上找到关于这些技术更多的信息和资料。

现在让我们进入第2章，聚焦到本书的核心，也是所有技术中最负盛名的技术：可用性测试。

第2章

可用性测试是什么

可用性测试通常泛指所有评估产品或系统的方法。大部分时候，说这话的人显然是指第1章提到的方法之一。全书中提到可用性测试指的是：招募用户作为测试被试者，他们代表目标受众来评估产品满足特定可用性标准的程度。被贴上可用性测试标签的专家评估和走查，等等，并不满足目标受众这一条件。

可用性测试是一个研究工具，它植根于经典实验方法论。一个测试的范畴非常宽泛，经典实验方法可以是大样本真实测试和精密设计的测试，也可以是针对一个个体非正式的定性研究。每一个测试方法都具有不同的目的，需要不同的时间和资源。本书更强调一种非正式、设计复杂度低的测试，旨在适应工业产品研发过程快速迭代的要求。

2.1 为什么测试？测试的目标

来自某些公司的观点，可用性测试是提升产品利润的众多尝试中的一环。确实有很多方面的工作可以达到这样的效果，最终可以让用户极大地受益：设计决策来源于对目标用户的测试结果，通过被试者揭露设计问题并加以补救，可以尽量减少或消除用户的挫败感。

2.1.1 指导设计

可用性测试的整体目标是：在产品发布之前，识别和纠正产品以及帮助文档中的可用性缺陷，整合这些数据来指导设计。其目的就是确保产品有如下特性。

- 对目标用户来说是有用的。
- 容易学习。
- 有效且高效地让用户完成他们想要做的。
- 用起来满意（甚至是愉悦）。