

类别不确定下的特征推理是基于类别 还是基于特征联结*

莫雷 陈琳

(华南师范大学心理应用研究中心, 广州 510631)

摘要 共有3个实验探讨归类不确定情况下的特征推理是基于类别进行还是基于特征联结进行。实验1在中文条件下重复了Verde等人2005的实验,得出了与之相符的结果,这个结果用基于类别的理性模型的设想或者是用基于特征联结的设想都可以解释。实验2考察被试在靶类别的类别特征频次并且特征结合出现频次高低不同的条件下特征推理的情况,实验2的结果表明,高集中与低集中两种条件下特征推理没有显著差异,不符合特征推理是基于类别进行的设想,而与特征推理是基于特征联结进行的设想吻合。实验3进一步考察被试在特征结合出现的总频次并且靶类别中特征结合出现的总频次高低不同的条件下特征推理的情况,结果表明,在高结合条件下进行特征推理要优于在低结合条件,支持了在归类不确定情况下的特征推理是基于特征联结进行的设想。据此可以认为,人们的特征推理是基于特征之间联结的频次进行,而不是基于类别进行。

关键词 单类说;理性模型;基于类别;基于特征联结;特征推理

分类号 B842

1 问题与目的

当前,心理学界关于类别的研究主要集中在探讨类别学习与类别运用问题(Markman & Ross, 2003; Ross & Warren, 2002; Yamauchi, Love & Markman, 2002; Yamauchi & Markman, 1998)。人们根据已有的类别知识对从未见过的物体归类,并且对该物体的某些特征进行预测和推理,这就是类别运用问题(Anderson, Ross & Chin-Parker, 2002; Erickson, Chin-Parker & Ross, 2005; Lin & Murphy, 2001; Ross, Gelman & Rosengren, 2005)。在许多情况下,人们不能完全确定某个物体应该归入哪一类,例如看到远处有一只动物,我们通常不能确定该动物是狗还是其他的动物,这种情境就被称为“归类不确定的情境”。然而,即使在这种无法确定归类的情境中,人们也会按照该物体最可能所属的类别对它的行为或者特征作出预测,我们将这种情况称为“归类不确定时的特征推理”。

在归类不确定的情境中人们是如何进行特征预测推理,成为了心理学界重点探讨的问题。目前主

要有两种解释:Anderson提出的理性模型和Murphy与Ross提出的“单类说”。

Anderson(1991)提出的理性模型认为,对于任何目标物F来说,人们预测其具有某特征j的可能性时,会考虑到目标物可能归属的每一类别,将遵从Bayesian规则做出预测,Bayesian规则可以用下面公式来表示: $P(j|F) = \sum_k P(K|F)P(j|K)$ 。在公式中,k代表学习者已经知道的类别。该公式表示,目标物F具有特征j的概率等于该目标物F属于某类别的概率乘上该类别的每一物体具有特征j的概率的总和,即人们综合多个类别的信息来对事物的特征作出预测。在各种类别中,如果存在着可能性最大的“靶类别”,人们虽然会考虑“靶类别”中的信息,但同时也会考虑其他“非靶类别”中的信息,其最后的决策是综合考虑多个类别的信息而得出的最优化结果。例如在森林中你看到一只动物,并且不能确定它是一只狗还是一只浣熊,在你预测这只动物有多大的可能性会伤害你时,按照理性模型,你既会考虑狗会伤害你的可能性,又会考虑浣熊会伤害你的可能性。如果你认为那只动物80%的可能性

收稿日期:2007-12-18

*教育部哲学社会科学重大课题攻关项目资助(05JZD00034)与教育部人文社会科学重点研究基地2004年度重大项目资助(05JJDXX005)。

通讯作者:莫雷, E-mail: molei@snu.edu.cn

是狗,20%的可能性是浣熊,同时你认为狗有10%的可能性会伤害你,浣熊有75%的可能性会伤害你,那么你预测这只动物会伤害你的可能性是 $0.8 \times 0.1 + 0.2 \times 0.75 = 0.23$ 。这就是Anderson理性模型的基本观点。

然而,Murphy和Ross等人(1994)的研究否证了理性模型,提出并验证了单类说。单类说认为,在归类不确定的情境中,人们会选择最可能的类别,即“靶类别”,然后只根据靶类别的信息进行特征预测,而不考虑非靶类别。单类说和理性模型的区别就在于,特征预测是否考虑非靶类别的信息。单类说认为特征预测只考虑靶类别的信息,而理性模型认为特征预测既会考虑靶类别的信息又考虑非靶类别的信息(常建芳,莫雷,2003;陈琳,莫雷,2007)。

Murphy与Ross等人从1994年的研究开始,一直对归类不确定条件下的特征推理进行研究。其研究表明,在归类不确定的情况下,不管是在人工类别条件下进行特征推理,还是在自然类别条件下进行特征推理,人们都是只考虑最可能的靶类别,而不考虑非靶类别。其系列研究结果均支持了单类说(Chin-Parker & Ross, 2004; Malt, Ross & Murphy, 1995; Murphy, 1982; Murphy, 2003; Murphy & Ross, 2005; Ross & Murphy, 1996; Lagnado & Shanks, 2003)。

然而,应该指出,Murphy等人以往所做的研究,主要是采用了同时呈现全部样例的方式,在特征推理的过程中,被试能够看到所有的类别及样例,并且特征预测的方式是进行概率判断。而在2005年的研究中,Verde和Murphy等人改变了原来的研究范式,采用了逐个呈现样例的研究范式(Verde, Murphy & Ross, 2005)。在逐个呈现样例的研究范式中,被试首先逐个对样例进行学习,然后对具有一个特征的样例是否最有可能具有另外一个特征进行预测,在推理的过程中,不给被试呈现样例所属的类别以及其他样例。

实验结果发现增加条件的特征推理在正确率,反应时和信心判断方面都显著优于基线条件。此表明,被试在进行特征推理时考虑了非靶类别的信息。显然,这个结果与Murphy等提出的单类说不相符。对于这个结果,Verde和Murphy等人也感到难以解释,他们只能提出,被试可能只是按照特征联结的频次来进行特征预测,而不是遵循理性模型进行思考。

