

# 性别和地域对概率判断准确性的影响

**摘要** 过分自信现象指个体的概率判断高于目标事件实际发生的概率，在不同文化、性别下普遍存在。本实验采用一般知识问题范式，使用两组单因素组间设计，研究性别（男性/女性）和地域（南方/北方）对平均概率分数 *PS*、偏差 *Bias*、斜率 *Slope*、离散 *Scat f* 的影响。结果显示，男性和女性的偏差 *Bias* 均显著大于 0，在四个因变量上均没有显著的性别差异和地域差异。本实验表明男性和女性均存在过分自信现象，且概率判断的准确性不受性别和地域的影响。

**关键词** 概率判断 过分自信 性别差异 地域差异

## 1 前言

概率判断（Probability Judgement）指面对决策中的随机因素时人们对事件发生的可能性的估计。概率判断在政策制定、风险评估、犯罪调查和医疗诊断等领域占有重要地位。（Keren, 1991）

Yates(1982)设计了研究概率判断的一般知识问题范式，即呈现有两个选项的一般知识问题，要求被试作答并报告他或她认为自己回答正确的可能性（50%到100%），计算概率判断的四个相关指标，即平均概率分数 *PS*（体现被试判断的总体准确性）、偏差 *Bias*（体现概率判断和事件实际发生概率的差异）、斜率 *Slope*（体现被试区分事件发生与不发生的辨别能力）、离散 *Scat f*（体现被试判断的随机变异性）。许多研究发现了“过分自信”现象，即人们关于一般知识问题的概率判断通常以偏高的方式偏离校准，校准是指如果一个人的校准是最理想的，那么在所有情形下他的平均概率判断将与目标事件实际发生的比例相等。（于窈，李纾，2006）有研究者认为过分自信是因为人们偏爱支持性的证据，更多地忽略反对性的证据。（Koriat et al., 1980）也有研究者认为过分自信是由不可靠的测量手段和随机误差造成的一种回归效应，并不能正确评估获得的信息。（Erev et al., 1994）过分自信现象在不同性别、文化下均存在。（Moore & Healy, 2008）

对于过分自信的性别差异的研究取得了不太一致的结果。Paulford 和 Colman(1997)的研究结果显示，男性被试普遍比女性被试有更高的过分自信。一项以中国被试为研究对象的实验却发现女性的过分自信比男性更高。（孟颖，2016）此外，Lunderberg 等人（2000）通过对来自多个国家的被试研究发现，在一个国家内过分自信的性别差异并不显著。有研究者发现，当实验材料为数学题时，男性表现出过分自信，女性却表现出过低自信，当实验材料为社会科学类题目时，男性和女性均表现为过分自信且没有显著性差异，所以他们提出，过分自信受到性别刻板印象的影响。（Jakobsson et al., 2000）因此，对于过分自信是否存在性别

差异，还有待进一步研究。

跨文化系列研究表明：面对常识和概率判断问题，集体主义文化成员（如中国人）比个体主义文化成员（如美国人）过分自信的程度更深。（于竊，李纾，2006）例如，有研究发现，中国被试的概率判断的偏差与美国被试的概率判断的偏差有显著差异，且中国被试几乎多一倍。（Yates et al., 1989）关于过分自信跨文化差异的解释之一是论点采纳模型(argument recruitment model)，当个体面对常识问题时，他首先要提出一些论点，然后权衡和评估这些论点，根据权衡评估的结果选择，并且论点越支持某一个选项，那么概率判断也就越高。（Yates et al., 1992）在论点采纳模型中，过分自信的跨文化差异可部分归因于教育传统的不同，因为它影响了论点采纳的习惯，研究分析得出两大教育策略，一是“直接指导”，强调信息的直接传递，二是建构式学习，强调辩证思维，集体主义文化较多“直接指导”式教育，而个体主义文化多是建构式。（Yang, 1997; Lee et al., 1998）

