

心理理论发展研究的一种新范式——矩阵博弈*

李晶^{1,2,3} 刘希平¹

(¹天津师范大学教育学院, 天津 300387) (²中国科学院心理研究所, 北京 100101)

(³中国科学院研究生院, 北京 100049)

摘要 心理理论指对他人心理状态的理解, 是一种毕生发展的能力。但之前的研究在研究对象和研究内容上存在一些缺陷。本文介绍了一种新范式——矩阵博弈(matrix game), 对利用其研究心理理论进行了可行性分析, 同时指出了两者相结合的理论基础和实证证据。未来研究可利用矩阵博弈来研究更广年龄范围人群心理理论的发展轨迹, 并借助 fMRI 等技术来测查矩阵博弈中心理理论推理的脑机制。

关键词 心理理论; 矩阵博弈; 决策; 思维

分类号 B844

1 心理理论经典研究范式的局限

心理理论(Theory of Mind)是个体对他人心理状态的认识。它是人们关于心理领域的内隐观念, 由一个抽象且连贯的因果解释系统构成, 使个体能够借助诸如信念、愿望、情绪等无法观测的心理状态来解释和预测行为。我们只能从个体在自然情境和实验情境中的言语和行为来推断他们的心理理论水平。心理理论研究是继皮亚杰关于儿童认知发展的研究和元认知研究之后, 又一个探讨儿童心理表征和心理理解的崭新角度(刘希平, 唐卫海, 方格, 2005)。

到目前为止, 研究心理理论的范式有许多种。目前, 探索儿童认识“错误信念”常用的研究方法有(罗杰, 卿素兰, 2005): (1)意外转移任务(unexpected transfer)。最经典的是“Maxi 和巧克力的故事”和“女孩和书的故事”。(2)欺骗外表任务(deceptive-appearance)范式, 经典的有“糖果盒”任务。(3)二级错误信念, 经典的任务是“约翰和玛丽的故事”。(4)失言识别任务。Baron-Cohen 及其同事首先提出了“失言识别任务(faux pas recognition)”的范式, 并用来测量 7~11 岁儿童的心理理论水平, 初步发现“失言识别任务”范式能够有效测量儿童的心理理论水平(Baron-Cohen,

O’Riordan, Stone, Jones & Plaisted, 1999; 刘希平, 安晓娟, 2010)。

上述研究范式虽然在一定程度上确实可以用来推断儿童的心理理论理解能力, 但是讲故事的方法使得结果较为主观, 易出现主试者效应, 此外, 在测试中需要儿童进行大量的言语加工, 因而就很难区分影响儿童表现的究竟是儿童的语言能力, 还是儿童对心理状态的理解能力(吴南和张丽锦, 2007; Flobbe, Verbrugge, Hendriks & Krämer, 2008)。因而使得实验结果反映的心理理论水平与真实具有的水平之间可能会存在一定的差距。一种可能的情况是理解另一个人一定级别信念和意图的能力本身存在, 但是在实验任务中为了能够将这种能力表现出来, 个体还必须还具备一些其他的能力, 如对情景知识的理解, 语言理解和表达能力等, 才能够对实验问题做出相应的反应。在讲故事的实验任务中, 不同的实验故事需要具备的情景知识有所不同, 而且儿童在理解和做出实验反应上都高度依赖于语言能力的发展, 这使得通过讲故事实验任务测查到的心理理论能力与其真实具备的心理理论能力之间可能存在一定的差距。同时呈现的刺激材料是静止的, 这与日常社会交往中的动态特征是相违背的, 因而这个任务不能模拟个体面对真实情景中所需要的认知需求(Chan, Gloria, Wong, Hui, Tang, Chan, Lam, Chiu, & Chen., 2010)。因此讲故事的研究范式具有一定的局限性。为此选用情景知识

收稿日期: 2010-4-22

* 教育部人文社会科学研究项目基金资助(09YJAXLX021)。

通讯作者: 刘希平, E-mail: lucy0579@yahoo.com.cn

较少,且较少依赖语言能力,呈现方式更客观,反应方式更简单的实验范式是非常有必要的。

此外,采用经典的错误信念任务范式对学前儿童心理理论进行探讨的研究结果大都表明大多数儿童在 4 岁左右就可以顺利通过心理理论任务。而李晓东、黄艳秋、刘萍和徐健(2007)用成人版的错误信念任务(意外转移任务)考察成人的表现,发现其在错误信念任务上的成绩并不理想,这与成人具有心理理论的事实相悖,可以认为将通过错误信念任务作为衡量心理理论的唯一评定标准是有待商榷的。目前探讨成人心理理论的任务范式主要有:欺骗任务、失言任务、隐喻任务、二级错误信念任务等(大都属于讲故事情境)。但就目前成人心理理论的实验任务来分析,并不能明确其与儿童范式的区别,已经发展较成熟的心理理论任务多数基于儿童个体,直接将考察儿童的任务用于成年人,可能会出现天花板效应。但是,使用与儿童期相同的任务,研究的结果仅在量的水平上对比不同个体心理理论水平的高低,似乎又默认了这样一个假设:在认识和解释心理状态的机制上,成人与儿童相比,仅有量变形式的发展。因此,真正适用于成人的实验任务有待探索,为揭示心理理论的毕生发展轨迹提供理论依据(林佳燕,傅根跃,刘文庆,2010)。