Verde和Murphy等人(2005)的研究结果,提出了一个值得认真考虑的新问题,促使人们对以往有

关这个问题的研究以及争论进行反思。本研究认为,综合考虑前人有关的研究结果,关于归类不确定条件下的特征推理的研究首先要解决的基本问题应该是:在归类不确定条件下的特征推理是基于类别进行还是基于特征的联结进行。

第一种设想是基于类别进行特征推理,这是心理学界比较认可的设想。该设想认为,人们推断某客体的缺失特征时,首先考虑该物体属于何种类别,然后根据该类别可能的特征状况对缺失特征进行预测,这种特征推理,称为“基于类别的特征推理”。在认可这种基于类别的特征推理的基础上,才有单类说与理性模型之争。单类说认为,人们在进行特征推理时只考虑靶类别的信息。理性模型则认为,人们在进行特征推理时会考虑所有类别的信息,二者虽然有重要区别,但是,它们都是首先将样例归入相应的类别,然后根据有关的类别(根据靶类别或者根据靶类别和非靶类别)所具有的要预测的特征的概率进行预测。因此,本质上都是基于类别的推理,其加工过程都是“将(缺失特征的)样例进行归类并基于有关类别进行特征推理”。

第二种设想是基于特征联结进行特征推理,这种设想认为,人们在推断某物体的缺失特征时,根本就不是先考虑该物体最可能属于什么类别,然后再根据类别的特征状况对缺失特征进行预测,而是直接根据呈现的样例的有关特征与预测特征(缺失特征)之间联结的频次进行推理,即根据特征的联结直接推断物体的缺失特征,可以称为“基于特征联结的特征推理”。这种基于特征联结所进行的特征推理,不是在类别的基础上进行的推理,不需要以类别为中介,既不会考虑样例属于靶类别的概率,也不会考虑样例具有要预测的特征的概率。

显然,类别不确定情况下的特征推理究竟是基于类别进行还是基于特征联结进行,这是首先要确定的问题,只有确定了人们是基于类别进行推理之后,才能够进一步探讨这种基于类别进行的推理是符合单类说还是符合理性模型。

总的来看,自从Murphy与Ross 1994年开创了关于归类不确定条件下的特征推理的研究以来,心理学界基本上是认可了基于类别的特征推理,并在这个前提下,开展单类说与理性模型的研究与论争。实际上,关于人们在类别不确定情况下的特征推理究竟是基于类别进行还是基于特征联结进行这个更重要、更基本的问题,一直没有进行过系统的、专门的研究,没有得到比较确切的解决。Verde和Mur-

phy 等人在 2005 年的研究中根据实验结果正式提出了可能存在基于特征联结进行推理的证据,但是,由于他们的实验并不是围绕着解决基于类别推理还是基于特征联结推理这个问题来设计的,其结果实际上既可以用基于特征联结推理进行解释,也可以用基于类别推理的理性模型进行解释,还不能真正有力地支持基于特征联结进行推理这个结论。

据上分析,本研究准备在 Verde 和 Murphy 等人 2005 年实验的基础上,进一步探讨在类别不确定情况下的特征推理是基于类别的推理还是基于特征联结的推理。本研究包括 3 个实验,实验 1 在中文条件下重复 Verde 和 Murphy 等人 2005 年的实验。实验 2 比较类别内特征高结合与低结合两种条件下的特征推理情况,考察在归类不确定情况下的特征预测是否会基于类别进行。实验 3 进一步比较类别特征整体上高联结与低联结两种条件下的特征推理情况,考察在类别不确定情况下的特征预测是否会基于特征联结进行。

2 实验 1

2.1 研究目的

在中文条件下重复 Verde 和 Murphy 等人 2005 年的实验,考察在类别不确定情况下的特征推理是否会得出与 Verde 和 Murphy 等人 2005 年的实验一致的结果。

2.2 研究方法

2.2.1 被试 35 名本科生参加本实验,男生 17 人,女生 18 人,文理科大致相当,均自愿参加,实验后给予一定报酬。

2.2.2 实验材料 实验材料由学习材料与测试材料组成。学习材料与 Verde 和 Murphy 等人 2005 年的研究相同,包括动物栖生环境的四个类别,分别是“森林”类、“沙漠”类、“沼泽”类与“草原”类。每个类别都包括 20 个动物样例,每个样例都有两个不同维度的特征,一个是表皮维度特征(A),分为 6 种水平:鳞(A1),毛(A2),羽(A3),皮(A4),刺(A5),壳(A6);另一个是器官维度特征(B),也分为 6 种水平:须(B1),爪子(B2),牙齿(B3),尾巴(B4),触手(B5),角(B6)。

在每个类别中,都会有一对特征高频率同时出现的,可以称为“主要特征对”。主要特征对中的单个特征称为“主要特征”,因此四个类别共有 8 个主要特征。类别 1 中高频率同时出现的特征对是 A1/B1(鳞,须),类别 2 中高频率同时出现的特征对是

A2/B2(毛,爪子),类别 3 中高频率同时出现的特征对是 A3/B3(羽,牙齿),类别 4 中高频率同时出现的特征对是 A4/B4(皮,尾巴)。但是,每个类别的主要特征对也有可能出现在其他的类别中,这就使得具有这些特征对的样例在归类时不确定。例如同时具有特征鳞和须的样例,更多地属于森林类,有时也会属于沼泽类。

上面 4 个类别的主要特征对出现在靶类别的次數都是 9 次,但是,类别 1 的主要特征对(A1/B1)在类别 3 中又出现了 4 次,类别 2 的主要特征对(A2/B2)在类别 4 中又出现了 4 次;而类别 3 与类别 4 的主要特征对(A3/B3 与 A4/B4)在其他类别中一次也没有出现。因此,类别 1 和类别 2 是增加条件,类别 3 和类别 4 是基线条件。增加条件和基线条件的主要特征对在靶类别中出现的频次是相同的,区别在于增加条件中主要特征对在非靶类别中出现的频次高于基线条件中主要特征对在非靶类别中出现的频次。

测试材料包括 24 个样例项目。每个样例项目都包括两个特征。其中 8 个样例项目的两个特征都是主要特征对,分别是:A1/B1、A2/B2、A3/B3、A4/B4、B1/A1、B2/A2、B3/A3、B4/A4;另外 8 个样例项目的两个特征是来自不同类别的主要特征随机组合构成的 8 个特征对,例如 A1/B2 等。剩余的 8 个测试项目分别由两个填充特征构成。每个测试项目都测试 3 轮,因此测试项目总共包括 72 个。

