

文章编号:1006-8309(2004)03-0003-04

信息的外部表征方式对贝叶斯推理成绩的影响

李晓明^{1,2},傅小兰¹,徐宇明¹

(1. 中国科学院心理研究所,北京 100101;2. 中国科学院研究生院,北京 100039)

摘要:基于分布式认知观点,考察信息的外部表征方式对人们解决贝叶斯推理问题的影响。实验采用 2(样例数:50、100)×2(呈现方式:频率信息集中呈现、样例信息逐个呈现)组间设计,并设置了概率信息呈现对照组。被试为 159 名大学生,均解决 3 个贝叶斯推理问题。结果表明,频率信息集中呈现组成绩最好,样例信息逐个呈现组成绩次之,概率信息呈现组成绩最差;样例数效应不显著;在样例信息逐个呈现条件下出现了明显的练习效应。

关键词:贝叶斯推理;分布式认知;频率信息集中呈现;样例信息逐个呈现;概率信息呈现

中图分类号:TB18;B849 **文献标识码:**A

The Effect of Informational External Representation on Bayesian Reasoning

LI Xiao-ming, FU Xiao-lan, XUAN Yu-ming

(Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China;

Graduate School, Chinese Academy of Sciences, Beijing 100039, China)

Abstract: The effect of informational external representation on Bayesian reasoning was studied from the perspective of distributed cognition. A 2 (Exemplar number: 50, 100) × 2 (Representation manner: frequency format, exemplar format) between-subjects design was used, and probability format was set as control condition. The results showed that: a) The performance of frequency format was the best among the three different external representation manners. b) exemplar number did not have any main effect. and c) There was significant practice effect in the exemplar format condition.

Key words: Bayesian reasoning; distributed cognition; frequency format; exemplar format; probability format

1 前言

分布式认知是一种新的认知观点^[1]。与传统的认知观点重视信息在头脑中的内部表征相比,该观点认为问题的外部表征可以与头脑中的内部表征相结合,在认知活动中共同起作用^[2]。外部表征方式对解决问题的绩效有很大的影响^[3]。如恰当的外部表征有助于解决河内塔问题^[4]。

以往研究表明,不同的外部表征方式会影响贝叶斯推理成绩^[5]。当以概率格式呈现贝叶斯推理问题时,只有 10%左右的被试能够正确解决问题,而当采用频率格式时,有 46%的被试能正确解决问题^[6]。研究者认为,被试成绩的提高是由于频率格式这种外部表征形式与人类进化发展出的内部推理规则系统相符合^[7]。

现实情景下,人们是通过自然取样获得频率信息的^[8]。但是前人研究往往是将事件的频率信息集中呈现给被试,而不是将信息以样例形式逐个呈现给被试,因此未能模拟现实情景中人们获得频率信息的方式。

本研究基于分布式表征的观点,在考察概率和频率两种信息呈现方式的同时,也考察样例逐个呈现信息条件下被试解决贝叶斯推理问题的成绩,试图进一步研究信息的外部表征方式对人们解决贝叶斯推理问题的影响。

2 研究方法

2.1 被试

中国农业大学 159 名本科生,82 名男生,77 名女生,被试年龄 18~25 岁,平均年龄 21.4 岁。

基金项目:国家自然科学基金资助项目(30270466);中国科技部 973 项目(2002CB312103);中国科学院心理研究所创新重点项目(0302037)

作者简介:李晓明(1980-),女,山东烟台人,硕士研究生,研究方向为认知心理学。

2.2 实验设计

实验采用 2(样例数:50、100) × 2(呈现方式:频率信息集中呈现、样例信息逐个呈现)组间设计,并且设置了概率信息呈现对照组。因变量为正确率。在样例信息逐个呈现条件下,因变量还有反应时间。

2.3 实验材料和程序

2.3.1 纸笔测试的实验材料

频率信息集中呈现和概率信息呈现均采用纸笔测试,都包括 1 个练习题目和 3 个正式题目。3 道正式题目的顺序采用拉丁方平衡。

频率信息集中呈现组材料示例:“已知全世界每 100 个人中就有 24 人患乙病。24 个患有乙病的人进行 X 化验时,有 16 人呈阳性反应。而没有患乙病的 76 个人在 X 化验中也有 24 人呈阳性反应。那么如果有一个人在 X 化验中呈阳性。问:他确实患有乙病的概率是多少?”(正确答案为 40%)

