


人机交互

Human-Computer Interaction

李伟欣, 夏榆滨
北航计算机学院

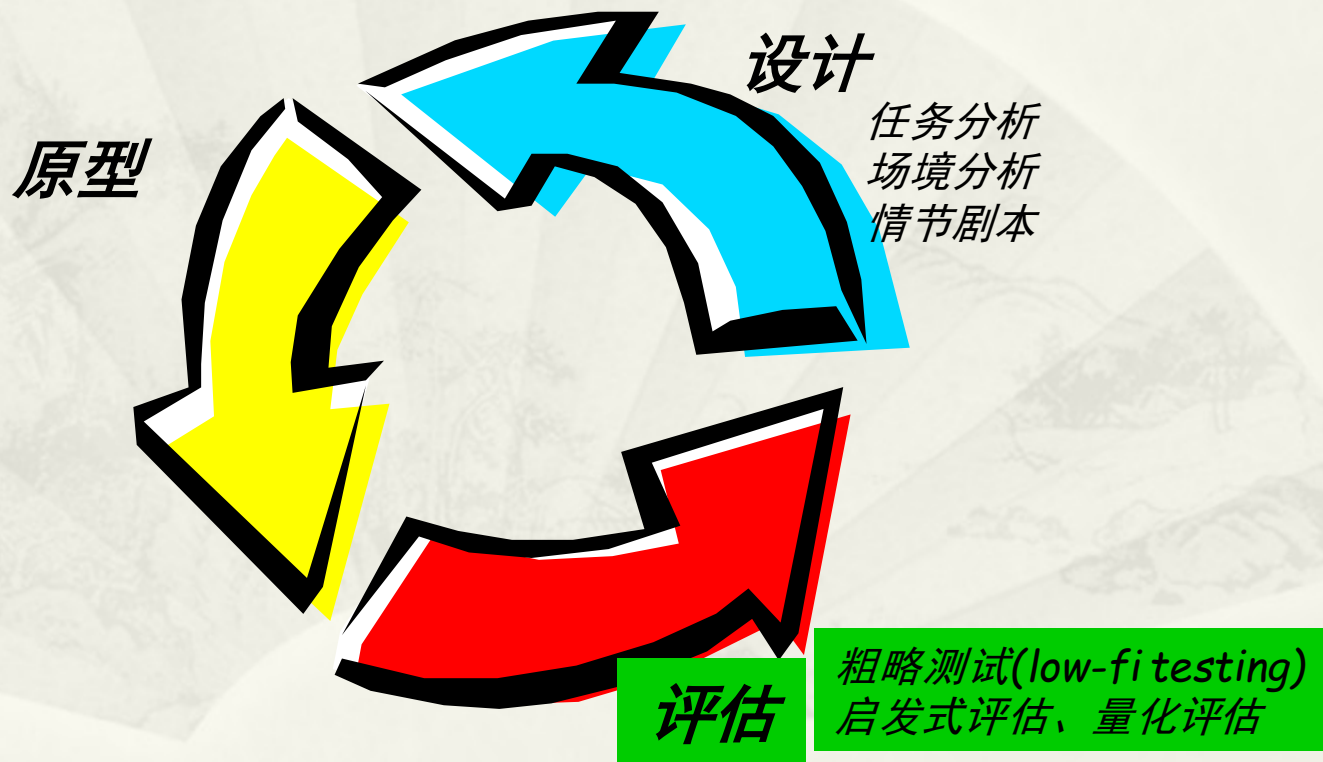


第七讲 可用性评估

天下难事必作于易

天下大事必作于细

迭代设计



1. 可用性评估(工程)范型

可用性



评估

工程化方法

1. 评估范型

- 较灵活的方法：

- ✓ 预测性评估：定性，简单、灵活

- 以实用的启发式原则和实践经验为基础，基于交互原理进行评估。

- ✓ 现场研究：定性，简单、灵活

- 面向实际应用的运行现场，进行研究与评估。

1. 评估范型

- 较正规的方法：

① 快速评估

② 可用性测试：对①的细化，规范化

2. 快速评估

2. 快速评估

快速评估→可基于：定性评估

✓ 低成本：

➤ 通常不需要专门的实验室或装备

✓ 快速

➤ 调整评估粒度，使其可以在短期内完成。

✓ 易用

➤ 简化组织与培训。

2. 快速评估

快速评估方法：

① 简化的可用性评估

具体可采用**认知走查**等形式，可快速和成本低廉地改进交互系统

② 启发式评估

具体可采用**启发式**评估流程

2. 快速评估

① 简化的可用性评估

认知走查(Cognitive Walkthrough , CW):

界面设计的**初期**，通过分析用户的心理加工过程来评价用户界面。
Polson et al. 1992提出

目标是，由**分析者**找出用户在完成任务的过程中是否会出现以下**问题**：

- a. 不知道下一步该做什么；
- b. 找不到解决问题的操作序列；
- c. 能发现操作的控制方式却不知道如何使用；
- d. 得不到合适的反馈，不知道他们的操作是否能顺利的完成任务。

发现问题之后，

下一步需要结合**认知心理学、设备特点和任务特点**找出原因，即：

- a. 评价界面如何支持用户的工作；
- b. 或说明为什么不能支持。

CW既可以由个人完成，也可以进行团队评价。

2. 快速评估

① 简化的可用性评估

认知走查(Cognitive Walkthrough , CW): 边做边说

HCI系统设计的初期, 通过分析用户的认知过程(Thinking aLoud)来评估交互系统

5个步骤(以交互界面为例)

- 1) 对交互界面的详细**描述**: 例如, 用户对菜单的位置的评判
- 2) 选择**任务**: 指定需评估的典型的界面任务, 确定一个或多个正确的操作序列。
- 3) 用**故事引导**用户了解完成任务所需要的操作: 用故事概念启示**用户选择**自己的行为
- 4) 观察记录用户的行为: 注意界面对用户自**引导作用**是否正常
- 5) 分析错误引导反映的界面上存在的问题: 分析原因
- 问题: 主观感受和**多义性**交互操作, 难以被分析

2. 快速评估

① 简化可用性评估

认知走查(Cognitive Walkthrough , CW):

走查的过程正是对用户问题解决过程的模仿

例如：

- I. 用户建立目标：检查目标的具体含义是否正确；
- II. 确定任务完成的操作：确定菜单条目、按钮、输入命令等操作；
- III. 用户完成选定的操作：执行操作action；
- IV. 认知：根据系统的反馈，检测自己的操作是否有利于目标的完成。

2. 快速评估

① 简化可用性评估

认知走查(Cognitive Walkthrough , CW):

- 具体执行时:

- ✓ 可由专业人员 “扮演” 用户，猜想用户的认知与操作。
- ✓ 逐步执行任务要求的行为序列
- ✓ **“关注是否容易学习”**
用户不是通过别人的“帮助”，而是自己通过探索性思考进行学习
- ✓ 需分析：交互系统对用户每一步的探索性学习，是否有较好的支持？

2. 快速评估

① 简化可用性评估

认知走查(Cognitive Walkthrough , CW):

在每一个行为点上，认知走查人员需要**判断并回答**:

- 行为结果和用户的目标是否一致?
 - 用户对活动的期望是否正确?
 - eg. 拖动文件到垃圾箱
- 用户可以看到提供的这个功能吗?
 - eg. 有些遥控器上的隐藏按键?
- 用户发现一个正确的功能，能知道该功能是所要的吗?
 - eg. 当用户看到隐藏按键，知道该按键的功能吗?
- 用户能理解行为执行后的反馈吗?

