

# 条件推理中的二元效应初探<sup>\*</sup>

张风华 杨 群 张庆林<sup>\*\*</sup> 汤永隆

(西南大学心理学院, 重庆, 400715)

**摘 要** 通过预备实验选取前件(或后件)是二元(或多元)的 5 种条件命题的 4 种推理形式作为实验材料, 以大学本科生为被试, 考察了五种不同条件命题类型对 4 种推理结果的影响。结果表明命题类型、推理规则以及两者的交互作用都非常显著。这个结果进一步验证了条件推理的集合映射模型。

**关键词:** 条件推理 集合映射模型 命题类型 推理规则

## 1 引言

条件推理是一种最简单的演绎推理形式, 在条件推理的过程中, 人们究竟是通过模型的建构和整合来得出结论, 再通过搜索反例来检验该结论的有效性, 从而最终做出正确的推断; 还是基于对命题所描述事实的发生频率(概率)进行一种更为直觉的判断和决策呢? 遗憾的是到目前为止, 人们的推理究竟遵循怎样的规律仍然没有一个统一可信的答案。

基于以往的一系列研究<sup>[1-5]</sup>, 张庆林等人于 1994 年提出了集合映射模型(也称问题空间模型)。该模型认为, 被试在推理任务中, 往往先对大前提进行语义加工, 并形成心理表征(即问题空间), 然后再依据小前提在这个问题空间中搜寻答案。大前提的命题语义加工(心理表征)决定了推理者在怎样的一个问题空间中搜索答案, 问题空间的性质(前件和后件之间的映射关系)决定了如何进行下一步的推理和推理的结果。这里所说的集合映射关系, 是指大前提的

前件集合的表征(分为 P、非 P 两个子集)、后件集合的表征(分为 Q、非 Q 两个子集)以及两者之间的映射(P 与 Q 的联系)。例如, 图 1 分别是“如果下雨, 那么打伞”、“如果不下雨, 那么去公园”、“如果下雨, 那么不去逛商店”、“如果 A, 那么 B”四个命题的在大多数人头脑中的集合映射关系图。这种集合映射关系形成了被试完成推理任务的一个问题空间。在图 1 中, 命题表征中的映射关系体现了被试在日常生活经验中所体验到的“前件”集(含肯定与否定两个子集)与“后件”集(含肯定与否定两个子集)之间关系的多少, 这将进一步影响推理的结果。例如, 对于命题“如果不下雨, 那么去公园”的表征而言(见图 1), 多数人会觉得, 说“不下雨去公园”就意味着“下雨就不去”, 因为他们觉得: 如果“下雨也去”, 那么“如果不下雨, 那么去公园”就是多余的了, 反正下雨、不下雨都去。在这种情况下, 集合映射关系中就没有“下雨”到“去公园”的映射。这种表征在其后的推理中成为一个搜索答案的问题空间。例如, 如果出现的小前提是“下雨了”(否

