

# 编码手册

## 1 编码手册说明

本册旨在规范编码流程，减少编码者之间的不一致性。

## 2 编码流程

编码流程主要分为编码阶段和校对阶段。

### 2.1 编码阶段

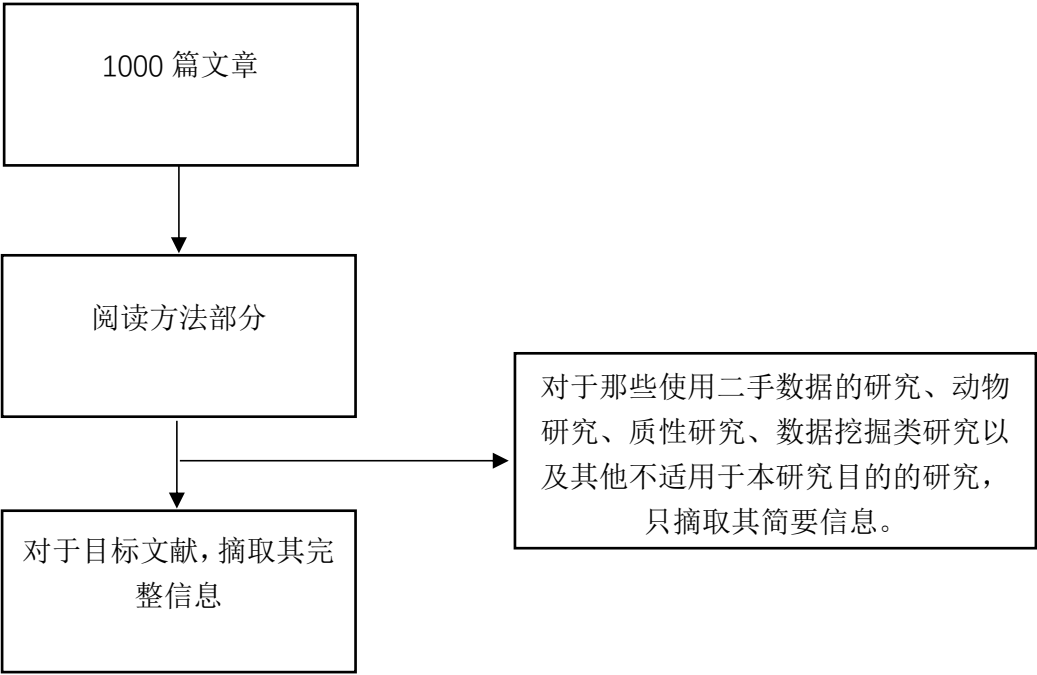


图 1 编码流程图

#### 2.1.1 方法部分<sup>1</sup>

首先，我们阅读方法部分以判断研究的数据是否是二手数据<sup>2</sup>、或者是不是

<sup>1</sup> 不管是摘取文章的详细信息，还是文章的简要信息。由于研究编号和被试组别所导致的同一研究在不同行记录时，即使信息是重复的也需要录入。

<sup>2</sup> 二手数据在数据来源是指通过公共或公开渠道取得的已存数据（柳武妹 等, 2022），其主要包括：政府数据、公共调查机构数据、商业机构调查数据以及其他研究机构或研究者的数据。

动物研究、质性研究、数据挖掘类研究以及其他不适用本研究目的的研究。如果是，我们仅记录其简要信息<sup>3</sup>。如果不是，我们将会根据本手册的第三部分摘录这篇文章的完整信息。

重复上步骤，直至将所有文献的编码完成。

## 2.2 校对阶段

(1) 编码阶段，每一篇文章的编码都由 2 人一组的编码者独立完成。

(2) 每篇文章编码完成后，借助代码和人工检查，对两人的编码内容是否一致进行打分（详见本文档“编码一致打分”）。对于出现分歧的编码内容，编码者会回到原论文进行核实，然后商定最终的编码结果；若编码者仍难以达成一致，项目负责人加入讨论，并决定最后编码结果。

(3) 编码过程中会定期（1~2 周）开展编码讨论会。讨论会中，编码者会对当前的编码情况进行讨论。

## 3 编码详解

- **文章编号**

每篇文章都有唯一的 ID（8 位数）。

- **题名**

文章的标题，摘取信息时请使用中文。

- **第一作者**

摘取信息时请使用中文。

- **期刊**

摘取信息时请使用中文。

---

<sup>3</sup> 简要信息包括：文章编号、题名、第一作者，期刊，研究编号、备注 1。

- 摘要

- 是否报告被试信息：0.未报告；1.若报告，填写报告的具体被试信息

- 如果被试包含大学生，在摘要中是否提及：0.未提及；1.提及

- 说明：被试信息包含：1.样本类型；2.受教育水平；3.职业；4.样本大小；5.性别；6.年龄；7.社会经济地位；8.民族；9.宗教信仰；a.被试生源地/籍贯/常住地/户籍区域；b.被试招募区域。