基于上述内容，本实验使用一般知识问题范式，采用两组单因素组件设计，研究性别和地域对平均概率分数  $PS$ 、偏差  $Bias$ 、斜率  $Slope$ 、离散  $Scatf$  四个因变量的影响。根据 Lunderberg 等人（2000）的研究结果，我们预期男性被试和女性被试均表现出过分自信，且偏差  $Bias$  没有显著性差异；有研究表明，由于传统上种植水稻比种植小麦更需要合作，所以中国的南方人比北方人的集体主义倾向更强（马欣然 等，2016），结合过分自信的跨文化研究结果，我们预期中国南方被试比北方被试表现出更多的过分自信；对于平均概率分数  $PS$ 、斜率  $Slope$  以及离散  $Scatf$  在性别和地域上的差异，本实验做出探索不提出假设。

## 2 方法

### 2.1 被试

剔除 3 个标准差以外的数据之后，被试为 74 名北京大学参加实验心理学实验课程的学生，年龄未记录，其中女性 43 名，来自北方（秦岭淮河以北）的被试 37 名。实验无报酬。

### 2.2 仪器和材料

实验材料为一张含有 40 道题问题的 A4 题纸，40 道题问题的内容为一般知识问题，涵盖了各个方面，包括年鉴问题（如，“秦统一中国的年代是”）、常识问题（如，“养兔子是否需要专门喂水”）、和科学历史问题（如，“长颈鹿的颈椎数目是”）等等。40 道提问均为两个选项的选择题，有且仅有一个选项为正确答案。

### 2.3 实验设计

本实验采用两组单因素组间设计。其中一个组间变量为地域，有两个水平（北方/南方），定义秦岭淮河以北为北方，秦岭淮河以南为南方；另一个组间变量为性别，有两个水平（男/女）。本实验的有 4 个因变量，其一为被试判断的整体准确性，用“平均概率分数”  $PS$  表示，公式如下，其中  $f_i$  为被试对第  $i$  题的概率判断， $d_i$  为第  $i$  题的结果指数（如果回答正确则  $d_i=1$ ，如果回答错误则  $d_i=0$ ）。

$$PS = \frac{\sum_{i=1}^{40} (f_i - d_i)^2}{40}$$

其二为被试的概率判断与事件实际发生概率的偏差大小,用偏差 *Bias* 进行衡量,其操作性定义为被试的平均概率判断 *f* 与被试的平均结果指数 *d* (即回答正确的比例)之差,公式如下。

$$Bias = f - d$$

其三为被试概率判断的准确性,用斜率 *Slope* 进行衡量,其操作性定义为回答正确时的平均概率判断 *f<sub>1</sub>* 与回答错误时的平均概率判断 *f<sub>0</sub>*之差,公式如下。

$$Slope = f_1 - f_0$$

其四为被试概率判断的随机变异,用离散 *Scatf* 进行衡量,其操作性定义为回答正确时的变异与回答错误时的变异的加权平均,公式如下,其中 *N<sub>1</sub>* 为被试回答正确的题目数量, *N<sub>0</sub>* 为被试回答错误的题目数量, *Var f<sub>1</sub>* 为被试回答正确时概率判断的方差, *Var f<sub>0</sub>* 为被试回答错误时概率判断的方差。

$$Scat f = \frac{N_1 Var f_1 + N_0 Var f_0}{N_1 + N_0}$$

本实验对潜在的额外变量进行了控制。首先,为了避免额外信息和结果反馈对结果造成影响,在答题过程中被试不得相互交流,也不能上网搜索,主试在所有被试均答完题目后再公布正确答案。其次,为了减少被试知识背景的个体差异对实验结果的影响,实验中采用的问题包括了多方面的知识。

## 2.4 实验程序

实验开始前,主试分发题纸,并宣读指导语,待被试没有以后之后开始。被试需要作答 40 道题目,要求被试从每道题所给的两个选项中选出他们认为正确的一项,并从 50%, 60%, 70%, 80%, 90%, 100% 中选出一个作为自己再这道题目上答对的概率判断(如果被试认为自己对该题完全没有把握则选 50%,如果被试十分肯定自己的选择是正确的则选 100%)。

待所有被试完成全部题目并对每道题都给出概率判断后,主试公布正确答案,统计每一种概率所对应的回答正确或错误的题目个数。

## 3 结果

本实验使用 SPSS 22 进行数据分析,剔除 3 个标准差之外的数据后,对被试在不同概率判断下的回答情况进行描述性统计,见表 1。基于全体被试在不同概率判断下回答正确或错误的题目数绘制全体被试的概率判断协方差图,见图 1。