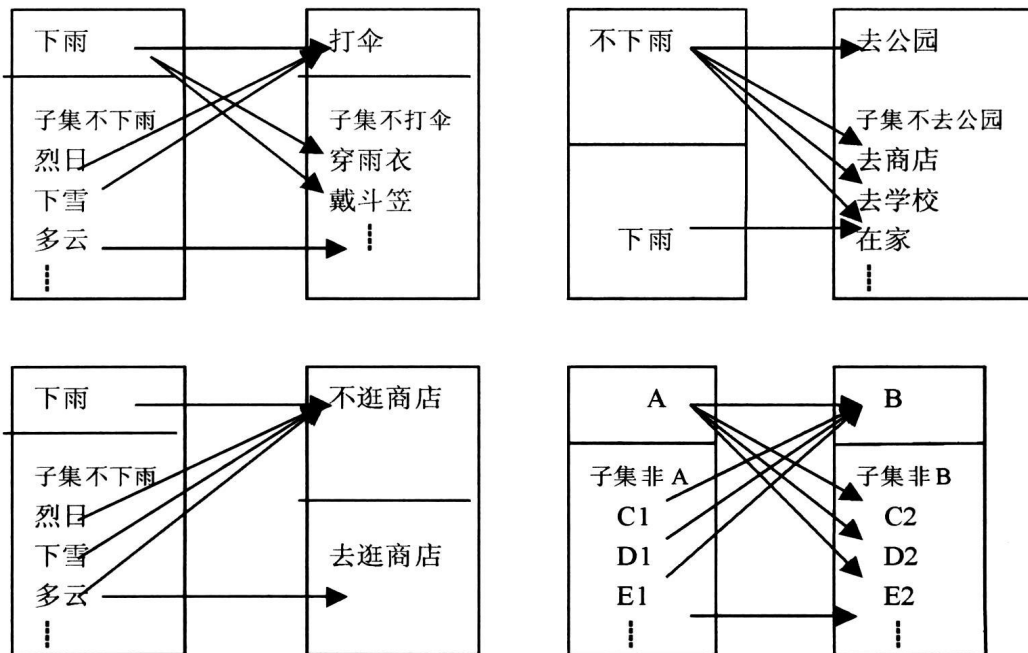


图 1 不同命题的不同表征(集合映射关系图)

国家重点学科西南大学基础心理学项目资助(西国重 04001)。

<sup>\*\*</sup>、通讯作者: 张庆林。E-mail: zhangql@swu.edu.cn

定前件),那么会得出“不去公园”(否定后件)的不符合逻辑的结论;如果出现的小前提是“去公园了”(肯定后件),那么,会得出“未下雨”(肯定前件)的不符合逻辑的结论。此时,出现阻碍效应。但是,就“如果下雨那么不去逛商店”而言,根据同样的原理,被试在小前提是“去逛了商店”(否定后件)的条件下更容易得出“未下雨”(否定前件)的符合逻辑的结论。此时,出现了促进效应。

虽然这一理论得到了张庆林等人 1994 年实验结果的证实<sup>[1]</sup>,但是,该实验中的大前提涉及到了否定词“不”,可能会产生“否定词效应”<sup>[6]</sup>,所以需要不含否定词的命题加以检验和发展。本实验采用前件(或/和后件)具有“二元”性质的命题材料来进一步检验集合映射理论。

## 2 关于条件命题的二元效应

在课堂上如果老师说:“如果是女生,那么举手”,一般人就会理解为“男生不举手”(违反逻辑的结论)。如果某个男生问“我是男生是否需要举手”,反而会被反问:“你是女生吗?”在这种情况下,符合逻辑的理解反而成了社会“不适应”的表现。在现实生活中,如果条件推理的大前提的前件具有二元的性质(例如“男”和“女”),那么,人们通常就会认为肯定一种可能就意味着否定另外一种可能,即存在着一种“二元表征思维”。反之,如果条件推理的大前提的前件不具有二元的性质,例如,对于大前提“如果下雨,那么打伞”而言,否定前件“下雨”,多数人会想到“不下雨”之外包括多种与“打伞”相关的状态,例如,“大太阳”、“下雪”、“下冰雹”等等,此时人们就不会从否定前件的信息中推理出否定后件的错误结论。原因可能是人们比较容易从否定前件子集(“未下雨”)中搜索到对“打伞”的映射的例子(如:出太阳打伞,下雪时打伞,见图 1),所以就不容易在否定前件的推理中得出否定的结论,而是得出“不一定”的结论。

上述的“二元表征思维”虽然不符合形式逻辑的规则,却是符合日常生活中的习惯化思维方式。而且,在日常生活中,并不会因为人们违反了形式逻辑规则而产生理解或交流上的问题。而且从生态学的角度看,人们之间的交流应该是最简洁的。人们一般会认为,如果“如果是女生,那么举手”和“如果是男生,那么也举手”同时正确,那么,“如果是女生,那么举手”,这句话就是多余的。所以,人们通常会认为,如果说“如果是女生,那么举手”,那么,“如果是男生,那么举手”就必须是错误的。

本研究假设,这种“二元表征思维”的逻辑的非理性和生态学的适应性,只能用集合映射理论来解释。按照集合映射理论的观点,条件命题中的前件(或后件)是二元(或多元),会影响条件命题表征中“前件”集与“后件”集之间的映射关系,并进一步影响推理的结果。本研究称这种影响为“条件命题的二元效应”,而且这种“二元效应”有时表现为阻碍效应,有时表现为促进效应。