2. 快速评估

① 简化可用性评估

CW具有五个特征：

- a) 由专业分析者操作的、反映的是分析者的判断，而不是用户测试；
- b) 分析特定的用户任务，而不是对整个界面特征作评价；
- c) 分析正确操作是否被用户采用，而不是进行用户行为的预测；
- d) 要找出原因，而不是仅仅发现界面中存在的问题；
- e) 追踪用户的心理加工过程来发现问题，而不是聚焦于界面本身。

2. 快速评估

② 启发式评估 (Heuristic Evaluation)

✓ 基于经验性的规则评估交互系统的可用性

启发式评估：只考虑重要问题，忽略次要问题

例如，可建立如下准则：

- 设计准则
- 评估准则

依据这些准则进行的**试探性**评估。

例如：“能否降低用户记忆的负担？”

启发式评估的特点：

- 简单、低成本
 - 可由专家进行操作
 - 易于发现较大的设计缺陷
- 评估工作化繁为简，有可能与实际交互情况不符
 - 简单检查清单，很难涵盖整体情况
 - 非定量评估，有时不够准确

2. 快速评估

② 启发式评估 (*Heuristic Evaluation*)

- Jacob Nielsen的启发式原则
- 用于UI界面或素描

3-5位专家评估UI

对249个可用性问题分析, 总结的10条启发性原则

H1.0 – 约 1990	H2.0 – 约 1994
简单而自然的对话	美观和最小化设计
采用用户的语言	系统和现实世界相匹配
使用户的记忆负担减至最低	运用再认而不是回忆
一致性	一致性和标准化
提供反馈	系统状态可视化
提供清晰的退出标志	用户控制和自由度
提供快捷方式	使用的灵活性和效率
提供良好的出错消息	帮助用户识别、诊断、恢复错误
预防错误	错误预防
帮助和文档	帮助和文档

2. 快速评估

② 启发式评估 (*Heuristic Evaluation*)

例—启发1: 可视化 (反馈)

Time Left: 00:00:19 searching database for matches



46%

FarmVille Timer - Timers

Today Tomorrow Later All Timers Compl

	Clover	Ready!
		03:45 AM on Sep 16, 2010
	Alpine Roses	02:29:54
		75%
	Carnival Squash	03:58:40
		50%

H2-1: 系统状态可视化

- 让用户始终清楚当前所发生的事情
 - ✓ 避免耽误用户的决策
- 例如:
 - 关于系统响应
 - 0.1 秒: 不需要特殊的指示
 - 1.0 秒: 用户通常会放弃对数据的追踪
 - 10 秒: 用户等待操作的最大的可忍耐时间
 - 延时较短时: 可以用小沙漏显示延迟
 - 延时较长时: 可以用百分比图像条进行指示
 - 显示的时长可以稍微估计得多一点较好

2. 快速评估

② 启发式评估 (*Heuristic Evaluation*)

例---启发1: 可视化 (反馈)

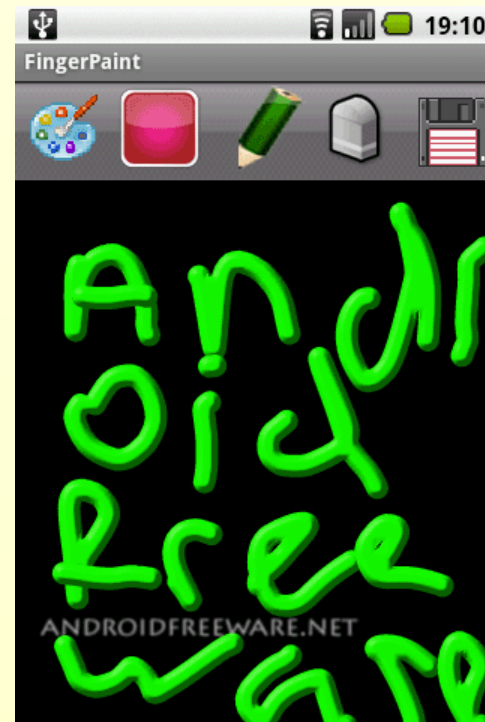
H2-1: 系统状态可视化

- 让用户始终清楚当前所发生的事情

✓ 可以提供冗余的信息



Feedback: Toolbar, cursor, ink



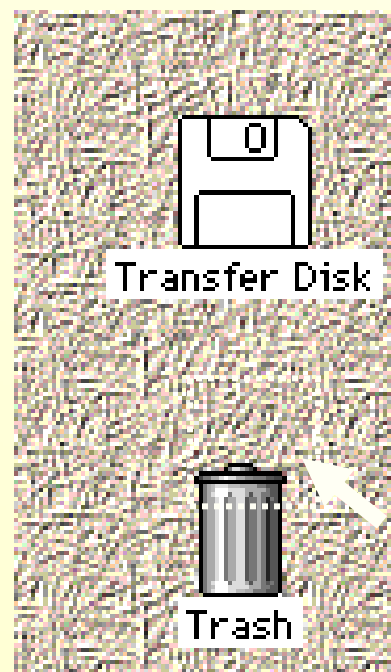
2. 快速评估

② 启发式评估 (*Heuristic Evaluation*)

例---启发2: 将系统和现实世界相匹配

H2-2: 系统和现实世界相匹配

- 使用用户熟练的语言
- 遵守现实世界的习惯
- 注意使用“隐喻”



