

# 采样假设背景下的归纳推理的典型性效应

卢丽洁\*, 舒睿

西南大学认知与人格教育部重点实验室, 重庆

收稿日期: 2022年11月4日; 录用日期: 2022年12月7日; 发布日期: 2022年12月20日

## 摘要

人们进行类别归纳的程度取决于它们关于观察到的实例是如何从世界上采样的信念。采样假设对归纳推理具有重要的意义, 而典型性效应是归纳推理中的一个关键现象。我们采用包含高典型性和低典型性的前提证据, 前提证据分别在不同的采样假设情况下呈现: 有意的(强采样)或随机的(弱采样)。根据贝叶斯模型理论我们预测, 强抽样比弱抽样导致更有限的泛化, 有更弱的推理强度。研究结果表明, 在两种条件下存在较强的典型性效应, 典型性效应依赖于样本证据的假设。我们的研究为归纳推理领域的研究提供了新的视角, 人们关于证据如何产生的信念对归纳推理有着很大的影响。

## 关键词

归纳推理, 采样假设, 典型性效应

# Typical Effects of Inductive Reasoning in the Context of Sampling Assumptions

Lijie Lu\*, Rui Shu

Key Laboratory of Cognition and Personality of the Ministry of Education, Southwest University, Chongqing

Received: Nov. 4<sup>th</sup>, 2022; accepted: Dec. 7<sup>th</sup>, 2022; published: Dec. 20<sup>th</sup>, 2022

## Abstract

When the subcategories of a category are often mentioned, they can also be called high typical, such as robins among birds, apples among fruits. Those that are less likely to come to mind or are less likely to be recalled are low typical, for example, penguins among birds, jujubes among fruits. The typicality effect of inductive reasoning means that when people infer the possibility of a conclusion in a high typical premise or a low typical premise, the probability of the conclusion in the

\*通讯作者。

high typical premise is greater than that in the low typical premise. Inductive reasoning also has its limitations. The data on which we make inferences based on experience are often incomplete and, therefore, the conclusions are inconclusive. However, more and more people believe that similarity alone is not enough to explain attribute induction. Sampling assumption means people's reasoning depends on how they believe an inductive argument is formed. One must make a key "sampling" assumption about how the available data was generated. Sampling assumption plays an important role in inductive reasoning, and typicality effect is a key phenomenon in inductive reasoning. The purpose of this study is to observe the characteristics of the typicality effect of inductive reasoning under sampling assumption, and to provide further theoretical basis for the study of sampling assumption and typicality. Based on the previous experimental paradigm, the typicality premise and sampling assumption are independent variables. There are two levels of typicality: high typicality and low typicality, and two levels of sampling assumption: deliberate (strong sampling) or random (weak sampling). Inference strength is the dependent variable. Each set of arguments consists of a premise and a conclusion (i.e. premise: "duck", conclusion: "bird"). The participants were asked to rate the likelihood of the premise's conclusion being true on a 4-point Likert scale based on the information obtained from the instructions and premise. The results showed that there is a strong typicality effect under both the strong condition and weak condition, and it indicated that the typicality effect depends on the sampling assumption of evidence. We predicted that strong sampling leads to more limited generalization and has weaker inference strength than weak sampling based on Bayesian model. Our study provides a new perspective for the field of inductive reasoning, that is how evidence is generated has a great influence on the strength of inductive reasoning.

## Keywords

Inductive Reasoning, Sampling Assumption, Typicality Effect

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

当人们被问到超市有什么水果时, 大多数人脑海里都会蹦出苹果、梨子等水果(苹果、梨子就是典型水果)。一个种类中被人们经常提及到的子类被称之为高典型, 例如水果中的苹果、鸟类中的麻雀等; 而那些不太容易出现在脑海或不易被回忆起来的子类则可以称为低典型, 例如, 水果中的枣子、鸟类中的企鹅。归纳推理的典型性效应是归纳推理中一个非常重要的效应, 它是指让人们在一个高典型前提或一个低典型前提中推论结论成立的可能性时, 高典型前提支持结论成立的可能性要大于低典型前提(金鑫, 陈云, 2014; Osherson, Smith, Wilkie, López, & Shafir, 1990; Sloman, 1993; Kiran & Thompson, 2003)。Osherson 等人(1990)的相似性覆盖模型提出高典型性前提的相似性要大于低典型性前提, 典型更大的前提被认为是更广泛的结论类别的代表, 因而具有更强的推理强度。

