

# 面向性别差异与任务复杂性的相关性判据研究<sup>\*</sup>

成 颖

南京大学信息管理系 南京 210093

[摘要] 通过194名研究对象相关性判断文本的内容分析,探讨任务复杂性与性别对相关性判据选择的影响。通过spearman相关系数检验发现,任务复杂性对相关性判据的选择存在显著影响;而性别仅仅对文献总体、作者和文献类型三类判据的选择存在影响,总体而言影响甚微。该结论提示检索系统的分析和设计需要考虑任务复杂性的影响,而无需考虑性别的差异。

[关键词] 相关性 判据 性别 任务复杂性

[分类号] G358

## On Gender and Task Complexity Oriented Relevance Criteria

Cheng Ying

Department of Information Management, Nanjing University, Nanjing 210093

[Abstract] By content analysis of text about relevance assessment of 194 subjects, the author explores the effects of task complexity and gender to relevance criteria selection. By spearman correlation test, results show that task complexity affects relevance criteria selection significantly; and gender only influences three relevance criteria, respectively literature collection, author and literature type. In general, gender has little effect. The conclusion prompts that the design and analysis of information retrieval systems need to consider the impact of task complexity.

[Keywords] relevance criteria gender task complexity

## 1 引言

图书情报学界以及计算机学界围绕相关性这一信息检索的核心领域进行了深入的研究,其中相关性判据(relevance criteria)是该领域的一个重要研究主题。相关性判据即影响用户相关性判断的因素。Mizzaro<sup>[1]</sup>列举了到1994年为止已有的24项相关性判据研究;Saracevic<sup>[2-3]</sup>综述了11项1995-2005年间的研究成果;笔者通过对Weily Interscience、EBSCO、Elsevier等数据库的检索,发现2005年至今有12篇相关论文。综合这些研究发现,相关性判据研究的主题已涉及学术信息源、网络资源、气象信息、网页、超链、医疗资源、日常生活、新闻博客以及在线问答平台等多个领域;部分研究还探讨了音乐、视频以及图片等多媒体信息资源的相关性判据;另有学者对信息搜寻过程(information search process, ISP)的不同阶段、不同人群、不同信

息资源类型的相关性判据进行了比较研究以及一些综述性的研究。但在现有研究中,尚未发现有从性别以及任务复杂性维度探讨其对相关性判据的影响者。

相关性判据研究对于相关性概念的拓展具有重要意义,在计算机学界完成的基于主题特征的信息检索系统的基础上,探索影响用户相关性判断的因素对于更具人性化以及个性化信息检索系统的分析和设计具有积极的价值。本文拟从任务复杂性以及性别角度探索其对相关性判据的影响,以为信息检索系统的设计和优化提供有价值的线索。

## 2 研究设计

### 2.1 研究的问题

本文围绕相关性判据主要探讨两个问题:①任务复杂性对信息用户相关性判据选择的影响;②信息用户的性别对相关性判据选择的影响。

<sup>\*</sup> 本文系国家社会科学基金项目“中文学术信息检索系统相关性集成研究”(项目编号:10CTQ027)和教育部人文社会科学规划基金项目“面向用户的相关性标准及其应用研究”(项目编号:07JA870006)研究成果之一。

2.2 数据集

研究对象为南京大学信息管理系 2002 级、2003 级、2004 级、2008 级本科生以及教育科学系 2008 级本科生的所有同学,共 194 人,样本量显著大于 Schamber<sup>[4]</sup> 以及 Barry<sup>[5]</sup> 等研究中的样本量(分别只有 30 与 18 人)。通过给学生布置一次“计算机网络”课程论文(作业),要求学生在提交课程论文中详细描述检索过程,后者主要包括检索策略、是否进行二次检索以及原因、相关文献相关性判断的判据。研究对象中男性多于女性,年龄段相对比较集中,在研究数据搜集的时候,大部分学生的年龄在 18-21 岁之间,数据的搜集时间分别是 2005 年 12 月、2006 年 11 月、2007 年 10 月以及 2010 年 12 月,分别在秋季学期的中后期进行。研究对象的人口特征如表 1 所示:

表 1 学生的人口特征

性别	男	123	教育程度	本科 3 年级	194
	女	71	年龄	18-21 岁	194
系科	信息管理	188	检索目的	课程论文	187
	教育学	6		课程作业	7