同时,心理理论水平不同的人,对他人主观世界猜测的层次不同,就如同下棋,有的棋手可以事先看出 3 步,有的棋手则可以看出 5 步,与人交往时对对方思维状况的猜测也会有这样的差异,这种差异在传统心理理论研究领域内也没有很好的办法进行测量。

尤其重要的是,对他人主观世界的认识不应该仅仅局限在与社会性交往相关联的范畴,在日常生活的各个领域,在政治、经济、军事、法律、国际关系和外交、环保、体育竞技等所有涉及到人与人之间的对抗与合作的领域,都涉及到对对方思维活动的判断,而传统的心理理论研究因为更多关注信念、愿望及情绪的内容,使得其研究领域、研究范式、研究的贡献等都被局限在与社会性成长相关的范畴,而忽视了与智慧活动相关的内容。

鉴于目前研究存在的不足,寻找一种更为客观的且适用于更广年龄范围的任务范式就显得尤为重要。行为决策中的博弈情景为心理理论的

进一步研究提供了可能。

2 决策中的博弈范式

决策博弈(Game of Decision Making)是经济学领域的研究手段之一(Kahneman & Tversky, 1982; Hedden & Zhang, 2002; Camerer, 2003)。在决策的过程中,每一个参与博弈的人的决策都会影响到其他参与者的选择(独裁者博弈除外,其中反应者只能无条件地接受提议者的提议而不能做出任何其它的选择)。因此,个体为了保证自己的利益最优化,就需要对他人的心理活动进行猜测,根据他人可能的决策来选择和调整自己的决策方式。这种猜测是多重交互的,最终获益最多的人应该是能够进行更多重猜测、并且对对手的水平进行准确判断的人。这种决策游戏可以通过直接的认知测量推断思考的等级,从而说明心理理论推理的等级。

决策游戏中比较经典的是选美竞赛(P-Beauty Contest)。它最初由 Keynes 于 1936 在思维实验中提到,后来研究者对它进行了系统的介绍(Camerer, 2003)。选美竞赛中博弈的目标是使博弈者自己的利益最大化,其过程如下:每一个博弈的小组一共有 n 个人($n \geq 2$);给被试提供一个数字空间,假设数字的大小代表价值的多少,例如 0~100;然后请被试在 0~100 的数字间给自己选择一个数字;大家选择之后,其选择数字的平均值就可以计算出来,而选择的数字最接近平均值 P 倍的博弈者获胜,赢得一个固定的已知奖励,失败者什么也得不到。在博弈开始之前告诉所有参与者 P 值的大小(Ohtsubo & Rapoport, 2006)。以数字范围为 0~100, P 为 0.7 为例,研究的假设是:如果参与者具有一级推理水平,则他应该可以舍弃 70 以上的所有数字,因为,即使所有人都选择 100,其平均值的 0.7 倍最多是 70,因此要是选择 70 以上的数字,就少有机会成为最接近平均值的 0.7 倍的人。所以对于有一级推理能力的人来讲,70 以上的数字是永远不会选的。如果参与者能够预见到其他对手也会有这样的推断,其他对手也都不选 70 以上的数字,那么其平均值的 0.7 倍最大就是 49,也就是说,一个具有二级推理水平的博弈者就会排除所有大于 49 的数字。而如果一个博弈者推断对手的推理水平是二级的话,他就可以预见到,对手的选择也会在 49 以下。因此其选择数据的平均值最大也就

是 $49 \times 0.7 = 34.3$, 即其选择数值的范围就会在 0~35 之间。根据理性人的假设, 这个思考过程必将一直持续下去, 其最后的纳什均衡值为 0, 但是在不同的被试群体中发现的结果却是人们选择的平均数字范围为 25~40, 有很大的标准差, 且这个结果在不同的国家及不同的年龄间都是非常稳定的(Camerer, 2003)。Camerer, Ho 和 Chong (2004)提出了泊松认知层次模型(Poisson Cognitive Hierarchy Model, 具体请参见理论基础部分的阐述)来解释多项实验结果, 以弥补纳什均衡理论在解释具体实验数据上的不足和无力。除此之外在投资博弈(Investment Game)和囚徒困境(Prison Dilemma)等决策博弈中博弈者要想自己获得最大的利益就必须对对方的心理状态进行一定的推测, 从而在这种交互决策过程中可以测量博弈者的心理理论推理等级。