归纳推理是一种基本的认知能力, 即人们超越所提供的已知事实来进行推理, 通过一个类别具有的属性从而推论出另一个类别或子类也有相同的属性, 人类可以自然而然的进行归纳它是学习、分类和决策的理论基础(Fernbach, 2006; Navarro, Dry, & Lee, 2012; Shepard, 1987; Tenenbaum & Griffiths, 2001)。归纳推理也有其局限性, 我们基于经验进行推论的数据往往是不完整的, 因此, 结论也不确定。然而, 有越来越多的人认为仅仅相似性不足以解释属性归纳(Kemp & Tenenbaum, 2009; Medin, Coley, Storms, & Hayes, 2003)。

采样假设则是指人们做出的归纳推理判断取决于人们关于论点是如何产生而形成的信念。人们必须做出一个关键的“采样”假设,即关于可用数据是如何生成的(Hayes, Navarro, Stephens, Ransom, & Di-levski, 2019; Navarro et al., 2012; Vong, Hendrickson, Perfors, & Navarro, 2013)。前提类别和结论类别间的关系、人们自身的知识结构(Medin et al., 2003)以及在实验操作中消息的提供者都会影响人们对论点形成的信赖程度进而影响他们的推理(Lawson, 2018)。之前的研究考虑了两种极端的可能性,即强采样和弱采样(Shepard, 1987; Tenenbaum & Griffiths, 2001)。例如,在 Hayes 等人(2019)的研究中,将强采样条件定义为“在每次试验中,你会看到三个生物具有特定特性的实例。请注意,这些例子是故意选择的,以最好地说明具有这种特性的生物的多样性。”或者“具有属性 x 的类别之所以被选择,这是一位老师在查阅专业的生物书籍后从具有属性 x 的生物中随机选择的。”弱采样条件定义为:“每次试验你都会看到三个生物具有特定特性的实例。我们让一名学生随机翻开一本关于植物和动物的书,记下他们看到的前三种生物,以及这些生物是否具有相关属性。这意味着你收到的信息可能不是对你做出判断最有帮助的——信息具有偶然性和随机性,学生有时会选择非常不同的项目,有时又非常相似的项目。”或者“随机选择了 2 种动物”。Navarro 等人(2012)认为强采样条件下的数据被认为是有意地生成一个概念的积极例子。总而言之,强采样具有选择性、逻辑性、可信赖性以及帮助性,而弱采样条件下的数据是在没有任何限制的情况下生成的,具有偶然性、随机性。

归纳推理在很大程度上是人们信念的构造过程。因此,人们必须依靠一些他自身对于证据价值的假设观察,进而建立起对数据是如何采样的理论知识,并通过一些方法将该理论与关于世界状态的信念联系起来(Navarro et al., 2012)。采样假设影响推理主要表现在概念学习任务(Navarro et al., 2012)、属性诱导任务(Ransom, Perfors, & Navarro, 2016),以及词汇学习任务(Xu & Tenenbaum, 2007)。证据价值的假设观察是由许多因素决定的,从而导致不同的采样信念,“强采样”和“弱采样”模型就是其中的例子。采样假设是内心信念与外界客观事实的桥梁。Hayes 等人(2019)在归纳推理的多样性效应中发现不仅归纳推理对采样假设敏感,而且采样假设对多样性效应存在一定程度上的削弱。因而采样假设在确定论点时哪些特征被认为与归纳推理问题相关方面有很大的作用。

