

界面可用性评价之启发式评价法

葛燕^{1,2},周荣刚^{1,2},张侃¹

(1. 中国科学院心理研究所,脑与认知科学国家重点实验室,北京 100101;
2. 中国科学院研究生院,北京 100039)

摘要:对界面进行可用性评价,是一个成功的界面设计的必经阶段。在众多的评价技术中,启发式评价法是一种非常经典的可用性诊查方法。文章详细介绍了该方法的特点、操作步骤及相关研究,并与可用性测试方法进行比较,结合一些研究进行了简单评价。

关键词:启发式评价法,可用性评价,界面

中图分类号:TB18 **文献标识码:**A

1 概述

启发式评价法(Heuristic Evaluation, HE)是最早使用的经典可用性评价方法之一。最初由Nielsen和Molich在1990年提出^[1],后获得广泛使用,至今仍是最流行的可用性诊察方法之一。

启发式评价是一种专家评审法,由几个评价者根据通用的可用性原则和经验来发现系统潜在的可用性问题。启发式评价以可用性启发式为基础,让少量评价者对用户界面进行系统的检查,找出存在的可用性问题,从而构建一个对系统的评价并试图找出解决的方案^[1]。由于启发式评估不需要用户参与,也不需要特殊设备,所以成本相对较低,而且使用快捷,被称为“简化的可用性工程方法(discount usability engineering)”^[2]。

2 启发式评价法的重要成分

专业的评价者和公认的启发式是启发式评价法不可缺少的两大成分^[3]。

2.1 评价者

在启发式评价中,评价者的绩效存在着很大的个体差异。具有一般的可用性知识和被评价的界面的相关经验的评价者发现的问题最多^[1]。因此评价者一般为两个领域的专家:软件工程领域或可用性和人的因素领域的专家。

在评价者人数上,单一的评价者就可以独自进行用户界面的可用性评价,但是,任何单一被试

都会遗漏界面中的大多数可用性问题,也有一些问题只能被少数人发现。人数越多,发现的可用性问题越多,一个评估人员可发现35%的可用性问题,5个可发现大约75%的可用性问题。但是在超过5人以后随着人数的增加,投入收益比率降低。因此,Nielsen建议采用3~5个评价者^[2]。

2.2 启发式可用性原则

启发式是进行启发式评价的基础,启发式指的是一种方针,或者说是一种一般的原则,把它们叫做“启发式”是因为它们比特定的可用性原则更加重要。

不同的研究者定义了不同的启发式。Nielsen根据对249个可用性问题进行因素分析提炼了一套具有最大解释力率的启发式,从而获得了一套修订的启发式^[1]。主要包括以下10条:

- (1) 系统应及时有效的提供其被使用过程中的状态
- (2) 系统应符合用户的认知加工习惯
- (3) 用户应可控制和自由使用系统
- (4) 系统设计的一致性和标准化
- (5) 移防错误的发生
- (6) 再认而不是回忆
- (7) 灵活和快捷的使用
- (8) 美观精炼的设计
- (9) 帮助用户识别、诊断和修正错误

基金项目:教育部“十五”规划课题(FBB011067);国家自然科学基金资助项目(30270465);中国科学院院长基金资助项目(JH02013)

作者简介:葛燕(1981-),女,云南玉溪人,工程心理学专业硕士研究生。

(10) 帮助文件的设置

启发式原则是进行启发式评价的关键。启发式的目的是帮助评价者找出可用性问题,而不是限制他们仅仅找出与这些启发式矛盾的问题。以上列出的启发式是比较典型的原则,在实际应用中,研究者往往根据需要提出特定的启发式,如 Sutcliffe 等人^[4]根据虚拟现实的特点,制定了适用于虚拟现实应用软件界面的启发式;Paddison^[5]等人也针对特殊领域开发出不同的启发式,以满足用户的要求;Sutcliffe^[6]提出新的启发式,将可用性的评价范围扩展到设计的动机方面。