总之, 博弈结构是用来研究人们思考过程的一个很有价值的方法。即使是年龄较小的儿童(4岁)也能很好地理解博弈的规则, 而且能够注意力集中地进行博弈(Perner, 1979)。同时博弈也是一个实验控制灵敏的工具(Ohtsubo et al. 2006), 呈现方式和反应方式都较讲故事方法客观, 对语言能力的依赖较少。从上面的例子中我们得知在博弈结构中通过直接的认知测量可以推断思考的等级, 因而利用这样较为客观的方式来探讨心理理论思维成分的发展情况是可能的。此外, 在博弈中也可以探讨如公平、合作、利他、信任等社会行为, 这样也可以进一步了解这些社会行为与心理理论之间的关系(Takagishi, Kameshima, Schug, Koizumi, & Yamagishi, 2010; Lucas, Wagner, & Chow, 2008; Paal & Bereczkei, 2007)。

矩阵博弈是在决策研究中经常使用的方法, 这种矩阵博弈一般主要是由三部分组成的, 包括博弈者, 所有可能的结果和各个博弈者在所有可能结果下得到的奖励值, 这样后两个部分就会组成一个或多个矩阵博弈。每个博弈者的目标是做出使自己利益最大化的决定(Ohtsubo et al. 2006)。这种形式尤其适用于两个玩家的博弈, 其中每一个博弈者都只能有一次移动, 而且其中博弈者同时并且独立地选择他们的移动方式。一个博弈者的策略以矩阵中的排呈现, 同时另一个博弈者的策略以列来呈现。矩阵中的每一个单元格都列出了当博弈结束时双方的奖励值。以矩阵的

大小来描述博弈: 一个 2×2 的博弈有两个博弈者, 且每个博弈者在两种可能的移动方式中进行选择。博弈的扩展形式可以表征为一棵树, 其中每一个节点代表博弈的一种可能的状态, 博弈在原始的节点上开始。每一个节点“从属于”一个固定的博弈者, 他在此节点上可能的移动方式间进行选择。当达到终止节点时博弈就结束了, 并且博弈者会各自得到终止节点上的奖励。当博弈者进行序列移动时, 博弈的扩展形式是很有用的, 博弈者在做出自己移动方式的选择之前能够了解到另一个博弈者的行动方案(Flobbe et al. 2008)。

在博弈中, 博弈者通过猜测对手会做什么, 从而帮助自己做出最好的选择, 但是这种猜测是交互的, 因为对手也会猜你将会做什么, 直到交互的一致性反应达到“一个平衡”为止, 此次博弈才会结束。这种推理是重复的(A 通过猜 B 将会猜 A 做什么来猜 B 将会做什么, 这样一直持续下去), 这样通过直接的认知测量可以推断思考的等级(Hedden et al. 2002)。在策略博弈中, 博弈者 I 会存在一个主导性策略, 即无论他的对手选择什么样的策略, 他的奖励都是较大的。由于对手也具有理性的, 他也会推测博弈者 I 的选择, 所以这个主导性策略永远不会是博弈者 I 最好的反应, 他将永远不会使用一个主导性策略, 而且如果博弈者 I 是理性的, 而且相信对手也是理性的, 那么就会出现这样一种状况, 即不仅博弈者 I 不会使用主导性策略, 而且其他的博弈者也将不会使用他们自己的主导性策略, 这样决策的过程就是一个重复排除主导性策略的过程(Camerer, 2003)。这个过程一共有两步, 第一步要求一个博弈者有充分的能力能够根据不同的选择去比较不同的结果, 第二步要求博弈者考虑其他博弈者也类似地会出现自己第一步所进行的推理过程。对于一个博弈者来说, 为了顺利进行第二步的推理, 他必须能够从对手的角度去考虑并且理解对手是如何看待这个博弈结构的。这样从博弈者之间进行的交互决策过程中我们就可以测量他们进行推理的思维等级。可见, 利用对博弈情境中决策过程的分析, 我们就可以获得个体内在思考等级情况的理解, 从而可以用博弈结构来探讨认知中的心理理论发展情况。

