

●钱佳平 刘兹恒

图书馆服务质量测评的影响因素及控制研究

摘要 在图书馆服务质量测评中,指标体系与服务质量的相关度,调查对象的选取和抽样方式以及问卷设计和数据分析处理技术等问题,都有可能影响评估数据的可靠性和有效性。要使服务测评指标体系与服务质量高度相关,需从两个方面加以控制。首先,评价指标的建立和优化要符合以下原则:现实服务评估和前瞻性预测的有机结合;指标内涵的单义性和具体化;指标体系稳定性和动态优化的有机结合。此外,还要鼓励用户发表意见,作为补充的反馈优化机制。表3。参考文献9。

关键词 服务质量测评 评价指标 可靠性检测 标准差 图书馆服务
分类号 G251

ABSTRACT In the evaluation of library service quality, the correlation between evaluation indicator system and service quality, the selection of survey objects and sampling methods, the design of questionnaires, the technologies for the processing of data and other factors may have impacts on the reliability and efficiency of the evaluation. In this paper, the authors propose some principle for the control of the correlation of the evaluation indicator system and service quality. 3 tabs. 9 refs.

KEY WORDS Evaluation of service quality. Evaluation indicator. Reliability test. Standard deviation. Library service.

CLASS NUMBER G251

20 世纪 80 年代中后期,美国服务市场营销学家 A. Parasuraman 等人提出了一个基于用户感知与期望差距的服务质量评价模型,并把他创建的评价体系命名为 SERVQUAL^[1]。从 90 年代中期美国一些大学图书馆率先将 SERVQUAL 应用于图书馆至今 10 多年来,国内外大量的研究揭示和证明了这种新型的测评工具体备代替传统图书馆评价体系的条件,而在创建卓越的图书馆文化方面,SERVQUAL 框架也蕴含着里程碑式的潜在价值^[2]。一种评价体系的生命力来自它在特定环境中应用的可操作性和获得数据结果的高度有效性、可靠性和鉴别(诊断)性。虽然从管理层到馆员都可以利用这种评价工具的测评获取丰富并有价值的数据和信息,但评价活动过程涉及的许多因素,都有可能影响到测评质量,制约着所获数据可靠性和有效性^[3-4]。因此有必要对这种基于差距模型的评估体系构建和操作实施进行规范化和科学化研究,以明确测评方法和过程与测评结果的相互关系,识别并有效控制其主要影响因素。本文试图通过基于实例的测评体系构建和应用进行探索,同时也期望以此唤起同仁对实证方法的关注,起抛砖引玉之作用。

1 影响服务质量测评效果的因素识别和确

认

基于差距模型的评价体系是一种建立在对服务质量属性进行分解基础上并通过精确测算获得数据的评价。操作过程的任何偏差,都有可能影响结论的准确性。

首先,基于差距模型的评价体系的实质是通过比较和计算用户对图书馆服务期望值与用户对图书馆提供服务的质量认知的差值来评价图书馆服务质量的,而服务质量是个抽象的概念,并非是机构行为中可观察层面上的东西,因而评价模型是采用一组假设来构建服务质量的评价指标,使服务质量与可观察的属性形成相关关系。就服务质量评价而言,评价中采用的指标集合和构成服务质量层面的合理性与服务质量的相关程度的高低,势必会影响评价结果的可靠性和有效性。事实上,在为数众多的研究结果表明原始模型中构成服务质量 5 个层面在大学图书馆环境中有必要进行适应性调整的情况下,研究者们提出了种种不同的层面构成模型,但至今还没有一个得到公认的适用于大学图书馆服务质量评价的服务质量构成层面和具体评价指标集合标准,包括参与图书馆有数百个之多的 LibQUAL⁺™ 项目^[5],在连续 5 年的 5 次调查中(目前正开始第 6 年度的评价活动),每次调查问卷中服务质量层面和调查问题都不同于前一

次。原因在于每一个图书馆及其目标用户各有其特点,不同用户群体对图书馆的需求又不尽一致,要从不同用户各种具体的需求或期望中总结出共性的东西作为评价指标,同时又要求这些共性的指标能最大限度地反映不同图书馆用户的观点,这项工作本身颇具难度。因而构建科学的评价指标需要在总结大量定性研究成果和实践探索经验的基础上有一个不断完善的过程。此外,评价指标中常见的对调查问题的描述不准确、存在歧义或没有明确界定等问题,容易使被调查者产生误解或无从判断,也会导致测评结果不能准确反映调查对象的属性,降低该测评数据的可靠性和有效性。