表 1 不同概率判断下回答正确或错误的题目数

	50%	60%	70%	80%	90%	100%	总计
正确	570	310	205	194	183	350	1812
错误	569	247	125	113	74	87	1215

所有被试  $PS=0.240$   $Bias=0.082$   $Slope=0.069$   $Scat f=0.029$

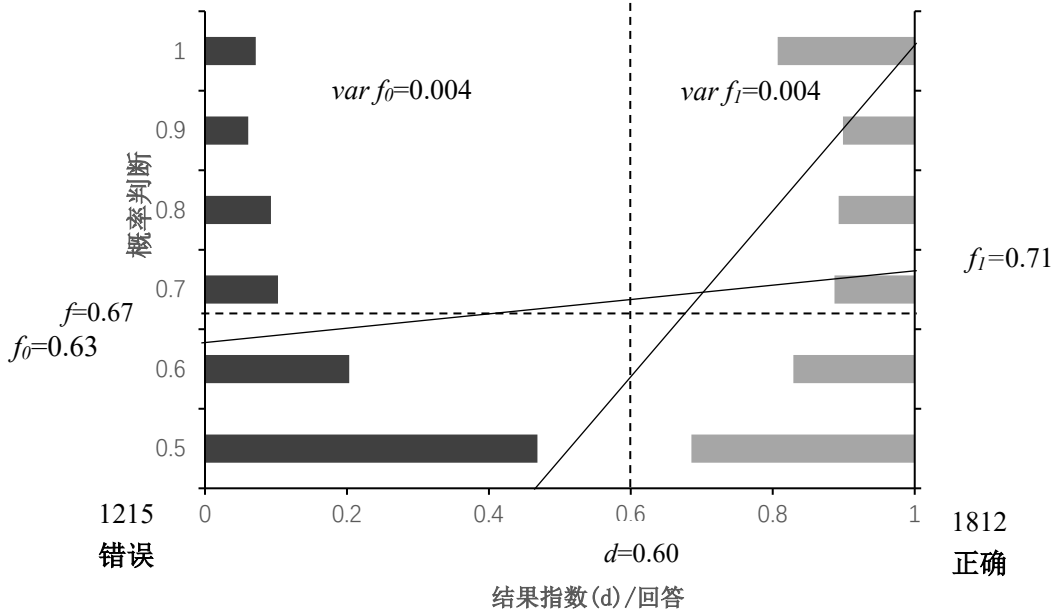


图1 全体被试的概率判断协方差图

根据被试在不同概率判断下的回答正确或错误的题目数量,计算出平均结果指数  $d$ , 回答正确时的平均概率判断  $f_1$ , 回答错误时的平均概率判断  $f_0$ , 所有题目的平均概率判断  $f$ , 进行描述性统计, 见表 2, 进而计算出被试的平均概率分数  $PS$ , 偏差  $Bias$ , 斜率  $Slope$  以及离散  $Scat f$ , 描述性统计见表 3。

表 2 不同被试的  $d, f_1, f_0, f$  ( $M \pm SD$ )

	$d$	$f_1$	$f_0$	$f$
男性	$0.61 \pm 0.08$	$0.74 \pm 0.06$	$0.64 \pm 0.06$	$0.70 \pm 0.05$
女性	$0.59 \pm 0.08$	$0.69 \pm 0.06$	$0.62 \pm 0.07$	$0.66 \pm 0.05$
北方	$0.60 \pm 0.08$	$0.71 \pm 0.07$	$0.62 \pm 0.07$	$0.68 \pm 0.06$
南方	$0.59 \pm 0.07$	$0.71 \pm 0.05$	$0.64 \pm 0.06$	$0.68 \pm 0.05$
总体	$0.60 \pm 0.08$	$0.71 \pm 0.06$	$0.63 \pm 0.06$	$0.68 \pm 0.05$

表 3 不同被试的各项因变量指标( $M \pm SD$ )

	$PS$	$Bias$	$Slope$	$Scat f$
男性	$0.24 \pm 0.04$	$0.09 \pm 0.10$	$0.09 \pm 0.05$	$0.03 \pm 0.01$
女性	$0.24 \pm 0.05$	$0.08 \pm 0.08$	$0.07 \pm 0.06$	$0.03 \pm 0.01$
北方	$0.24 \pm 0.05$	$0.08 \pm 0.10$	$0.09 \pm 0.05$	$0.03 \pm 0.01$
南方	$0.24 \pm 0.04$	$0.09 \pm 0.08$	$0.07 \pm 0.07$	$0.03 \pm 0.01$
总体	$0.24 \pm 0.04$	$0.08 \pm 0.09$	$0.08 \pm 0.06$	$0.03 \pm 0.01$