基于以上的分析, 我们看出在博弈推理的过程中可以很好地考察交互思考的级别。

3 利用矩阵博弈探讨心理理论推理情况的可行性分析

3.1 矩阵博弈中研究心理理论的优势

在构建的矩阵博弈中研究心理理论推理水平有诸多优势(Camerer & Fehr, 2002)。首先,它比传统研究心理理论的讲故事范式在呈现和反应方式上更加客观,在一定程度上减少了对语言能力的依赖,提供了对内在心理理论推理过程的一种较为准确的测量方式,同时也将研究对象扩展到一个更为广的范围。其次,与错误信念任务相比,博弈是非常有趣的,因为它们都是应用型的任务。参加者能够意识到使用心理理论会带来一定的好处,但是他们并不是明显地被要求使用心理理论,从而为更加真实地调查实际现实生活中心理理论的发展情况提供了可能,同时在博弈这种应用型任务中也可以用来调查成人的心理理论推理情况,因为 Keysar, Lin 和 Barr (2003)的研究显示个体在应用型任务上的表现是非常差的,从而可以避免成人的表现出现天花板效应。第三,博弈比故事任务会有更大的多样性和重复性,从而可以操纵更多的项目,而且可以测量到个体表现间更多的变异性。第四,这种范式所探讨的思维策略推理情况将心理理论的研究内容扩展到思维领域,从而使心理理论的研究领域更为广泛。总体而言,如果年幼儿童能够成功地使用博弈矩阵(Perner, 1979),那么它将是一个比过去的方法更严格,而且能够定量测量心理理论发展和策略推理的手段。这样就可以使用这个范式来更为客观地研究心理模型的发展情况。

3.2 矩阵博弈中研究心理理论的理论基础

使用矩阵博弈来研究心理理论具有一定的理论基础。Camerer 等人(2004)提出了泊松认知层次模型。在此模型中,每一个博弈者都相信他能比另一个博弈者更好地理解博弈情境,并且假定博弈者能再现策略思维过程的决定规则。认知层次模型由以下两部分组成,一个是做 k 阶段思考的博弈者可以重复的决定规则,另一个是 k 阶段博弈者的频率分布 $f(k)$ (假设是泊松分布)。可重复性过程开始于“0 阶段型”,博弈者并没有对他们的对手做出任何假设,而且仅根据某种可能性分布来进行选择(Camerer 等假设可能基于简单性选择)。“ k 阶段”思考者根据标准的泊松分布认为对手分布在 0 阶段到 $k-1$ 阶段之间,即他们能够

准确地预测到比他们思考等级低的博弈者的相对频率,但是却忽略了一些博弈者可能做与他们一样多或更多阶段思考的可能性(Camerer et al, 2004)。总之,在认知层次模型中,博弈者使用了一系列的回溯思考,假设较高等级的博弈者能够根据他们知觉到的其他思考者将会做什么的分布情况而选择使自己利益最大化的策略。例如在选美竞赛博弈中,做 0 级思考的博弈者在所有的数字上平等的随机化,1 级博弈者认为他们与 0 级博弈者做博弈,所以他们预期平均值是 50,并会选择 50 的 $2/3$ (p 值),即 33。2 级博弈者认为他们是与 0 级和 1 级博弈者做博弈。如果 2 级博弈者正确猜测到较低等级思考者的相对比率,他们的最好反应是 26,而且由于工作记忆的限制,人们很难达到更高水平的推理等级。认知层次模型对博弈结构的预测与真实的实验结果相一致(Camerer, 2003)。Camerer et al.(2004)对泊松认知层次模型和纳什模型的匹配性进行了分析,发现泊松认知层次模型本质上比纳什模型能够更好地匹配实验性数据,当纳什预测是好的接近值时,认知层次模型也能做出较为准确的预测,而当纳什预测值是比较差的接近值时,认知层次模型也能做出准确的预测。

3.3 矩阵博弈中研究心理理论的实证支持

Perner (1979)曾使用 2×2 矩阵博弈来研究儿童对自己和他人高奖励的关注情况,对儿童使用的策略进行了分析。虽然文章并没有明显地讨论心理理论推理的级别,但是这个 2×2 的矩阵博弈结构也可以理解为是一个心理博弈。博弈的表征形式为 4 个单元格(2×2),其中每一个单元格包含了每个博弈者的奖励。儿童和对手(一个成人研究者)各自独立地选择一个列或排。在他们公布选择后,他们所选的交叉的那个单元格决定了两个博弈者的奖励值。当儿童扮演列或排角色时,它们使用的策略以及反应的心理理论推理情况是不一样的。列博弈者拥有一个主导性策略,这个博弈者能在不需要考虑对手行动的情况下找到自己利益最大化的策略。而排博弈者没有主导性策略,而且只能通过预测列博弈者将会怎么做而找到自己利益最大化的策略。实验中所有的儿童既扮演列也扮演排的角色,要求一半的儿童在选择自己的策略之前先预测对手的选择而另一半儿童在选择自己的策略之后再预测。