实验的任务是,分别呈现样例及其两个特征,被试需要判断具有第一个特征的样例是否最可能具有另一个特征。如果人们在归类不确定下的特征推理是按照单类说,只考虑靶类别,那么,四个类别的主要特征对的特征预测概率都应该是: $(12/16) \times (12/20) = 9/20$;这样,在增加条件下(即类别 1 与类别 2 的特征预测条件下)被试对预测特征预测的准确率、反应时和信心判断都应该与基线条件(即在类别 3 与类别 4 的特征预测的条件下)相同。而如果人们在归类不确定情况下的特征推理是按照理性模型,既会考虑靶类别的基本概率,同时也考虑非靶类别的概率,那么,在基线条件下被试对主要特征对的特征预测概率还是等于 $9/20$,但是,在增加条件下,被试对主要特征对的特征预测的概率就应该等于靶类别的预测特征概率加上非靶类别中的预测特征概率,即: $(12/16) \times (12/20) + (4/16) \times (4/20) = 10/20$ 。这样,在增加条件下被试对预测特征预测的准确率、反应时和信心判断都应该优于基线

条件。

2.2.3 实验设计 本实验是单因素被试内设计,被试接受两种条件的实验处理。自变量是增加条件和基线条件两个水平;因变量是特征推理的正确率、反应时与信心判断。正确特征推理是指呈现具有一个特征的样例时,被试能正确地判断该样例是否具有另一个特征;反应时指被试进行正确推理所需要的时间;信心判断则是指被试进行正确推理时自我感觉的把握程度。

2.2.4 实验程序 本研究沿用了 Verde 和 Murphy 等人 2005 年的范式,实验过程包括学习和测试两个阶段。

学习阶段分为四步。第一步学习材料,被试学习四个类别的 80 个样例,每个类别 20 个样例。电脑会按照类别随机逐个呈现样例,同时包括其栖息地(类别)与两个特征,均采用文字方式呈现,而不是用图形表示。呈现时间 4000ms,要求被试认真学习各种类别的成员及其特征。

第二步,练习反馈,电脑逐个呈现学习过的样例及其特征,要求被试判断这个样例属于四个类别中的哪个类别,根据每个类别的编号按键,电脑立即对被试的答案给予反馈,如果被试选择的类别是错误的,那么需要重新选择,直到选择了正确的类别。每个样例只测验一次,顺序随机。

第三步,重复第一步学习。

第四步,重复第二步练习反馈。

测验阶段:包括 8 个练习项目和 72 个测验项目。在每个项目测验之前,首先出现红色的十字,然后电脑呈现样例以及它的一个特征(例如,一种生物具有鳞),500ms 之后再呈现第二个特征,被试需要预测该样例是否最有可能具有第二个特征(那么它是否最有可能具有须)。被试要尽快做出“是”“否”的反应,并且要在四点量表上对自己反应的正确程度做信心判断。1 表示没有信心,2 表示比较没有信心,3 表示比较有信心,4 表示有信心。

2.3 结果与分析

统计增加条件和基线条件两种条件下被试对 8 个主要特征对的测试项目的特征推理的正确率、反应时与信心判断的数据,并进行统计分析,其中反应时与信心判断的数据只包括正确推理的数据。 t_1 是以被试为随机变量进行的统计检验(被试检验), t_2 是以项目为随机变量进行的统计检验(项目检验)。结果见表 1。

表 1 增加条件与基线条件下特征推理的比较

实验条件	正确率	反应时(ms)	信心判断
增加条件	95.9%	1039.50 ± 399.81	3.91 ± 0.28
基线条件	87.9%	1619.70 ± 684.58	3.68 ± 0.41

表 1 的结果表明,增加条件下特征预测的正确率显著高于基线条件下特征预测的正确率,被试检验差异非常显著: $t_{(34)} = 3.27, p < 0.005$;项目检验差异非常显著: $t_{(2)} = 7.00, p < 0.01$ 。增加条件下特征预测的反应时也显著短于基线条件下特征预测的反应时,被试检验差异非常显著: $t_{(34)} = -6.40, p < 0.001$;项目检验差异非常显著: $t_{(3)} = -10.51, p < 0.005$ 。并且,在增加条件下特征预测的信心判断显著高于基线条件下特征预测的信心判断,被试检验差异非常显著: $t_{(34)} = 3.18, p < 0.005$;信心判断项目检验差异非常显著: $t_{(2)} = 6.70, p < 0.01$ 。这个结果与 Verde 和 Murphy 等人 2005 年的研究结果完全一致。

实验 1 验证了 Verde 和 Murphy 等人 2005 年的研究结果,被试在增加条件下和基线条件下对样例的特征预测存在显著差异,据此可以否定单类说。但是,如前所述,根据这个实验的设计,该结果可能有两种解释:一种是用理性模型进行解释,也就是认为,在归类不确定情况下的特征推理遵从考虑多个类别的理性模型,被试在进行特征预测时既考虑了靶类别的特征情况,同时也考虑了非靶类别的特征情况,综合两者的信息做出预测推论;另一种是基于特征联结推理的解释,也就是认为,被试在进行特征预测时只是考虑测试样例中的特征与各个备择特征联结的概率,选择出联结概率最大的备择特征作为预测特征。显然,根据这个实验的设计逻辑,上述结果只能否定单类说,而无法确定被试实际上是按照理性模型还是基于特征联结来进行特征预测。实验 2 准备通过改变实验材料中类别主要特征在靶类别与非靶类别的结合比例,以考察在类别不确定情况下的特征推理是否会基于类别进行。

3 实验 2

3.1 研究目的

比较靶类别中类别特征出现以及特征结合出现的频次不同的条件下,被试进行特征推理的情况,考察在归类不确定情况下的特征预测是否会基于类别进行。

本实验控制特征出现的总频次与特征结合出现

的总频次,操作了特征在靶类别出现的频次以及特征在靶类别结合出现的频次,形成靶类别内特征出现频次以及特征结合出现频次高低不同的“高集中条件”和“低集中条件”,通过比较这两种条件下特征预测的情况,验证被试进行特征推理是否会基于类别进行。