- 目标人群

- 目标人群：1.陈述的特定人群；2.陈述的一般人群；3.推断的一般人群。

- 编码依据：记录编码者做出上一条编码的依据，即摘取原文。

说明：目标人群的编码主要阅读文章的讨论部分，尤其是结论、研究局限性等地方。若未能找到，可以阅读摘要中的结论或文章的其他地方。下面提供可视化图和例子供大家理解目标人群的编码。

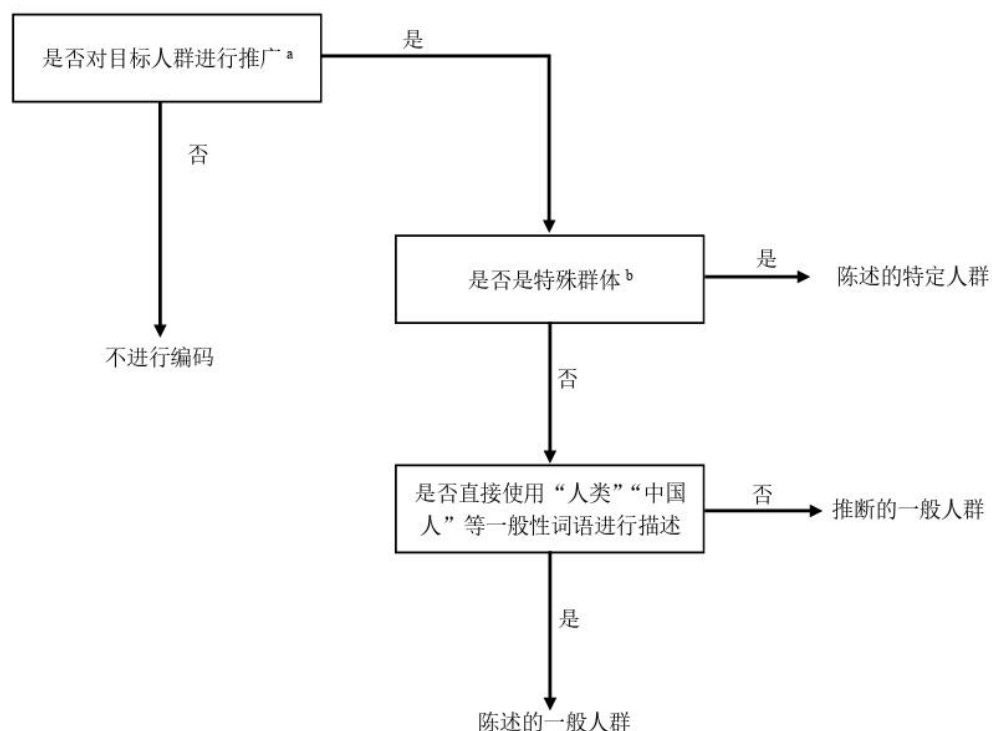


图 2 目标人群编码决策树

a. 以下是两个例子，用以阐明结论的推广情况：

例子 1：两个研究团队调查了中国青少年的孤独感，他们以上海青少年为样本。两个团队都发现了上海青少年中存在高水平的孤独感。团队 1 认为这项研究表明中国青少年普遍感到更高水平的孤独感。相反，团队 2 则认识到其样本的局限性，指出这项研究发现上海青少年孤独感较高，同时建议进一步研究以确定其他地区青少年是否也有相同的心理状态。在这种情况下，团队 1 将其发现从上海青少年推广到所有中国的青少年；而团队 2 则没有进行这样的推论。

例子 2：两个团队研究注意力的认知加工过程。团队 1 得出结论，并隐含地将其推广为人类注意力认知加工机制为 xxx。而团队 2 则将其结论限制在研究的具体发现上，指出在这项研究中，注意力认知加工机制为 xxx。团队 2 还承认，需要进一步探索才能确定研究中的该被试群体得到的结论能否