Perner 发现相较于预测对手将会选择他们的主导性策略来说,扮演列角色的儿童能够更成功地选择他们自己的主导性策略。在这个博弈中,当提问儿童列博弈者将会怎么做时需要一级推理能力和当提问儿童排博弈者将会怎么做时需要二级推理能力。在最年幼的 4~6 岁组中只有 50% 的参加者能够正确预测,这与机会水平相一致。稍大一点的儿童能够做出正确的预测:当扮演排角色时,74% 的儿童能够做出正确的预测。然而当扮演列角色时正确预测的比例接近 50%。Perner 认为儿童对他们对手的观点并不感兴趣,它对他们并没有帮助,因为作为列博弈者拥有主导性策略,不需要预测的情况下就可以找到。然而,当作为列博弈者进行预测时需要的是二级推理而不是一级推理。因此,研究者假设二级推理的难度也使得预测分数较低。这个研究初步证实了利用矩阵博弈结构来探讨心理理论的发展情况是可能的。

Stahl 和 Wilson (1995)采用一对多同时决策的方式考察了人们对其他博弈者推理深度的假设。他们将人们的心理理论水平分为了 0 级水平、1 级水平、2 级水平、朴素纳什型(naïve Nash-type)、全面型(worldly)和理性预期。根据他们的定义,0 级水平的人随机选取一个有效的策略;1 级水平的人假设所有的其他博弈者都是 0 水平,且根据这一假设而选择一个最好的策略;2 级水平的人假设博弈者是由 1 级水平和 0 级水平的人组成的;朴素纳什型的人相信其他的博弈者按照理论进行博弈,呈现纳什平衡;全面型的人假设人们是由 0 级水平、1 级水平、2 级水平和朴素纳什型的人组成的;理性预期型的人相对于前面几种有限理性类型而言的,他们具有无限理性。他们在实验中对 48 个大学生参加者使用了 12 个 3×3 的矩阵博弈进行了研究,一共有两个实验阶段。最后结果发现 48 个被试中 8 个是 0 级水平,9 个是 1 级水平,1 个是 2 级水平,7 个是朴素纳什型,19 个是全面型。这个结果说明在所有的 12 个博弈上,大多数被试都可以看作是使用一种思维过程,或者是使用一种关于其他博弈者的模型。在剩余的 4 个被试中,3 个大致相当于朴素纳什型或全面型,剩余的一个被试最有可能是 1 级水平类型,但是精确性较低。所有的被试中并没有人是属于理性预期型。这个结果一方面说

明在矩阵博弈中探讨心理理论推理的等级是可能的,另一方面也说明在交互决策中人们大都是有限理性的,无限理性的人是不存在的。

Ohtsubo (2002)的研究证实选美竞赛博弈的选择与心理理论能力测量的成绩之间存在正相关。在他的研究中,选美竞赛博弈的参数值是 p 为 0.50,间隔是[0, 100],参加博弈的人数在 8 到 16 之间,这样 1 级水平的博弈者应该选择小于 50 的数字,而 2 级水平的博弈者应选择小于 25 的数字。结果显示选择低于 25 数字的参加者在 Imposing Memory Task (IMT)测验上的成绩要显著高于那些选择高于 25 数字的参加者。其中 IMT 测验可以用来对心理能力进行测量。这个任务由描述社会情境的四个故事组成,每个故事都要读出来,并通过投影仪以书面形式呈现出来,听完每个故事后,要求被试回答几个二选一的迫选问题,要正确回答其中的一些问题需要被试具备一定的心理理论能力(如理解他人的期望、信念等)。这个研究说明博弈中进行决策的能力和心理理论能力之间存在很大的相关。

Hedden 和 Zhang (2002)采用巧妙的矩阵博弈的方式探讨了序列进行的博弈中参加者的策略推理情况,从而反映了他们的心理理论推理水平。他们使用的实验材料是 2×2 的矩阵博弈,即有两个博弈者,而且每个人都有两种策略。四格博弈中被试和对手(计算机)轮流做出走或是留的选择,被试作为第一个博弈者在起始格做出选择,他拥有行动的操纵权,比对手拥有更大的控制最后结局的能力。被试和对手选择走还是留是基于四个格内各自的奖励值。奖励值的不同组合使得不同心理理论推理水平的被试在起始格的选择是不一样的。这样通过被试在起始格内的选择就可以考察其心理理论推理水平。研究结果显示参加者存在一个默认的心理模型,即参加者总是以 1 级推理开始,默认对手是 0 级水平。但是当经过一段时间后与 1 级对手进行博弈的参加者的心理模型会转移到 2 级水平,但是这种转变并不是很完全。同时当对手在不同的试验系列之间发生策略转变时,参加者也能够相应的进行策略的转移,虽然这种转变并不是很完全。这说明参加者在进行博弈的过程中进行了学习。此研究为在矩阵博弈结构中研究心理理论推理能力提供了实证性支持。