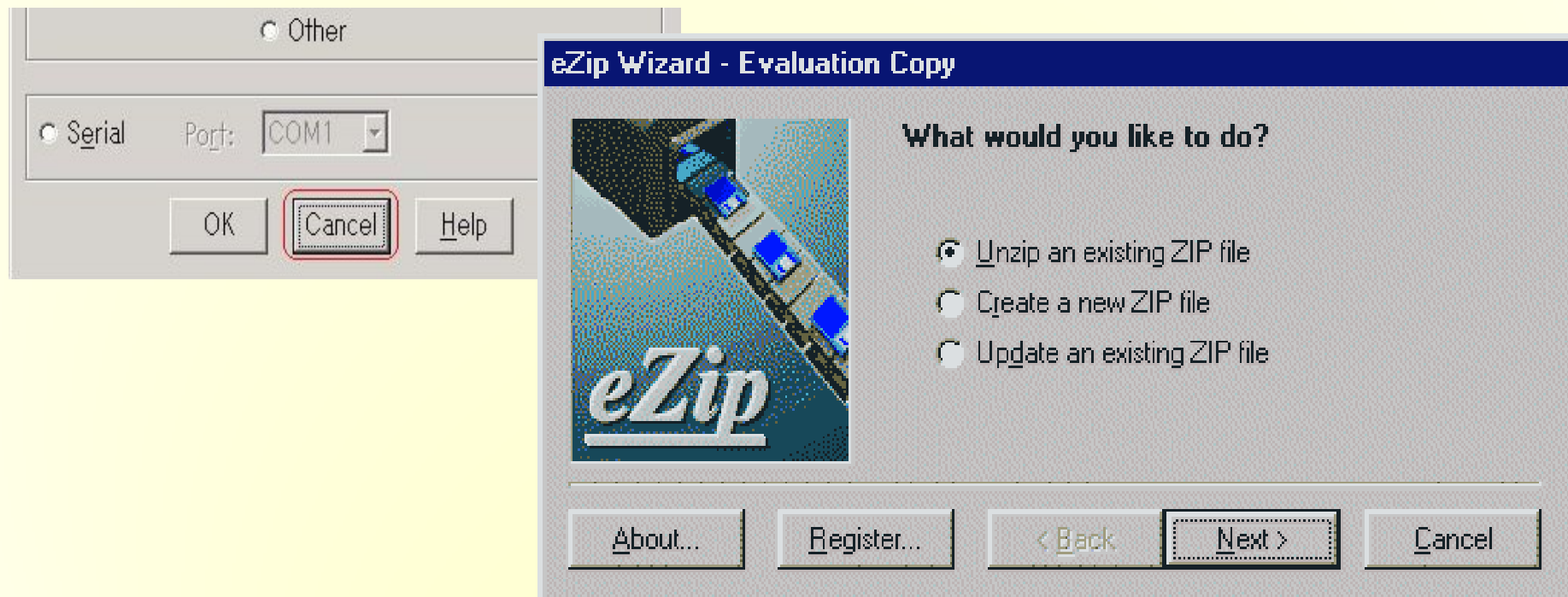
例如：早期的Mac桌面：将磁盘拖入垃圾箱，是删除，而不是弹出

2. 快速评估

② 启发式评估 (*Heuristic Evaluation*)

例---启发3: 控制 & 自由度

H2-3: 用户控制和自由度

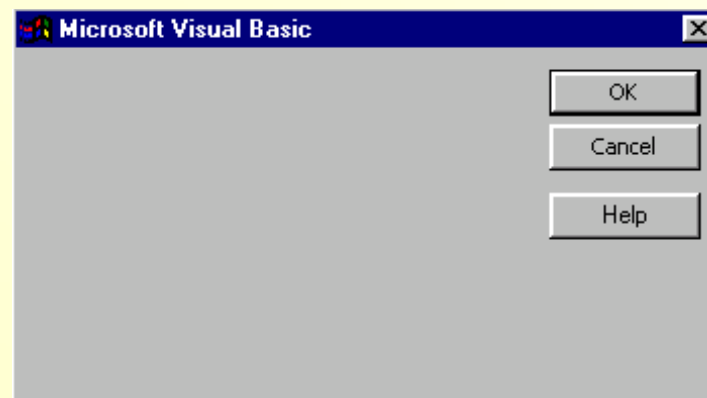
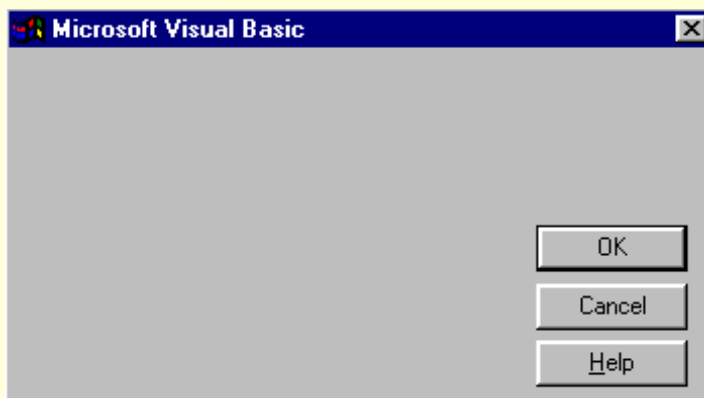
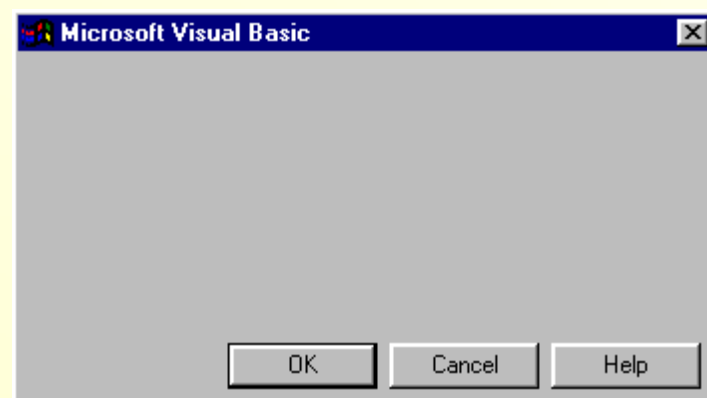
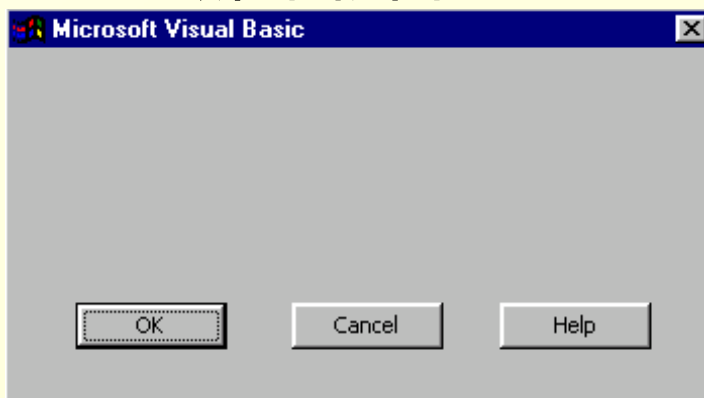


2. 快速评估

② 启发式评估 (*Heuristic Evaluation*)