以前的研究也发现否定证据与采样假设有着一定的联系,Voorspoels 等人(2015)的研究中发现了非单调性效应,也就是负面的证据提高了人们概括的意愿,这与人们推理时做出的采样假设密切相关(Voorspoels, Navarro, Perfors, Ransom, & Storms, 2015)。非单调性效应会在当人们相信推理的证据来源于一个权威信息提供者时出现,而当推理的信息是随机选择时负面证据的推理强度就会不及积极证据。非单调性效应出现在强采样情况下,而弱采样情况下不存在。同样的,采样假设也被证实与归纳推理的样本大小(Fernbach, 2006; Navarro et al., 2012; Vong et al., 2013)有关。Fernbach (2006)发现,贝叶斯模型强调的“大小原则”(Kemp & Tenenbaum, 2003)在强采样假设的情况下更容易被遵循,而在弱采样或较模糊的假设情况下不容易被遵循,这暗示着,虽然人们有时对基于贝叶斯模型下的采样假设敏感,但是基于强采样假设的模型无法推广到一般的属性归纳单调性效应情况中去。

相似性覆盖模型(Osherson et al., 1990)和基于特征的模型(Sloman, 1993)是归纳推理的经典理论模型,但它们都对采样假设不敏感(Fernbach, 2006)。相似性覆盖模型没有对数据采样重新加权说明(Navarro et al., 2012)。另一方面,基于传统的理性分析的模型旨在寻求计算中存在的问题并试图用系统的方法解决并提出一个最优的方案。这些模型共同的主题就是以贝叶斯模型为基础。在以前的研究中,出现了一种基于采样假设来进行计算的方法,这类模型称为采样敏感贝叶斯模型(Sampling Sensitive Bayesian model, Fernbach, 2006)。这类模型依赖于基于数据是如何进行采样的假设的似然计算方法。具体来说,采样敏感贝叶斯模型一般假设人们使用强采样假设来进行的属性归纳(Kemp & Tenenbaum, 2003; Sanjana & Tenenbaum, 2003),以概念为中介来比较前提和结论的。

先前的行为研究表明归纳推理的多样性效应, 单调性效应都会受采样假设的影响, 而典型性效应是否同样也会受采样假设影响以及是什么程度的影响是未知的, 因此本研究的目的是探索在不同采样假设下, 归纳推理的典型性效应具有什么样特征, 对采样假设的研究以及典型性的研究进一步提供理论依据。本研究在 Hayes 等人(2019)的实验范式基础上进行改进, 前提的典型性和采样假设为自变量, 典型性有两个水平: 高典型和低典型, 采用假设也有两个水平: 强采样和弱采样; 推理强度为因变量。每组论点都有一个前提和一个结论组成(前提: “鸭子”, 结论: “鸟”)。实验要求被试根据从指导语中获取的信息来对前提推断结论成立的可能性进行评分, 评分采用的是李克特 4 点量表。根据贝叶斯泛化梯度模型: 一个强抽样模型的泛化梯度随样本数量的增加而增加, 呈现为较陡的曲线, 以及 Hayes 等人(2019)的和采样敏感性贝叶斯模型(Kemp & Tenenbaum, 2003; Sanjana & Tenenbaum, 2003), 因此我们预测在本研究中: 1) 高典型性前提的推理强度大于低典型前提; 2) 弱采样假设组的推理强度大于强采样假设组。

## 2. 方法

### 2.1. 被试

有偿招募 40 名在校大学生, 年龄范围 18~26 岁。所有被试视力或矫正视力正常, 均为右利手, 无神经症或精神病史, 母语为汉语。8 名被试由于数据缺失或无效而被排除在外。最终 32 名在校大学生(女: 22), 平均年龄 21.63 岁( $SD = 2.38$ )纳入统计分析。其中, 强采样组(年龄:  $Mean = 21.38, SD = 2.47$ )、弱采样组(年龄:  $Mean = 21.88, SD = 2.18$ )各 16 名被试。