Yoshida, Dolan 和 Friston (2008) 基于最优化控制和博弈理论提出了心理理论的一个新的模型“心理博弈理论”(game theory of mind)。首先,他们通过以效用表示的价值函数来表征目标,在博弈过程中博弈者间的交互作用过程(表征为心理理论)可以通过函数来表示,并提出了相应的公式。其次,他们提出了模型,认为在博弈过程中预测对手的行为是可行的,而且在系列博弈中基于选择来判断思考等级的数量也是可能的。同时他们通过使用来自于“stag-hunt”博弈的模拟数据和真实数据进行了模型比较。第三,他们的研究发现通过最优化效用函数本身和生成非熟练的但明显利他的代理者都能产生相同熟练程度的行为。这在等级博弈理论和共同进化上存在行为性上的相关。此外,他们还提出在研究心理理论模型的核心子成分是否能预测大脑中心理理论区域的脑电活动方面,神经成像是很有用的方法,并且这种方法有助于将相竞争的观点结合起来。这个研究用数据模拟,公式推理和神经成像的方法为在博弈中研究心理理论提供了理论模型上的支持,并提出了“心理博弈理论”。

Verbrugge 和 Mol (2008) 的研究采用了一个拥有不完全信息的博弈来调查了成人使用和获得心理理论复杂技巧和策略的程度以及相关的语言使用技巧情况。对于参加者来说,在这个博弈中拥有关于对手心理状态的好模型是非常有益的,尤其使用二级心理理论是非常有帮助的。结果发现,当呈现许多不同系列的博弈时,大多数参加者在实验中并没有获得这些复杂的技能。然而,有一些参加者确实能够使用如二级心理理论这样的高级认知技能。

博弈结构不仅可以用来测量心理理论的认知成分,还可以用来测量心理理论的其他成分,如意图,以及心理理论与社会性间的关系。Sutter (2003) 使用最后通牒博弈来研究小学生、中学生和大学生心理理论和公平概念发展情况,结果发现所有的被试都能够考虑到提供者的意图,并且在没有其他更公平选择的情况下接受一个不是太公平的选择。然而,小学生和中学生比大学生更容易拒绝不公平的提议,这说明他们判断一个最后通牒博弈的提议时不仅考虑到对方的意图,也考虑到了提议带来的结果。这个研究说明矩阵博弈情景不仅可以研究如思维这样的高级

认知过程,也可以用来探讨如意图这样的社会认知成分,这也为传统研究心理理论的方法注入了新的血液。

Rilling, Sanfey, Aronson, Nystrom 和 Cohen (2004) 使用 fMRI 扫描来考察两个不同的人际交互任务是否会激发相似的心理理论神经网络,这两个任务中对手都能够通过获得一定的反馈来推测意图。这两个任务是最后通牒博弈和囚徒困境博弈,其中被试面对的有人类对手和电脑对手。结果发现两个博弈中大脑区域有很大程度上覆盖,既包括熟知的心理理论区域,也包括先前研究中没有报告的几个区域,这几个区域可能与被试在真实的社会交往中的表现有关。这个结果说明被试在这两个博弈中进行选择时都同样使用了心理理论能力,或是与心理理论相关的能力,当然在不同的博弈中还涉及到相应的社会性能力,这个研究为我们在博弈结构中研究心理理论与社会性发展间的关系提供了一定的基础。

Sally 和 Hill (2006) 的研究以 6~10 岁的正常儿童和相同年龄的孤独症 (autistic spectrum disorders, ASD) 儿童为研究对象调查了儿童的心理理论发展情况。研究任务包括囚徒困境博弈、最后通牒博弈和独裁者博弈,并测量了儿童的一级和二级信念理解能力。结果发现通过二级错误信念任务的儿童比未通过的儿童在囚徒困境博弈中出现更多的合作行为; ASD 儿童和没有通过二级错误信念任务的儿童在最后通牒博弈中会做出更少的提议,而大多数通过二级错误信念的儿童会做出一半的提议;在最后通牒博弈中作为接受者来说,6 岁儿童和 ASD 儿童比年长儿童更容易接受小的提议。

Rachelle (2009) 采用了博弈范式来探讨 4~8 岁儿童心理理论的发展,博弈范式的设计思路是一个社会性对手通过观察个体的行为从而对其意图有所了解,而个体也可以通过观察对手的行为而了解其意图,这被认为是后期发展的心理理论,相对于缺乏这种意识的儿童来说,拥有这种意识的儿童将会是更为有效的竞争者。在博弈过程中需要儿童能够利用对手的行为作为线索来指导自己的选择,这种能力暗示着可以利用非语言信号作为信息源来推断对手隐藏的意图。儿童检测这种信号的能力存在发展趋势,相对于 4 岁和 6 岁的儿童来说,8 岁儿童能够意识到行为线索