检验男女性被试在概率判断时是否存在过度自信倾向。将男性被试的 *Bias* 与 0 进行单样本 *t* 检验, 结果显示, 男性被试的 *Bias* 显著大于 0,  $t(32)=5.31$ ,  $p<.001$ 。将女性被试的 *Bias* 与 0 进行单样本 *t* 检验, 结果显示, 女性被试的 *Bias* 显著大于 0,  $t(42)=6.00$ ,  $p<.001$ 。

由于非参检验不需要满足对数据分布的假设, 本实验参照 Yates(1989)的统计分析方法, 使用 *Mann-Whitney U* 检验。

对不同性别的被试的平均概率分数 *PS*、偏差 *Bias*、斜率 *Slope* 以及离散 *Scat f* 进行独立样本 *Mann-Whitney U* 检验。结果显示, 男性和女性在平均概率分数 *PS* 上没有显著性差异, *Mann-Whitney U*=703.00,  $p=.946$ ; 男性和女性在偏差 *Bias* 上没有显著性差异, *Mann-Whitney U*=630.50,  $p=.408$ ; 男性和女性在斜率 *Slope* 上没有显著性差异, *Mann-Whitney U*=536.50,  $p=.07$ ; 男性和女性在离散 *Scat f* 上没有显著性差异, *Mann-Whitney U*=512.50,  $p=.092$ 。男性被试和女性被试的概率判断协方差图见图 2。

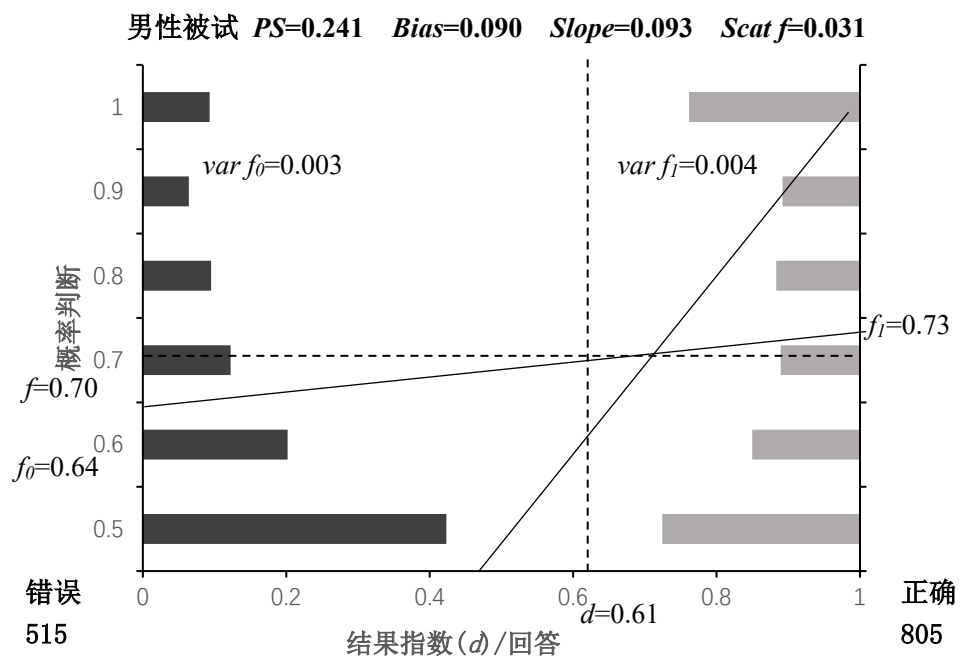
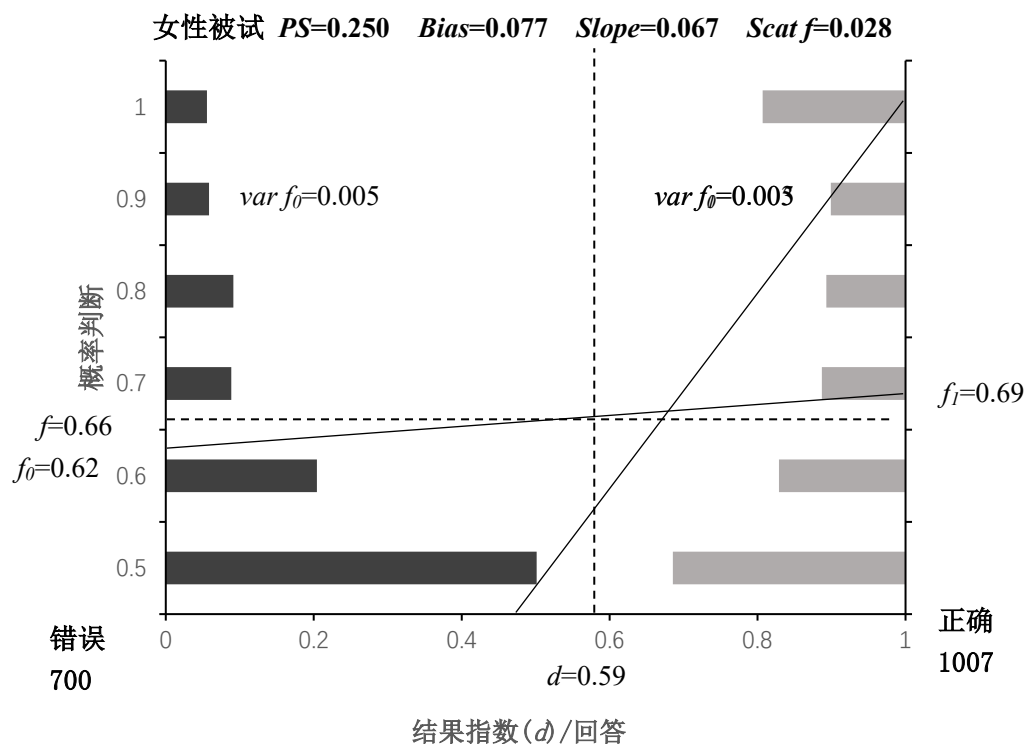