## 3 实验方法

### 3.1 被试

选取西南大学本科生 35 人,其中男生 17 人,女生 18 人,他们均未专门学过逻辑学,视力或矫正视力正常,年龄在 19

—23 岁。

### 3.2 实验设计

实验采用  $5 \times 4$  的两因素被试内实验设计。因素 1 为 5 种类型的条件命题:命题 I 为前件二元、后件二元(BB);命题 II 为前件二元、后件多元(BN);命题 III 为前件多元、后件二元(NB);命题 IV 为前件多元、后件多元(NN);命题 I—IV 的内容均是被试熟悉的具体内容。命题 V 为抽象命题,前件多元、后件多元(NN)。每一种命题分别有 3 个同质的条件命题,总共 15 个条件命题。所有命题都经过预备实验证实符合实验要求。下面分别给出了 5 种条件命题的一个例子。

命题 I:如果是女生,那么站在左边。

命题 II:如果比赛赢了,那么奖励 100 元。

命题 III:如果是老人,那么就享受特殊待遇。

命题 IV:如果去北京,那么坐汽车。

命题 V:如果是字母“A”,那么是红色。

因素 2 为 4 种推理规则:肯定前件(MP)、否定前件(DA)、肯定后件(AC)、否定后件(MT)。

因变量为被试的得分。记分标准为选择正确(符合形式逻辑的规则)记 1 分,错误记 0 分。因此每种命题下每种推理规则的最高分为 3 分,最低分为 0 分。

### 3.3 实验程序

用计算机对被试进行个别测试。整个实验采用 E-prime 程序进行编程,五种大前提和四种推理规则均以随机的方式出现。每个被试需要做 60 个题目,为了防止疲劳效应,实验分两个阶段完成,做完一半题目后中间进行休息。每个被试总历时 10min 左右。实验结束后要求被试填写一份相关的问卷,要求对实验中的 15 种命题的前件与后件进行“二元”评定。目的是检验被试所认可的“两元”、“多元”是否与实验中所划分的“两元”、“多元”一致。

### 3.4 结果

实验结束后,共收集到 33 个完整的数据。对这 33 个被试的问卷进行统计,发现被试对实验材料中“二元与多元”的划分与正式实验中对实验材料“二元与多元”的划分基本上是一致的,证明该实验的材料是有效的。因此对 33 名被试的推理结果进行重复测量的二元方差分析。

#### 3.4.1 命题类型与推理规则的主效应

被试在五种命题类型下四种条件推理的得分以及平均正确率,见表 1。对表 1 的数据结果做方差分析,见表 2。结果表明命题类型的主效应显著,  $F(4, 659) = 3.688$ ,  $p = 0.006$ ;推理规则的主效应非常显著,  $F(3, 659) = 33.725$ ,  $p = 0.000$ ;两者的交互作用也非常显著,  $F(12, 659) = 11.596$ ,  $p = 0.000$ 。

对命题类型的进一步配对  $t$  检验结果表明:类型 I 与类型 IV、类型 V 在 0.005 的水平上存在着显著差异,类型 II 与类型 I 在 0.05 的水平上存在着差异,类型 III 与类型 V 在 0.05 的水平上存在着差异;对推理规则进一步的配对  $t$  检验结果表明,肯定前件与否定前件、肯定后件、否定后件之间都存在着非常显著的差异;否定后件与否定前件、肯定后件之间存在着非常显著的差异;而否定前件与肯定后件之间不存在着显著差异。

#### 3.4.2 命题类型与推理规则的交互作用

表 1 不同命题的条件推理的成绩 (n=33)

命题 类型	MP			DA			AC			MT		
	M	SD	%	M	SD	%	M	SD	%	M	SD	%
I	2.727	0.664	<b>90.909</b>	0.394	0.736	<b>13.131</b>	0.606	0.649	<b>20.202</b>	2.394	0.736	<b>79.800</b>
II	2.606	0.736	86.900	1.182	0.716	39.393	1.424	0.922	47.474	1.697	0.969	49.494
III	2.000	0.492	<b>53.535</b>	1.061	1.043	34.343	1.121	1.037	37.373	2.303	0.758	76.770
IV	2.576	0.653	88.890	1.515	0.988	<b>51.515</b>	1.606	0.851	<b>53.535</b>	1.424	1.181	<b>47.474</b>
V	2.667	0.804	89.900	1.303	1.000	43.434	1.394	1.043	46.464	1.909	1.083	63.636
Total	2.515	0.670	82.027	1.091	0.897	36.363	1.230	0.900	41.010	1.945	0.945	63.435