3.2 研究方法

3.2.1 被试 大学一年级本科生 35 名,男生 15 名,女生 20 名,均自愿参加,实验后给予一定报酬。

3.2.2 实验材料 实验材料由学习材料与测试材料组成,本实验测试材料与实验 1 相同,学习材料根据本研究目的作了改变。

学习材料包括四个类别,每个类别 20 个样例,类别 1 的主要特征是 A1/B1,类别 2 的主要特征是 A2/B2,类别 3 的主要特征是 A3/B3,类别 4 的主要特征是 A4/B4。在本实验材料中,每个主要特征出现的总频次相同(16 次),各类别的两个主要特征结合出现总频次也相同(13 次),但是各类别的主要特征在靶类别(即本类别)内出现的频次与结合出现的频次不同,形成靶类别内特征出现频次高并且特征结合出现频次高的“高集中条件”和靶类别内特征出现频次低并且特征结合出现频次低的“低集中条件”。

类别 1 与类别 3 属于“高集中条件”,类别 1 的两个主要特征(A1 与 B1)在靶类别(类别 1)中出现频次为 28,其中结合出现 13 次、分离出现 2 次(B1 为 2 次),在非靶类别(类别 3)中分离出现 4 次(A1 为 3 次,B1 为 1 次);类别 3 的两个主要特征(A3 与 B3)出现情况也是如此;类别 1 与类别 3 构成了“高集中条件”。类别 2 与类别 4 属于“低集中条件”,类别 2 的两个主要特征(A2 与 B2)在靶类别(类别 2)中出现频次为 24,其中结合出现 9 次、分离出现 6 次(A2 与 B2 各 3 次),在非靶类别(类别 4)中结合出现 4 次;类别 4 的两个主要特征(A4 与 B4)出现情况也是如此;类别 2 与类别 4 构成了“低集中条件”。

3.2.3 实验设计 本实验是单因素被试内设计,自变量是靶类别中类别特征出现以及特征结合出现的频次状况,分为高集中条件与低集中条件两个水平。因变量是特征预测的正确率,反应时,信心判断。

3.2.4 实验程序 包括学习训练阶段与测试阶段,具体做法与实验 1 相同。

按照本实验的设计,根据实验结果就可以确定被试是基于类别进行特征推理还是基于特征联结进

行特征推理。

第一情况,假定被试是基于类别进行特征推理,那么,无论他们遵循的是单类说还是理性模型,高集中条件下特征预测应该优于低集中条件下的特征预测。

因为,假定人们在推理中遵循的是单类说,那么,在高集中条件下,类别 1 中具有特征 A1 的样例同时具有特征 B1 的概率是: $(13/16) \times (15/20) = 39/64$;同理,类别 3 中具有特征 A3 的样例同时具有特征 B3 的概率也是 $39/64$ 。而在低集中条件下,类别 2 中具有特征 A2 的样例同时具有特征 B2 的概率是: $(12/16) \times (12/20) = 9/20$;同理,类别 4 中具有特征 A4 的样例同时具有特征 B4 的概率也是 $9/20$ 。这样,人们在高集中条件下特征预测的各项指标就会优于低集中条件。

如果人们在推理中遵循的是理性模型,那么,根据贝叶斯规则的公式,在高集中条件下,无论靶类别是类别 1 或者类别 3,在所有类别中具有特征 A1(或 A3)的样例同时具有特征 B1(或 B3)的概率是: $(13/16) \times (15/20) + (3/16) \times (1/20) = 99/160$;而在低集中条件下,无论靶类别是类别 2 或者类别 4,在所有类别中具有特征 A2(或 A4)的样例同时具有特征 B2(或 B4)的概率是: $(12/16) \times (12/20) + (4/16) \times (4/20) = 1/2$ 。因此,人们在高集中条件下特征预测的各项指标同样会优于低集中条件。

还有一种可能,被试是基于类别进行推理,但是他们注意的是靶类别中具有特征 A_k 的样例同时具有特征 B_k 的概率,这种可能性是前人研究中没有考虑过的。在本实验条件下,由于高集中条件下两个主要特征在靶类别中结合出现的频次为 13,而在低集中条件下,两个主要特征在靶类别中结合出现的频次是 9,因此,人们在高集中条件下特征预测的各项指标还是会优于低集中条件。

总之,只要被试是基于类别进行推理,那么,无论遵循单类说还是遵循理性模型,或者是只注重靶类别内特征结合的比例,高集中条件下对缺失特征的预测都应该优于低集中条件下对缺失特征的预测。

第二种情况,假定被试是基于特征联结进行推理,那么,由于在高集中条件下与低集中条件下,各个类别的主要特征结合呈现的总频次相等(都是 13 次),因此人们在高集中条件下(即类别 1 与类别 3 的条件下)对缺失特征的预测与低集中条件下(即类别 2 与类别 4 条件下)对缺失特征的预测应该没

有差异。

3.3 结果与分析

分别统计高集中与低集中两种条件下被试特征推理的正确率、反应时与信心判断的数据并进行统计检验,其中反应时与信心判断的数据只包括正确推理的数据。 t_1 是以被试为随机变量进行的统计检验(被试检验), t_2 是以项目为随机变量进行的统计检验(项目检验)。结果见表2。

表2 高集中与低集中条件下特征推理的比较

实验条件	正确率	反应时(ms)	信心判断
高集中条件	93.4%	1021.90 ± 493.38	3.84 ± 0.29
低集中条件	95.8%	1186.50 ± 730.68	3.81 ± 0.33

表2的结果表明,被试在高集中和低集中两种条件下进行特征推理的正确率、反应时与信心判断的差异都不显著。正确率被试检验结果是: $t_{(34)} = -0.90, p > 0.05$;项目检验结果是: $t_{(2)} = -1.39, p > 0.05$ 。反应时被试检验结果是: $t_{(34)} = -1.94, p > 0.05$;项目检验结果是: $t_{(2)} = -2.59, p > 0.05$ 。信心判断被试检验结果是: $t_{(34)} = 0.58, p > 0.05$;项目检验结果是: $t_{(2)} = 1.35, p > 0.05$;