被推广到更为广泛的人群中。在这种情况下，团队 1 隐含地将其发现推广到所有人类，而团队 2 则没有进行直接的推广。

b. 特殊人群：这一类别包括但不限于婴儿、青少年、老年人、精神病患者、囚犯等人群。

#### 研究编号<sup>4</sup>:

说明：1.研究编号指摘录的编码信息来自文章的第几个研究；2.如果文章中有预研究，需要根据实际情况讨论后决定是否录入。若录入，预研究编码为 0，其他正式研究研究编号不变。如果有多个预研究，那么第一个预研究为 0，第二个为 0.1，第三个为 0.2，以此类推。3.如果一个大研究中出现若干小研究，比如：研究 2a，研究 2b，研究编号记为 2a，2b，并且使用单独行来记录信息。4.如果一篇文章中不同研究使用同一批被试，只需要记录使用该批被试的第一个研究，其余使用该批被试的研究仅需要记录研究编号，并且在备注 2 备注为：被试重复使用。

#### ● 备注 1:

说明：二手数据研究、质性研究、动物研究、挖掘数据类研究以及其他不适用于本研究目的的研究，并分别备注为：secondary, qualitative, animal, mining, inapplicable。其他不适用本研究目的的研究可以在 Remark3 写下你认为的原因。

#### ● 研究类型:

■ 1.实验；2.问卷；3.其他（填写文字）

说明：1.概念上讲，实验研究是指采用实验法进行的研究，实验法指的是在控制条件下对某种心理现象进行观察的方法。问卷研究是采用问卷法进行的研究，问卷法是指通过严格的书面调查表收集心理学变量数据的一种研究方法。2.若编码者无法判断使用了哪种方法，可以通过阅读

---

<sup>4</sup> 为提高数据的一致性，减少因研究编号和被试组别的不同而产生的数据冗余和数据不一致，在编码过程中，这两个条目同一编码小组内请务必保持一致。

论文的方法、摘要或文章的其他部分，根据作者的描述判断。

- **研究组别：**

说明：如果同一研究中被试的人口学信息按照不同组别进行报告（比如：青年组 vs 老年组），请分别在不同行记录被试的信息。如果有多层分组，请结合文章实际情况记录，务必保证被试信息的记录最完整、分组形式最简洁且组别之间是独立的。

- **样本类型：**

- 1.大学生（含硕博研究生）；2.非大学生的学生；3.婴儿、幼儿或学龄前儿童；4.非学生的成人；5.其他

- 摘取文章中关于描述被试类型的句子或文字。

说明：1.若有研究同时涵盖各类群体或未明确说明，则记为其他。2.若为追踪研究，被试在不同追踪时间上身份有所转变，样本的类型为研究所陈述的主旨人群为主。3.在职硕博研究生需编码为“4.非学生的成人”。4.摘取原文信息时请务必保证是完整的句子且句子能体现出被试的类型。

- **受教育水平：**

- 0.未报告；1.报告

- 受教育水平信息：编码者摘取原文关于被试受教育水平的描述。

说明：只有当样本类型为4或5时才填写该项编码。

- **职业：**

- 0.未报告；1.报告

说明：只有当样本类型为4或5时才填写该项编码。

- **样本大小：**

说明：1.样本大小优先记录研究的有效被试量。若论文未提供有效被试

量，请记录未排除无效被试的样本量。2.编码时若文章同时具有表格和文字描述，以文字描述为准。

● **性别：**

■ 0.未报告；1.报告

■ 男性人数（占比）/女性人数（占比）

说明：1.如果文章描述上同时拥有男性和女性人数以及男性和女性占比，请优先记录人数。2.若文章仅提供其中一类性别的人数或人数占比，请计算出未报告的另一类性别人数或人数占比并填写。

● **年龄：**

■ 0.未报告；1.报告

■ 如果报告，请摘录年龄的详细信息。

说明：1. 如果文章同时报告年龄的多种指标，尽可能都进行记录。2. 如果文章的年龄信息以月、日（或天）为单位，请转化以年为单位进行填写，转化后的结果请保留两位小数。

● **社会经济地位：**

■ 0.未报告；1.职业；2.受教育程度；3 收入；4.主观社会经济地位；5.其他

■ 如果报告类别为其他，请摘取文内相关信息。

说明：1.社会经济地位，也称为社会阶层、社会分层（王淑燕 等, 2023; 杨沈龙 等, 2020）。社会经济地位可以分为客观的和主观的。其中客观的社会经济地位指的是高社会经济地位的个体拥有更多的金钱、更高的受教育程度以及职业声望。而主观的社会经济地位指的是个体感知到的自己在社会中相对于他人所处的社会等级。2.学生或未成年人的社会经济地位常常使用父母的社会经济地位进行表示。所以，当被试为学生或未成年人，文章中又报告

了父母的社会经济地位信息，此时请录入相应的类别。3.若被试类型为4或5时，该处编码内容可能会与编码条目“受教育水平”“职业”重复，此时仍需要进行编码。4.若研究报告多个类别，此时类别间使用英文状态下的分号隔开（比如：同时报告了职业和受教育程度，则为：1;2）。

