

两种实验范式下三段论推理格效应对比*

姜海霞¹, 杜国平²

(1. 曲阜师范大学马克思主义学院, 济宁 273165; 2. 中国社会科学院哲学研究所, 北京 100732)

摘要:三段论推理的格效应是指三段论的中项位置会影响三段论推理的难度以及结论端项的顺序。为对比结论生成范式和结论评价范式下三段论推理的格效应差异,并探究其他形式因素对推理难度的影响,采用48个三段论形式对275名中学生进行了实验。结果表明,在两种实验范式下,四个格的难度差异主效应均显著。结论生成范式比结论评价范式格的难度差异更显著。格效应并不能充分解释不同三段论的推理难度差异。在同一格内以及不同格之间不同三段论的推理难度受前提性质和顺序以及弱式和标准式的差异影响。研究给出了一个解释三段论推理难度差异的新分析框架,有助于推动解决三段论推理复杂性这一问题。

关键词:三段论推理;格效应;推理难度;结论评价范式;结论生成范式

中图分类号: B842.5

文献标识码: A

文章编号: 1003-5184(2024)01-0018-07

1 问题的提出

逻辑学的规范性研究其核心是揭示正确思维,特别是有效推理的应然状态,即揭示正确思维的基本方法和应该遵守的思维规则;逻辑学的描述性研究则着重探究主体自然语言有效沟通或合理思维中,特别是有效推理的实然状态,揭示正确思维的产生、发展、养成和提高了的认知规律。这两种类型的研究对于探究智能的本质和一般发展规律都具有不可或缺的重要价值。作为逻辑学最基本也是最核心的要素之一——三段论的研究也是如此。三段论指的是由两个包含着一个共同项的性质判断推出一个新的性质判断的推理。其中结论中的主项称为小项,用S表示,结论中的谓项称为大项,用P表示。在两个前提中都出现的共同词项叫“中项”,通常用M表示。含有小项的前提称为小前提,含有大项的前提称为大前提。根据中项在两个前提中的位置不同,三段论可以分为四个不同的格。其中,中项作为大前提的主项、小前提的谓项的为第一格,中项在两个前提中均作为谓项的为第二格,中项在两个前提中均作为主项的为第三格,中项作为大前提的谓项、小前提的主项的为第四格。将三段论的前提和结论排列组合可形成256种不同的形式。传统形式三段论中的直言命题分为四种:全称肯定命题(简称A命题)、全称否定命题(简称E命题)、特称肯定命题(简称I命题)和特称否定命题(简称O命题)。根据前提和结论中所涉及的性质命题的量和质,三段论又分为不同的式。通常用前提性质和格表示三段

论的式。比如AEO3式表示:第三格AEO式。其中,大前提为A命题,小前提为E命题,结论为O命题。

由于三段论形式多样,在日常生活中经常使用,且推理判定规则明晰,使其成为心理学领域采用最早的推理任务之一,可追溯到Störrings在1908年发表的关于关系推理和三段论推理的实验研究(Politzer, 2004)。多年来,三段论推理的实证研究关注的一个核心话题是三段论的推理复杂性问题,其中重要的分支议题之一是三段论的推理难度问题。在规范性研究中,基于主谓项非空的假设,三段论四个格共有24个有效的推理形式。在描述性研究中,可以发现,三段论四个格在人们的实际推理过程中,存在着难度差异(Johnson-Laird & Bara, 1984; Hardman & Payne, 1995; Politzer et al., 2017)。以气氛效应和概论启发式模型(PHM)为代表的表层特征理论(surface feature theories)和以心理模型、双重加工理论为代表的分析型理论(analytic theories)分别从三段论的表层特征和推理隐含的逻辑结构表征方面对三段论推理的前提解释、信息组合和加工过程进行了理论解释,但是尚无一个理论能够充分解释造成不同形式的三段论难度差异的原因是什么。

三段论的格效应(figural effects)最早由Laurence Frase于1968年提出。他指出,三段论的中项位置会影响三段论的推理难度。实验发现:三段论从第一格到第四格的难度逐步递增,不同格的三段