根据本实验的设计逻辑,上述结果表明:第一,被试没有基于靶类别的有关特征出现频次进行特征预测,不符合单类说;第二,被试没有综合靶类别与非靶类别的有关特征出现频次进行特征预测,不符合合理性模型;第三,被试也没有基于靶类别的有关特征结合出现的频次进行特征预测,同样不符合基于类别推理的设想。也就是说,本实验结果并不符合基于类别进行推理的设想,而是符合基于特征联结进行推理的设想。

本实验的这个结论虽然是建立在无差异的统计结果之上,但是,如果结合实验1的设计与结果来看,这个结论也是可信的。因为,在实验1,增加条件中具有特征 A_k 的样例同时具有特征 B_k 的概率是 10/20,而基线条件中具有特征 A_k 的样例同时具有特征 B_k 的概率是 9/20,但是,两种条件下特征推理各项指标就有了非常显著的差异。而本实验的高集中条件与低集中条件的概率分别是 99/160 与 80/160,差异程度比实验1的要大,但是,两种条件下特征推理各项指标没有显著差异。可见,根据本实验结果得出被试并没有基于类别进行推理的结论是可信的。

然而,仅仅根据本实验的结果就得出被试是基于特征联结进行推理的结论还是不够的。首先,本

实验这种无差异的统计结果尽管可信,但也只是符合了基于特征联结进行推理的设想,并没有直接证明被试就是基于特征联结进行特征推理。只有变化特征联结的条件,得出联结频次高条件下的特征推理优于联结频次低条件下的特征推理,才能证明基于特征联结进行推理的设想。

其次,具体分析本实验两种条件下三项指标的结果,第一个准确率指标,两种条件下虽然差异不显著,但是成绩都很高(93%以上),可能有天花板效应;第二个反应时指标,两种条件下虽然差异尚未达到显著性水平,但是也达到边缘显著。因此,有必要进一步设计实验验证本研究的设想。

4 实验3

4.1 研究目的

比较类别特征结合出现以及在靶类别内特征结合出现的频次不同的条件下,被试进行特征推理的情况,考察在归类不确定情况下的特征预测是否会基于特征联结进行。

本实验控制特征出现的总频次与特征在靶类别出现的频次,操作了各类别的两个特征结合出现的总频次以及两个特征在靶类别结合出现的频次,形成了特征结合出现频次以及在靶类别内特征结合频次高低不同的“高结合条件”和“低结合条件”,通过比较这两种条件下特征推理情况,验证被试在归类不确定时进行特征推理是否基于特征联结进行。

4.2 研究方法

4.2.1 被试 大学本科生共 34 人,其中男生 16 人,女生 18 人,均自愿参加,实验后给予一定报酬。

4.2.2 实验材料 实验材料由学习材料与测试材料组成,本实验测试材料与实验1相同,学习材料根据本研究目的作了改变。

本实验的学习材料也是包括四个类别,每个类别 20 个样例,类别 1 的主要特征是 A1/B1,类别 2 的主要特征是 A2/B2,类别 3 的主要特征是 A3/B3,类别 4 的主要特征是 A4/B4。但是,在本实验材料中,每个主要特征出现的总频次相同(16 次),各类别的两个主要特征在靶类别(即本类别)内出现的频次也相同(各 12 次),但是各类别的主要特征结合出现的总频次以及在靶类别内结合出现的频次不同,形成特征的高结合与低结合两种条件。

类别 1 与类别 3 属于“高结合条件”,这两个类别的主要特征 A1/B1 和 A3/B3 结合出现的总频次都是 16,在靶类别中结合出现都是 12 次, A1/B1 两

个特征在靶类别(类别 1)中结合出现 12 次,在其他类别(类别 3)中出现 4 次;A3/B3 两个特征在靶类别(类别 3)中结合出现 12 次,在其他类别(类别 1)中结合出现 4 次;类别 1 与类别 3 构成了特征“高结合条件”。类别 2 与类别 4 属于“低结合条件”,这两个类别的主要特征 A2/B2 和 A4/B4 结合出现的总频次都是 13,在靶类别中结合出现都是 9 次,A2/B2 两个特征在靶类别(类别 2)中结合出现 9 次,分离出现各 3 次,在其他类别(类别 4)中结合出现 4 次;A4/B4 两个特征对在靶类别(类别 4)中结合出现 9 次,分离呈现各 3 次,在其他类别(类别 2)中结合出现 4 次。类别 2 与类别 4 构成了特征“低结合条件”。

4.2.3 实验设计 本实验是单因素被试内设计,自变量是各类别主要特征结合出现的频次,分为特征高结合与特征低结合两个水平。因变量是特征预测的正确率,反应时,信心判断。

4.2.4 实验程序 包括学习训练阶段与测试阶段,具体做法与实验 1 相同。

按照本实验的设计,根据实验结果就可以确定被试是基于类别进行特征推理还是基于特征联结进行特征推理。

第一情况,假定被试是基于特征联结进行推理,那么由于高结合条件下特征结合出现的频次是 16,而低结合条件下特征结合出现的概率只是 13,因此高结合条件下的特征推理就应该优于低结合条件。

第二种情况,假定被试是基于类别进行特征推理,那么,无论他们遵循的是单类说还是理性模型,高结合条件与低结合条件的特征预测都不会有显著差异。

因为,假定人们在推理中遵循的是单类说,由于各个类别中具有第一个主要特征的样例同时具有第二个特征的概率都是: $12/16 \times 12/20 = 9/20$,因此,高结合条件(类别 1 与类别 3)与低结合条件(类别 2 与类别 4)的特征推理应该没有差异。

如果人们的特征预测遵循理性模型,根据理性模型计算公式,各个类别中具有第一个主要特征的样例同时具有第二个主要特征的概率都是: $12/16 \times 12/20 + 4/16 \times 4/20 = 9/20 + 1/20 = 10/20$,因此,高结合条件(类别 1 与类别 3)与低结合条件(类别 2 与类别 4)的特征推理也应该没有差异。

4.3 结果与分析

分别统计高结合与低结合两种条件下被试特征推理的正确率、反应时与信心判断的数据并进行统

计检验,其中反应时与信心判断的数据只包括正确推理的数据。 t_1 是以被试为随机变量进行的统计检验(被试检验), t_2 是以项目为随机变量进行的统计检验(项目检验)。结果见表 3。

表 3 高结合与低结合条件下特征推理的比较

实验条件	正确率	反应时(ms)	信心判断
高结合条件	99.0%	855.00 ± 273.48	3.91 ± 0.28
低结合条件	84.2%	1455.60 ± 504.34	3.68 ± 0.41