例---启发4: 一致性

H2-4: 一致性和标准化



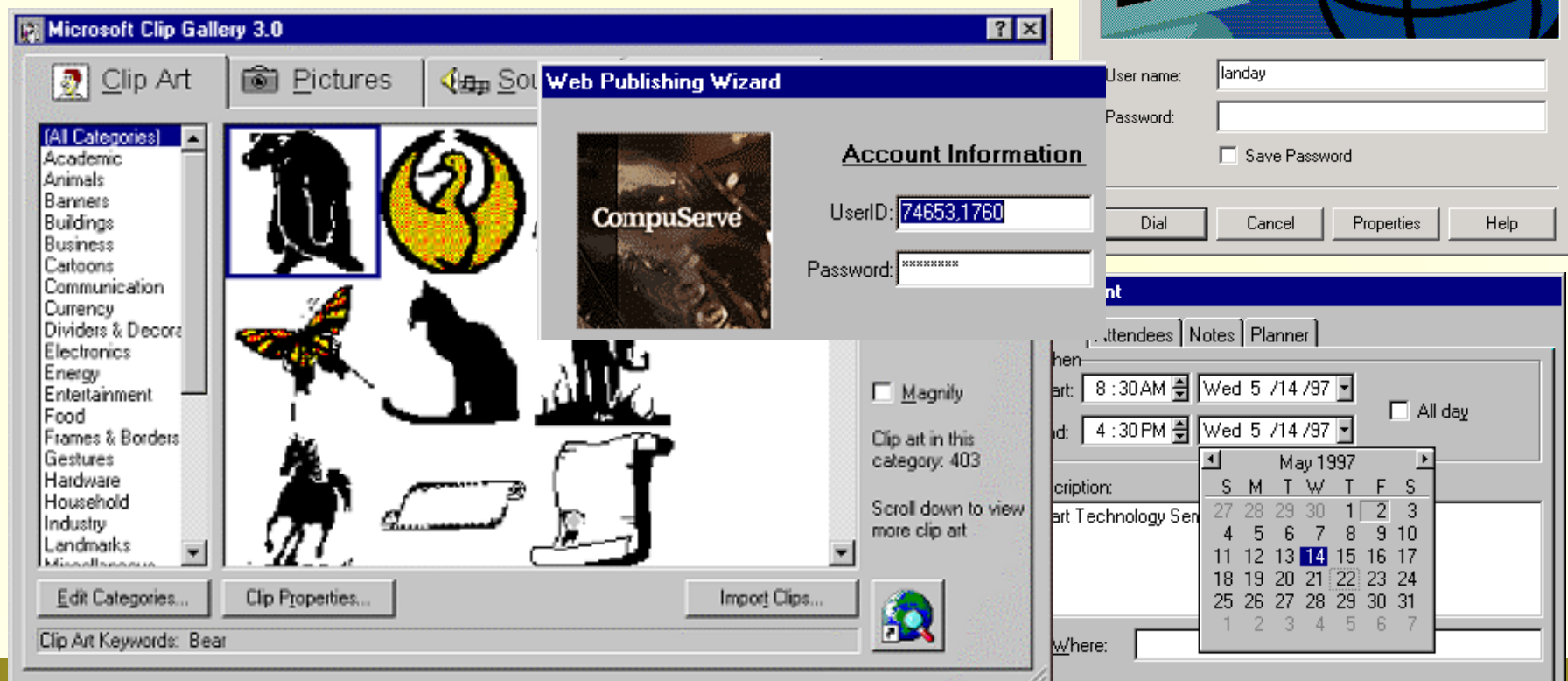
2. 快速评估

② 启发式评估 (*Heuristic Evaluation*)

例---启发5、6: 错误和记忆

H2-5: 避免错误

H2-6: 运用再认而不是回忆



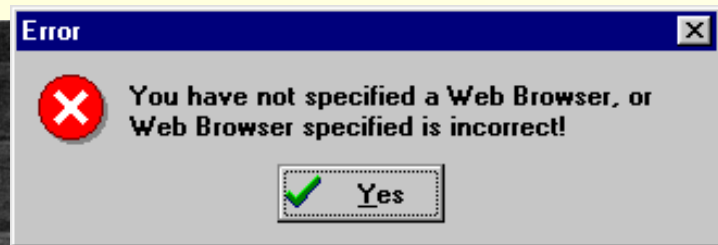
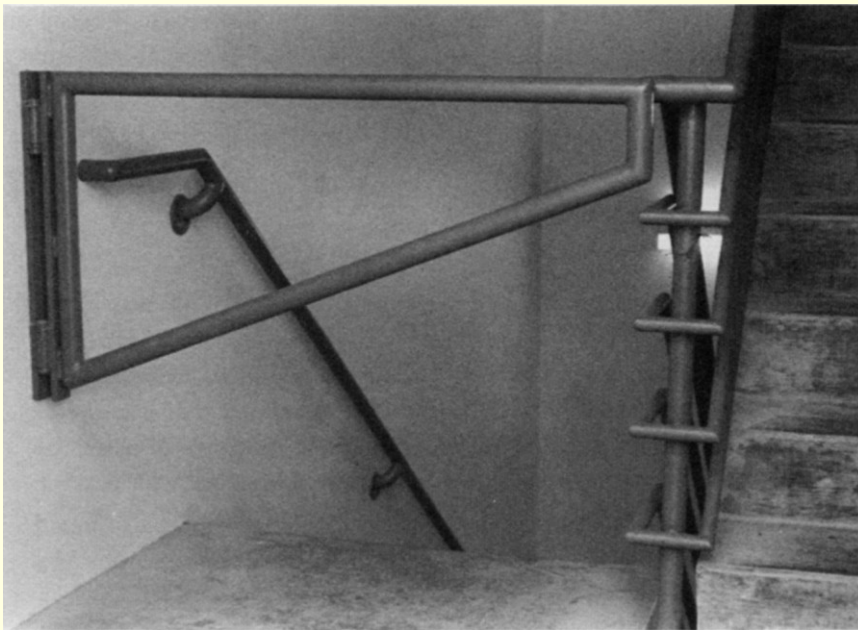
2. 快速评估

② 启发式评估 (*Heuristic Evaluation*)

例---启发5、6: 错误和记忆

H2-5: 避免错误

H2-6: 运用再认而不是回忆



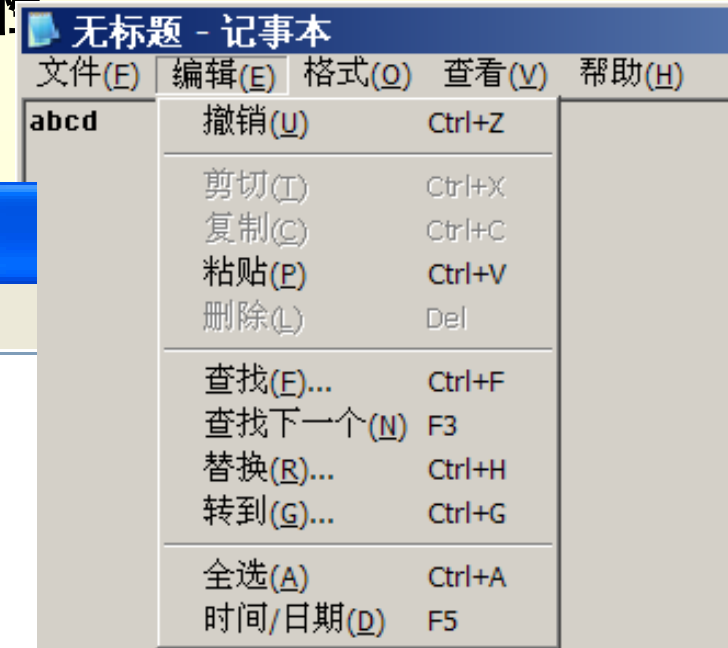
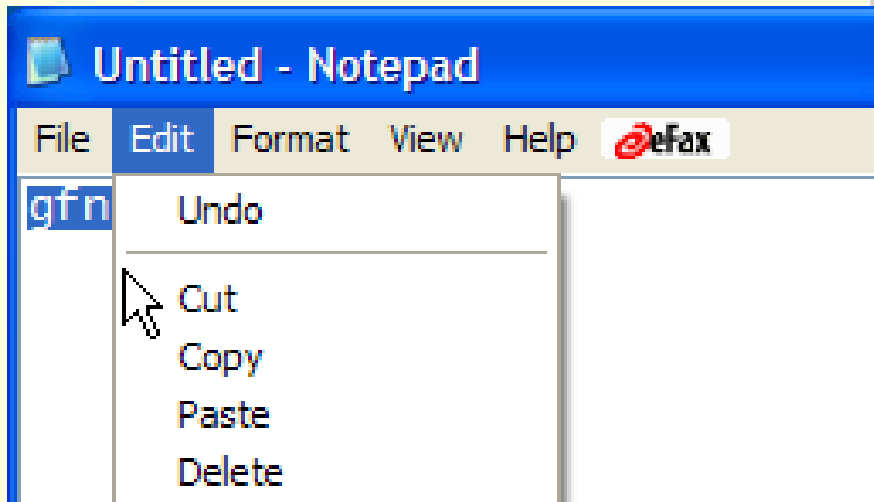
强制方法: 避免错误