3 启发式评价的操作步骤

作为一种非正式的可用性诊察技术,启发式评价并没有一种非常标准操作程序。Nielsen 自己也指出,研究者在使用时会根据实际情况进行变通。但 Nielsen 也提出了启发式评价需要具备的一些基本过程^[11]:

3.1 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可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价需要 1~2h。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可用性问题的优先级。

3.2 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:每一个评价者都独立完成一份书面报告,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。使用观察者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。

3.3 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,报告主要的可用性问题的设计方面存在的一般问题,并提出一些设计和改进的建议。

3.4 评价的成果

使用启发式评价方法的产物是可用性问题的列表,每一个问题后面还注明了评价者认为它所违

反的可用性原则。启发式评价不是提供一种系统的方法来对产生的可用性问题进行修正,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可用性问题,当可用性问题被确认时,它们也同时得到相当明显的修正。

4 启发式评价法的应用领域

启发式评价法可用于界面的设计、编码、测试和使用阶段,它不需要完成特定的任务或执行步骤,特别适合界面设计的早期^[7]。

根据启发式评价的特点,Englefield 总结出在以下领域可使用启发式评价:对典型用户限制访问的领域;需要快速转变的领域;评价者不能在同时间地工作的领域;注意的宽度比敏感度更重要的领域;强烈的客户关系降低了对从真正的用户收集支持性建议的需要的领域^[8]。

近年来,启发式评价在网站原型^[9]、医疗设备^[10]和软件界面^[11]等方面应用非常广泛,并取得了显著的成果。

5 启发式评价存在的问题

5.1 所发现的可用性问题

有研究者将启发式评价与其他可用性评价方法进行了比较,发现虽然启发式评价报告了大量低优先级的问题,但是在所比较的四种方法中,它是需要花费最少努力就能获得最多严重问题的方法^[12];在相同诊察时间内,与认知走查法和界面原则相比,启发式评价发现了最多比例的真实问题^[13]。但是,也有研究者认为通过启发式评价所找到的可用性问题并不真的成为问题^[14]。Nielsen 也指出启发式评价更容易发现次要问题,因此需要等级评定作为补充^[11]。也有研究者建议使用工效学检查表来弥补这一不足^[15]。

与用户测试相比,启发式评价发现问题的类型不同^[16]。在识别与基于技能的绩效水平结合的可用性问题和与基于规则的绩效水平结合的可用性问题,启发式评价的效果要优于用户测试,而在发现与基于知识的绩效水平结合的可用性问题,用户测试的效果要优于启发式评价。专家仍然不能取代真正的用户,所以在使用时应两者结合^[1,15]。

5.2 评价者效应

在启发式评价方法上,由于目标、评价过程和可用性问题的标准都不是很清晰,导致了评价者效应的存在^[17],解决的办法是成倍地增加评价者。同时 Sears 发现,评价者的经验也会影响可用

性问题的发现^[18]。因此,如何选择好的评价者也是启发式评价面临的重要问题。

6 启发式评价的新发展

基于启发式评价存在的问题,一些研究者提出了改进的方法。Muller 根据 Nielsen 的启发式评价提出了可供分享的启发式评价法,在两个方面进行了扩展:为启发式列表增加了新的领域;在评价者中增加了用户(即工作领域的专家)^[19]。另外,Sears 提出的启发式走查法综合了启发式评价、认知走查法和可用性走查法的优点,是一种比启发式评价更具结构性,具有有效的任务导向的评价方法。可用性问题的诊断也是一种在完全性、有效性和效率性等方面优于启发式评价的新方法^[20]。

7 小结

启发式评价作为一种“简化的可用性工程学”方法,非常适用于时间、金钱等资源短缺的情况。但是启发式评价不能确保提供“完美”的结果或找出界面最终的所有可用性问题。它有自己的优点,也存在一些不足,因此最好能与其他相应的技术结合使用。

参考文献:

- [1] Nielsen J. Heuristic evaluation [A]. Nielsen J, Mack R. Usability inspection methods [M]. New York: John Wiley and Sons, 1994. 25-62.
- [2] Nielsen J. Usability engineering [M]. San Francisco: Morgan Kaufmann, 1994. 115-164.
- [3] Rajesh V. Heuristic evaluation [EB/OL]. <http://www.cc.gatech.edu/classes/cs675197winter/Topics/heuristic/>,2004-11-30.
- [4] Sutcliffe A, Gault B. Heuristic evaluation of virtual reality applications[J]. *Interacting with Computers*, 2004, 16(4): 831-849.
- [5] Paddison C, Englefield P. Applying heuristics to accessibility inspections[J]. *Interacting with Computers*, 2004, 16(3): 507-521.
- [6] Sutcliffe AG. Assessing the reliability of heuristic evaluation for website attractiveness and usability [EB/OL]. <http://csdl.computer.org/comp/proceedings/hicss/2002/1435/05/14350137.pdf>,2004-11-30.
- [7] Avouris NM, Tselios N, Tatakis EC. Development and evaluation of a computer-based laboratory teaching tool[J]. *Computer Applications in Engineering Education*, 2001, 9(1): 8-19.
- [8] Englefield P. A pragmatic framework for selecting empirical or inspection methods to evaluate usability[EB/OL]. [http://www-306.ibm.com/ibm/easy/eou/ext.nsf/publish/50?OpenDocument&./publish/1118/\\$File/paper1118.pdf](http://www-306.ibm.com/ibm/easy/eou/ext.nsf/publish/50?OpenDocument&./publish/1118/$File/paper1118.pdf), 2004-11-30.
- [9] Levi MD, Conrad FG. A heuristic evaluation of a World Wide Web Prototype [EB/OL]. <http://stats.bls.gov/ore/html/paper/st960160.htm>,2004-11-30.
- [10] Graham MJ, Kubose TK, Jordan D, et al. Heuristic evaluation of infusion pumps: implications for patient safety in Intensive Care Units[J]. *International Journal of Medical Informatics*, 2004, 73(11-12): 771-779.
- [11] Johnson D, Wiles J. Effective affective user interface design in games[J]. *Ergonomics*, 2003, 46(13-14): 1332-1345.
- [12] Jeffries R, Miller JR, Wharton C, et al. User interface evaluation in the real world: A comparison of four techniques[EB/OL]. <http://www.miramontes.com/writing/uevaluation/uevaluation.html>,2004-11-30.
- [13] Cuomo DL, Bowen CD. Understanding usability issues addressed by three user-system interface evaluation techniques[J]. *Interacting with Computers*, 1994, 6(1): 86-108.
- [14] Stanton NA, Stevenage SV. Learning to predict human error: Issues of acceptability, reliability and validity[J]. *Ergonomics*, 1998, 41(11): 1737-1747.
- [15] 吴昌旭, 张侃. 诊断型用户界面可用性评价方法(M)(上)——简介与评价[J]. *人类工效学*, 2000, 6(3): 54-57.
- [16] Fu LM, Salvendy G, Turley L. Effectiveness of user testing and heuristic evaluation as a function of performance classification[J]. *Behavior and Information Technology*, 2002, 21(2): 137-143.
- [17] Hertzum M, Jacobsen NE. The evaluator effect: a chilling fact about usability evaluation methods[J]. *International Journal of Human-Computer Interaction*, 2001, 13(4): 421-433.
- [18] Sears A. Heuristic walkthrough: finding the problems without noise[J]. *International Journal of Human-Computer Interaction*, 1997, 9(3): 213-234.
- [19] Muller MJ, Matheson L, Page C, et al. Methods & tools: participatory heuristic evaluation[J]. *Interactions*. 1998, 5(5): 13-18.
- [20] Andre TS, Hartson H, Williges RC. Determining the effectiveness of the usability problem inspector: A theory-based model and tool for finding usability problems[J]. *Human Factors*. 2003, 45(3): 455-482.

[收稿日期]2004-12-08

[修回日期]2005-04-25