3 数据分析

数据分析采用内容分析法完成,编码在借鉴 Schamber、Barry 等研究的基础上,以探索性的、归纳式的、迭代的方式完成。通过内容分析,本研究解析出了 9 种类型的相关性判据,分别是传播特征、文献内容、情境、文献使用、系统特征、愉悦感知、文献质量、文献总体以及文献特征<sup>[6]</sup>。本研究中,2002、2003 级学生的任务是撰写当时还处于热点的课程论文,2004 级学生为一道难度适当的作业,2008 级学生虽也是课程论文,但是主题已经是比较成熟的内容。2002、2003 级学生的任务复杂性远高于 2004 以及 2008 级学生,因此本研究将 2002 与 2003 级学生的任务归结为高复杂性任务,而 2004 与 2008 级学生的任务归结为低复杂性任务。学生的性别数据见表 1。本研究采用非参数检验方法分析任务复杂性与性别对相关性判据选择的影响。

3.1 分析方法

探讨任务与性别两个维度对相关性判据选择的影响,可供选择的检验方法有 t 检验、卡方检验以及 MWU 检验等。由于 t 检验的依据是单一变量的均值,而本研究的数据不具备该特征,从而被排除;MWU 检验的依据是中位数,与 t 检验类似,也被排除在外;卡方检验应该也是比较合适的检验方法,但是其适用条件

之一是理论频数不能太小,多数教科书认为,频数小于 5 的不能多于 20%<sup>[7]</sup>,但是本研究中的频数数据有很多突破了 20% 的限制,因此也只能放弃;可以考虑的还有 Wilcoxon 符号秩和检验,但是本研究中的数据并不符合配对的要求,从而也只能舍弃;最终本研究选取 spearman 相关系数进行检验。spearman 等级相关是根据等级资料研究两个变量间相关关系的方法。它是依据两列成对等级的各对等级数之差来进行计算,所以又称为“等级差数法”。spearman 等级相关对数据条件的要求没有积差相关系数严格,只要两个变量的观测值是成对的等级评定资料,或者是由连续变量观测资料转化得到的等级资料,不论两个变量的总体分布形态、样本容量的大小如何,都可以用 spearman 等级相关来进行研究。根据 spearman 相关系数的这些不太苛刻的条件,本研究的数据使用该方法进行检验,采用 SPSS 13 完成相关数据的分析。

3.2 大类相关性判据

大类相关性判据结果如表 2 所示:

表 2 大类相关性判据分布

任务 复杂性	高		低		性别	女		男	
	学生数	百分比	学生数	百分比		学生数	百分比	学生数	百分比
愉悦	2	0.40	0	0.00	情境	2	0.87	3	0.77
情境	4	0.81	1	0.80	愉悦	2	0.87	0	0.00
传播特征	7	1.41	1	0.80	传播特征	4	1.73	4	1.02
系统特征	21	4.23	2	1.60	系统特征	10	4.33	13	3.32
使用	38	7.66	12	9.60	使用	14	6.06	36	9.21
质量	127	25.60	41	32.80	总体	16	6.93	36	9.21
总体	49	9.88	3	2.40	内容	24	10.39	45	11.51
内容	49	9.88	20	16.00	质量	62	26.84	106	27.11
文献特征	199	40.12	45	36.00	文献特征	97	41.99	148	37.85
合计	496	100	125	100.00	合计	231	100.00	391	100.00

由表 2 可见,不同复杂性的任务类型,学生的相关性判据存在差异,比如高复杂性任务组中有学生提到了愉悦类别的相关性判据,而低复杂性任务组则未出现。文档特征判据高复杂性任务组占 40.12%,而低复杂性任务组则占 36.00%,其他每组都存在一定的差异。采用 spearman 相关系数检验的结果显示秩相关系数为 0.966,P 值为 0.000,因此接受原假设。

不同性别的学生在选择相关性判据时也存在一定差异,比如女同学提到了愉悦认知,而男同学组则没有涉及。其他不同类别的相关性判据在百分比方面都有或多或少的不同,比如文献特征,女生组所占比例为 41.99%,而男生组则为 37.85%。采用 spearman 相关系数检验的结果显示秩相关系数为 0.992,P 值为

0.000,因此也接受原假设。

### 3.3 子类别相关性判据分析

通过大类相关性判据的分析,可以得到一个基本的结论,即学生在进行相关性判据的选择时,在任务复杂性与性别维度方面没有显著性差异,那么在子类别方面有没有显著性差异呢?