其次,调查对象的选取和抽样方式不当也会导致样本不能代表总体或代表性不够的结果,造成调查结果出现偏差;而问卷设计的合理与否,关系到对调查数据的偏差和误差的控制是否有效,以及有效信息量的大小。值得注意的是,利用该模型调查返回的数据还需要采用有效的处理方法和分析技术,才能确认所得数据的可靠性和有效性,并实现数据利用的最大化。

显然,评价中涉及的指标体系与服务质量的相关度、调查对象的选取和抽样方式、以及问卷设计和数据分析处理技术等方面的问题,都有可能构成制约评估数据可靠性和有效性的因素。本文将着重指标体系与服务质量相关度的控制研究。

2 指标体系与服务质量相关度的控制

要使指标体系与服务质量高度相关,还要保证获得数据的价值,需从两个方面加以控制。首先,指标体系的构建要符合特定要求,具体而言就是要使评价指标的确立和优化符合以下原则。

原则一:现实服务评估与前瞻性预测的有机结合。基于差距分析的评价工具是图书馆用来引出、跟踪、理解用户对服务质量的观点,以此驱动图书馆服务并引导(西方图书馆界称之为“经营”)用户期望的系统化工具。服务质量评价不仅仅是针对现实服务的评估,更重要的是要能揭示用户对图书馆服务的需求趋向,并从中获得决策所需的预测信息。

原则二:指标内涵的单义性和具体化。评价指标内涵的单义性指的是指标的内涵只能代表一个概念或意义,而不能有第二种意义。只有这样,才能做到表意明晰,以杜绝含糊其词,防止因歧义而造成误解,增强可受性和可行性。评价指标的具体化则是指对

评估指标的描述必须明确和具体,能揭示所对应服务的具体特点和实质性内容,使用户对指标所描述的内涵和功能理解明确无误。倘若评价指标的条目表述不够具体,或内涵不够明确,就会造成因用户理解不同而产生对同一问题给分依据不一,或者伸缩性太大,从而扩大数据的偏差,直接影响结果的有效性和可靠性。文献调研中,我们发现国内相关的研究明显存在评价指标过于抽象和笼统、指标内涵不具备单义性、对指标的描述不够具体等缺陷。在问卷调查中,对过于抽象或笼统的调查问题,因没有独特性而缺乏明确的判断依据,用户势必只能进行模糊而随意的判断,反映的数据也不可能有高度的可靠性和有效性。

原则三:指标体系稳定性和动态优化的有机结合。构成服务质量的层面和指标集必须具有相对的稳定性,这是利用这一工具进行重复测评的机构得以观察和发现用户对服务质量的认知和期望发展趋势的前提条件^[6]。指标体系必须动态优化,首先是因为,若是单个图书馆开展基于差距模型的测评以获取图书馆管理决策所需信息,则没有也不宜采用通用标准的评估指标,而应因馆制宜,围绕测评的具体目标构建。适用于别的图书馆的指标集,即使这个指标集是建立在对用户信息行为和需求进行系统调研和准确识别的基础上,也应先修改再使用。因此,识别和确认目标用户群体的信息行为特点和信息需求特征,获得用户判别图书馆服务质量的关键因素,是构建符合良好测评条件的指标体系的前提。其次,图书馆正处在一个变化发展的环境,用户对图书馆的需求和期望也在变化发展。体现用户对图书馆服务质量的期望与感受(认知)的服务质量层面和核心指标也不是固定不变的,必须根据发展的态势进行动态调整。第三,人类对任何事物的认识都有一个不断完善和深化的过程,我们对服务质量概念的理解也一样,不可能一步到位。随着经验的积累和与用户沟通的增加,会更加充分和逐步深化。已经进入第6年的LibQUAL⁺™调查项目,每年的评价指标(调查问题)都要根据与各地不同类型的用户访谈等定性研究和定量测试的结果,进行不同程度的修改和调整就是明显例证。