2. 快速评估

② 启发式评估 (Heuristic Evaluation)

例---启发7: 灵活性

H2-7: 使用的灵活性和效率



附加快捷键,增加灵活性和效率

2. 快速评估

② 启发式评估 (*Heuristic Evaluation*)

例---启发8: 美观

H2-8: 美观和最小化设计

Form Title -- (appears above URL in most browsers and is used by WWW search engines)		Background Color:
Q&D Software Development Order Desk		FFFBF0
Form Heading -- (appears at top of Web page in bold type)		Text Color:
Q&D Software Development Order Desk <input checked="" type="checkbox"/> Center		000080
E-Mail responses to (will not appear on Web page)	Alternate (for mailto forms only)	Background Graphic
dversch@q-d.com		
Text to appear in Submit button	Text to appear in Reset button	<input type="radio"/> Mailto
Send Order	Clear Form	<input checked="" type="radio"/> CGI
Scrolling Status Bar Message (max length = 200 characters)		
WebMania 1.5b with Image Map Wizard is here!		
<< Prev Tab		Next Tab >>

不应有大量无关信息

2. 快速评估

② 启发式评估 (*Heuristic Evaluation*)

例---启发8: 美观

H2-8: 美观和最小化设计



2. 快速评估

② 启发式评估 (*Heuristic Evaluation*)

例---启发9: 帮助用户

H2-9: 帮助用户发现/诊断/修复错误



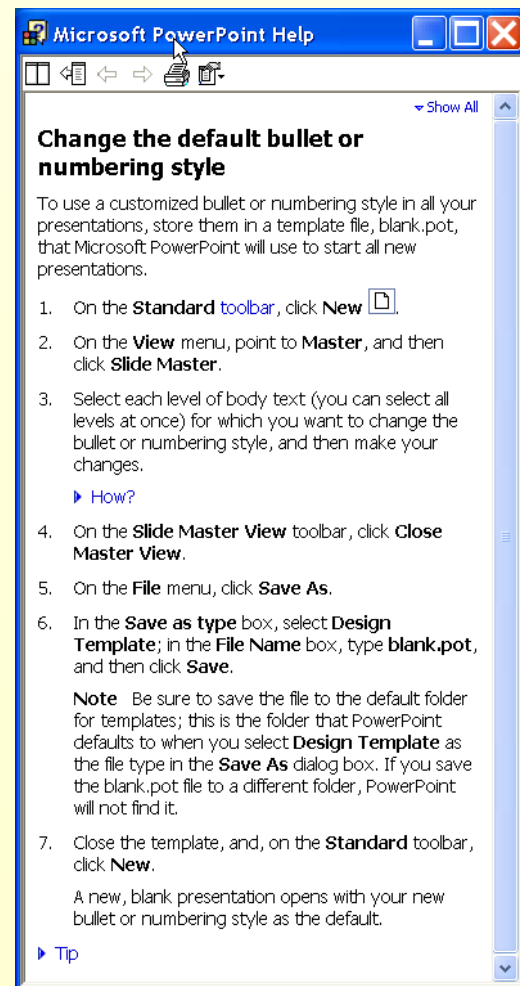
2. 快速评估

② 启发式评估 (Heuristic Evaluation)

例---启发10: 文档

H2-10: 帮助和文档

- 易于查找
- 提示用户当前工作相关的问题
- 列出易于执行的具体步骤
- 不要过于庞大
- 体谅用户



2. 快速评估

② 启发式评估 (*Heuristic Evaluation*)

启发评估流程

2. 快速评估

② 启发式评估 (*Heuristic Evaluation*)

启发评估流程

启发式评估阶段(1-2)

1) 评估前的必要知识（规则等）培训

2) 评估

- 每人单独做界面评估，然后汇总结果
- 两轮工作
 - 第一轮: 熟悉流程和范围
 - 第二轮: 有所聚焦
- 每人提供问题列表
 - 解释是否与启发式规则相匹配
 - 列出问题的清单

2. 快速评估

② 启发式评估 (*Heuristic Evaluation*)

启发评估流程

启发式评估阶段(3-4)

3) 问题等级评价

- 划分问题等级
 - 美观性，较小的，主要的，很严重的
- 先每个人单独划分，然后按小组划分

4) 形成任务报告

- 与设计团队讨论所得到的结果
- 给出潜在的解决方案建议
- 评估解决问题的难度

3. 可用性测试(评估)

3. 可用性测试(评估)

① 什么是可用性评估？

可用性定义（ISO9241-11）：

产品在特定环境下特定用户用于特定用途时所具有的：

效果、效率和用户主观满意度。

✓ 可用性测试与快速评估的对比：

- 前者是对后者评估过程与内容的一种细化。
- 前者需要制定不同主题的评估计划及其费用。如，人员、场地、道具、后勤等。后者是一种简化方法，成本较低。
- 前者需要评估团队花费大量的时间，后者可以控制在极短时间内完成。

3. 可用性测试(评估)

② 为什么要做评估？

●研究导向：发明与发现

- ✓ 验证性评估：发明了一个全新的技术，我想知道用户对这个创新技术的反应，以确认它是否有价值。
- ✓ 对比性评估：发明了一个可替代现有技术的新技术，我想知道它是否比现有技术更有价值（对比）。