概率信息呈现组材料示例:“已知全世界患乙病的人占总人数的 24%。如果一个人患有乙病,他进行 X 化验时呈阳性反应的概率是 66.67%。一个人即使没有患乙病,他在 X 化验中也有 31.58% 的概率呈阳性。那么如果有一个人在 X 化验中呈阳性。问:他确实患有乙病的概率是多少?”(正确答案为 40%)

2.3.2 计算机呈现下的实验程序

样例逐个呈现条件的实验程序用 E-Prime 编写。材料也包括 1 个练习题目和 3 个正式题目。正式题目的顺序采用拉丁方平衡。

在练习阶段,给被试依次呈现 10 个体检者的化验和诊断结果,每个样例呈现 4 s;然后要求被试根据第 11 个体检者的化验结果作出诊断。例如,其中一个样例为“体检者:1 号;W 化验结果:- (阴性);诊断结果:没有甲病”;测试阶段为“体检者:11 号;W 化验结果:+ (阳性);诊断结果:请你判断该人实际患有甲病的概率有多大?(请按相应键做选择,1 键对应 10%,2 键对应 20%,以此类推。)”

在正式实验阶段,样例呈现和测试都与练习阶段相同,只是样例数为 50 或 100。

每个正式题目之后让被试看 2 min 的风景画进行休息。完成样例数为 50 的实验大约需 25 min;完成样例数为 100 的实验大约需 35 min。

3 实验结果分析

3.1 不同条件下贝叶斯推理的成绩分析

不同条件下被试解决 3 道贝叶斯推理问题的正确率见表 1。对 3 道题目下的正确率分别进行 2(呈现方式:频率信息集中呈现、样例信息逐个呈现) × 2(样例数:50、100) 的 Logistic 回归分析,结果表明,在 A 题中呈现方式的主效应显著 ($P < 0.001$),频率信息集中呈现条件下的正确率 (51.7%) 要明显高于样例逐个呈现条件下的正确率 (21.4%);在 B 题中呈现方式有边缘主效应 ($P = 0.057$),频率信息集中呈现下的成绩 (50%) 略微高于样例逐个呈现条件下的成绩 (33.5%);在 C 题中呈现方式的主效应都不显著。所有的交互作用均不显著。

将 3 种信息呈现方式下不同样例数和 3 道题目的结果合并后进行 χ^2 检验,结果表明,呈现方式主效应显著, $\chi^2 = 24.29, P < 0.005$ 。进一步的 χ^2 检验表明,频率信息集中呈现下的成绩 (51.7%) 显著地好于样例逐个呈现条件和概率信息呈现下的成绩 (分别为 33.2% 和 23.3%),但样例逐个呈现和概率信息呈现条件下的成绩之间没有显著差异。

对被试在样例逐个呈现条件下做判断的反应时间进行 3(题目:A、B、C) × 2(样例数:50、100) 方差分析,结果表明,项目主效应、样例数主效应、两者交互作用均不显著。

3.2 样例逐个呈现条件下的练习效应分析

先后呈现的 3 道题目的正确率(图 1)和反应时间(图 2)均显示,在两种样例数条件下可能都存在着一定的练习效应。对样例逐个呈现条件下的正确率进行 χ^2 检验,结果表明:样例数为 50 条件下,有练习效应的趋势,但并不显著;而样例数为 100 条件下, $\chi^2 = 11.78, P < 0.005$,出现了明显的练习效应。

对样例逐个呈现条件下被试做判断的反应时间进行 3(问题顺序) × 2(样例数:50、100) 方差分析,结果表明:问题顺序的主效应显著, $F(2, 134) = 14.98, P < 0.05$,即在样例逐个呈现条件下存在着明显的练习效应;样例数的主效应不显著;交互作用不显著。

对概率信息呈现和频率信息集中呈现下的正确率进行分析,结果表明在这两种实验条件下,问题顺序的效应均不显著。

表1 不同信息呈现方式下被试解决贝叶斯推理问题的正确率(%)

问题	呈现方式						—
	频率信息集中呈现			样例信息逐个呈现			
	50	100	平均值	50	100	平均值	
A	46.7	56.7	51.7	12.1	30.6	21.4	26.7
B	50.0	50.0	50.0	36.4	30.6	33.5	23.3
C	53.3	53.3	53.3	42.4	47.2	44.8	20.0
平均值	50.0	53.3	51.7	30.3	36.1	33.2	23.3