3.3.1 质量相关性判据 由表3可见,质量类别在任务维度方面都存在显著的差别:

表3 质量类别相关性判据分布

任务 复杂性 相关性 判据	高		低		相关性 判据	女		男	
	学生 数	百分 比	学生 数	百分 比		学生 数	百分 比	学生 数	百分 比
美感	1	0.30	0	0.00	美感	1	0.57	0	0.00
直觉	1	0.30	0	0.00	知识水平	1	0.57	1	0.36
知识水平	2	0.61	0	0.00	直觉	1	0.57	0	0.00
系统性	3	0.91	4	3.23	准确性	1	0.57	7	2.50
结构清晰	4	1.21	3	2.42	系统性	2	1.15	5	1.79
准确性	5	1.52	3	2.42	先进性	4	2.30	7	2.50
专指性	8	2.42	8	6.45	结构清晰	5	2.87	2	0.71
可靠性	9	2.73	3	2.42	专指性	6	3.45	10	3.57
先进性	9	2.73	2	1.61	可靠性	7	4.02	5	1.79
实时性	13	3.94	0	0.00	学术性	6	3.45	14	5.00
学术性	18	5.45	2	1.61	深度	9	5.17	15	5.36
简明性	15	4.55	12	9.68	实时性	8	4.60	5	1.79
深度	21	6.36	3	2.42	简明性	9	5.17	18	6.43
权威性	25	7.58	19	15.32	新颖性	13	7.47	22	7.86
可理解性	25	7.58	18	14.52	可理解性	13	7.47	30	10.71
广度	27	8.18	13	10.48	权威性	18	10.34	26	9.29
新颖性	31	9.39	4	3.23	广度	18	10.34	22	7.86
详细	35	10.61	29	23.39	主题	26	14.94	53	18.93
主题	78	23.64	1	0.81	详细	26	14.94	38	13.57
合计	330	100.00	124	100.00	合计	174	100.00	280	100.00

首先,高复杂性任务组提到的美感、直觉、知识水平以及实时性等判据在低复杂性任务组中没有涉及。在高复杂性任务组中最重要的主题相关性判据在低复杂性任务组中只有一位同学提及,处于可有可无的位置。在高复杂性任务组中,对学生有重要影响的相关性判据为主题、详细、新颖性、广度、可理解性、权威性以及深度等。而在低复杂性任务组中详细、权威性、可理解性、广度、简明性以及专指性则成为最主要的相关性判据。数据显示:高复杂性任务组更倾向于新颖、深度、主题性等判据,表现为其更青睐于研究型文献,比如研究性论文;而低复杂性任务组更倾向于广度、简明性以及专指性等相关性判据,这些判据直接指向了教材、综述以及标准等主要反映比较成熟学术信息的文献类型。采用 spearman 相关系数检验的结果显示,秩相关系数为 0.546, P 值为 0.016,在 0.05 置信度水平下接受原假设,该结果出乎笔者的意料。

数据显示女生组比男生组多了一些感性的相关性判据,比如美感、直觉,而男生组没有提及这些判据。除此之外,两组的相关性判据在高频次组相似,都有简明性、新颖性、可理解性、权威性、广度、主题、主题以及详细等判据,如果说有差别,也仅仅是细微的,各个判据的位次和比例稍有区别,如主题在男生组占 18.93%,而在女生组则占到 14.94% 等。spearman 相关系数检验的结果显示秩相关系数为 0.898, P 值为 0.000,因此接受原假设。

3.3.2 总体相关性判据 表4呈现了总体类别相关性判据在任务复杂性层面的分布结果:

表4 总体类别相关性判据分布

任务 复杂性 相关性 判据	高		低		相关性 判据	女		男	
	学生 数	百分 比	学生 数	百分 比		学生 数	百分 比	学生 数	百分 比
顺序	2	2.67	0	0.00	搜索引擎	1	3.23	2	4.26
外部证实	2	2.67	0	0.00	外部证实	1	3.23	1	2.13
质量	2	2.67	0	0.00	质量	1	3.23	1	2.13
存取	3	4.00	0	0.00	顺序	2	6.45	0	0.00
搜索引擎	3	4.00	0	0.00	文献计量	2	6.45	2	4.26
查准率	4	5.33	0	0.00	存取	2	6.45	1	2.13
文献计量	4	5.33	0	0.00	查准率	3	9.68	1	2.13
查全率	5	6.67	0	0.00	查全率	4	12.90	1	2.13
学术数据库	7	9.33	0	0.00	同质性	4	12.90	16	34.04
同质性	18	24.00	2	66.67	学术数据库	4	12.90	3	6.38
数量	25	33.33	1	33.33	数量	7	22.58	19	40.43
合计	75	100.00	3	100.00	合计	31	100.00	47	100.00