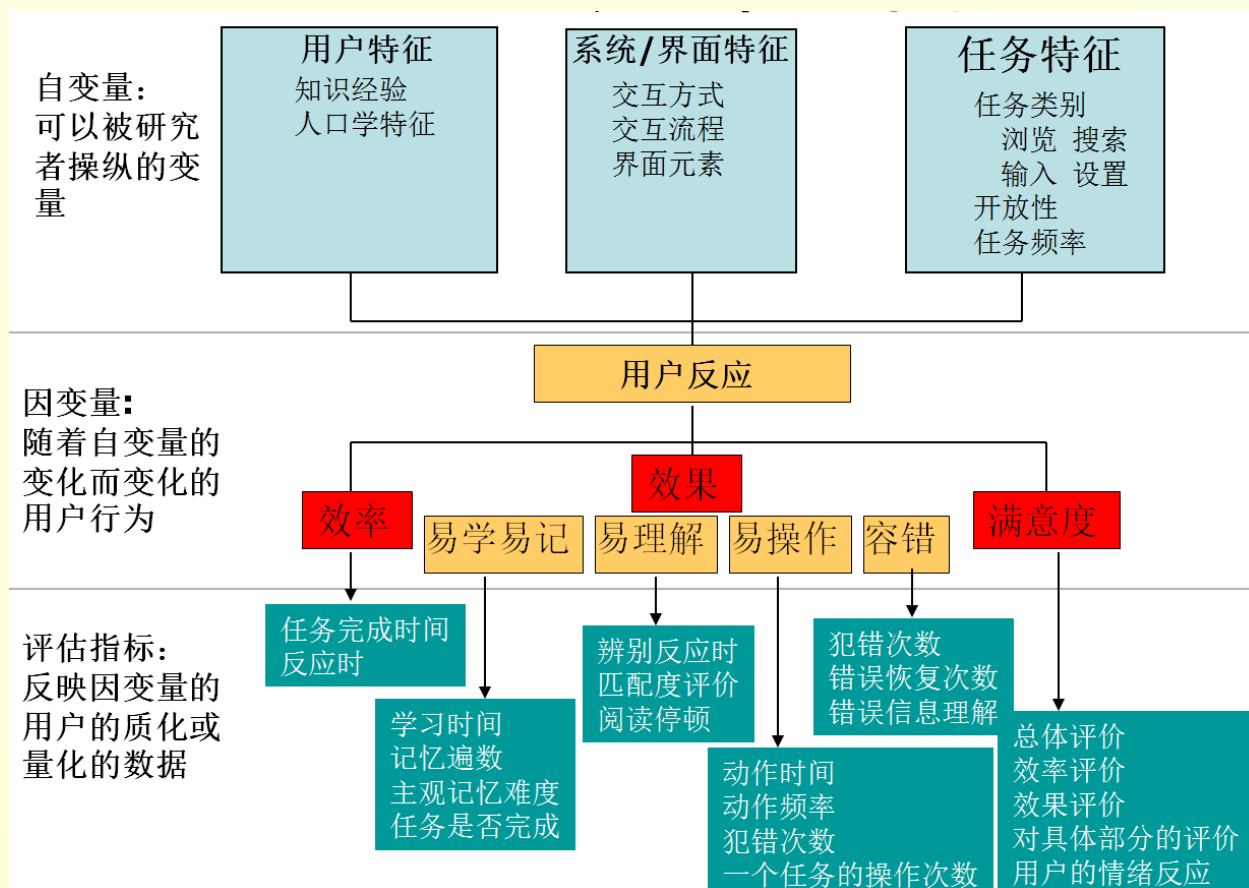
●产品导向：改进与创新

- ✓ 战略上的目标
 - a) 使我的产品所提供的功能用户真正“想要”和“想用”，建立起清晰的产品定位。
 - b) 使我的产品在同类产品中更具核心竞争力。
- ✓ 具体目标
 - a) 建立可用性标准
 - b) 控制开发成本
 - c) 降低开发风险
 - d) 降低技术支持和维护成本
 - e) 提高销售额

3. 可用性测试(评估)

③ 要评估哪些方面？

——确定评估指标



3. 可用性测试(评估)

④ 选择哪类评估方法？

1、用户参与的评估

(1) 对比实验：用户数：

- 多组设计的可用性比较；
- 用户数：例如，每组12人以上。

(2) 可用性观察测试：

- 通过观察找出界面可用性问题，提出设计改进建议；
- 可以是正式的实验室观察，也可以是非正式的观察。
- 用户数：2人以上，一般6-12人。

2、用户不参与的评估

(1) 启发式评估

- 可用性专业人员根据已有的可用性原则，对界面进行逐一评估

(2) 认知走查法

- 可用性专业人员将自己“扮演”成为用户，通过一定的任务对界面进行检查评估。

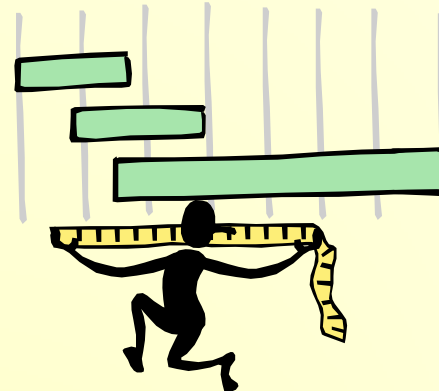
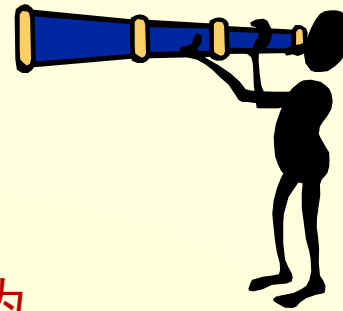
(3) 其他方法

3. 可用性测试(评估)

④ 选择哪类评估方法？

如何搜集实验数据

- 两种类型的数据
 - 过程数据
 - 观察记录用户的行为
 - 底层数据
 - 记录发生的各种数据 (time, errors, success)
 - i.e., 各种因变量



3. 可用性测试(评估)

④ 选择哪类评估方法？

出声思考 (Thinking Aloud)---

IBM公司Clayton Lewis 在1982年在《以任务为中心的界面设计》书中被阐述

- 需要知道用户所思,而不仅仅是所为
- 要求被试者在交互时说出:
 - ✓ 在想什么
 - ✓ 想要做什么
 - ✓ 引起问题的原因
 - ✓ 所看到的事物
- 记录
 - ✓ 描述用户在做什么,说明为什么这样做



3. 可用性测试(评估)

⑤ 评估前需要哪些准备？

- (1) 人员组成——用户参与的评估
 - 一个可用性评估组中最好有两位以上的可用性专家.
 - 专家工作：
 - 照看好测试用户 (使其自然放松)
 - 观察测试中发生的事情，管理测试过程，若需要的话提出问题
 - 观察时要做好笔记
 - 管理测试环境与设备，例如，录音设备。

3. 可用性测试(评估)

⑤ 评估前需要哪些准备？——评估准备

- (1) 人员组成——用户参与的评估（续）
 - 例如，对两位可用性专家管理的小组：
 - 测试结束时，能结合两种视角对测试结果进行分析
 - 其中一人可以照看测试用户并坐在测试用户的旁边，另外一个人可以做记录，并照看好记录设备
 - 两个人同时对被测试用户进行观察及交流
 - 更多的可用性相关专家加入测试小组（如项目经理、开发人员等），则有利于细化和深化相关的观察和交流内容。

3. 可用性测试(评估)