注:样例信息逐个呈现条件下,样例数为50和100时被试数分别为33和36,其余条件下的被试数都为30。

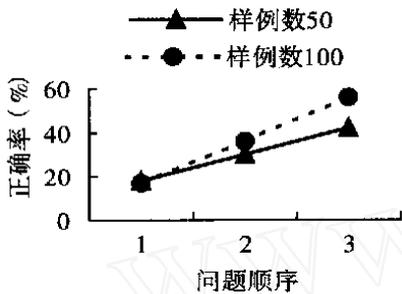


图1 样例逐个呈现条件下先后3道题目的正确率

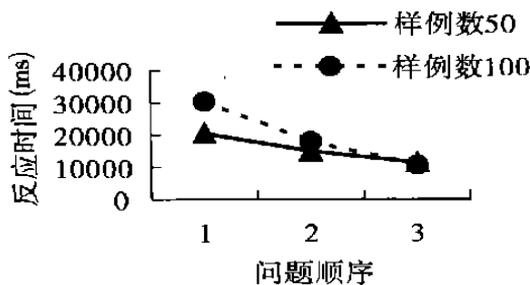


图2 样例逐个呈现条件下先后3道题目的反应时间

4 讨论

4.1 信息的外部表征方式对推理成绩的影响

本研究发现,样例数对解决贝叶斯推理的成绩没有显著影响,但信息的外部表征方式对推理成绩有显著影响。频率信息集中呈现条件下的成绩要远远高于样例信息逐个呈现和概率信息呈现下的成绩,而且高于前人实验研究所获得的正确率^[5]。这可能是因为在在本研究中,频率信息集中呈现条件采用纸笔测验,没有任何时间限制,因此被试有较充足的时间理解和解决问题。

样例逐个呈现条件下的成绩高于概率信息呈现条件下的成绩,可能是由于样例逐个呈现条件模拟了人们平时获得信息的情景,在这种条件下人们通过自然取样获得了频率信息,因此成绩较好。但是,在样例逐个呈现条件下,被试需要先收集一系列单个的样例信息,然后进行额外的认知加工,才能获得推理过程所需的频率信息,所以样例逐个呈现条件下的成绩明显低于频率信息集中呈现下的成绩。本研究结果表明,频率信息集中呈现是最有利于人们解决贝叶斯推理问题的外部

表征方式。

4.2 样例逐个呈现条件下的练习效应

对样例逐个呈现条件下的正确率和反应时所做的分析,都表明存在着练习效应,而对概率信息呈现和频率信息集中呈现条件下的正确率进行分析并没有发现此现象。这可能是因为,在样例逐个呈现条件下,随着实验的进行,被试逐渐意识到要完成贝叶斯推理任务,只需注意有症状及疾病的样例数和有症状但无疾病的样例数^[7],使认知资源的分配更为合理,从而提高了完成推理任务的绩效,表现出明显的练习效应。

上述发现有实际的应用价值。由于样例逐个呈现条件模拟了真实环境中人们获得信息的方式,这预示着在解决现实环境中的贝叶斯推理任务时,随着练习的增加和对信息环境的熟悉,人们解决贝叶斯推理问题的能力会逐渐提高。

5 结论

本研究的主要结论是: 频率信息集中呈现条件下的成绩最好,样例逐个呈现条件下的成绩次之,概率信息呈现条件下的成绩最差,说明频率信息集中呈现是最有利于人们解决贝叶斯推理问题的外部表征方式; 样例数对贝叶斯推理成绩没有显著影响; 样例逐个呈现条件下存在着明显的练习效应。

参考文献:

- [1] 周国梅,傅小兰. 分布式认知——一种新的认知观点[J]. 心理科学进展,2002,10(2):147-153.
- [2] Zhang J, Norman DA. Representations in distributed cognitive tasks[J]. Cognitive Science,1994,18(1):87-122.
- [3] Zhang J. The nature of external representations in problem solving[J]. Cognitive Science,1997,21(2):179-217.
- [4] Zhang J. The interaction of internal and external representations in a problem solving task[A]. Proceedings of the 13th Annual Conference of the Cognitive Science Society [C]. Hillsdale, NJ:Lawrence Erlbaum Associations,1991. 954-958.