* 基金项目:国家高水平研究生公派留学基金项目(201804920060)。

通讯作者:姜海霞, E-mail: hjang253@126.com。

论推理难度存在显著差异(Frase,1968)。此后有一些研究得出了相同的结论(Roberge,1971;Erickson,1974)。Dickstein(1978)指出,Frase的实验并不能充分解释三段论格的难度差异。他从信息处理方向的角度解释四个格的难度差异,认为逆向处理信息(从大项到小项)比正向处理信息(从小项到大项)更难,这是第四格最难,第一格最简单的原因。由于第二格和第三格可以正向或者逆向处理信息,因此,两者的难度在第一格和第四格之间。1984年,Johnson-Laird和Bara指出,Frase只发现了部分三段论格效应,除了格的难度差异,三段论中项的位置差异还会造成结论端项顺序的偏差效应,这一效应主要体现在第一格和第四格上:第一格的三段论更容易得出S-P结论,而第四格的三段论更容易得出P-S结论(标准三段论中,第一格和第四格得出的结论都是S-P)。其他两格的结论没有显示出方向偏好。第一格和第四格的结论端项顺序偏差现象在多项研究中都得到了证实(Johnson-Laird & Bara,1984;Morley et al.,2004;Stuppel & Ball,2005)。

以上实验关于三段论四个格的难度差异得出了不同结论。有研究指出,实验范式会对结果产生影响(Dickstein,1978;Espino et al.,2000;Stuppel & Ball,2007)。比如,Espino等人(2000)认为,格效应仅出现在结论生成范式下的实验。三段论格效应的研究主要有两种实验范式:结论生成范式(conclusion production paradigm)和结论评价范式(conclusion evaluation paradigm)。结论生成范式要求被试根据给出的三段论的两个前提推断出正确结论,一般采用简答题的方式。结论评价范式则是给出完整的三段论形式,要求被试评价结论的有效性。结论评价范式的推理任务可以采用简答或者多项选择的方式。一般认为,两种范式的区别在于被试处理三段论前提和结论的顺序不同。三段论格效应的大部分实验研究都采用了结论生成范式,结论评价范式并没有得到重视(Markovits & Nantel,1989;Quayle & Ball,2000;Evans et al.,2001;Morley et al.,2004)。Morley等人(2004)指出,结论评价范式或许比结论生成范式下的格效应小,因为结论评价范式下的推理主要受结论驱动的信息处理过程的影响。Stuppel和Ball(2005)采用眼动实验对比了两种实验范式下第一格和第四格的认知负荷。研究表明,第一格的认知负荷高于第四格。特称命题的前提注视时间显著高于全称命题。Espino等人(2000)采用结论生成范式的实验得出了相同结果。Jia等人(2009)采用结论评价范式对第一格和第四格的认知加工时间进行了眼动实验,结果表明,第四格比第一格需要更多的认知加工时间。上述研究多数采用了结论评价

范式或者结论生成范式中的一种,研究对三段论四个格的难度得出了不同的结论。格效应是否受到实验范式的影响?在两种范式下格效应是否能充分解释推理的难度差异问题?这些问题仍尚待解决。由于第一格和第四格的端项顺序差异已经得到多项研究证实,此次研究将聚焦两种实验范式下三段论四个格的推理难度差异,系统探究格效应是否能对不同三段论形式的推理难度差异给出充分解释这一问题。具体研究问题包括:(1)在两种实验范式下,三段论四个格的难度差异是否显著?(2)在同一格内以及不同格之间前提的顺序、性质以及弱式和标准式等因素是否对推理难度有影响?

系统对比两种实验范式下三段论格效应的差异将进一步厘清已有研究关于第一格和第四格的难度差异问题的争议,检验实验范式对格效应的影响,深入分析格效应是否能够充分解释不同的三段论形式推理难度差异,并在此基础上分析造成格效应的原因是什么。其次,通过分析前提性质(全称/特称、肯定/否定)、顺序、弱式和标准式等对三段论推理难度的影响,研究希望给出一个解释三段论推理难度差异的新的分析框架,以推动三段论推理复杂性这一问题的理论研究。

2 研究方法

2.1 研究对象

实验共有275名初中学生参加。其中,初一学生100名(女生40人,男生60人),初三学生175名(女生84人,男生91人)。初一学生平均年龄13.02岁, $SD=0.57$ 。初三学生平均年龄14.85岁, $SD=0.55$ ($N=174$,有1名被试没有报告年龄)。将被试随机分为两组:一组($N=135$)作答结论生成范式的三段论题目,另外一组($N=140$)作答结论评价范式的三段论题目。有7名被试无法给出有效推理,删除这些无效作答。