图2-1 男性被试的概率判断协方差图



对不同地域的被试的平均概率分数  $PS$ 、偏差  $Bias$ 、斜率  $Slope$  以及离散  $Scat f$  进行独立样本  $Mann-Whitney U$  检验。结果显示，北方被试和南方被试在平均概率分数  $PS$  上没有显著性差异， $Mann-Whitney U=642.00$ ， $p=.406$ ；北方被试和南方被试在偏差  $Bias$  上没有显著性差异， $Mann-Whitney U=652.50$ ， $p=.470$ ；北方被试和南方被试在斜率  $Slope$  上没有显著性差异， $Mann-Whitney U=593.00$ ， $p=.180$ ；北方被试和南方被试在离散  $Scat f$  上没有显著性差异， $Mann-Whitney U=558.50$ ， $p=.175$ 。北方被试和南方被试的概率判断协方差图见图 3。

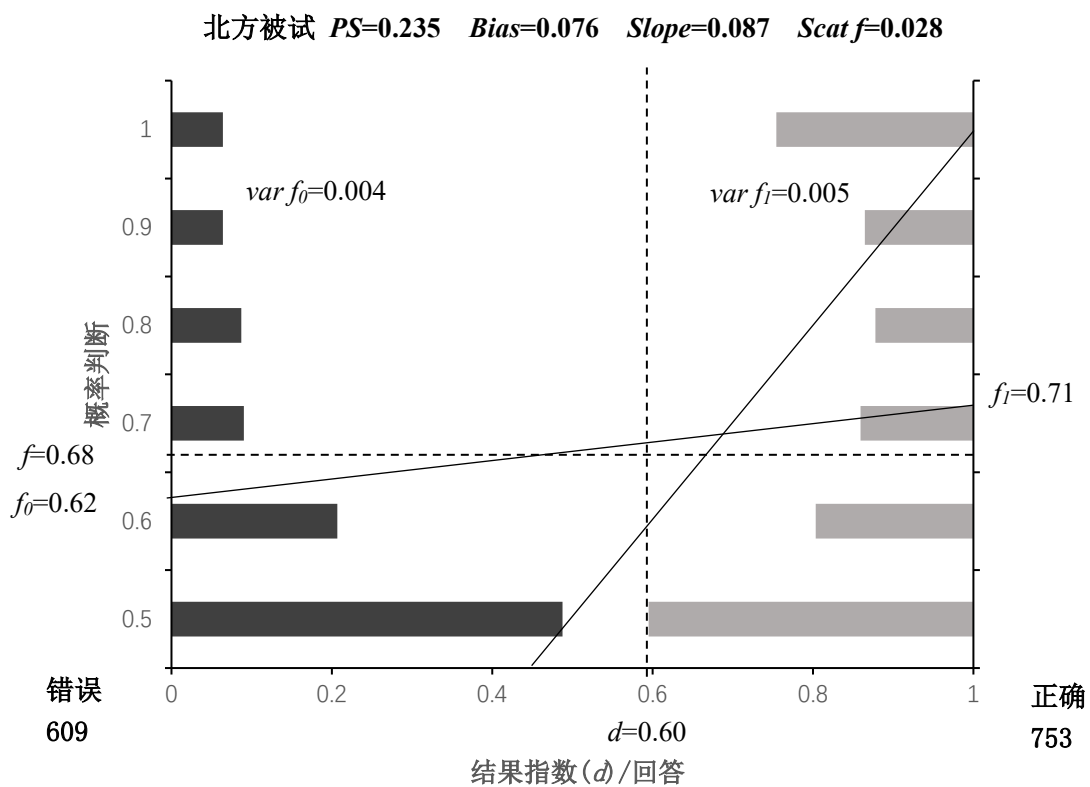


图3-1 北方被试的概率判断协方差图

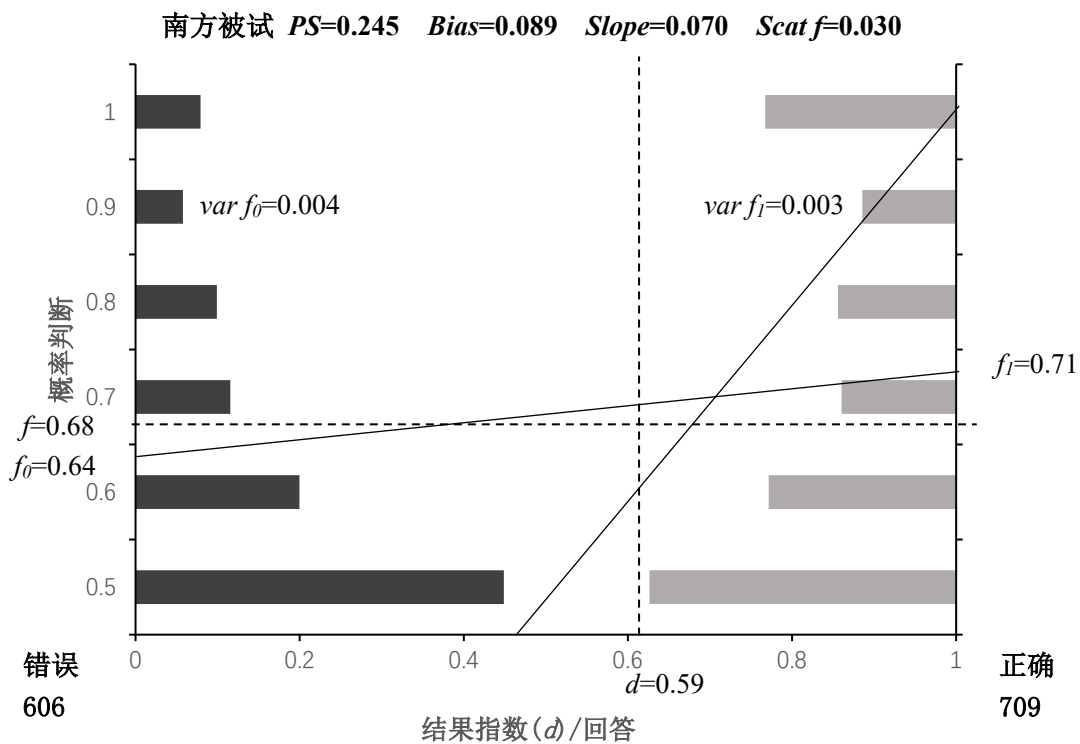


图3-2 南方被试的概率判断协方差图



## 4 分析与讨论

本实验使用一般知识问题范式,采用两组单因素组内设计,研究性别和地域对四个因变量的影响。结果显示,男性被试和女性被试的偏差 *Bias* 均显著大于 0 且没有显著性差异,证明了男性和女性均存在过分自信现象且没有发现过分自信的性别差异,与实验预期相符。但北方被试和南方被试的偏差 *Bias* 没有显著性差异,与实验预期不符。在不同性别和不同地域上平均概率分数 *PS*、斜率 *Slope*、离散 *Scatf* 均没有显著性差异。

首先,对于男性和女性均体现出过分自信现象,且偏差 *Bias* 没有显著性差异,与 Lundeberg 等人(2000)的研究结果相符,也为偏差 *Bias* 是由性别刻板印象造成的这一理论提供了支持(Jakobsson et al., 2000)。因为本实验所选用的一般知识题目包含多个方面的内容,减少了性别刻板印象的影响,而且本实验是在被试作答结束之后再收集被试的性别地域信息,没有启动被试的性别刻板印象,因此实验中并没有因为性别刻板印象造成偏差 *Bias* 上的性别差异。