### 2.2. 实验设计与实验程序

实验采用 2 (采样假设: 强采样, 弱采样)  $\times$  2 (典型性: 高典型, 低典型) 的两因素混合实验设计, 其中采样假设为被试间变量, 典型性为被试内变量。被试随机分配到强、弱采样假设组。强、弱采样假设组都包括了高典型前提条件和低典型前提条件。不同采样假设组的被试会收到不同的指导语。在强采样假设组, 被试会被暗示论点的前提条件是精心选择, 有助于结论的成立; 在弱采样假设组, 被试会被暗示论点的前提条件是随机选择的, 它是否符合逻辑需要被试自行判断。

在强采样假设组的指导语:

“在每个题目中, 你将会看到一个具有特定分子结构的生物个体。请注意, 我们精心地挑选了这个个体, 使它是具有该分子结构的最典型的个体。例如, 下表呈现的‘鸭有分子结构 D1’, 那么‘鸭’即为那个精心挑选的典型个体。”

在弱采样假设组的指导语:

“在每个题目中, 您将会看到一个具有特定分子结构的生物个体。我们让一位学生随机打开一本有关动植物的书, 并记录下他们遇到的第一个生物个体, 但是这个个体是否具有某种特定分子结构是不确定的。这意味着您收到的信息对您做出判断并不是最有帮助的。请注意, 学生有时会选择差异很大的个体, 有时会选择差异很小的个体。例如, 下表呈现的‘鸭有分子结构 D1’, 那么‘鸭’即为那个随机选择的典型个体。”

随后被试会对 188 个试次(示例见表 1)进行推理判断(94 个高典型试次, 94 个低典型试次)。每个试次都包含 1 个基本水平的前提类别和一个上位水平的结论类别(例如, 前提: 苹果, 结论: 水果), 前提类别蔬菜、水果、昆虫、鸟和哺乳动物, 结论类别词语都在蔬菜、水果、昆虫、鸟和哺乳动物这 5 个之一。前提类别和结论都具有相同的属性结构, 如: 鸭有属性 W1, 鸟有属性 W1。在实验中, 我们要求被试“请仔细阅读下列指导语句子, 参考获取的信息, 根据前提判断结论成立的可能性大小, 并进行 4 点评分, 分值 1 表示‘可能性非常小’, 分值 4 表示‘可能性非常大’, 请用‘√’选出你的评分”。

**Table 1.** Part of the experiment sample  
**表 1.** 部分实验试次

序号	前提	结论	可能性			
			1	2	3	4
1	鸭有分子结构 D1	鸟有分子结构 D1				
2	狮子有分子结构 X2	哺乳动物有分子结构 X2				
3	银耳有分子结构 A2	蔬菜有分子结构 A2				
4	西瓜有分子结构 M3	昆虫有分子结构 M3				
5	翠鸟有分子结构 E3	蔬菜有分子结构 E3				
6	花生有分子结构 B3	蔬菜有分子结构 B3				
7	芒果有分子结构 E5	昆虫有分子结构 E5				
8	油麦菜有分子结构 K1	蔬菜有分子结构 K1				
9	西兰花有分子结构 N2	蔬菜有分子结构 N2				
10	啄木鸟有分子结构 X3	鸟有分子结构 X3				

### 2.3. 实验材料

在正式实验招募了额外的被试评估了前提材料的典型性程度和熟悉性程度(见表 2), 采用李克特 6 点量表评估了前提类别的典型性程度(1, 最不典型; 6, 最典型)和采用李克特 5 点量表评估熟悉性程度(1, 最不熟悉; 5, 最熟悉), 其中, 高典型前提类别的平均典型性程度为 5.59 ( $SD = 0.18$ ), 平均熟悉性程度为 4.06 ( $SD = 0.48$ ); 低典型前提类别的平均典型性程度为 2.96 ( $SD = 0.72$ ), 平均熟悉性程度为 3.90 ( $SD = 0.60$ )。对前提包含的基本类别的典型性评价和熟悉性评价结果分析显示, 高低典型基本类别的典型性程度存在显著差异( $t[54] = 22.278, p < 0.001, \text{Cohen's } d = 6.88$ ), 高低典型基本类别的熟悉性程度不存在显著差异( $t[54] = 1.04, p = 0.305, \text{Cohen's } d = 0.32$ ), 表明了本实验的实验材料只在典型性程度上有差异, 而在熟悉性上每个被试对高低典型性的前提感受是一样的, 本实验前提典型性控制有效。