2.2 实验设计

实验采用两因素完全随机实验设计,第一个因素是两种范式:结论评价范式和结论生成范式。第二个因素是三段论的四个格,因变量为三段论推理成绩。两因素均为被试间变量。将两个范式下的48个三段论各分为4组题目,每位被试作答随机分配的12道题目。

实验采用了22个有效三段论(其中包括5个弱式)和26个无效三段论。在四个格之间选取了前提相同的三段论,比如下表1中的第一行AA前提在第1格到第4格之间相同。同一格内,选取了前提相同但顺序不同的三段论,比如第一格的EIO式和IEO式。为了避免语料内容对推理的影响,两种范式下的同一个三段论采用相同的语料,前提性质

和顺序相同的三段论的语料相同。实验材料的话题涉及职业、动物、植物等生活中学生较为熟悉的中性话题。题目呈现顺序为随机的。下表 1 列出了实验选用的标准三段论形式。由于结论生成范式和结论评价范式都是给出了两个前提,分别让被试推出结论、让被试选择正确的结论,被试给出的可能不是标准形式。比如第一格的 AA 前提,有些被试可能会推出 AAI1 或者 AAA1,在统计时根据被试具体得出的结论确定三段论的式。这里为了更明晰地表达实验选取的标准形式,将完整形式呈现。

表 1 实验选用的三段论形式

第一格	第二格	第三格	第四格
AAA1	AAA2 *	AAA3 *	AAA4 *
AII1	AII2 *	AII3	AII4 *
IAI1 *	IAI2 *	IAI3	IAI4
AOO1 *	AOO2	AOO3 *	AOO4 *
OA01 *	OA02 *	OA03	OA04 *
OE01 *	OE02 *	OE03 *	OE04 *
EIO1	EIO2	EIO3	EIO4
IE01 *	IE02 *	IE03 *	IE04 *
(EA01)	(EA02)	EA03	EA04
AEO1 *	(AEO2)	AEO3 *	(AEO4)
EAE1	EAE2	EAE3 *	EAE4 *
(AAI1)	AAI2 *	AAI3	AAI4

注: * 为无效三段论,括号内的形式为弱式。

在结论生成范式下,被试需要根据给出的两个前提,推出一个正确的结论。为了告知被试如何推出正确结论,在指导语后给出两个三段论推理的例子(有效式 AEE2 和无效式 I11)。该范式的指导语如下:

同学你好!

这是一个测试人们如何进行推理的小测验。本测验一共有 12 道题目,每道题目都有两个前提,假设前提的内容都是正确的,你需要根据这两个前提推出一个正确的结论。如果根据两个前提无法推出正确的结论,请回答“无有效结论”,结论的形式可以从以下五种形式选择其中之一:

- ……都是/都有……
- 有些……是/有……
- ……都不是/都没有……
- 有些……不是/没有……
- 无有效结论

比如以下两个例子:

1. 前提:偶蹄目动物都是脊椎动物,
前提:角鲸都不是脊椎动物,
结论:因此,角鲸都不是偶蹄目动物。

2. 前提:有些逻辑学家是数学家,
前提:有些数学家是菲尔兹奖获得者,
结论:因此,无有效结论。

请在每道题下面的空白处简要写下你的分析思路。

根据给出的两个前提,有些被试会给出不同的结论从而形成不同的式,这种情况主要出现在第一格和第四格。在评分时根据被试实际得出的三段论来确定结论是否有效。比如,给出第一格 EA 前提,被试有可能得出 EAE1 式、EAO1 式、AEE4 式或者 AEO4 式四种有效式:

- 前提:六条腿的动物都不是小美喜欢的宠物,
前提:甲虫都是六条腿的动物,
结论 1:甲虫都不是小美喜欢的宠物。
结论 2:有些甲虫不是小美喜欢的宠物。
结论 3:小美喜欢的宠物都不是甲虫。
结论 4:小美喜欢的宠物有些不是甲虫。