- **民族：**

- 0.未报告；1.报告
- 如果报告，摘录其信息，比如：汉族;藏族

- **宗教信仰：**

- 0.未报告；1.报告
- 若报告，摘录其详细信息。

说明：宗教信仰的信息类别分为：1.佛教；2.基督教；3.伊斯兰教；4.其他宗教信仰。记录时只需要记录相对应的数字即可。

- **被试区域<sup>5</sup>：**

- 0.未报告；1.户籍；2.籍贯；3.常住地；4.生源地
- 若报告，请摘录其详细信息。摘录格式为：户籍(填写原文信息);籍贯(填写原文信息)

说明：由于某些信息暗含着被试的区域信息，所以编码者在编码过程中只有当文章出现上述区域字眼并且有具体的信息时才算报告并录入信息，本条尽量不要推测。

- **被试招募区域：**

- 0.未报告；1.省级行政区域/大区（填写文字）；2.非中国样本

说明：1.省级行政区域的编码只需要记录简要名称，比如：香港特别行政区记为香港，湖北省记为湖北，广西壮族自治区记为广西。2.如果一个研究包含多

---

<sup>5</sup> 在区域编码中，不同区域之间使用英文状态下的分号隔开。



个省级行政区域，每个省级行政区域都要记录，并在同一单元格内以英文状态下的分号隔开。3.若研究描述多个省级行政区域，但是仅仅列举其中一些，比如：“被试来源包括广东、北京、上海、四川、浙江、湖南、湖北、陕西、重庆等全国 26 个省份/地区”，这种情况仅记录被试提及的省级行政区域，并在备注 2 备注：区域报告不完整。4.若研究以大区字样进行描述<sup>6</sup>，比如：华南地区、华北两个省份、东部地区、东部某省等，可以填写作者的描述。

- **抽样方法：**

- 0.未报告；1.非概率抽样；2.概率抽样

说明：1.非概率抽样指的是研究者有目的挑选一部分个体组成样本或是指总体中某些元素没有被选择的机会，亦或是无法准确确定选择概率的任何抽样方法。其包括：方便的、随意的、偶然的抽样；滚雪球抽样、配额抽样、连续抽样等（各类抽样方法的含义，可参考网页：[https://en.wikipedia.org/wiki/Nonprobability\\_sampling](https://en.wikipedia.org/wiki/Nonprobability_sampling)）。2.概率抽样指的是总体中每个个体被抽中的概率是已知的。概率抽样包括：简单随机抽样、系统抽样、分层抽样、整群或多阶段抽样、概率与大小成比例的抽样等（各类抽样方法的含义，可参加网页：[https://en.wikipedia.org/wiki/Sampling\\_\(statistics\)](https://en.wikipedia.org/wiki/Sampling_(statistics))）

- **被试招募及数据收集方式：**

- 0.未报告；1.非线上方式；2.线上方式

- 若报告为线上方式，请记录线上平台名。多个平台名之间用英文状态下的分号隔开。若无平台名，则不进行记录。

说明：只有当被试使用线上渠道招募，并且数据采集也使用线上方式时（比如使用 Mturk、问卷星、脑岛等平台，通过平台的系统进行数据收集，主试与被试完全没有接触），才记录为 2。如果被试招募方式和数据采集其中只要有一个能判断为非线上，就记录为 1。排除以上两种情况，则记为 0。

---

<sup>6</sup> 以地理区域划分，中国一般分为七大地理区域，分别是：华东、华南、华北、华中、东北、西北、西南。以经济划分，中国一般分为四大区域，分别是：西部、东部、中部、东北。

- **备注 2:**

说明：1.备注 2 填写本手册除备注 1 以外规定的其他必须备注的内容。2.多条备注请在每条备注前添加序号。

- **备注 3:**

说明：备注 3 中编码人员可以备注自己认为编码时要记录的事项。2.多条备注请在每条备注前添加序号。

## 参考文献

杨沈龙, 喻丰, 胡小勇, 郭永玉. (2020). 心理学研究中社会阶层的操作化界定及其衍生问题. *心理科学*, 43(2), 505–511.

柳武妹, 杨巧英, 康娜. (2022). 二手数据用于多方法研究的建议——以消费者行为研究为例. *外国经济与管理*, 44(11), 121–137.

王淑燕, 刘启鹏, 江伟. (2023). 社会经济地位的测量与研究操纵：现状与问题——基于心理学领域的研究. *心理研究*, 16(1), 12–20.