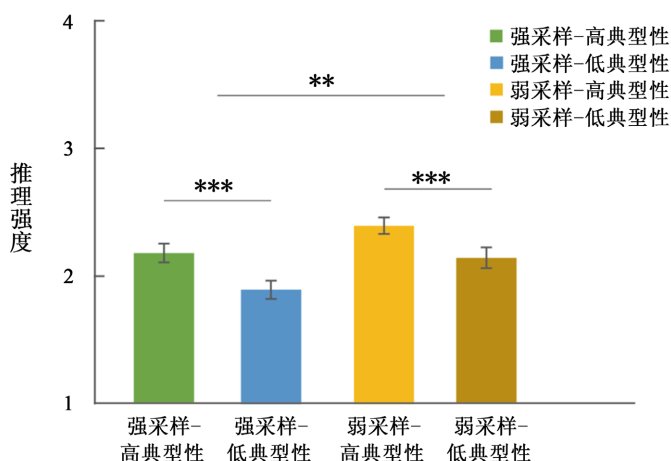
**Table 2.** Scores of typicality and familiarity of some experimental premise categories  
**表 2.** 部分实验前提类别典型性和熟悉性评分

前提类别	典型性	熟悉性
老虎	5.67	3.6
狼	5.61	3.4
石榴	5.44	4.32
橙子	5.72	4.56
芒果	5.5	4.36
猕猴桃	5.53	4.4

### 3. 结果

分别计算了强、弱采样组每个被试高、低典型性前提推论结论成立的可能性评分(见表 3), 以及 4 个条件下每组的平均推理强度及其组内标准误差(见图 1)。





**Figure 1.** Mean inference strength and within-group standard errors in the strongly sampling-high typical condition, strongly sampling-low typical condition, weakly sampling-high typical condition, and weakly sampling-low typical condition. \* $p < 0.05$ , \*\* $p < 0.01$ , and \*\*\* $p < 0.001$

**图 1.** 强采样 - 高典型条件、强采样 - 低典型条件、弱采样 - 高典型条件和弱采样-低典型条件下的平均推理强度及其组内标准误差, \*表示  $p < 0.05$ , \*\*表示  $p < 0.01$ , \*\*\*表示  $p < 0.001$

对四个条件进行重复测量方差分析后发现, 采样假设和典型性二者之间的交互作用不显著,  $F(1, 15) = 0.18, p = 0.678, \eta_p^2 = 0.01$ 。典型性的主效应显著,  $F(1, 15) = 92.79, p < 0.001, \eta_p^2 = 0.86$ 。高典型的前提推论结论的强度远大于低典型(高典型:  $Mean = 2.29, SD = 0.31$ ; 低典型:  $Mean = 2.02, SD = 0.32$ ;  $F(1, 31) = 100.92, p < 0.001, \eta_p^2 = 0.77$ )。采样假设的主效应显著,  $F(1, 15) = 6.11, p = 0.026, \eta_p^2 = 0.29$ , 弱采样假设组的推论强度要显著高于强采样假设组(弱采样假设组:  $Mean = 2.28, SD = 0.33$ ; 强采样假设组:  $Mean = 2.04, SD = 0.31$ ;  $F(1, 31) = 11.69, p = 0.002, \eta_p^2 = 0.27$ )。

为了进一步验证典型性效应是存在于强采样假设组还是弱采样假设组, 分别在强采样假设组和弱采样假设组进行配对样本  $t$  检验, 结果发现, 无论是在强采样下,  $t(1, 15) = 5.46, p < 0.001, Cohen's d = 1.37$ , 还是弱采样下,  $t(1, 15) = 10.47, p < 0.001, Cohen's d = 2.62$ , 都存在典型性效应。

**Table 3.** Scores of high and low typical premise inference conclusions for each subject in the strong and weak sampling groups