Wason 和 Johnson - Laired (1972) 指出,与简答题一样,选择题也属于结论评价范式。一般的结论评价范式要求被试判断一个完整的三段论中的结论是否能由前提有效地推出。这种范式得到的关于被试的作答信息较少,为此实验在结论评价范式中采用了选择题的形式,要求被试在评价五个选项后选出正确的结论。选项由全称肯定命题、全称否定命题、特称肯定命题、特称否定命题以及“以上都推不出”组成。弱式仅有 3 个选项(特称肯定命题、特称否定命题和“以上都推不出”),原因在于如果给出所有的性质判断,被试可能会选择全称结论,从而测不出被试对弱式的判断。该实验范式的指导语如下:

同学你好!

这是一个测试人们如何进行推理的小测验。本测验一共 12 道题目,每个题目的题干都由两个前提组成,假设前提的内容都是正确的,请从选项中选出一个能够由前提推出的正确结论,并用圆圈圈出。注意:有些题目根据前提推不出正确的结论,请选择“以上都推不出”。请在每道题下面的空白处简要写下你的分析思路。

该实验范式的例子如下:

第二服装厂的员工都是女性,
这辆车上的乘客有些是第二服装厂的员工,
因此,_____。

- A. 这辆车上的乘客都是女性
- B. 这辆车上的乘客都不是女性
- C. 这辆车上的乘客有些是女性
- D. 这辆车上的乘客有些不是女性

* 弱式是指可以推出全称结论的,却只推出了特称结论。

E. 以上都推不出

3 实验结果

3.1 两种实验范式下三段论四个格的难度差异对比

对在四个格的作答数据进行方差分析,结果表明四个格的主效应显著, $F(3, 1594) = 14.23, P < 0.05, \eta_p^2 = 0.026$ 。LSD 事后检验结果表明,第一格与第二格的均值无显著差异 ($P = 0.984$),第一格的均值显著高于第三格 ($P < 0.05$) 和第四格 ($P < 0.05$)。第二格均值显著高于第三格 ($P < 0.05$)。第三格与第四格的均值无显著差异 ($P < 0.05$)。表 2 统计了结论生成范式下四个格的有效三段论、无效三段论以及总平均答对率。

表 2 结论生成范式下三段论推理的答对率

	有效三段论	无效三段论	总平均答对率
第一格	60.09%	21.15%	40.62%
第二格	52.68%	27.90%	40.29%
第三格	24.94%	35.32%	30.13%
第四格	37.16%	19.18%	28.17%
总平均答对率	43.72%	25.89%	34.80%

在结论生成范式下,分析被试的作答发现了以下现象:(1)在四个格中,当两个前提均为全称命题时,相较于特称结论,被试更倾向于得出全称结论。比如,第一格的 AA 前提,仅有 10.53% 的被试得出了特称结论。第二格的 EA 前提,100% 的被试得出了全称结论。第三格的 AA 前提,73.5% 的被试得出全称结论。第四格的 AE 前提,87.34% 的被试得出了全称结论。(2)被试推出的结论的主项和谓项的顺序受到前提中主项和谓项出现的顺序影响。其中最显著的是第四格的被试更倾向于推出 P-S 结论,从而使得三段论成为第一格。第二格和第三格也有此现象。比如,AE3 式的前提“水盆中的鱼都是老陈钓的,水盆中的鱼都不是鲫鱼”,有 81.82% 推出了“老陈钓的鱼都不是鲫鱼”这一结论。(3)除了第三格,其他几格的无效三段论答对率显著低于有效三段论。被试倾向于给无效三段论推出一个结论。(4)弱式的答对率低于标准式。

表 3 统计了结论评价范式下四个格的有效三段论、无效三段论以及总平均答对率。方差分析的结果显示,四个格的均值差异主效应显著, $F(3, 1705) = 5.922, P < 0.05, \eta_p^2 = 0.01$ 。LSD 事后检验结果表明,第一格与第四格的均值存在显著差异 ($p < 0.05$),与其他两格均值差异不显著。第二格与第三格均值不存在显著差异。第二格与第四格均值差异不显著。第三格与第四格均值存在显著差异 ($p < 0.05$)。