注：黑体字表示各种推理形式下的最高与最低正确率(%)。

命题类型与推理规则之间存在着非常显著的交互作用(见表 2)。进一步的简单效应检验表明,就命题类型 I 而言,对于否定前件与肯定后件,与其它四种命题类型相比都达到了显著差异,表现出了阻碍效应;对于否定后件,与命题类型 II、IV、V 相比都达到了显著差异,表现出了促进效应。就命题类型 II 而言,对于否定后件,与命题类型 I、III 相比都达到了显著差异,表现出阻碍效应。就命题类型 III 而言,对于肯定后件,与命题类型 I 相比达到了显著差

表 2 两因素方差分析组内效应的检验结果

变异来源	平方和	自由度	均方	F	P
命题类型	7.400	4	1.850	3.688**	0.007
推理规则	217.180	3	72.393	33.725***	0.000
命题类型×推理规则	69.691	12	5.808	11.596***	0.000
命题类型×被试	64.200	128	0.502		
推理规则×被试	206.070	96	2.147		
命题类型×推理规则×被试	192.309	384	0.501		
总计	795.786	659			

注：\*  $p<0.05$ , \*\*  $p<0.01$ , \*\*\*  $p<0.001$ 。

异,表现出了促进效应;对于否定后件,与命题类型 II、IV 相比达到了显著差异,也表现出了促进效应。就命题类型 IV 与命题类型 V 的比较而言,对于否定后件,达到了显著差异,表现出了阻碍效应。

4 讨论

4.1 对集合映射模型的检验

本实验中的因素 1 主效应显著,且与因素 2 的交互作用显著,这一结果用以往有关条件推理的主要理论,都是很难解释的。本研究从条件推理的命题的前件与后件是否是“二元”这一个新的角度来解释本实验条件下的结果,验证了集合映射模型。

第一,命题类型 I(BB)的结果最引人注目,从表 1 可以看到,其否定前件和肯定后件的正确率是五种命题中最低的,存在着很强的阻碍效应,而否定后件正确率是五种命题中最高的,存在很强的促进效应。对此,集合映射的解释是(以“如果是女生,那么站在左边”为例):大多数被试对命题“如果是女生,那么站在左边”所形成的表征可以直观地表示为图 2,其特点是在集合映射关系中只有“女生”到“站在左边”和“男生”到“站在右边”的映射,而没有“女生”到“站在右边”和“男生”到“站在左边”的映射。正是有了这样的表征,所以在否定前件(男生)的情况下 86.87%的被试选择了否定后件(站在右边);在肯定后件(站在左边)的情况下 79.8%的被试选择了肯定前件(是女生),表现出了很强的阻碍效应。同样的原因,在否定后件(站在右边)的情况下,79.8%的被试选择了否定前件(是男生)的结论,表现出明显的促进效应。

第二,就命题类型 II (BN)而言,被试在头脑中的表征容易缺少从否定前件到肯定后件的映射,例如,被试看到“如果比赛赢了,那么奖励 100 元”,就自动会认为“比赛输了就不

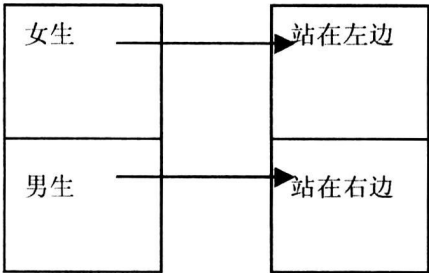


图 2 命题的前件二元,后件二元

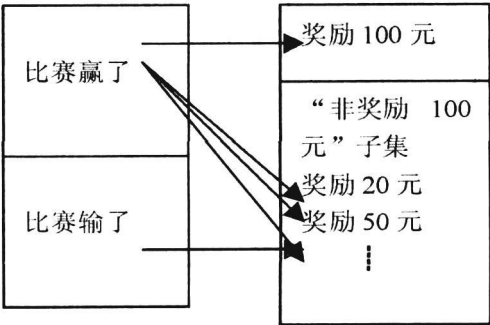


图 3 命题的前件二元,后件多元

能奖励 100 元”(见图 3)。但是,被试在头脑中的表征容易有从肯定前件到否定后件的映射,例如,被试会认为“比赛赢了也可能奖励 20 元或 50 元等”。因此在否定后件中表现出阻碍效应。