**表 3.** 强、弱采样组每个被试高、低典型性前提推论结论成立的可能性评分

强采样组被试		弱采样组被试	
高典型性前提	低典型性前提	高典型性前提	低典型性前提
2.53	2.13	2.72	2.52
2.43	2.13	2.69	2.00
2.61	2.32	2.48	2.33
2.46	2.05	2.12	1.71
1.98	1.96	2.36	2.34
1.86	1.46	2.03	1.88
2.53	2.22	2.57	2.20
2.28	1.90	2.76	2.43
1.98	1.80	2.68	2.40

## Continued

2.00	1.80	2.61	2.21
1.89	1.76	1.97	1.64
1.65	1.38	2.55	2.64
2.01	1.71	1.96	1.71
2.02	1.80	2.37	1.90
2.21	1.84	2.59	2.51
2.49	2.15	2.02	1.88

## 4. 讨论

推理的核心特征之一是归纳推理, 它不仅受到推理内容和推理者的先验知识的限制, 而且还受到他们如何利用信息推理的限制(Voorspoels et al., 2015)。先前的行为研究表明归纳推理的多样性效应, 单调性效应都会受采样假设的影响, 而典型性效应是否同样也会受采样假设影响以及是什么程度的影响是未知的, 本研究采用与 Hayes 等人(2019)基础上进一步改进实验探究归纳推理的典型性效应在采样假设背景下的特征表现。研究结果表明, 弱采样假设组比强采样假设组有更强的推理强度, 高典型前提的推理强度是要显著高于低典型前提。

正如贝叶斯模型预测的那样(Hayes et al., 2019; Tenenbaum & Griffiths, 2001), 当人们相信推理的证据是强采样时, 会出现了稳健的多样性效应。然而, 当背景提示证据是随机产生的(弱抽样), 多样性效应减弱。值得注意的是, 这种衰减意味着弱采样的总体支持度高于强采样下(Hayes et al., 2019)。采样假设对多样性效应最大的影响在于弱采样假设下人们进行推理的强度大于强采样下, 即使是多样性较低的证据也可能会有较高的推理强度, 强采样比弱采样导致更多的限制属性泛化。如果学习者相信前提论证由一系列随机事实组成, 弱采样对信念的修正会局限于对不一致假设的修正, 而强采样模型预测, 人们应该转向小的信念假设(Sanjana & Tenenbaum, 2003; Tenenbaum & Griffiths, 2001; Voorspoels et al., 2015; Ransom, Hendrickson, Perfors, & Navarro, 2018)。基于贝叶斯模型的预测, 采样假设捕获一些对采样的敏感性。那么, 在强采样条件下, 人们进行归纳推理会表现得更为严格, 有研究发现人们会基于从封面故事中获取的信息来收紧他们对一个类别的泛化标准, 并且不会扩展泛化的类别边界(Vong et al., 2013; Xu & Tenenbaum, 2007; Ransom et al., 2016)。因为在本实验中强采样下归纳推理得分显著低于弱采样条件下, 人们在强采样条件下的推理表现得更为严格, 对于从前提推论结论成立的可能性上会更严格的遵守从指导语中获取的信息。弱采样条件下比强采样条件下有更强的推理强度, 表明采样假设对归纳推理有影响, 但在弱采样情况下典型性效应会有所衰减。

在归纳推理领域, 前提的典型性对属性归纳有显著的影响(Liang, Chen, Lei, & Li, 2016)。它表现为当人们在被要求说出某个类别的成员时, 主要提供的是典型的例子, 这类例子也被称为高典型例子, 而容易被人们忽视的类别成员也被称为低典型。高典型前提的推理强度是要显著高于低典型前提, 这表明归纳推理的典型性效应存在。但是我们并没有发现采样假设和典型性之间的交互作用, 无法说明在强或弱采样下, 高、低典型性的表现如何。同时, 我们的预测是基于贝叶斯模型模拟计算建构出来的理论进行的预测, 并没有通过实际的数学模型运算, 从而进行假设的预测是我们的不足。Hayes 等人(2019)的研究中运用贝叶斯模型模拟计算出, 在强采样假设下多样化的论点被评为更强推理的可能性更大, 表现为更陡曲线。而在弱采样下, 多样性效应存在衰减的情况, 即在弱采样情况下, 多样性前提的论点被评为更强推理的可能性减少, 表现为较缓的曲线。