表 3 结论评价范式下三段论推理答对率

	有效三段论	无效三段论	总平均答对率
第一格	61.98%	17.21%	39.60%
第二格	52.60%	16.50%	34.55%
第三格	50.37%	18.97%	34.67%
第四格	34.80%	17.21%	26.01%
总平均答对率	49.94%	17.47%	33.71%

分析结论评价范式的作答,发现以下现象:(1)前提均为全称命题的有效三段论,被试更倾向于得出全称结论。比如,第一格的 AA 前提(44.44% 得出全称结论,33.33% 得出特称结论)、第四格的 EA 前提(50% 得出全称结论,38.89% 得出特称结论)。这一点在弱式上体现地更加明显。弱式 AAII 仅有两个特称选项和“以上都推不出”这三个选项,有 47.22% 的被试得出了特称肯定结论,有 44.44% 的被试认为推不出特称结论。(2)无效三段论的答对率显著低于有效三段论。(3)前提顺序影响推理答对率。

从上述分析结果可知,两种实验范式下的四个格均值差异的主效应均显著。结论生成范式比结论评价范式的效应量更大,差异更显著。这说明格对三段论推理难度差异具有较强的解释力。通过对被试作答结果分析发现,同一格内有些三段论存在较大的难度差异,不同格之间较容易的格的部分三段论难度高于较难的格的三段论。影响同一格内三段论形式差异的因素有前提性质和顺序、标准式和弱式。以下将具体分析这些因素对推理难度的影响。

3.2 两种实验范式下前提性质对推理难度的影响

前提性质是指三段论的两个前提的质(肯定/否定)和量(全称/特称)。表 4 列出了结论生成范式下前提性质相同的几组三段论的总平均答对率。

从总平均答对率来看,前提性质在四个格之间的差异不大,这与被试给出的结论有关,有些三段论可以在两个格之间互相转化。这与结论评价范式让被试评价固定的几个选项不同,结论生成范式下的被试可以得出不标准的有效三段论从而降低了四个格之间的总体难度差异。具体到每个格之间的难度差异,第一格中具有全称肯定前提 AA 和 EA/AE 前提的三段论答对率显著高于其他几个前提的三段论。第二格中的 EA/AE 和 AO/OA 前提的三段论答对率显著高于其他几个前提。第三格的 AI/IA 前提和 AO/OA 前提的三段论答对率显著高于其他。第四格的 AA 前提的三段论答对率显著高于其他。

表 4 结论生成范式下前提性质对推理难度的影响

	第一格	第二格	第三格	第四格	总平均答对率
AA	67.92%	13.69%	5.88%	71.23%	39.68%
EA/AE	53.77%	68.14%	5.94%	17.65%	36.38%
AI/IA	34.14%	38.86%	45.31%	20.86%	34.79%
AO/OA	16.54%	50.85%	43.94%	25.18%	34.13%
EI/IE	40.00%	31.33%	23.66%	25.90%	30.22%
总平均答对率	42.47%	40.57%	24.95%	32.16%	35.04%

表 5 统计了结论评价范式下前提性质相同的几组三段论的总平均答对率。从总平均答对率来看,具有两个全称前提的 AA 式和 EA 式的总平均答对率高于其他特称和全称组合的前提。在前提均为全称的三段论中,AA 式比 EA 式简单。在前提为全称和特称组合的三段论中,具有两个肯定命题 AI 和 IA 的三段论总平均答对率高于具有否定前提的三段论。这说明肯定前提比否定前提简单,全称前提

比特称前提简单。具体到每个格的难度差异,第一格中的 AA 式和 EA/AE 式的答对率显著高于其他几个前提。第二格的 EA/AE 和 EI/IE 式的答对率显著高于其他几个前提。第三格中的 AA 和 AI/IA 式的答对率显著高于其他几个前提。第四格的 AA 式答对率显著高于其他几个前提。这可能与前提中的主项和谓项的位置有关,经分析,被试更倾向于将前提中的主项和谓项作为结论的主项和谓项。

表 5 结论评价范式下前提性质对推理难度的影响

	第一格	第二格	第三格	第四格	总平均答对率
AA	66.04%	14.81%	60.52%	57.90%	49.82%
EA/AE	58.31%	46.72%	32.34%	23.41%	40.20%
AI/IA	35.71%	13.89%	60.98%	29.78%	35.09%
AO/OA	2.78%	39.76%	24.63%	19.07%	21.56%
EI/IE	31.94%	45.15%	36.43%	18.69%	33.05%
总平均答对率	38.96%	32.07%	42.98%	29.77%	32.61%