此外,除了构建指标要遵循以上原则,还要鼓励用户发表意见,作为补充的反馈优化机制。一种有效的方法是在问卷中引导并留出空间让用户发表评论、建议或各种意见。作为一种聆听用户的新方式^[7],让用户直接用自己的语言表达对图书馆服务的观点、

意见和建议,应该成为问卷的必要组成部分。这些评论和意见既是调查数据的印证和解释,也是进一步完善和优化测评指标的重要依据。

3 基于实例的测评有效性和可靠性检测

为探索具有高可靠性、有效性和可操作性的大学图书馆服务质量测评体系,检测上述控制策略的可操作性及效果,从 2004 年 1 月开始,我们进行了基于实例的测评体系建构与求证研究^[8]。

3.1 实例研究概述

首先进行了基于差距模型的服务质量测评体系的框架设计,建立了包括评估目标的确立、指标集的构建与优化、数据处理分析和调查结果的可靠性与有效性检测等评估关键环节的操作规程。然后,以所在图书馆为实验基地,按照设计框架中的操作规程进行了测评体系的建构和测评实验。实例研究根据为图书馆决策和资源配置提供科学数据依据这一评估目标设定,结合图书馆现在提供的服务和正在酝酿的工作计划,将测评的重点定位于图书馆信息资源建设与整合、虚拟参考咨询、个性化服务、用户教育和图书馆网站建设等方面的服务,将服务质量具体分解为信息资源的可获取性、信息获取的便捷性、馆员知识能力和修养(态度)、以及图书馆环境等层面。根据这些层面,收集和选择了包括 40 余个测评指标的数据集合,经过选择、修改和调整,最后确定了一个 29 项指标的数据集合。设计问卷,采用典型调查和抽样调查相结合,通过电子邮件和现场发放问卷的方式,调查了近期利用过图书馆的 73 名不同群体的用户对图书馆服务的感受与期望,获得 65 份有效问卷。对回收的有效问卷进行数值统计,并采用象限分析结合差值排序法进行数据分析,获得从用户的角度揭示的测评图书馆服务质量情况。最后,对问卷的内容结构和获得的数据进行可靠性和有效性检测和确认。

3.2 对测评结果的有效性和可靠性确认

对回收的有效问卷进行统计,首先得到用户对 29 个核心问题的平均期望值、感受值及其差值的排序表。其中,被调查用户对问卷所调查的核心问题的平均期望值和感受值分别为总值 7 分中的 5.99 分和 4.99 分,总体上用户对图书馆服务的感受值低于期望值,平均差值为 -1。表 1、表 2 和表 3 分别列出的是用户对测评图书馆服务的期望值和感受值最高,以及两者差值最大的 6 个核心问题。在对统计数据运用象限图示法分析之前,我们对数据的可靠性和有效

性进行了分析和确认。鉴于调查数据的可靠性首先取决于指标体系与服务质量之间的相关度,然后才是获得数据的准确程度,故实例研究中进行的有效性和可靠性检测由问卷内容结构和评价结果(数据)两部分组成。

表 1 用户期望值最高的 6 项评价指标

序号(*)	期望的评价问题描述	平均期望值	标准差
1(7)	在家中或办公室即可检索并获得电子资源	6.72	0.54
2(3)	电子和网络资源能满足我学习/研究的需要	6.66	0.61
3(29)	图书馆环境标志清晰,能很快找到资料所在位置	6.63	0.65
4(2)	期刊的收藏完整	6.57	0.66
5(15)	使我能检索各类电子信息资源	6.52	0.73
6(27)	有利于安静地学习和思考	6.51	0.79
	总平均(共 29 项指标)期望值	5.99	1.02

*表示问卷内问题序号。

表 2 用户感受值最高的 6 项评价指标

序号(*)	服务感受的评价问题描述	平均感受值	标准差
1(25)	给予用户个别的关注	5.64	1.35
2(27)	有利于安静地学习和思考	5.57	0.99
3(23)	图书馆员乐意帮助用户	5.55	1.10
4(29)	图书馆环境标志清晰,能很快找到资料所在位置	5.37	1.25
5(12)	提供的导航使我能方便找到所需的信息资源	5.35	1.17
6(24)	工作人员谦逊有礼	5.32	1.18
	感受值(共 29 个问题)总平均	4.99	1.32