分析发现,低复杂性任务组仅仅有3名同学分别提到了数量与同质性相关性判据,而顺序、外部证实、质量、存取、搜索引擎、查准率、文献计量、查全率以及学术数据库等相关性判据则没有学生提及,从而提示低复杂性任务组对于信息检索结果总体的关注远远不及高复杂性任务组,其着眼点主要在于只要能够找到解决简单问题的信息即可,对于是否查准与查全等没有太多关注,从而提示不同程度的任务复杂性对于信息检索系统的设计存在非常大的影响。spearman 相关系数检验的结果显示秩相关系数为 0.67, P 值为 0.024,在置信度为 0.05 时接受原假设。

从表4可以发现,男生组主要的相关性判据是数量与同质性,二者合计占到了 74.47% 的比例;而女生组则分布相对均匀,查准率、查全率、同质性、学术数据库以及数量是影响女生相关性判断的主要判据。spearman 相关系数检验结果显示秩相关系数为 0.526, 而 P 值为 0.097 > 0.05,因此接受备择假设。

3.3.3 内容类别相关性判据 表5的数据显示,低复杂性任务组希望文献内容是综合性的,最好有图解

以及比较分析,三者占到了低复杂性组的 91.17%;高复杂性任务组判据的分布比较均匀,学生对文献信息内容的选择更希望是综合性的,占到了 36.67%,但是专题性的文献也受到青睐,占到了 17.78%,对于文献中的微观信息则希望出现实践性的、图解、比较分析以及实验等内容。这些结果显示低复杂性任务组更期望直接找到答案的文献信息,而高复杂性任务组则需要更多的原始文献中通过归纳、分析以及提炼等需要更多认知努力的工作才能完成。spearman 相关系数检验结果显示秩相关系数为 0.517,P 值为 0.126>0.05,接受备择假设。

表 5 内容类别相关性判据分布

任务 复杂性 相关性 判据	高		低		相关性 判据	性别		女		男	
	学生 数	百分 比	学生 数	百分 比		学生 数	百分 比	学生 数	百分 比	学生 数	百分 比
数据	2	2.22	0	0.00	方法	0	0.00	2	2.60		
方法	2	2.22	0	0.00	数据	0	0.00	2	2.60		
技术实现	2	2.22	2	3.58	类别	0	0.00	3	3.90		
类别	3	3.33	0	0.00	技术实现	0	0.00	4	5.19		
实验	5	5.56	0	0.00	实验	2	4.26	3	3.90		
比较分析	6	6.67	4	11.76	比较分析	4	8.51	6	7.79		
图解	10	11.11	9	26.47	实践	5	10.64	6	7.79		
实践	11	12.22	0	0.00	图解	6	12.77	13	16.88		
专题性	16	17.78	1	2.94	专题性	6	12.77	11	14.29		
综合性	33	36.67	18	52.94	综合性	24	51.06	27	35.06		
合计	90	100.00	34	100.00	合计	47	100.00	77	100.00		

女生对方法、数据、类别以及技术实现等文献未予以关注;实验、比较分析、实践、图解、专题性以及综合性等成为男女生判别文献相关与否的主要文献内容依据,且类别间的变化趋势相似。spearman 相关系数检验的结果显示秩相关系数为 0.921,P 值为 0.000,接受原假设。

3.3.4 使用类别相关性判据 表 6 的结果显示,潜在利用对低复杂性任务组没有影响。该结果可以得到合理的解释,由于任务复杂性比较低,不需要耗费很多的认知努力就可以解决问题,因而学生也就无需下载更多的文献以备将来使用。其他几个类别中可发现任务复杂性的存在一定差别,比如,低复杂性任务组中,提供背景知识及有效性的占据了 85%,解释了绝大部分使用类别的相关性判据。而对于高复杂性任务组,架构借鉴、有效性以及提供背景知识等几个判据的比例相似,都在 30% 左右。spearman 相关系数检验的结果显示秩相关系数为 0.872,P 值为 0.054>0.05,因此接受备择假设。

在性别方面,使用类别的相关性判据体现了比较好的一致性趋势,虽然性别组在具体的相关性判据方

表 6 使用类别相关性判据分布

任务 复杂性 相关性 判据	高		低		相关性 判据	性别		女		男	
	学生 数	百分 比	学生 数	百分 比		学生 数	百分 比	学生 数	百分 比	学生 数	百分 比
读者对象	1	1.52	1	5.00	读者对象	1	4.35	1	1.59		
潜在利用	4	6.06	0	0.00	潜在利用	2	8.70	2	3.17		
架构借鉴	20	30.30	2	10.00	架构借鉴	5	21.74	17	26.98		
有效性	20	30.30	4	20.00	有效性	5	21.74	19	30.16		
提供背景知识	21	31.82	13	65.00	提供背景知识	10	43.48	24	38.10		
合计	66	100.00	20	100.00	合计	23	100.00	63	100.00		