- [5] 赵晓东,傅小兰. 贝叶斯推理的改进方法——以频率格式代替概率格式进行信息表征[J]. 心理科学, 2002, 25(1): 96-98.
- [6] Gigerenzer G, Hoffrage U. How to improve Bayesian reasoning without instruction: Frequency formats[J]. Psychological Review, 1995, 102(4): 684-704.
- [7] Gigerenzer G, 包燕. 生态学智力: 人的推理算法对频

率的适应[J]. 心理学动态, 2001, 9(4): 325-329.

- [8] Shanks D. A connectionist account of base - rate biases in categorization[J]. Connection Science, 1991, 3(2): 143-162.

[收稿日期] 2004-01-13

[修回日期] 2004-05-09

(上接第 2 页)

4 讨论

4.1 实验结果表明, Polar-star 信息显示方式在提高驾驶员异常信息判定的准确性方面优势明显。仪表组反应时较快则说明传统的仪表显示方式在预测趋势功能上具有优势。

4.2 本实验对驾驶员在不同显示形式下进行信息判定的准确性和反应时间测定的结果虽然不能完全反映 SA 水平指标,但从信息预测能力的角度衡量 SA 水平的重要性看来,实验结果初步反映出 Polar-star 显示方式在未来的驾驶舱设计中具备重要的应用潜力。

4.3 未来的研究工作应侧重于两个方面。一是借助现代仪器设备如脑电图测试仪(EEG)等,深入探讨状态警戒水平的评定指标和方法。二是基于这些方法和手段对 Polar-star 显示内容及方式进行进一步的优化设计,例如,怎样进行两种显示方式的综合设计;如何实现两者之间优势互补等

问题。

参考文献:

- [1] Trujillo A. Vertex movement for mission status graphics: A polar star display [R]. NASA report. TMF2002-211414. 2002.
- [2] Trujillo A, Alenka S. Human factors and information operation for a nuclear power space vehicle[R]. NASA report. TMF2002-211409. 2002.
- [3] Bartolone A, Trujillo A. Glass-cockpit pilot subjective ratings of predictive information, collocation, and mission status graphics: An analysis and summary of the future focus of flight deck research survey[R]. NASA report. TMF2002-211419. 2002.
- [4] 杨治良. 实验心理学[M]. 杭州:浙江教育出版社, 1998.

[收稿日期] 2003-08-13

[修回日期] 2004-05-09

(上接第 25 页)

欢的,这与常用的操作界面(Windows)一致。前景色也存在一定的规律,白色、黄色系和绿色系是被试比较偏好的前景色,但这一趋势仅经过数据观察得到,是否存在显著差异有待进一步的研究证明。搭配上来看,白色、黄色系和绿色系与蓝色、紫色和灰蓝色背景色搭配比较受被试欢迎。

总体来说,在操作系统界面的颜色设计中,背景色建议采用蓝色、紫色、灰蓝色和青绿色也是值得考虑的颜色。在蓝色、紫色和灰蓝色背景下,前景色建议采用白色、黄色系和绿色系。

参考文献:

- [1] 曾虹文. 中国学生颜色爱好特点及其发展[J]. 心理科学通讯, 1986, 9(1): 16-22.
- [2] 王益荣. 中国大学生颜色爱好的实验研究[J]. 心理科学, 1997, 20(6): 559-560.
- [3] Ellis L, Fricke C. Color preferences according to gender

and sexual orientation [J]. Personality Individual Differences, 2001, 31(8): 1375-1379.

- [4] 李文馥. 中国城市人爱好颜色的特点——颜色爱好与几种因素的关系[J]. 心理科学通讯, 1990, 13(5): 13-19.
- [5] 吴筱珍,王骥业. 关于大学生的《颜色——情绪》的调查[J]. 心理科学通讯, 1986, 9(1): 38-40.
- [6] 黄希庭,黄巍,李小融. 关于中国人颜色情调的研究[J]. 心理科学, 1991, 14(6): 1-7.
- [7] 孔燕,葛列众,王勇军. 黑白背景下 4 种颜色突显工效的比较研究[J]. 人类工效学, 1999, 5(4): 12-14.
- [8] 曹立人,朱祖祥. 对比度因素对彩色 CRT 视觉工效的影响[J]. 人类工效学, 1995, 1(1): 32-36.
- [9] 朱祖祥,曹立人. 目标——背景色的配合对彩色 CRT 显示工效的影响[J]. 心理学报, 1994, 26(2): 128-135.

[收稿日期] 2003-08-08

[修回日期] 2004-05-20