⑤ 评估前需要哪些准备? ——评估准备

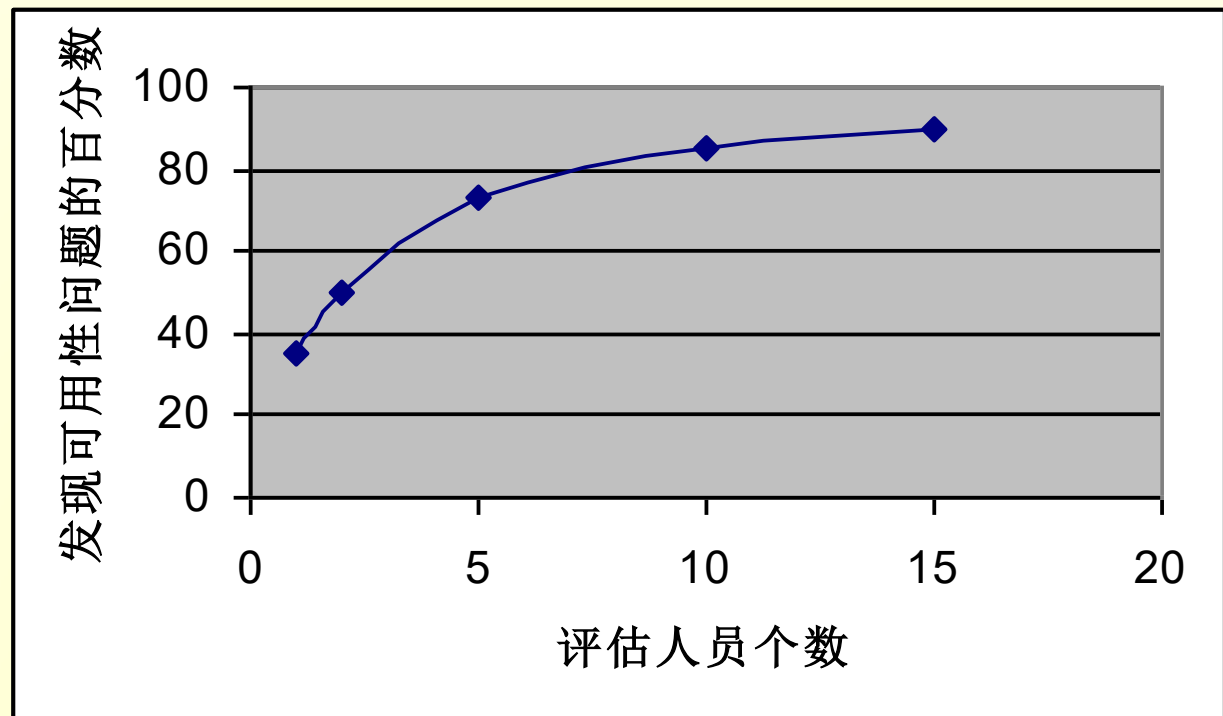
- (1) 人员组成——用户不参与的评估

- 最好有多位可用性相关的专家

3. 可用性测试(评估)

⑤ 评估前需要哪些准备？ ——评估准备

- 评估人员个数与发现问题的可能性：



3. 可用性测试(评估)

⑤ 评估前需要哪些准备？——评估准备

• (2) 人员培训——用户参与的评估

- 评估小组中至少有一个成员是可用性专业人员。所有的其他成员必须知道如何完成如下工作：
 - 在可用性测试期间如何与用户进行沟通：考虑用户背景特点
 - 客观观察的方法：考虑系统功能
 - 聆听用户真实意图的方法：考虑用户认知特点
 - 将观察（所看到的和所听到的）与推论（对观察做出假设和结论）区别开来：不要主观臆断
 - 有效地作笔记的方法：不丢失相关信息
 - 询问时，不引导或责备用户：保证信息与实际运行情况匹配

3. 可用性测试(评估)

⑤ 评估前需要哪些准备？——评估准备

• (2) 人员培训——用户不参与的评估

- 评估小组中所有成员均为可用性相关人员。所有的成员必须知道如何完成如下工作:
 - 理解可用性评估原则
 - 了解评估实施过程
 - 共同形成评估任务，分工好
 - 形成规范化的可用性评估的报告

3. 可用性测试(评估)

⑤ 评估前需要哪些准备？——评估准备

• (3) 观察者安排——用户参与的评估

— 如何安排观察者？

- 在典型的可用性测试中，防止过度分散相关用户的注意力，例如，与用户待在同一个房间的观察者(专家)尽量不多于1人。
 - ✓ 可用性实验室中通常具有一个独立的房间，在这个房间里，其他的非现场观察者可以通过单向玻璃或显示器来观察可用性测试过程。
 - ✓ 如果把测试安排在一个会议室，则可以在隔壁为观察者(专家)搭建一个视频或投影装置。

3. 可用性测试(评估)

⑤ 评估前需要哪些准备？——评估准备



3. 可用性测试(评估)

⑤ 评估前需要哪些准备？ ——评估准备



3. 可用性测试(评估)

⑤ 评估前需要哪些准备？ ——评估准备

- (4) 评估计划

- 评估范围：评估内容
- 评估维度：评估哪些方面
- 测试用户：用户类型
- 评估环境：实验室？会议室？其他地方？远程？
- 评估预算：评估成本

3. 可用性测试(评估)

⑤ 评估前需要哪些准备？——评估准备

- (5) 评估任务设计

- 选择需要评估的相关交互任务
- 准备、试验并精炼为这些任务所设计的场景。
- 分配测试时间。
- 如果是对比实验
 - 被对比任务内评估、被对比任务间评估、混合评估
 - 无关变量的消除与平衡
 - 顺序效应的平衡
 - 被对比任务间的“练习”（熟悉）效应的消除
- 满意度问卷

3. 可用性测试(评估)

⑤ 评估前需要哪些准备？——评估准备

- (6) 招募测试用户——用户参与的评估
 - 招募那些当前的或潜在的用户来参加测试
 - 组建可用性评估团队
 - 建立“用户群”

3. 可用性测试(评估)

⑤ 评估前需要哪些准备？——评估准备

- 预测试（用户参与的评估）

- 尽量找与测试用户背景相似的人
- 例如，2个人左右
- 评估过程预演
 - 检查软硬件是否有冲突
 - 任务描述是否有问题
 - 实验者的口头说明是否清晰
 - 完整过一遍相关任务。

3. 可用性测试(评估)

⑥ 如何实施评估？——评估实施

- 可用性观察测试——用户参与的评估
 - （1）评估专家概括描述测试目的
 - 自我介绍
 - 介绍目的
 - 参观被测系统，功能介绍
 - 保密声明

3. 可用性测试(评估)

⑥ 如何实施评估？——评估实施

- 可用性观察测试——用户参与的评估

- （2）出声思维训练

- 在做什么？
 - 看到了什么？
 - 怎么想的？
 - 有什么疑问？

3. 可用性测试(评估)

⑥ 如何实施评估？——评估实施

- 可用性观察测试——用户参与的评估

- (3) 任务介绍

- 向用户介绍有关系统的必要信息
 - 书面和口头告知用户任务：一次一个任务。
 - 答疑：回答用户关于目标、过程或任务的问题。但不能回答类似于“我该怎样做这项任务？”的问题，可以告知他们：“那正是我需要观察的情况，我想看看系统是怎样帮助你完成这个任务的。”

3. 可用性测试(评估)

⑥ 如何实施评估？——评估实施

• 可用性观察测试——用户参与的评估

— (4) 开始测试

- 用户采用出声思维法完成任务。如果用户沉默5-10秒，告诉他“请一直说”。
- 不要详细解释，如不要说“你在想什么”或“请解释你在做什么”
- 如果用户开始解释他们的步骤而不是报告工作记忆区的内容，先让其停止，然后说：“请不要向我解释你正在做什么。表现得就好像你一个人在解决问题时一样自言自语。”
- 适时让用户休息

3. 可用性测试(评估)

⑥ 如何实施评估？——评估实施

- 可用性观察测试——用户参与的评估

- (5) 结束测试

- 询问用户关于系统或研究的问题。回答可以回答的问题或寻求答案的途径。
 - 询问用户关于所测试产品的意见和改进建议。
 - 致谢

3. 可用性测试(评估)