面存在频数的差别,但是总的比例都是随着读者对象、潜在利用、架构借鉴、有效性以及提供背景知识呈现递增,该特点在前面已经分析的相关性判据类别中还是首次出现。spearman 相关系数检验结果显示秩相关系数为 0.975,P 值为 0.005,接受原假设。

3.3.5 文献特征类别相关性判据 文献特征具有丰富的内涵,本研究中学生主要使用学术信息数据源,文献特征主要涵盖作者、来源期刊、机构、文献类型、文献格式、出版单位、基金、参考文献、篇幅、语种、出版时间、关键词、全文、题名以及文摘等特征,这方面结果如表 7 所示:

表 7 文献特征类别相关性判据分布

任务 复杂性 相关性 判据	高		低		相关性 判据	性别		女		男	
	学生 数	百分 比	学生 数	百分 比		学生 数	百分 比	学生 数	百分 比	学生 数	百分 比
文献格式	2	0.63	0	0.00	文献格式	1	1.18	1	1.10		
出版单位	2	0.63	2	13.33	基金	1	1.18	3	3.30		
基金	4	1.26	0	0.00	参考文献	2	2.35	5	5.49		
参考文献	5	1.58	2	13.33	出版单位	2	2.35	2	2.20		
篇幅	6	1.89	5	33.33	出版时间	5	3.58	9	9.89		
语种	12	3.79	5	33.33	篇幅	7	8.24	4	4.40		
出版时间	13	4.10	1	6.67	语种	9	10.59	8	8.79		
关键词	16	5.05	0	0.00	全文	11	12.94	10	10.99		
全文	21	6.62	0	0.00	关键词	11	12.94	5	5.49		
题名	36	11.36	0	0.00	题名	18	21.18	18	19.78		
文摘	44	13.88	0	0.00	文摘	18	21.18	26	28.57		
作者	31	9.78	3	3.37	来源期刊	9	5.20	22	9.44		
来源期刊	31	9.78	0	0.00	机构	13	7.51	20	8.58		
机构	32	10.09	1	1.12	作者	19	10.98	15	6.44		
文献类型	62	19.56	15	16.85	文献类型	47	27.17	85	36.48		
合计	317	100.00	34	100.00	合计	85	100.00	91	100.00		

表 7 显示低复杂性任务组主要受文献类型、篇幅和语种的影响,其他文献特征的影响甚微。高复杂性任务组则主要受到题名、文摘、作者、来源期刊、机构以及文献类型的影响。该结果显示文献的各种特征都会影响信息用户的相关性判断,因此在信息检索系统的设计中,应尽可能地将文献的各种特征考虑到检索系统中。spearman 相关系数检验结果显示秩相关系数为 -0.2,P 值为 0.944>0.05,接受备择假设。

性别方面,文献特征类别的相关性判据也体现了比较好的一致性趋势。女生组中,文献类型、机构、作者、题名以及文摘占据了文献特征类别相关性判据的前 5 强,而男生组则是题名、文摘、来源期刊、机构以及文献类型位于前列,二者的区别在于女生组中作者与男生组中作者的位次发生了细微的变化,其他 4 个则基本相同。文献特征中的篇幅、基金、出版单位以及文献格式居于相对次要的位置。spearman 相关系数检验结果显示秩相关系数为 0.838,P 值为 0.000,接受原假设。

3.3.6 作者类别相关性判据 表 8 的数据显示,低复杂性任务组仅关注作者的声誉和职称,且仅有 3 名学生提到这两个判据。高复杂性任务组则是职称、客观以及声誉占据主要的位置,不过合计也仅 31 人分别提到了这些判据,总量偏低。个中原由,可能是学生作为初学者,对学界作者还不了解,因此更多地依据其他判据类别确定文献的相关性,仅少部分同学了解问题域的作者情况。spearman 相关系数检验结果显示秩相关系数为 0.686,P 值为 0.132 > 0.05,接受备择假设。

表 8 作者类别相关性判据分布

任务 复杂性	高		低		相关性 判据	性别		女		男	
	学生 数	百分 比	学生 数	百分 比		学生 数	百分 比	学生 数	百分 比	学生 数	百分 比
高产	2	6.45	0	0.00	群体	1	5.26	1	6.67		
群体	2	6.45	0	0.00	高产	2	10.53	0	0.00		
研究领域	4	12.90	0	0.00	客观	3	15.79	4	26.67		
职称	5	16.13	1	33.33	研究领域	3	15.79	1	6.67		
客观	7	22.58	0	0.00	职称	4	21.05	2	13.33		
声誉	11	35.48	2	66.67	声誉	6	31.58	7	46.67		
合计	31	100.00	3	100.00	合计	19	100.00	15	100.00		