总之,通过以上分析,在两种实验范式下,前提性质对推理难度的影响主要表现为:在同一格内,第一格具有全称肯定前提的三段论、第二格 EA/AE 前提的三段论和第三格 AI/IA 前提的三段论难度显著低于其他前提的三段论。在不同格之间,不同的前提性质对推理难度也有较为显著的影响。比如,在结论生成范式下,AA 前提在第一格和第四格的难度显著低于其他两格。这与中项在前提中的位置以及前提中的主项和谓项出现的顺序有关。比如,第一格 AAA1 式的前提:所有的玫瑰花都是杂交品种,所有的盆栽植物都是玫瑰花,中项“玫瑰花”将两个前提连接起来,被试可以从“所有的盆栽植物都是玫瑰花,所有玫瑰花都是杂交品种”这一顺序很简单地推出“所有盆栽植物都是杂交品种”。AA 前提第二格和第三格的难度显著高于其他两格,原因在于第二格和第三格的中项分别位于前提的谓项和主项,被试很难通过中项将两个前提连接。在同一格内,三段论的难度差异受到前提性质的影响,比如 AA1 式和 EI1 式,这体现了量词的复杂性对推理难度的影响。

3.3 两种实验范式下前提顺序对推理难度的影响

前提顺序是指三段论两个前提的顺序关系,标准三段论一般将大前提放在第一句,小前提放在第二句。表 6 列出了两种实验范式下前提顺序不同的

三段论的总平均答对率。

表 6 两种实验范式下前提顺序不同的三段论总平均答对率

前提	结论生成范式	结论评价范式
AI	41.53%	37.93%
IA	28.05%	32.25%
AO	38.11%	22.22%
OA	30.15%	20.09%
EI	21.04%	41.70%
IE	39.41%	24.40%
AE	31.66%	25.01%
EA	38.73%	47.79%

从表 6 可以看出,在结论生成范式下的 AI 和 IA、EI 和 IE 前提顺序对推理难度有较大影响。另外几个前提顺序不同的形式对推理难度没有显著差异。在结论评价范式下的某些前提组合中,前提顺序对推理难度有着较为显著的影响。比如 AE 和 EA、EI 和 IE 前提。在 AO 和 OA 前提、AI 和 IA 前提的三段论中,前提顺序对推理难度的影响不大。

3.4 两种实验范式下弱式和标准式的难度对比

三段论的有效式中共有 5 个弱式,分别为 AA11、EA01、AE02、EA02 和 AE04。表 7 对比了弱式与其对应的标准式之间的难度差异。由于将 AEE2 作为指导语的示例,在结论评价范式下没有使用 AEE2 和 AEE4。在结论生成范式下给出了

AE2 前提和 AE4 前提,这里只统计结论生成范式下两者的弱式和标准式的答对率对比。

表 7 结论评价范式下弱式和标准式的难度对比

三段论	结论评价范式 总平均答对率	结论生成范式 总平均答对率
AAA1	84.85%	63.64%
AAI1	47.22%	4.41%
EAE1	88.57%	66.15%
EA01	69.70%	1.54%
EAE2	81.82%	43.94%
EA02	27.78%	0
AE4	/	52.94% (EAE1)
AE04	28.57%	0
AE2	/	41.18%
AE02	30.56%	0

从表 7 可以看出,结论评价范式下,弱式的答对率显著低于标准式,在具有两个全称前提的三段论中,被试倾向于给出全称结论。在结论生成范式下,被试一般会根据给出的全称前提得出不固定的结论,这里根据 AA1、EA1、EA2、AE2 和 AE4 这个几个前提推出的特称和全称结论占比分析其平均答对率。结论生成范式下标准式的答对率显著高于弱式。其中 AE4 式所有有效结论中被试都给出了 P-S 结论,使得 AE4 式转化为 EAE1 式。综上,在两种范式下弱式的答对率都显著低于标准式。