*表示问卷内问题序号。

表 3 服务质量差值最大的 6 项评价指标

序号(*)	评价问题描述	感受值(P)	期望值(E)	差值(P-E)
1(1)	提供的参考书能满足我学习/教学的需要	4.72	6.43	-1.71
2(3)	电子和网络资源能满足我学习/研究的需要	5.00	6.66	-1.66
3(16)	(图书馆网站)提供在线咨询	3.98	5.63	-1.65

续表

序号 (•)	评价问题描述	感受值 (P)	期望值 (E)	差值 (P-E)
4(2)	期刊的收藏完整	4.97	6.57	-1.60
5(15)	(图书馆网站)使我能检索各类电子信息资源	4.92	6.52	-1.60
6(13)	图书馆网站提供的链接总是有效和及时更新的	4.86	6.45	-1.59
	总平均分(共29个问题)	4.99	5.99	-1.00

*表示问卷内问题序号。

3.2.1 问卷结构的可靠性检测

实例调查问卷由3部分组成。第一部分是问卷的主体,包含了用来分解确认服务质量的29项指标(核心问题)。每一调查问题后分设两栏,要求用户就该指标对于构成“理想图书馆的重要程度”和“您感受的图书馆实际服务质量”分别赋值,共设7个等级,分值随服务品质的重要性或感受质量从低到高递增。最后,回答者还可以提出他认为重要的期望或给出简短意见。问卷的第二部分,要求用户对第一部分中的服务质量层面的重要性进行赋值,并征询用户对5个层面是否涵盖构成服务质量评价要素的看法和意见,检测用户对服务质量层面的合理性和全面性的认定。问卷的第三部分要求用户给出“总体上,图书馆提供的服务,在何种程度上符合或满足您对图书馆服务的期望”这一问题的评分,取值范围为1~10分,用来记录用户对图书馆服务的整体评价。由于指标体系与服务质量的相关度取决于测评内容的代表性程度及覆盖面,根据以上问卷结构,比较问卷第一与第三两个部分的数据总值就可以获得一个一致性差值,代表指标体系与服务质量的相关度;而通过收集用户在第一和第二部分中的补充意见,可以检测调查测评的内容是否包含了图书馆行为目标的全部内容。具体方法是:首先,通过比较整体评价得分和经加权处理分解确认的感受得分之间的差异,获得反映指标的代表性和全面性的数据。用户感受总加权平均值由以下公式求出:

$$\bar{P}_{\text{加权}} = \frac{\sum_{i=1}^n P_i \times E_i}{N \times \bar{E}}$$

公式中, $\bar{P}_{\text{加权}}$ 为用户感受的总加权平均值, P_i 为第*i*个调查问题的感受值, E_i 为第*i*个调查问题的期望值, $\bar{E}$ 为各调查问题的总平均期望值。该公式对各

项指标的感受得分按用户对相应服务的期望值进行加权处理,获得用户对图书馆服务感受的总加权平均值。所得总加权平均值即可与第三部分总体得分的平均值进行比较。

结果,我们得到的用户对图书馆服务总体评分的平均值是10分中的6.29分,得出的感受加权总平均值为5.00。相当于10分制的7.14分,二者的差值为0.85(<1.0),可以认为是接近一致。对于总体平均得分低于分解确认的平均值这一结果,可能与指标的采集侧重于服务功能,而对其构成服务质量的全面性考虑不够有关,这也可以从用户对图书馆的补充建议和意见得到印证。另外,用户对图书馆在寒假期间开放时间太少的意见较为集中,而调查又正值假期将要结束之时,用户对此感受特别强烈,总体得分低于分解问题的得分可能与此有关。这也是提示图书馆的管理必须是全方位的,一个看似孤立的事件,实际上必然影响用户对图书馆的总体评价。

实例研究发现,92.3%的用户认为问卷中29个问题已经包括了图书馆服务质量评估需要测评的问题,没有重要的指标被遗漏。对于服务质量层面的构成,有96.8%的用户表示没有其他重要的层面被遗漏。

在整理用户补充的其他期望与意见时,我们发现这些补充的内容,与其说是期望,不如说是用户对图书馆服务的评论和意见集。限于篇幅,仅列举3条:①网站的功能如资源导航是最关键的因素,页面美观与否是无关紧要的东西;②图书馆应重视学习和研究气氛的营造,希望能提供一些壁画之类能催人奋进的装饰;③延长开放时间,开放时间的延长应包括闭架书库。