第三,就命题类型 III (NB)而言,被试在头脑中的表征容易有否定前件到肯定后件的映射,例如,被试看到“如果是老人,那么就享受特殊待遇”,就会自动激活“军人、残疾人也会享受特殊待遇”(见图 4)的事例,因此在肯定后件和否定后件

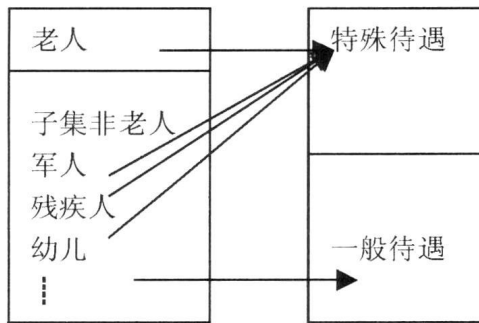


图4 前件多元, 后件二元

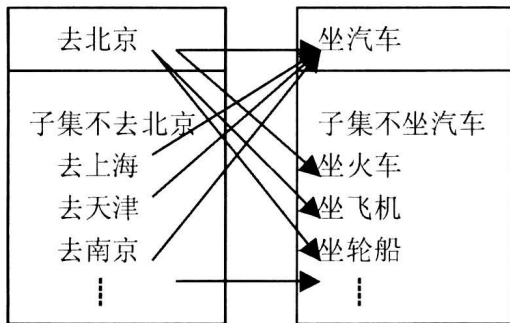


图5 前件多元, 后件二元

第四,就命题类型Ⅳ(NN)而言,从表1可以看到,其否定前件和肯定后件正确率是五种命题中最高的,而否定后件正确率是五种命题中最低的。集合映射理论的解释是,大多数被试对命题类型Ⅳ的表征容易有从否定前件到肯定后件的映射(见图5)。以“如果去北京,那么就坐汽车”为例,对否定前件的推理而言,“不去北京集”容易使被试激活“去上海”、“去济南”、“去南昌”坐“汽车”的事例。同样的,对“不坐汽车集”也容易激活“坐飞机”、“坐火车”去“北京”的事例。因此前、后件均为多元的情况下,大多数被试都能够激活P到非Q以及非P到Q的映射关系,既有“去北京”到“坐汽车”,“去北京”到“不坐汽车集”的映射,也有“不去北京集”到“坐汽车”,“不去北京集”到“不坐汽车集”的映射。正是因为有这种表征,所以在否定前件(不去北京)的情况下,51.51%的被试选择了不一定(可能坐汽车也可能不坐汽车)的正确答案,在肯定后件(坐汽车)的情况下53.53%的被试也选择了不一定(可能去北京也可能不去)的正确答案。同样的原因,在否定后件(没有坐汽车)情况下,仅有47.47%的被试选择了否定前件(没有去北京)的正确答案。

第五,就命题类型Ⅴ而言,虽然和命题类型Ⅳ同是“前、后件多元”,但是,命题类型Ⅴ是抽象命题,而类型Ⅳ是具体命题。结果表明,两者只有在否定后件的情况下,命题类型Ⅴ的正确率高于类型Ⅳ。其原因可能是因为抽象命题的情况下,比较少的被试会想到P到非Q的映射,以“如果是字母A,那么是红色”为例,被试不太容易想到“如果是字母A,那么是非红色”也是正确的。但是在具体命题的情况下,比较多的被试会想到P到非Q的映射,以“如果去北京,那么坐汽车”为例,被试比较容易想到“到北京不坐汽车”的例子。