带来更多的信息价值，并且能更成功地通过任务。

以上的研究都为将心理理论和博弈理论结合起来进行探讨奠定了一定的基础，使得我们对心理理论和博弈理论各自的研究领域都有了丰富的认识，尤其是为利用矩阵博弈范式研究心理理论提供了一定的实证支持。

总体看来，博弈理论最初主要用于经济学领域，研究经济行为中如何使个体利益最大化。但是现在也越来越多地应用于各个进行决策的领域，同时研究者也逐渐认识到心理因素对博弈过程和结果产生了很大的影响(Gummerum, Hanoch 和 Keller, 2008)。同时一些先前的研究中也确实发现利用博弈结构来研究心理理论推理的等级情况是可行的(Hedden et al., 2002)。目前心理理论领域主要是关注儿童对心理理论中社会性方面的理解，如多关注儿童对他人愿望、意图和情绪状态的认识，而缺乏对思维等心理理论的自然认知方面以及对成人领域心理理论发展情况的研究。因而利用决策中的矩阵博弈来研究心理理论将可能是一个非常好的途径。

既然利用决策博弈来探讨心理理论是可行的，那么这种结合必将具有深远的意义：第一，创造性地把经济学中的决策博弈理论引入心理理论发展研究范畴，使其研究具有更加客观的研究范式，同时把心理理论的研究理念引入经济学的范畴，也会对经济学的研究提供一定的参考。第二，把心理理论研究的路线，从单纯关注与儿童社会化成长相关的研究思路，拓展到与智力活动息息相关的决策活动中，从此，心理理论的研究领域可以沿着两个方向同时展开：一个是让个体与周围的人群更加和谐地相处，另一个是在斗智斗勇中获得最大利益；第三，解决了传统心理理论研究中令人困惑的若干问题，特别是目前研究对象主要是儿童，而通过决策博弈情境研究心理理论可以将被试的范围扩展到成年人，从而能够从更加广阔的范围上去考察心理理论的发展机制，此外，博弈情景中对结果的反应上较少依赖于语言能力的发展，这也为尽可能排除语言能力对实验结果的干扰提供了可能。

4 研究展望

虽然利用矩阵博弈结构来研究心理理论是可行的，而且无疑会对扩展心理理论的研究提供

一个更广阔的平台，但是目前研究者使用的矩阵博弈结构还具有很大的变化性，并没有发展出类似于标准“错误信念任务”那样经典的研究范式。同时何种形式的矩阵博弈结构能够更好地考察心理理论推理的思维成分，目前还没有定论。

未来的研究可以着眼于探讨更适宜考察心理理论思维成分的矩阵博弈结构应该具备什么样的特征，以及利用矩阵博弈结构来研究更广年龄范围人群的心理理论发展轨迹，如可以利用矩阵博弈情境来探讨对手推理水平不同以及发生不同转变的情况下青少年(中学生)、青年(大学生)、中年以及老年人群心理理论推理的预测适应性和决策适应性情况。同时也可以借助于 fMRI 等生理学仪器来测查矩阵博弈结构中涉及的心理理论推理的脑机制情况，并可以同传统心理理论任务中的脑机制情况进行比较，从而对心理理论推理的脑机制有更清晰和全面的认识。此外，在矩阵博弈中还可以考察社会性发展与心理理论之间的关系，从而从自然认知和社会认知两方面对儿童及青少年的心理发展特点有更清晰的认识和理解。

参考文献

李晓东, 黄艳秋, 刘萍, 徐健. (2007). 成人在错误信念任务上的表现. *心理学探新*, 27(3), 44-47.

林佳燕, 傅根跃, 刘文庆. (2010). 成人心理理论的研究回顾与展望. *心理科学进展*, 18, 456-463.

刘希平, 安晓娟. (2010). 研究心理理论的新方法—失言识别任务. *心理科学进展*, 18, 450-455.

刘希平, 唐卫海, 方格. (2005). “儿童对主观世界认识的发展”研究的热点. *心理科学*, 28, 192-196.

罗杰, 卿素兰. (2005). 心理理论研究的起源与进展. *湖北大学学报(哲学社会科学版)*, 32, 578-582.

吴南, 张丽锦. (2007). 心理理论和语言发展的关系. *心理科学进展*, 15, 436-442.

Baron-Cohen, S., O'Riordan, M., Stone, V. E., Jones, R., & Plaisted, K. (1999). Recognition of Faux Pas by normally developing children and children with Asperger syndrome or High-Functioning Autism. *Journal of Autism and Developmental Disorders*, 29, 407-418.

Camerer, C. F., & Fehr, E. (2002). Measuring social norms and preferences using experimental games: A guide for social scientist. *Institute for Empirical Research in Economics University of Zurich, Working Paper No. 97*, 1-38.