采样假设具有个体差异性(Vong et al., 2013; Hayes et al., 2019; Navarro et al., 2012; Ransom et al., 2018)。有研究发现,不同受试者对采样假设的敏感程度存在显著的个体差异,并且个体的生物学背景知识储备程度也会对推理判断产生影响(Proffitt, Coley, & Medin, 2000; Bailenson, Shum, Atran, Medin, & Coley, 2002),不同的个体从典型的或者不典型的证据中自然会产生归纳假设方式的差异。我们研究以及实验结果表明了采样假设对归纳推理中重要的归纳效应的影响,如典型性效应。先前对归纳推理的典型性效应的理论解释主要关注不同的样本间的相似性或熟悉性是如何促进属性归纳的。例如,相似性覆盖模型认为跟类别间的相似覆盖率有关(Osherson et al., 1990),因此,典型成员与该类别的其他成员或结论间有更高特征覆盖率,而不典型成员与该类别中的其他成员共享较少特征,有较少的覆盖率。相反,我们的方法表明,典型性效应的推理强度不仅取决于前提证据的典型性程度,也取决于一个人做出的关于前提类别是如何选择出来的假设。此外,我们的研究结果也为基于采样假设的重要研究提供了支持,与以前研究结论一致(Navarro et al., 2012; Ransom et al., 2016; Hayes et al., 2019),采样假设的核心问题是强采样和弱采样之间的区分和对比,这是一个非强即弱的问题。这也表明了社会知识背景的重要性。

## 5. 结论

采样假设影响归纳推理的典型性效应,通过在强采样和弱采样下都呈现高典型性和低典型性的前提证据来进行归纳推理,我们发现,无论是强采样还是弱采样条件都存在较强的典型性效应,典型性效应依赖于样本证据的采样假设,但强采样下的典型性效应有所衰减。我们的研究为归纳推理领域的研究提供了新的视角,采样假设对归纳推理的推理强度也有着很大的影响。

## 致 谢

非常感谢国家自然科学基金(61431013; 31671158)、中央高校基本业务费(SWU1709248)提供的资金支持,非常感谢龙长权教授对本研究提供的指导和帮助,非常感谢舒睿在实验设计、实验数据收集等方面的帮助。

## 参考文献

- 金鑫, 陈云(2014). 儿童归纳推理心理效应研究综述. *心理技术与应用*, (9), 21-24+28.
- Bailenson, J. N., Shum, M. S., Atran, S., Medin, D. L., & Coley, J. D. (2002). A Bird's Eye View: Biological Categorization and Reasoning within and across Cultures. *Cognition*, 84, 1-53. [https://doi.org/10.1016/S0010-0277\(02\)00011-2](https://doi.org/10.1016/S0010-0277(02)00011-2)
- Fernbach, P. M. (2006). Sampling Assumptions and the Size Principle in Property Induction. In R. Sun, & N. Miyake (Eds.), *Proceedings of the 28th Annual Conference of Cognitive Science Society* (pp. 1287-1292). Cognitive Science Society.
- Hayes, B. K., Navarro, D. J., Stephens, R. G., Ransom, K., & Dilevski, N. (2019). The Diversity Effect in Inductive Reasoning Depends on Sampling Assumptions. *Psychonomic Bulletin & Review*, 26, 1043-1050. <https://doi.org/10.3758/s13423-018-1562-2>
- Kemp, C., & Tenenbaum, J. (2003). Theory-Based Induction. In *The Proceedings of the Twenty-Fifth Annual Conference of the Cognitive Science Society* (pp. 658-663). Cognitive Science Society.
- Kemp, C., & Tenenbaum, J. B. (2009). Structured Statistical Models of Inductive Reasoning. *Psychological Review*, 116, 20-58. <https://doi.org/10.1037/a0014282>
- Kiran, S., & Thompson, C. K. (2003). Effect of Typicality on Online Category Verification of Animate Category Exemplars in Aphasia. *Brain and Language*, 85, 441-450. [https://doi.org/10.1016/S0093-934X\(03\)00064-6](https://doi.org/10.1016/S0093-934X(03)00064-6)
- Lawson, C. A. (2018). Knowing When to Trust a Teacher: The Contribution of Category Status and Sample Composition to Young Children's Judgments of Informant Trustworthiness. *Journal of Experimental Child Psychology*, 173, 380-387. <https://doi.org/10.1016/j.jecp.2018.04.003>
- Liang, X., Chen, Q., Lei, Y., & Li, H. (2016). How Types of Premises Modulate the Typicality Effect in Category-Based Induction: Diverging Evidence from the P2, P3, and LPC Effects. *Scientific Reports*, 6, Article No. 37890. <https://doi.org/10.1038/srep37890>