由以上的分析可以看出,不同的命题类型在四种规则下,表现出了十分复杂的差异。对这种复杂差异,只能够用被试在对命题进行语义加工的时候,形成的前、后件之间的集合映

射关系的差异来解释。本实验结果基本上证实了前面的实验假设,在一定程度上验证了张庆林等提出的集合映射模型<sup>[1,4]</sup>。

#### 4.2 对其它理论模型的质疑

集合映射模型能很好地解释条件推理中的概率效应。例如,就条件命题,“如果去逛街,就会丢失钱包”(低条件概率)和“如果经常参加体育锻炼,身体就会很健康”(高条件概率)而言,很显然丢失钱包的原因有多种,可能是在逛街的时候,也可能是在做公共汽车的时候,或者是上下班的时候等;而身体很健康的原因则很可能是经常参加体育锻炼。因此,后者很容易从“身体健康”推测出“经常参加体育锻炼”的结论,而前者则不容易从“丢失了钱包”推出“去逛了街”的结论。所以,与其说是条件概率对推理产生了影响,倒不如说是被试对条件命题进行表征时所形成的前、后件之间的联系,直接决定了个体对结论可信度的直观判断。

按照双加工理论,对于具体的命题,被试倾向于采用系统1,根据已有的知识经验进行快速启发;而对于抽象的命题,更倾向于采用抽象的推理规则和假设思维模式,因此在具体命题的条件下更应该得出不符合逻辑的判断。但实验结果表明,抽象命题的条件下,4种推理规则的正确率与具体命题相比,都不是最高的。在抽象命题的条件下,4种推理规则的推理结果差异显著,在4种具体命题的4种推理规则下,也呈现出复杂的推理结果,有些条件下出现促进效应,有些情况下出现阻碍效应。因此不能把人类的推理过程简单地归为“两类加工”。

#### 4.3 人类遵循着一种朴素的生活逻辑

人类究竟是采用形式逻辑进行思维,还是遵循了某种“生活逻辑”?在以往的研究中<sup>[7,8]</sup>,大家对此争论不休。针对本研究的结果,人类的思维活动似乎遵循着某种朴素的生活逻辑,以“如果考试作弊,就会受到处罚”为例,个体在进行条件推理时,首先会根据自己以往的知识经验对命题进行适当的语义表征,“考试作弊,会受处罚”、“考试中受到处罚,就是作弊了”似乎已经是一种经验性的知识,个体会根据激活的这种日常生活经验进行判断,因此前、后件之间的这种相关联系的多少,即集合映射关系的多少,直接影响着人类的思维活动。由此可见,人类并不是单纯的采用形式逻辑进行思维,在进化过程中,人类为了适应生存,更学会了用生活逻辑进行思维。

#### 4.4 未来研究方向

综上所述,条件推理的过程就是个体在对命题进行语义表征之后,对问题空间的搜索过程,其推理的结果取决于日常生活经验中积累的前、后件之间相关联系的性质。但是,本实验并没有回答关于集合映射关系的所有问题,例如,对于条件命题的表征,本研究并没有真正测量被试对前、后件之间的映射关系,包括这种联系的有无、多少、容易激活程度、容易被接受的程度。关于条件推理的认知机制还有许多值得探讨的问题,例如,个体是根据搜索到的一个“映射”就作出判断,还是尽可能搜索到更多的“映射”关系才最后作出判断?“映射”关系的激活是否是自动的,被激活这种“映射”关系的进一步加工是否是控制加工?如何对这种自动加工与控制加工进行分离?因此在后续的研究中,需要进一步探

索,找出真正影响条件推理的因素。

## 5 结论

本研究发现,不同的推理规则之间确实存在着差异,肯定前件与其它三种推理规则相比,比较容易掌握。不同的命题类型下,被试的推理结果也不同。这些结果证明了被试对条件命题的表征确实影响了人们的推理,而人们的生活经验决定了对各种不同类型的条件命题的不同表征。