4 讨论与展望

研究系统对比了两种实验范式下三段论四个格的难度差异,并在此基础上分析了造成同一个格内和不同格之间三段论推理难度差异的原因。实验结果发现,在两种范式下均存在较为显著的格效应,结论生成范式比结论评价范式的格效应更显著。两种范式下,第一格的推理难度均显著低于第四格。研究发现,不同三段论的难度差异除了受到格效应的影响,还受其他形式因素的影响。因此,格效应并不能充分解释不同三段论之间的难度差异。在同一格内以及不同格之间,三段论的推理难度差异受到前提性质、顺序、前提中主项和谓项位置以及弱式和标准式的影响。对比前提性质对推理难度的影响发现:在同一格内,第一格具有全称肯定前提的三段论、第二格 EA/AE 前提的三段论和第三格 AI/IA 前提的三段论难度显著低于其他前提的三段论。在不同格之间,不同的前提性质对推理难度也有较为显著的影响。一般而言,全称比特称更为简单,肯定比否定更为简单。这与中项在前提中的位置、前提中的主项和谓项出现的顺序以及量词复杂性有关。对比前提顺序对推理难度的影响发现:两种范式下,前提顺序对部分前提的三段论推理难度影响较大。对比弱式和标准式发现,两种实验范式下弱式的难度

均显著高于标准式,被试倾向于对前提均为全称命题的三段论推出一个全称结论。研究还发现,在两种范式下,有效三段论的答对率显著高于无效三段论。

格效应反映了被试在推理过程中处理前提信息和结论信息的信息加工机制。研究仅在表层结构上分析了影响三段论推理难度的差异。对于不同三段论之间的深层加工机制以及认知负荷差异,需要更为精密的仪器去系统测量,比如采用眼动仪测量被试在不同三段论上的注视时间、注视轨迹以及推理方向等以获得更为精确和丰富的研究数据。两种实验范式反映了两种推理方向,从前提到结论的正向推理(forward processing)与从结论到前提的逆向推理(backward processing)是否分别发生在结论生成范式和结论评价范式中仍需要检验。两种实验范式中,被试是否会采用不同的推理策略?推理策略对于推理难度的影响如何?这些问题可以在今后的研究中进一步探讨。根据研究问题和设计,研究仅选取了 48 个三段论形式,今后的研究可以测试更多的三段论形式,从而将研究结论推广到更多的三段论。为了获得更多的作答信息,两种实验范式分别采用了简答题的结论生成范式和选择题的结论评价范式。同时,限于研究问题仅对比推理难度差异,不对比结论端项顺序差异,结论评价范式将结论限制为标准三段论的结论,今后的研究根据研究设计和研究问题可设置更多选项,以测试被试对于结论端项顺序的选择。结论生成范式采用了开放的简答题形式,这使得某些实验预想的三段论在一定情况下会转化成其他格的三段论,今后的研究根据研究设计可采用填空题等形式对结论端项的顺序进行限制。此外,已有三段论的推理研究更多关注有效三段论,对无效三段论缺乏关注,无效三段论的答对率为什么显著低于有效三段论?今后的研究可以更多地探究无效三段论的推理复杂性问题。

综上,研究通过对比两种实验范式下的三段论推理差异,检验了实验范式对格效应的影响,厘清了已有研究关于第一格和第四格的难度差异问题的争议。在此基础上,得出了格效应并不能充分解释不同的三段论形式推理难度差异这一结论。通过分析前提性质(全称/特称、肯定/否定)、顺序、弱式和标准式等对三段论推理难度的影响,研究给出了一个解释三段论推理难度差异新的分析框架,这有助于推动解决三段论推理复杂性这一问题。

参考文献

Dickstein, L. S. (1978). The effect of figure on syllogistic reasoning. *Memory & Cognition*, 6, 76-83.
Erickson, J. R. (1974). A set analysis theory of behavior in for-