女生组中,作者的声誉、职称影响其相关性判断,男生组则主要关注作者的声誉。spearman 相关系数检验结果显示秩相关系数为 0.779,P 值为 0.068 > 0.05,接受备择假设。

3.3.7 文献类型相关性判据 表 9 显示,低复杂性任务组的文献类型相关性判据比较集中,主要是教材、标准与网络百科。高复杂性任务组的文献类型主要是技术白皮书、标准、新闻、教材以及综述,但是更为大量的是期刊论文,该数据没有在表中反映出来。从文献类型的分析可以得到基本的结论,即低复杂性任务组倾向于采用比较成熟的文献类型解决问题,比如教材以及标准;而高复杂性任务组则需要更多地依赖学术论文等文献才能解决问题,教材以及标准等非常成熟的文献则被边缘化。spearman 相关系数检验结果显示秩相关系数为 0.208,P 值为 0.516 > 0.05,接受备择假设。

表 9 文献类型相关性判据分布

任务 复杂性	高		低		相关性 判据	性别		女		男	
	学生 数	百分 比	学生 数	百分 比		学生 数	百分 比	学生 数	百分 比	学生 数	百分 比
学位论文	0	0.00	3	4.29	访谈	0	0.00	1	1.18		
网络百科	0	0.00	12	17.14	讲座	0	0.00	1	1.18		
访谈	1	1.61	0	0.00	新闻	0	0.00	10	11.76		
讲座	1	1.61	0	0.00	学位论文	1	2.13	2	2.35		
文献类型	2	3.23	1	1.43	文献类型	1	2.13	1	1.18		
官方网站	3	4.84	2	2.86	技术白皮书	1	2.13	4	4.71		
会议文献	4	6.45	0	0.00	官方网站	2	4.26	3	3.53		
技术白皮书	5	8.06	0	0.00	网络百科	2	4.26	10	11.76		
标准	5	8.06	18	25.71	会议文献	3	6.38	2	2.35		
新闻	9	14.52	1	1.43	标准	7	14.89	16	18.82		
教材	15	24.19	30	42.86	综述	7	14.89	13	15.29		
综述	17	27.42	3	4.29	教材	23	48.94	22	23.58		
合计	62	100.00	70	100.00	合计	47	100.00	85	100.00		

性别维度的 spearman 相关系数检验结果显示秩相关系数为 0.523,P 值为 0.081 > 0.05,接受备择假设。

3.3.8 机构相关性判据 表 10 的数据显示,低复杂性任务组对于来源期刊的声誉几乎没有关注,对于机构的其他特征完全没有关注,而高复杂性任务组则主要关注机构的声誉。从性别层面的分析可以发现,女生与男生对机构声誉的关注没有显著差异。由于只有三组数据,进行 spearman 相关系数检验的意义已经不大。

表 10 机构类别相关性判据分布

任务 复杂性	高		低		相关性 判据	性别		女		男	
	学生 数	百分 比	学生 数	百分 比		学生 数	百分 比	学生 数	百分 比	学生 数	百分 比
高产	1	3.13	0	0.00	客观	1	7.69	3	15.00		
客观	4	12.50	0	0.00	高产	1	7.69	0	0.00		
声誉	27	84.38	1	100.00	声誉	11	84.62	17	85.00		
合计	32	100.00	1	100.00	合计	13	100.00	20	100.00		

3.3.9 系统特征相关性判据 从表 11 的数据可见,低复杂性任务组对于系统特征不是非常在意,而高复杂性任务组则主要关注系统的存取情况。通过性别层面的分析可以发现,女生与男生对系统特征的关注没有显著差异。出于与机构组同样的原因,不再进行 spearman 相关系数检验。

表 11 系统特征类别相关性判据分布

任务 复杂性	高		低		相关性 判据	性别		女		男	
	学生 数	百分 比	学生 数	百分 比		学生 数	百分 比	学生 数	百分 比	学生 数	百分 比
响应速度	1	4.17	0	0.00	著录详细	0	0.00	1	7.69		
著录详细	1	4.17	0	0.00	响应速度	1	7.69	0	0.00		
存取	22	91.67	2	100.00	存取	12	92.31	12	92.31		
合计	24	100.00	2	100.00	合计	13	100.00	13	100.00		

3.3.10 传播特征相关性判据 低复杂性任务组没有将传播特征纳入视野,而高复杂性任务组则主要关注

传播特征中的引用情况、点击率以及转载等特征。从性别层面的分析可以发现,女生与男生对传播特征的关注没有显著差异。同理也不再进行 spearman 相关系数检验。如表 12 所示:

表 12 传播特征类别相关性判据分布

相关性判据	任务复杂性		性别	
	高	低	女	男
	学生数	百分比	学生数	百分比
响应速度	转载	1 12.50	0	0.00
著录详细	点击率	2 25.00	0	0.00
存取	引用	5 62.50	4 100.00	2
合计	合计	8 100.00	4 100.00	5

3.3.11 来源期刊、情境以及愉悦相关性判据 低复杂性任务组中没有学生提到来源期刊的声誉,高复杂性任务组有 31 位同学提到了来源期刊的声誉,分别是 9 位女生、22 位男生。情境与愉悦类型的相关性判据样本量分别为 5 和 2,继续分析其在任务层面以及性别层面的分布和统计学显著性已经意义不大,因此不再展开。

4 研究结论

结合本文的研究目的,对于性别以及任务复杂性对相关性的判据的选择是否存在显著影响,本文通过 spearman 相关系数分析给出了答案。

4.1 判据大类

任务复杂性和性别维度都接受原假设,即任务复杂性和性别维度在大类相关性判据的选择方面具有显著相关关系,从而表明任务复杂性高低和性别在信息用户的大类相关性判据的选择方面具有一致的行为。该结论表明检索系统的分析与设计无需考虑任务复杂性和性别维度,因而应该把重点放在两组中高百分率的因素上面,比如应考虑文献特征与质量这两个明显重要的大类相关性判据,然后再考虑总体、内容、使用以及系统特征,最后也不应该放弃传播特征、情境以及愉悦等相关性判据。表 2 中的数据显示,如果能够设计出更具愉悦感知的系统,则更容易受到女生的欢迎。

4.2 文献质量

在 0.05 的置信水平下,文献质量相关性判据在任务复杂性和性别维度都接受原假设;如果在 0.01 的置信水平下,则在任务复杂性维度接受备择假设;表 3 的数据也表明,研究对象在任务复杂性维度的相关性判据方面存在明显的差别,因此本文更倾向于接受备择假设。从而也就意味着在检索系统的分析和设计方面

无需考虑性别的差异,但还是有必要充分考虑任务复杂性的影响,即信息检索系统最好能够根据任务复杂性的不同输出不同的检索结果,比如任务复杂性低的查询请求输出综述、标准等比较成熟的文献,而高复杂性任务组则输出原创性强的研究性文献。

4.3 文献内容

在任务复杂性维度接受备择假设,在性别维度接受原假设,即检索系统的分析和设计在内容层面无需考虑性别的影响,而应该考虑任务复杂性的影响。根据表 5,对于高复杂性任务组,检索系统应该尽可能提供实践性强、专题性以及综合性的信息内容,低复杂性任务则提供综合性、有图解以及有比较分析的检索结果。该分析结果说明,高复杂性任务需要付出更多的认知努力,而低复杂性任务组则倾向于尽可能少地付出认知努力。

4.4 文献使用

在性别维度接受原假设,在任务复杂性维度接受备择假设。该结论的实践意义在于:在信息检索系统的分析与设计中有必要考虑任务复杂性对使用的影响。不过,使用过于广泛,不同的领域都具有不同类型的使用类别,本研究分析得出的使用类别可能仅仅是挂一漏万,因此也意味着信息检索系统如果要考虑该层面的差异,难度非常大。

4.5 文献特征

任务复杂度对于文献特征具有显著的影响,而性别层面的影响则甚微。通过对学术信息检索系统特征的调研发现,检索系统在检索提问的输入方面已经充分地考虑了几乎所有的文献特征,而在检索结果的输出方面尤其是相关性排序方面还有待改进,在这些方面有必要考虑任务复杂性的影响。

4.6 作者与机构

由于本文的研究对象为领域新人,对领域中的作者以及研究机构的了解甚少,因此提到这类判据的学生比例偏低。作者判据在复杂性与性别维度都接受备择假设,表 10 的数据表明任务复杂性维度对于机构判据的影响比较显著。这些结论提示信息检索系统的分析与设计,尤其在相关性排序方面有必要充分考虑作者和机构声誉等判据。

4.7 文献类型

任务复杂度和性别对文献特征判据的选择都具有显著的影响,提示信息检索系统的分析与设计应考虑二者的影响,对于高复杂性任务多提供研究论文与综述,而对于低复杂性任务则多提供教材、标准等比较成