3.2.2 数据的可靠性检测

测评数据的可靠性,是指测评的准确程度。一般而言,当调查避免或排除了随机因素的干扰,不同的读者对于同一问题的回答趋于一致时,调查本身引起的误差(如回答者因心理干扰或对问题的理解产生歧义等问题)极小,此时的调查被认为是可靠的;反之,当调查对象对同一问题的回答大相径庭时,调查的可靠性就降低。本文采用标准差^[9]表示内部数据的一致性。统计结果见表1和表2,表明数据的离差在正常范围内。

参考文献

- 1 Parasuraman, A., Valerie A. Zeithaml, and Leonard L. Berry.

●张正强

电子文件管理元数据中时间元素的语义结构研究*

摘 要 电子文件管理元数据中,时间元素存在4种语义结构:时间元素的前端模式语义结构、后端模式语义结构、连续体模式的语义结构、分面语义结构。建立电子文件管理元数据标准时,最好以连续体模型为依据来建构,同时采用分面语义结构。表6。图3。参考文献5。

关键词 电子文件管理 元数据 时间元素 语义结构

分类号 G250.76

ABSTRACT In the metadata for the management of electronic documents, date elements have four semantic structures. The author thinks that when we establish a metadata standard for the management of electronic documents, we had better use the continuum-model and the facet semantic structure. 6 tabs. 3 figs. 5 refs.

KEY WORDS Electronic document management. Metadata. Date element. Semantic structure.

CLASS NUMBER G250.76

时间元素是元数据集中一个核心的元素,也是任何元数据标准中都不缺少的元素之一。在电子文件管理元数据中,时间元素也具有如此重要的地位。但是电子文件管理元数据中的时间元素在其语义结构上却要比通用元数据如“都柏林元数据集”中的时间元素的结构更复杂,而这种复杂性又是由电子文件管理所需历经其整个生命周期的性质所决定的,因此研究、理解电子文件管理元数据中的时间元素,必须研究、理解电子文件的生命周期,并从中得出规律性的认识,对于更科学地建立我国电子文件管理元数据

标准将是十分有益的。

1 电子文件的生命周期与通用元数据集中时间元素的语义结构

在著名的《都柏林核心元数据集》中和以此为基础制定的国际标准《ISO 15836:2003(E)信息和文献——都柏林核心元数据集》中^[1],时间元素被定义为“与资源生命周期中某一个事件相关的时间”。由于这两个元数据集是国际上跨学科、跨专业领域的通用标准,因此在这个定义中的“资源”也就具有高度

- A Conceptual Model of Service Quality and Its Implications for Future Research. Journal of Marketing, 1985, 49(4)
- 2 Nitecki. Changing the Concept and Measure of Service Quality in Academic Libraries. Journal of Academic Librarianship, 1996, 22
- 3 Thompson, B., & Cook, C. Stability of the reliability of LibQUALTM scores: A “Reliability Generalization” meta-analysis study. Educational and Psychological Measurement, 2002, 62
- 4 Thompson, B., Cook, C., & Thompson, R. L. Reliability and structure of LibQUALTM scores. portal: Libraries and the Academy, 2002, 2
- 5 <http://www.libqual.org/>
- 6 Brinley Franklin and Terry Plum. Library usage patterns in the electronic information environment. Information Research, 2004, 9(4)
- 7 New Ways of Listening to Library Users: Tools for Measuring Service quality. <http://www.arl.org/libqual/events/listen/index2.html>
- 8 钱佳平. 基于差距分析的大学图书馆服务质量评价体系建构: 研究与实践. 北京大学硕士学位论文, 2004
- 9 朱德全. 现代教育统计与测评技术. 重庆: 西南师范大学出版社, 1998
- 钱佳平 浙江大学图书馆副研究馆员, 管理学硕士。通信地址: 杭州市。邮编 310027。
- 刘兹恒 北京大学信息管理系教授。通信地址: 北京。邮编 100871。
- (来稿时间: 2005-04-25)

* 本文为国家社科基金重点资助项目“基于XML电子文件管理元数据标准研究”系列研究成果(项目批准编号04ATQ002)。