- mal syllogistic reasoning tasks. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola Symposium* (pp. 78–90). John Wiley & Sons.
- Espino, O., Santamaria, C., & Garcia – Madruga, J. A. (2000). Figure and difficulty in syllogistic reasoning. *Current Psychology of Cognition*, 19, 417–428.
- Espino, O., Santamaria, C., Meseguer, E., & Carreias, M. (2005). Early and Late Processes in Syllogistic Reasoning: Evidence for Eye – movements. *Cognition*, 98, B1–B9.
- Evans, J. St. B. T., Handley, S. J., & Harper, C. N. J. (2001). Necessity, possibility, and belief: A study of syllogistic reasoning. *Quarterly Journal of Experimental Psychology*, 54A, 935–958.
- Frase, L. T. (1968). Associative Factors in Syllogistic Reasoning. *Journal of Experimental Psychology*, 76, 407–412.
- Hardman, D., & Payne, S. (1995). Problem difficulty and response format in syllogistic reasoning. *Quarterly Journal of Experimental Psychology*, 48A, 945–975.
- Jia, X. Q., Lu, S. F., Zhong, N., & Yao, Y. Y. (2009). *Figural effects in syllogistic reasoning with evaluation paradigm: An eye – movement study*. International Conference on Brain Informatics.
- Johnson – Laird, P. N., & Steedman, M. J. (1978). The psychology of syllogisms. *Cognitive Psychology*, 10, 64–99.
- Johnson – Laird, P. N., & Bara, B. (1984). Syllogistic inference. *Cognition*, 16, 1–61.
- Markovits, H., & Nantel, G. (1989). The belief bias effect in the production and evaluation of logical conclusions. *Memory and Cognition*, 17, 11–17.
- Morley, N. J., Evans, J. St. B. T., & Handley, S. J. (2004). Belief bias and figural bias in syllogistic reasoning. *Quarterly Journal of Experimental Psychology*, 57A, 666–692.
- Politzer, G. (2004). Some precursors of current theories of syllogistic reasoning. In K. Manktelow & M. – C. Chung (Eds.), *Psychology of reasoning. Theoretical and historical perspectives* (pp. 214–240). Hove: Psychology Press.
- Politzer, G., Christelle, B., & Emmanuel, S. (2017). Preadolescents solve natural syllogisms proficiently. *Cognitive Science*, 41(5), 1031–1061.
- Quayle, J. D., & Ball, L. J. (2000). Working memory, metacognitive uncertainty, and belief bias in syllogistic reasoning. *Quarterly Journal of Experimental Psychology*, 53A, 1202–1223.
- Roberge, J. (1971). Further examination of mediated associations in deductive reasoning. *Journal of Experimental Psychology*, 87, 127–129.
- Stuppel, E. J. N., & Ball, L. J. (2005). An inspection – time analysis of figural effects and processing direction in syllogistic reasoning. Proceedings of the 27th annual meeting of the cognitive science society. *Sheridan Printing, Alpha, New Jersey*, 27, 2092–2097.
- Stuppel, E. J. N., & Ball, L. J. (2007). Figural Effects in a Syllogistic Evaluation Paradigm. *Experimental Psychology*, 54(2), 120–127.
- Wason, P. C., & Johnson – Laird, P. N. (1972). *Psychology of reasoning: Structure and content*. Cambridge: Harvard University Press.

Comparison of the Figural Effect of Syllogistic Reasoning between Two Experimental Paradigms

Jiang Haixia¹, Du Guoping²

(1. School of Marxism, Qufu Normal University, Jining 273165;

2. Institute of Philosophy, Chinese Academy of Social Sciences, Beijing 100732)

Abstract: The figural effect of syllogistic reasoning refers to the effect that the position of the middle term in syllogisms influences the direction of end terms in the conclusion and the difficulty of the syllogism. To understand the difference of the figural effect of syllogistic reasoning tasks between the conclusion production paradigm and the conclusion evaluation paradigm and other factors influencing the difficulty of reasoning, an experiment was conducted on 275 middle school students using 48 syllogisms. The results showed that under both experimental paradigms, the main effect of difficulty differences in the four figures was significant. The difference in difficulty of the conclusion production paradigm is more significant than the conclusion evaluation paradigm. The figural effect cannot fully explain the differences in reasoning difficulty among different syllogisms. The difficulty of reasoning different syllogisms within the same figure and between different figures is influenced by the properties and order of premises, as well as the differences between weak and standard forms. The study provides a new analytical framework to explain the differences in the difficulty of syllogistic reasoning, which helps to promote the resolution of the complexity of syllogistic reasoning.

Key words: syllogistic reasoning; figural effects; the difficulty of reasoning; conclusion evaluation paradigm; conclusion production paradigm