其次,对于北方被试和南方被试在偏差 *Bias* 上没有显著性差异这一结果,我们认为有两点可能的原因。其一, Lee 等人(1998)认为集体主义文化和个体主义文化的过分自信程度的差异是由教育方式引起的,虽然有研究表明中国南方的集体主义倾向高于北方(马欣然 等, 2016),但中国的教育方式在南北方均是“直接指导”式占主导地位,这可能导致了过分自信的地域差异不显著。其二,认为集体主义文化过分自信更高的研究大多使用的是来自不同国家的被试,本实验使用的均为中国被试,有研究表明,与中国和美国之间的集体主义倾向差异相比,虽然中国南北存在集体主义倾向的差异,但这种差异要小得多。(Talhelm et al., 2018)所以可能是中国南北方的集体主义倾向差异比较小,导致本实验的实验结果与跨文化研究的结果不符。

综上所述,本实验的结果表明,男性和女性被试均表现出显著的过分自信,在偏差 *Bias*, 平均概率分数 *PS*、斜率 *Slope*、离散 *Scatf* 上均没有显著的性别差异和地域差异。后续研究可对概率判断中的性别、地域差异做进一步的探究并考虑刻板印象的作用。

## 参考文献

- Dahlbom, L., Jakobsson, A., Jakobsson, N., & Kotsadam, A. (2011). Gender and overconfidence: are girls really overconfident?. *Applied Economics Letters*, 18(4), 325-327.
- Keren, G. (1991). Calibration and probability judgements: Conceptual and methodological issues. *Acta Psychologica*, 77(3), 217-273.
- Lundeberg, M. A., Fox, P. W., Brown, A. C., & Elbedour, S. (2000). Cultural influences on confidence: Country and gender. *Journal of Educational Psychology*, 92(1), 152.
- Lundeberg, M. A., Fox, P. W., Brown, A. C., & Elbedour, S. (2000). Cultural influences on



- confidence: Country and gender. *Journal of Educational Psychology*, 92(1), 152.
- Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological review*, 115(2), 502.
- Talhelm, T., Zhang, X., & Oishi, S. (2018). Moving chairs in Starbucks: Observational studies find rice-wheat cultural differences in daily life in China. *Science advances*, 4(4), eaap8469.
- Yates, J. F. (1982). External correspondence: Decompositions of the mean probability score. *Organizational Behavior and Human Performance*, 30(1), 132-156.
- Yates, J. F., Zhu, Y., Ronis, D. L., Wang, D. F., Shinotsuka, H., & Toda, M. (1989). Probability judgment accuracy: China, Japan, and the United States. *Organizational behavior and human decision processes*, 43(2), 145-171.
- 马欣然, 任孝鹏, 徐江. (2016). 中国人集体主义的南北方差异及其文化动力. *心理科学进展*, 24(10), 1551-1555.
- 孟颖.(2016).不同框架下的管理者过度自信对风险决策的影响研究(硕士学位论文,山西师范大学).
- 于窈, 李纾. (2006). “过分自信”的研究及其跨文化差异.