表 3 的结果表明,高结合条件下特征预测在正确率、反应时与信心判断三项指标均显著优于低结合条件下的特征预测。正确率方面,被试检验差异非常显著: $t_{(33)} = 3.78, p < 0.005$;项目检验差异非常显著: $t_{(2)} = 43.15, p < 0.001$ 。反应时方面,被试检验差异非常显著: $t_{(33)} = -6.67, p < 0.001$;项目检验差异非常显著: $t_{(2)} = -12.65, p < 0.005$ 。信心判断方面,被试检验差异非常显著: $t_{(33)} = 3.00, p < 0.005$;项目检验差异也非常显著: $t_{(2)} = 7.41, p < 0.005$ 。

上述结果表明,被试在归类不确定时特征推理是基于特征联结进行,直接受到特征联结频次的影响。本实验同时也得出了不支持基于类别进行推理的证据,再次证明了实验 2 的结论。

5 讨论

Verde 和 Murphy 等人 2005 年的研究采用了逐个呈现样例的范式,对归类不确定下的特征推理进行了进一步的研究,结果表明,增加条件下的特征推理优于基线条件下的特征推理。这个研究结果显然与他们以往采用集中呈现样例的范式所得到的结果不同。对此,Verde 和 Murphy 等人感到难以解释。他们提出,在逐个呈现样例的研究范式中,被试的推理可能是按照特征联结的频次进行。

本文分析了 Verde 和 Murphy 等人的研究,提出了归类不确定情况下的特征推理首先要解决的最基本问题是,特征推理是基于类别还是基于特征联结进行。本研究设计了 3 个实验探讨这个问题。

实验 1 在中文条件下重复了 Verde 和 Murphy 等人 2005 年的实验,获得了与之一致的结果:增加条件下的特征推理明显优于基线条件下的特征推理。但这个实验结果,既可以用基于类别进行推理的理性模型观点进行解释,也可以用基于特征联结进行推理的观点进行解释。实验 2 和实验 3 进一步对这两种关于归类不确定时特征推理的实质的不同

设想作出检验。

实验2 控制特征出现的总频次与特征结合出现的总频次,操作了特征在靶类别出现的频次与特征在靶类别结合出现的频次,形成靶类别内特征出现频次高并且特征结合出现频次高的“高集中条件”和靶类别内特征出现频次低与特征结合出现频次低的“低集中条件”,通过比较这两种条件下的特征推理情况,直接对归类不确定时特征推理是否基于类别进行作出检验。结果表明,高集中条件与低集中条件下的特征推理没有显著差异,否证了特征推理是基于类别进行的设想,符合特征推理是基于特征联结进行的设想。

实验3 控制特征出现的总频次与特征在靶类别出现的频次,操作了各类别的两个特征结合出现的总频次与两个特征在靶类别结合出现的频次,形成了特征结合出现频次高并且在靶类别内特征结合频次高的“高结合条件”和特征结合出现频次低并且在靶类别内特征结合频次低的“低结合条件”,通过比较这两种条件下特征推理情况,直接对归类不确定时特征推理是否基于特征联结进行的设想作出检验。结果表明,高结合条件与低结合条件下的特征推理差异显著,证实了被试的特征推理是基于特征联结进行的设想。

表4 列出了两个实验的设计及结果与结论。

表4 实验2与实验3的设计与结果

实验条件	特征总频次	靶类别特征频次	特征结合总频次	靶类别特征结合频次	基于类别			基于特征联结	
					单类说	理性模型	结合说		
实验2	高集中	相同	高	相同	高	否定	否定	否定	符合
	低集中	相同	低	相同	低				
实验3	高结合	相同	相同	高	高	不符	不符	不符	证实
	低结合	相同	相同	低	低				

由此可见,本研究结果明确地解答了 Verde 和 Murphy 等人 2005 年研究提出的疑问,验证了在归类不确定条件下的特征推理是基于特征联结进行这个基本设想。

值得注意的是,本研究与 Verde 和 Murphy 等人 2005 年研究尽管采用的是逐个呈现类别样例的方式,但是,根据本研究的结果与结论,同样可以对以往采用类别材料同时呈现的方式所得出的支持单类说与支持理性模型的研究结果作出合理的解释。

以往心理学界对归类不确定的特征推理的研究,主要采用了类别材料同时呈现的方式,所得出的结果虽然有争议,但都是以基于类别的单类说与理性模型之争。Murphy 和 Ross1994 年用同时呈现的方式对人们在归类不确定情况下特征预测所进行的研究,在实验中同时呈现的材料如图1所示。

这些图形是4名儿童的图画,根据图1,如果要求被试估计“三角形是黑色”的概率时,Bob为靶类别,按 Bayesian 规则的计算方法,靶类别预测特征的基本概率是 $(3/5) \times (3/4) = 0.45$,其他三个非靶类别预测特征的基本概率为 $(1/5) \times (1/4) + (1/5) \times (1/4) + 0 \times 0 = 0.1$,被试的预测概率等于靶类别预测特征的基本概率与其他三个非靶类别预测特征的基本概率之和,即等于0.55。而如果要求被试估计“正方形是白色”的概率时,John为靶类别,按

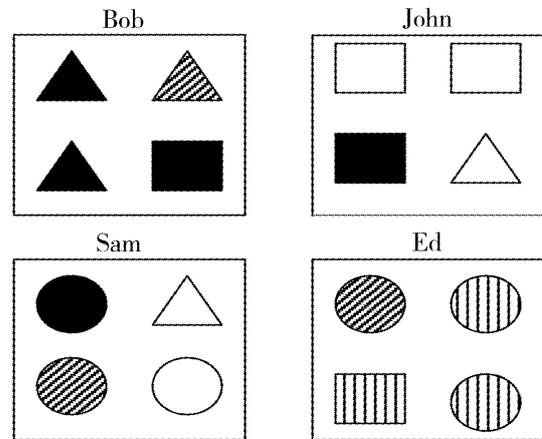


图1 Murphy1994年研究所用材料样例

Bayesian 规则公式的计算方法,靶类别预测特征的基本概率也是 $(3/5) \times (3/4) = 0.45$,与前面的相同,但其他三个非靶类别预测特征的基本概率则是 $0 \times (2/4) + (1/5) \times 0 + (1/5) \times 0 = 0$ 。被试的预测概率等于靶类别预测特征的基本概率与其他三个非靶类别预测特征的基本概率之和,即等于0.45。研究者将第一种情况(估计“三角形是黑色”的概率)即非靶类别中预测特征的基本概率大于0的情况称为“增加条件”,将第二种情况(估计“正方形是白色”的概率)即非靶类别中预测特征的基本概率等于0的情况称为“基线条件”。由于上述两种条件