熟的文献。

#### 4.8 文献总体

在 0.01 的置信水平下,任务复杂度和性别对文献总体判据的选择都具有显著影响,对于高复杂性任务需要保证文献的查全率、学术性,减少同质性并保证数量。

总之,面对不同复杂性的任务时,在 0.01 置信水平下检索系统的分析和设计方面有必要考虑文献质量、文献总体、文献内容、文献使用、文献特征、作者以及文献类型的影响;虽然没有进行检验,表 10-12 的数据清晰地表明机构、系统特征、来源期刊以及传播特征也都受到任务复杂性的影响,因此可以得到如下结论:任务复杂性对相关性的判据选择存在显著影响,也表明检索系统的分析和设计要充分考虑任务复杂性的影响。

spearman 相关系数分析显示文献总体、作者和文献类型受到性别的影响,表 10-12 的数据表明性别对于机构、系统特征、来源期刊以及传播特征的影响甚微。因此可以得出结论:性别对于相关性判据选择的影响很小,检索系统的分析和设计无需考虑性别的影响。

〔作者简介〕成 颖,男,1971 年生,副教授,博士,发表论文 40 篇。

响。

#### 参考文献:

- [1] Mizzaro S. Relevance: The whole history. *Journal of the American Society for Information Science*, 1997, 48(9): 810-832.
- [2] Saracevic T. Relevance: A review of the literature and a framework for thinking on the notion in information science. Part II: Nature and manifestations of relevance. *Journal of the American Society for Information Science and Technology*, 2007, 58(3): 1915-1933.
- [3] Saracevic T. Relevance: A review of the literature and a framework for thinking on the notion in information science. Part III: Behavior and effects of relevance. *Journal of the American Society for Information Science and Technology*, 2007, 58(13): 2126-2144.
- [4] Schamber L. Users' criteria for evaluation in a multimedia information seeking and use situation [dissertation]. Syracuse: Syracuse University, 1991.
- [5] Barry C L. The identification of user relevance criteria and document characteristics: Beyond the topical approach to information retrieval [dissertation]. Syracuse: Syracuse University, 1993.
- [6] 成颖. 信息检索相关性判据及应用研究 [学位论文]. 南京: 南京大学, 2011.
- [7] 贾俊平. 统计学. 北京: 清华大学出版社, 2006.

(上接第 102 页)

- [2] MacMillan J, Paley M, Entin E B, et al. Questionnaires for distributed assessment of team mutual awareness//Stanton N A, Hedge A, Brookhuis K, et al. *Handbook of human factors and ergonomics methods*. Boca Katon: CRC Press, 2004.
- [3] Langan-Fox J, Anglim J, Wilson J R. Mental models, team mental models and performance: Process, development and future directions. *Human Factors & Ergonomics in Manufacturing*, 2004, 14(4): 331-352.
- [4] Kirkman B L, Mathieu J E. The dimensions and antecedents of team rituality. *Journal of Management*, 2005, 31(5): 700-718.
- [5] Kraut R, Streeter L. Coordination in large scale software development. *Communications of the ACM*, 1995, 38(3): 69-81.
- [6] Edmondson A C. Speaking up in the operating room: How team leaders promote learning in interdisciplinary action teams. *Journal of Management Studies*, 2003, 40(6): 1419-1452.
- [7] Dane E, Pratt M G. Exploring intuition and its role in managerial decision making. *Academy of Management Review*, 2007, 32(1): 33-54.
- [8] Gibson C B, Earley P C. Collective cognition in action: Accumula-

tion, interaction, examination, and accommodation in the development and operation of group efficacy beliefs in the workplace. *Academy of Management Review*, 2007, 32(2): 438-458.

- [9] John E, et al. Interactive effects of team and task shared mental models as related to air traffic controllers' collective efficacy and effectiveness. *Human Performance*, 2010, 23(1): 22-40
- [10] Bettenhausen K L, Murnighan J. Kn. The emergences of norms in competitive decision-making groups. *Administrative Science Quarterly*, 1985, 30(3): 350-372
- [11] Dougherty D. Interpretive barriers to successful product innovation in large firms. *Organization Science*, 1992, 3(2): 179-202.
- [12] Jackson S E, Joshi A, Erhardt N L. Recent research on teams and organizational diversity: SWOT analysis and implications. *Journal of Management*, 2003, 29(6): 801-830.
- [13] Saavedra R, Earley P C, Van Dyne L. Complex interdependence in task performing groups. *Journal of Applied Psychology*, 1993, 78(1): 61-72.
- [14] 常涛, 廖建桥. 团队性绩效考核对知识共享的影响模型研究. *科研管理*, 2011, 32(1): 109-121.

〔作者简介〕张桂平,女,1975 年生,教师,博士研究生,发表论文 10 余篇。

廖建桥,男,1957 年生,教授,博士生导师,发表论文 160 余篇。