- Medin, D. L., Coley, J. D., Storms, G., & Hayes, B. L. (2003). A Relevance Theory of Induction. *Psychonomic Bulletin & Review*, *10*, 517-532. <https://doi.org/10.3758/BF03196515>
- Navarro, D. J., Dry, M. J., & Lee, M. D. (2012). Sampling Assumptions in Inductive Generalization. *Cognitive Science*, *36*, 187-223. <https://doi.org/10.1111/j.1551-6709.2011.01212.x>
- Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-Based Induction. *Psychological Review*, *97*, 185-200. <https://doi.org/10.1037/0033-295X.97.2.185>
- Proffitt, J. B., Coley, J. D., & Medin, D. L. (2000). Expertise and Category-Based Induction. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *26*, 811-828. <https://doi.org/10.1037/0278-7393.26.4.811>
- Ransom, K. J., Hendrickson, A. T., Perfors, A., & Navarro, D. J. (2018). Representational and Sampling Assumptions Drive Individual Differences in Single Category Generalisation. In C. Kalish, M. Rau, J. Zhu, & T. Rogers (Eds.), *Proceedings of the 40th Annual Conference of the Cognitive Science Society* (pp. 930-935). Cognitive Science Society.
- Ransom, K. J., Perfors, A., & Navarro, D. J. (2016). Leaping to Conclusions: Why Premise Relevance Affects Argument Strength. *Cognitive Science*, *40*, 1775-1796. <https://doi.org/10.1111/cogs.12308>
- Sanjana, N. E., & Tenenbaum, J. B. (2003). Bayesian Models of Inductive Generalization. In M. I. Jordan, Y. LeCun, & S. A. Solla (Eds.), *Advances in Neural Information Processing Systems* (pp. 59-66). Cambridge, MA: MIT Press.
- Shepard, R. N. (1987). Toward a Universal Law of Generalization for Psychological Science. *Science*, *237*, 1317-1323. <https://doi.org/10.1126/science.3629243>
- Sloman, S. A. (1993). Feature-Based Induction. *Cognitive Psychology*, *25*, 231-280. <https://doi.org/10.1006/cogp.1993.1006>
- Tenenbaum, J. B., & Griffiths, T. L. (2001). Generalization, Similarity, and Bayesian Inference. *Behavioral and Brain Sciences*, *24*, 629-641. <https://doi.org/10.1017/S0140525X01000061>
- Vong, W. K., Hendrickson, A. T., Perfors, A., & Navarro, D. J. (2013). The Role of Sampling Assumptions in Generalization with Multiple Categories. In *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 3699-3704). Cognitive Science Society.
- Voorspoels, W., Navarro, D. J., Perfors, A., Ransom, K., & Storms, G. (2015). How Do People Learn from Negative Experience? Non-Monotonic Generalizations and Sampling Assumptions in Inductive Reasoning. *Cognitive Psychology*, *81*, 1-25. <https://doi.org/10.1016/j.cogpsych.2015.07.001>
- Xu, F., & Tenenbaum, J. B. (2007). Sensitivity to Sampling in Bayesian Word Learning. *Developmental Science*, *10*, 288-297. <https://doi.org/10.1111/j.1467-7687.2007.00590.x>