## 6 参考文献

- 1 张庆林,王振勇,蒲晓容. 条件推理中的命题语义加工. 心理科学, 1994, 17(6): 346—353
- 2 张庆林,杨雄. 大学生解决四卡问题的叙述理由效应. 心理学报, 1998, 20(3): 348—352
- 3 张庆林,杨雄. 四卡问题的内容促进效应的实验研究. 心理科学, 1997, 20(4): 311—383
- 4 邱江,张庆林. 有关条件推理中概率效应的实验研究. 心理科学, 2005, 28(3): 554—557
- 5 张庆林,王永明,张仲明. 假设检验思维过程中的启发式策略研究. 心理学报, 1997, 29(1): 29—36
- 6 刘志雅,赵冬梅,郑雪. 假言推理的否定词效应. 心理科学进展, 2002, 10(4): 395—402
- 7 Daniel G. Goldstein, Gerd Gigerenzer. Models of Ecological Rationality: The Recognition Heuristic. Psychological Review, 2002, 109(1): 75—90
- 8 Evans B T. Logic and Human Reasoning: An Assessment of the Deduction Paradigm. Psychological Bulletin, 2002, 128(6): 978—996

# A Research on the Binary Effect of Conditional Reasoning

Zhang Fenghua, Yang Qun, Zhang Qinglin, Tang Yonglong

(Department of Psychology, Southwest University, Chongqing 400715)

**Abstract** Five types of proposition were chosen as the experiment material in four inference forms whose antecedents and consequents could be binary or non-binary by prepared experiments on university students. The influence of five types of proposition on the performances of the four inferences was investigated in the experiments on university students. The results indicated that the main effects of the type of proposition and inference forms were significant. Furthermore, the interaction among them was significant, too. These results further verified the sets mapping model of conditional reasoning.

**Key words:** conditional reasoning, sets mapping model, type of proposition, inference form

(上接第 1467 页)系数会丢失大批数据信息。

Maki 等人<sup>[5]</sup>发现个体元认知监控的绝对精确性与任务难度、学习者的知识水平等有关,而相对精确性则相对独立,这说明二者测量的可能是元认知监控的不同方面。针对 Gamma 系数的第一点不足,我们建议在研究中不仅要考察个体 JOL 的相对精确性,也应考察绝对精确性,并对之进行比较,以求更深入的揭示个体元认知监控的机制与作用。

Gamma 系数只利用同序对和异序对的数据,而不将,和等 3 种等值对纳入计算。本文认为,前两种等值对数据不能为个体的 JOL 精确性提供十分有用的信息,但包含的信息却显示了被试 JOL 精确性的信息,它表示被试在预测判断阶段对两个项目赋予了同样等级的值,而在实际测试阶段其测试成绩的值恰又是相等的。所以我们建议在计算时考虑这一指标,比将 Gamma 系数修正为等形式。修正的指数是否具有良好的统计属性,还有待用数据模拟(simulation)的方法加以考察验证。

# The methodology for research on the relative JOL accuracy

Chen Qishan

(Department of Educational Psychology, The Chinese University of Hong Kong, Hong Kong)

**Abstract** Judgment of learning (JOL) is a person's predictions about the likelihood of correctly retrieving recently studied items on an upcoming test. It has become one of the most intensely investigated of the metacognitive judgments. In this paper, the methodology for research on JOL, its accuracy and the relative accuracy index of JOL were introduced and discussed. Finally, some suggestions and new directions were proposed.

**Key words:** Judgment of learning, relative accuracy, LJR, PRAM, Gamma correlation

## 5 参考文献

- 1 Nelson T O, Narens L. Why investigate metacognition? In Metcalfe J, Shimamura A P. Eds. Metacognition: Knowing about knowing. Cambridge, Mass: MIT Press, 1994, 1—25
- 2 Goodman L A, Kruskal W H. Measures of association for cross classifications. Journal of the American Statistical Association, 1954, 49: 732—764
- 3 Nelson T O. A comparison of current measures of the accuracy of feeling-of-knowing predictions. Psychological Bulletin, 1984, 95: 109—133
- 4 Nelson T O, Narens L, Dunlosky J. A revised methodology for research on metamemory: Pre-judgment recall and monitoring (PRAM). Psychological Methods, 2004, 9(1): 53—69
- 5 Maki R H, Shields M, Wheeler A E, Zaccilli T L. Individual differences in absolute and relative metacomprehension accuracy. Journal of Educational Psychology, 2005, 97(4): 723—731