下靶类别的预测特征的基本概率相同,因此,如果人们在预测特征时符合理性模型,就会综合考虑靶类别和非靶类别的信息以作出特征预测,那么增加条件下的预测概率应该比基线条件下的预测概率有显著的提高;如果被试遵从了单类说的观点,就会只考虑靶类别中的信息,那么被试在两种条件下的预测就没有明显差异。其实验结果表明。被试在两种条件下的特征预测差异不显著,从而支持了单类说。

然而,莫雷与赵海燕(2002)的研究提出,在 Murphy 和 Ross1994 年的研究中,没有考虑到非靶类别中目标特征与预测特征两个特征的结合或分离这个重要因素,其研究材料中增加条件下的非靶类别的两个特征都是分离的,如图 1,增加条件下的非靶类别的“三角形”(目标特征)与“黑色”(预测特征)是分离的,这或许是其得出支持单类说的结果的原因。根据这个考虑,莫雷引进了特征结合与分离这个变量,探讨了非靶类别中目标特征与预测特征的结合与分离是否影响被试进行特征预测对非靶类别信息的利用。莫雷 2002 年的研究也采用同时呈现的方式,材料见图 2。

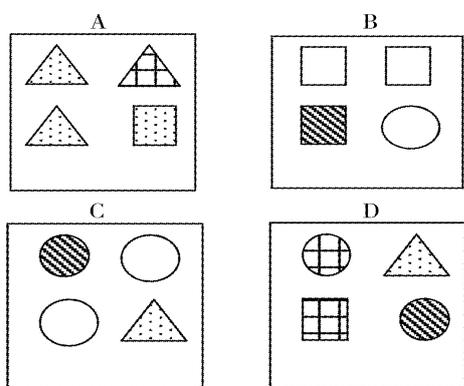


图 2 莫雷 2002 年所用材料样例

从图 2 可见,如果预测“三角形 - 圆点”的概率时,最可能的靶类别是 A,如果预测“正方形 - 空白”的概率时,最可能的靶类别是 B,在这两种情况下靶类别的特征基本概率及维度结合情况是一样的,但非靶类别预测特征的基本概率不相同,对“正方形 - 空白”的特征预测是基线条件(其非靶类别预测特征的基本概率是: $1/4 \cdot 0 + 0 \cdot 0 + 1/4 \cdot 0 = 0.00$),而对“三角形 - 圆点”的特征预测是属于增加条件(其非靶类别预测特征的基本概率是: $0 \cdot 0 + 1/4 \cdot 1/5 + 1/4 \cdot 1/5 = 0.10$)。图 2 与 Murphy 和 Ross1994 年研究的材料(图 1)不同点在于,在图 2 中,非靶类别的目标特征与预测特征(三角形与圆

点)是结合的,如 C 和 D 中各有一个带有圆点的三角形,因而,对“三角形 - 圆点”的特征预测构成了“增加 - 结合”条件。如果按单类说的观点,被试的决策过程只受到靶类别信息的影响,那么在“增加 - 结合”条件下,非靶类别中两种维度部分结合应该不会对被试的特征预测概率产生影响,因此,对特征的概率与基线条件没有显著差异;反之,如果“增加 - 结合”条件下被试的特征预测概率显著高于基线条件,则说明在非靶类别维度结合的情况下,被试在决策时受到了非靶类别信息的影响,符合 Bayesian 规则。实验结果表明,在非靶类别维度结合的情况下,被试在预测特征时确实受到了非靶类别信息的影响,表现出对 Bayesian 规则的符合。据此,莫雷提出了将特征结合因素加进 Bayesian 规则的修正公式。

然而,上述两项研究,虽然一个得出的是单类说的结论,另外一个得出的是理性模型的结论,似乎都是基于类别的推理,但是,实际上它们的不同结果都可以统一用基于特征联结推理的观点进行解释。因为,在 Murphy 和 Ross1994 年的研究中,增加条件与基线条件下靶类别中目标特征与预测特征结合出现的频次相同;而在增加条件下,非靶类别的目标特征与预测特征虽然在同一个类别中出现,但是两者是分离的(见图 1)。这样,如果是基于特征联结进行推理,那么无论是基线条件还是增加条件,它们的非靶类别的特征联结的频次都为 0,因此,所得出的增加条件与基线条件下特征预测概率差异不显著的结果,表面上看是表明特征概率的预测只是受靶类别的影响,实质上是表明特征概率的预测受目标特征与预测特征结合频次的影响。反之,莫雷等 2002 年的研究中,增加条件下非靶类别的目标特征与预测特征在同一类别中是结合的(见图 2),这样,如果是基于特征联结进行推理,那么由于增加条件下非靶类别的特征联结的频次大于基线条件,也就是总的特征联结频次高于基线条件,因此,得出了增加条件下的特征预测概率高于基线条件的结果。

可见,归类不确定条件下的特征推理主要是受特征联结的频次的影响这个设想,不仅在本研究以及 Verde 和 Murphy 等人的研究结果中得到验证,而且可以解释原先采用同时呈现方式对归类不确定情况下特征推理的研究所得出的各种不同结果。因此,可以初步认为,归类不确定条件下的特征推理主要是基于特征联结进行。当然,这个结论还有待于进一步的研究进行验证,尤其应该注意的是,以往有

关研究无论是逐个呈现样例的范式还是同时呈现样例的范式,每个类别成员都只有 2 个类别特征,如果增加类别成员的特征数量,为了更便于进行特征推理,被试是否会改用基于类别进行推理的策略?换言之,在归类不确定条件下的特征推理,是默认基于特征推理的策略,还是如人们提出来的根据认知经济性原则灵活采用基于特征联结策略还是基于类别的策略,还需要进一步进行探讨(陈战胜,莫雷,2003)。