Camerer, C. F. (2003). Behavioral studies of strategic thinking in games. *Trends in Cognitive Sciences*, 7, 225-231.

Camerer, C. F., Ho, T. H., & Chong, J. K. (2004). A cognitive hierarchy model of games. *The Quarterly of Economics*, 8, 861-898.

- Chan, K. S., Wong, H. Y., Hui, L. M., Tang, Y. M., Chan, K. W., Lam, M. L., Chiu, P. Y., & Chen, Y. H. (2010). Game theoretical approach to theory of mind deficits in schizophrenic patients with delusion(s) of reference. *Schizophrenia Research*, 117(2-3), 286.
- Flobbe, L., Verbrugge, R., Hendriks, P., & Krämer, I. (2008). Children's Application of Theory of Mind in Reasoning and Language. *Journal of Logical Language Information*, 17, 417-422.
- Gummerum, M., Hanoch, Y., & Keller, M. (2008). When Child Development Meets Economic Game Theory: An Interdisciplinary Approach to Investigating Social Development. *Human Development*, 51, 235-261.
- Hedden, T., & Zhang, J. (2002). What do you think I think you think? Strategic reasoning in matrix games. *Cognition*, 85, 1-36.
- Kahneman, D., & Tversky, A. (1982). *Judgment under Uncertainty: Heuristics and Biases*. Britain: Cambridge University Press.
- Keysar, B., Lin, S., & Barr, D. (2003). Limits on theory of mind use in adults. *Cognition*, 89, 25-41.
- Lucas, M., Wagner, L., & Chow, C. (2008). Fair game: The intuitive economics of resource exchange in four-year olds. *Journal of Social, Evolutionary, and Cultural Psychology*, 3, 74-88.
- Ohtsubo, Y. (2002). *Strategy learning in two-person constant-sum game and theory of mind*. Poster session presented at the 4th annual meeting of the Human Behavior and Evolution Society of Japan, Hokkaido University, Japan.
- Ohtsubo, Y., & Rapoport, A. (2006). Depth of reasoning in strategic form games. *The Journal of Socio-Economics*, 35, 31-47.
- Paal, T., & Berezkei, T. (2007). Adult theory of mind, cooperation, Machiavellianism: The effect of mindreading on social relations. *Personality and Individual Differences*, 3, 541-551.
- Perner, J. (1979). Young children's preoccupation with their own payoffs in strategic analysis of 2x2 games. *Developmental Psychology*, 15, 204-213.
- Rachelle, S. (2009). Development of theory of mind from age's four to eight. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, 70(6-B), 3813.
- Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *NeuroImage*, 22, 1694-1703.
- Sally, D., & Hill, E. (2006). The development of interpersonal strategy: Autism, theory-of-mind, cooperation and fairness. *Journal of Economic Psychology*, 27, 73-97.
- Stahl, D. O., & Wilson, P. W. (1995). On players' models of other players: theory and experimental evidence. *Games and Economic Behavior*, 10, 218-254.
- Sutter, M. (2003). On the nature of fair behavior and its development with age. Max Planck Institute for Research into Economic Systems, Jena. *Working paper*, 1-12.
- Takagishi, H., Kameshima, S., Schug, J., Koizumi, M., & Yamagishi, T. (2010). Theory of mind enhances preference for fairness. *Journal of Experimental Child Psychology*, 105, 130-137.
- Verbrugge, R., & Mol, L. (2008). Learning to Apply Theory of Mind. *Journal of Logic, Language and Information*, 17, 489-511.
- Yoshida, W., Dolan, R. J., & Friston, K. J. (2008). Game Theory of Mind. *PLoS Computational Biology*, 4(12), 1-14.

A New Paradigm to Study the Development of Theory of Mind: Matrix Game

LI Jing^{1, 2, 3}; LIU Xi-Ping¹

(1 School of Education of Tianjin Normal University, Tianjin, 300387, China)

(2 Institute of Psychology, Chinese Academy of Sciences, Beijing, 100101, China)

(3 Graduate School of Chinese Academy of Sciences, Beijing, 100049, China)

Abstract: Theory of Mind (ToM) is an ability, which develops throughout one's life, to make inferences about the mental states of others. Studies about ToM have quite a few limits on the participants and its research topics. This review introduces a new paradigm, matrix game, and analyzes the feasibility of adopting it to the study of ToM. Furthermore, the theoretical basis and empirical evidences of combining ToM study and matrix game are proposed in this review. It is believed that the matrix game paradigm can be applied in a wide spectrum of ages to study the developmental trajectory of ToM in future studies. Besides, the neuro-imaging techniques such as fMRI can be introduced to explore the neural mechanism of ToM involved in matrix game.

Key words: theory of mind; matrix game; decision making; thinking