⑥ 如何实施评估？——评估实施

- 启发式评估、认知走查——用户不参与的评估（参考前面：快速评估相关内容，细化相关内容）

— 评估过程

- 一个小组 (例如，3-5人) 评估一个交互任务：可以无用户
- 每个评估者根据启发式原则独立地地进行检查评估
- 合并问题：不同的评估者可能会发现不同的问题
- 共同估计问题的严重性
- 任务报告/头脑风暴 → 如何修正问题，给出解决方案




3. 可用性测试(评估)

⑦ 如何撰写评估报告? ——评估报告

- 评估报告因不同的目标和阅读对象而不同。
- 举例：
 - ✓ 以启发式评估报告为例——用户不参与的评估
 - 可用性问题的事实（关键事件的证据）
 - 用户的目标。
 - 用户行动的效果。
 - 对问题的解释
 - 问题的严重性或优点的好处
 - 可能的解决方案/其他可能方案

3. 可用性测试(评估)

⑦ 如何撰写评估报告? ——评估报告

<p>(范例)</p> <p>标识符 SunLine-1 问题</p> <p>问题描述 功能模块混淆，导致用户容易出错理解和交互错误。</p> <p>评估依据 启发式原则：信息组织关系的清晰性和可理解性。</p> <p>解释说明 下面是网站页面截图。</p>  <p>这个页面的优点是：设计者很好地利用了 Tab 的概念组织功能模块，使得功能模块之间的视觉组织比较清晰。但是，我们看到上述 7 个模块之间明显不是一个并列关系，而设计者却将其设计成并列关系。这会导致用户在网站功能的理解上出现问题，在交互上也会出现问题。</p> <p>事实上，“登录”和“注销”只是用户第一次进入或退出的时候才使用一次，其与“编辑日志”、“日志管理”等可能需要来回切换的模块放在一起显然会增加用户的误操作。如不小心点击了“注销”，此时用户不得不重新登录，这会让用户感到厌烦。</p> <p>另外，“编辑日志”中的“编辑”一词太计算机化，应该改为“日志写作”。</p> <p>解决方案 该页面拆分成两个页面，第一个页面为登录页面，第二个页面为日志写作和管理页面。两个页面的原型如下图所示。</p>  <p>登录页面</p>	 <p>日志写作和管理页面</p> <p>另外，“注销”也在“日志写作和管理页面”中，但放在整个网页的右上角处（在该原型中已有标出）。</p> <p>与其他可用性方面的关系 无</p>
---	--

(范例)

标识符

SunLine-1 问题

问题描述

功能模块拥挤，导致用户容易出现理解和交互错误。

评估数据

启发式原则：信息组织关系的清晰性和可理解性。

解释说明

下面是网站页面截图。



这个页面的优点是：设计者很好地利用了 Tab 的概念组织功能模块，使得功能模块之间的视觉组织比较清晰。但是，我们看到上述 7 个模块之间明显不是一个并列关系，而设计者却将其设计成并列关系。这会导致用户在网站功能的理解上出现问题，在交互上也会出现问题。

事实上，“登录”和“注销”只是用户第一次进入或退出的时候才使用一次，其与“编辑日志”，“日志管理”等可能需要来回切换的模块放在一起显然会增加用户的误操作，如不小心点击了“注销”，此时用户不得不重新登录，这会让用户感到厌烦。

另外，“编辑日志”中的“编辑”一词太计算机化，应该改为“日志写作”。

解决方案

该页面拆分成两个页面。第一页面为登录页面，第二个页面为日志写作和管理页面。两个页面的原型如下图所示。



登录页面

标识符

问题描述

评估数据

解释说明

解决方案

日志写作

日志管理

评论管理

引用管理

系统设置

内容管理

日志写作和管理页面

另外，“注销”也在“日志写作和管理页面”中，但放在整个网页的右上角处（在该原型中没有画出）。

与其他可用性方面的关系

与其他可用性方面的关系

无

3. 可用性测试(评估)

参考资料：美国国家标准与技术研究院
《Common Industry Format for Usability Test Reports》

Produced by the NIST Industry USability Reporting project:
www.nist.gov/iusr

总结：评估范型

评估范型	预测性评估	现场研究	快速评估	可用性测试
用户角色	通常不参与	自然行为	自然行为	执行测试任务集
控制权	评估人员为专家	评估人员与用户合作	评估人员实施最低限度控制	评估人员密切控制
评估地点	类似实验室的环境，通常在客户处进行	自然工作环境	自然工作环境或实验室	实验室
适用情形	使用 启发式评估 的 原型测试 ，可在 任何阶段 进行：模型可用于评估潜在设计的特定方面	常用于 设计初期 ，以检查设计是否满足用户需求，发现问题，发掘应用契机	快速了解 设计反馈 。可使用其他交互范型的技术，如 启发式评估	测试 原型或产品
数据类型	专家们列出 问题清单 ，由模型导出 量化数据 (如两种设计的任务执行时间)	应用草图、场景、例证等的 定性 描述	通常是 定性 的非正式描述	定量数据 ，有时是统计数据。可采用 问卷 调查或访谈搜集用户意见
反馈到设计	专家列出一组 问题 ， 通常附带解决方案建议 。为设计人员提供根据模型计算出的时间值	通过描述性的 例证 、 草图 、 场景 和 工作日志	通过 草图 、 例证 、 报告	通过 性能评估 、 错误统计报告 等为未来版本提供设计标准
基本思想	专家检查以实用的启发式原则和实践经验为基础，采用基于理论的分析模型	可以是客观观察或现场研究	以用户为中心、非常实用	基于试验的实用方法，即可用性工程

可用性测试(评估)技术总结

方法	生命周期阶段	所需用户人数	主要优点	主要缺点
启发式评估	设计早期、迭代的“内部循环”	无	能发现单个可用性问题和解决专家用户碰到的问题	不涉及真实用户，故无法在用户需求方面有“惊人发现”
边做边说(认知走查)	任务分析、后续研究	3-5人	查明用户的误解、测试费用低	用户感到不自然，即便专家用户也很难用语言表述
观察	任务分析、后续研究	3人或以上	揭示用户的真实任务、对系统功能与特征提出建议	很难安排，实验人员无法控制
问卷调查	任务分析、后续研究	至少30人	发现用户主观偏好、容易重复	需要进行小规模试验(避免出现误解)
访谈	任务分析、后续研究	5人	灵活、可以深入了解用户观点和用户体验	耗时，难以进行分析比较
焦点小组	任务分析、参与式设计	每组6-9人	自发响应、团体动力学	难以分析，有效性差
可用性测试	最终测试、后续研究	至少20人	发现经常使用(或很少)使用的特征、能够重复进行	对应用的分析需要大量数据，面临用户隐私问题困扰
交互记录和用户反馈	后续研究	上百人	跟踪用户需求和观点的变化	需要专门部门来处理回复

· 作业

1. 什么是启发式评估，如何撰写启发式评估报告？
2. 为微软powerpoint程序设定一个快速评估小任务（找到一个*相对*不完美之处），自己做认知走查和问题记录，并拟定一个评估报告提纲。
(1~2页纸)



再见