6 结论

本研究的结果表明,在本实验采用逐个呈现样例的研究范式的条件下,人们在归类不确定时的特征推理是基于特征之间的联结频次进行,原先研究所得出的基于类别进行特征预测的各种结果,实际上都可以用基于特征联结进行特征预测的观点做出更有说服力的解释。

参 考 文 献

- Anderson, A., Ross, B. H., & Chin-Parker, S. (2002). A further investigation of category learning by inference. *Memory & Cognition*, 30, 119-128.
- Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review*, 98, 409-429.
- Chang, J. F., & Mo, L. (2003). The review of prediction when categorization is uncertain (in Chinese). *The Chinese Journal of Psychological Science*, 26, 159-160.
- [常建芳,莫雷.(2003).类别不确定时的预测推理研究述评. *心理科学*, 26, 159-160]
- Chen, L., & Mo, L. (2007). The development of the single-category view when categorization is uncertain (in Chinese). *Advances in Psychological Science*, 15, 64-70.
- [陈琳,莫雷,刘志雅.(2007).归类不确定时特征推理的单类说理论. *心理科学进展*, 15, 64-70]
- Erickson, J. E., Chin-Parker, S., & Ross, B. H. (2005). Inference and classification learning of abstract coherent categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 86-99.
- Lagnado, D. A., & Shanks, D. R. (2003). The influence of hierarchy on probability judgment. *Cognition*, 89, 157-178.
- Lin, E. L., & Murphy, G. L. (2001). Thematic relations in adults' concepts. *Journal of Experimental Psychology: General*, 30, 3-28.
- Malt, B. C., Ross, B. H., & Murphy, G. L. (1995). Predicting features for members of natural categories when categories is uncertain. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 646-661.
- Markman, A. B., & Ross, B. H. (2003) Category use and category Learning. *Psychological Bulletin*, 29, 592-613.
- Mo, L., & Chen, Zh. Sh. (2003). Rule-based categorization strategy and example-based categorization strategy in categorization. *Acta Psychologica Sinica*, 35, 29-36.
- [莫雷,陈战胜.(2003).规则策略和样例策略在归类过程中的运用. *心理学报*, 35, 29-36.]
- Mo, L., & Zhao, H. Y. (2002). Influence of association and separation in the dimensions on the predictions in uncertain categorization. *Acta Psychologica Sinica*, 34, 470-479.
- [莫雷,赵海燕.(2002).维度的结合与分离对归类不确定性预测的影响. *心理学报*, 34, 470-479.]
- Murphy, G. L. (2002). The big book of concepts. *Introduction* (pp. 1-10). Cambridge, MA: MIT Press.
- Murphy, G. L. (2003). The downside of categories. *Trends in Cognitive Science*, 12, 513-514.
- Murphy, G. L. (1982). Cue validity and levels of categorization. *Psychological Bulletin*, 91, 174-177.
- Murphy, G. L., & Ross, B. H. (1994). Predictions from uncertain categorization. *Cognitive Psychology*, 27, 148-193.
- Murphy, G. L., & Ross, B. H. (2005). The two faces of typicality in category-based induction. *Cognition*, 95, 175-200.
- Ross, B. H., Gelman, S. A., & Rosengren, K. S. (2005). Children's category-based inferences affect classification. *British Journal of Developmental Psychology*, 23, 1-24.
- Ross, B. H., & Murphy, G. L. (1996). Category based prediction: Influences of uncertainty and feature associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 736-753.
- Ross, B. H., & Warren, J. L. (2002). Learning abstract relations from using categories. *Memory & Cognition*, 30, 657-665.
- Verde, M. F., Murphy, G. L., & Ross, B. H. (2005) Influence of multiple categories in inductive inference. *Memory & Cognition*, 33, 479-487.
- Yamauchi, T., Love, B. C., & Markman, A. B. (2002). Learning nonlinearly separable categories by inference and classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 585-593.
- Yamauchi, T., & Markman, A. B. (1998). Category learning by inference and classification. *Journal of Memory and language*, 39, 124-149.

Which One Is Better? Based on Categories or Based on Feature Association When Categorization Is Uncertain

MO Lei, CHEN Lin

(Center for Studies of Psychological Application, South China Normal University, Guangzhou 510631, China)

Abstract

When categorization is uncertain, the inner mechanism of people's feature induction induced some arguments. In terms of predicting features, Rational Model indicated that people would take all categories into consideration while single-category theory advanced that people only paid attention to the target category. The essential difference between Rational Model and single-category theory is whether people would take the alternative categories into consideration when predicting features. But neither Rational Model nor single-category theory could explain the results of the research by Verde et al in 2005. So Verde et al concluded that the feature prediction in their research might be based on feature association. Therefore, three experiments were designed in this research to investigate how people predicted features when categorization was uncertain.

Experiment 1 replicated the experiment of Verde et al (2005) in Chinese circumstance so that the effect of language would be excluded. In Experiment 2, the frequency of the feature association in order to find out whether categories were considered. Experiment 3 the variable of category was controlled in order to find out whether people's induction was based on feature association.

Seventy subjects were evenly grouped into experiment 1 and experiment 2, and thirty-four subjects in experiment 3. All of the subjects were individually tested on computers, which controlled list generation, stimulus presentation, and response recording. The procedure comprised five phases: (1) study, (2) a training test, (3) study, (4) a training test, and (5) a final test. During the study phase, all exemplars were shown once. Each exemplar was shown for 4000msec. During the training test phase, subjects should categorize all exemplars into their correct categories. During the final test phase, subjects should decide whether or not the to-be-predicted feature was the one that most likely to appear with the given feature and then to make a yes-no judgment.

Experiment 1 gained the same results as the experiment conducted by Verde et al (2005). By controlling the variable of feature association, Experiment 2 found that there was no significant difference in feature prediction between two conditions, which indicated that people's induction was not category-based. Via controlling the variable of category, Experiment 3 found that there was significant difference in feature prediction between two feature association conditions, which indicated that people's induction was based on feature association.

The three experiments illustrate that people are inclined to predict features on the basis of feature association when categorization is uncertain. In contrast to the category-based prediction, prediction based on feature association is an economical strategy, which can be adopted to explain many results of the previous studies.

Key words single-category theory; Rational Model; based on categories; based on feature association; feature prediction