**The Coding Manual**

**1 Coding Manual Instruction**

The primary aim of this coding manual was to standardize the coding procedure and eliminate inconsistencies across coders.

**2 Coding Procedure**

The whole process of coding is divided into two stages: coding and proofreading stages.

**2.1 Coding Stage**

145 articles published in *Psychological Science* in 2022 and XXX articles of big-team science

For studies using large datasets and secondary data, animal studies, case studies, and data mining studies, public data on demographic variables are not searched or extracted, and only brief information is extracted during coding.

Reading the method section

Check to see if the article makes the data publicly available. For articles that do not provide original data, obtain the data by emailing the author.

Filter the articles that have access to the raw data for the demographic variables

Merge data from demographic variables

Extract demographic variables from these articles

Extracting the target articles’ complete information

Figure 1. The procedure of the coding stage

2.1.1 Reading the methods section

First, we read the methods section to determine whether the research data come from large datasets or secondary data, or whether they come from animal studies, case studies, or data mining studies. If so, we will only record their short information (see the second code in the first table of the "code\_example" file). If not, we will extract the full indicators of this article from the third part of this manual (see the first table in the "code\_example" file and the first code in the second table).

2.1.2 Extraction of raw demographic variable data

For articles with access to raw data (publicly available or obtained by emailing the authors), we extract data on demographic variables from each article, organize a data file for each study containing only the subject ID and demographic variables, and finally merge data from all articles published in *Psychological Science* in 2022 and big-team science articles into a consolidated dataset for subsequent data analysis.

2.2 Code proofreading

(1) To further increase the reliability and objectivity of the coding content, we will redistribute the articles for proofreading after the coding stage is completed.

(2) The coders will be divided into groups, with two coders in each group proofreading the same documents. If there is an inconsistency between two independent coders, we will go back to the article and resolve the inconsistency through discussion. If an agreement cannot be achieved between two coders, a third coder, a senior author, will join the discussion to make a final decision.

**3 The Detailed Information of Code Dimensions[[1]](#footnote-1)**

* **Article IDs:** Each article has a unique ID
* **Article Title:**
* **Study Number:**

*Note: 1. The study number refers to this number of studies reported in the article; 2. If the article includes a pre-test, the pre-test is only recorded if it is written in the form of a formal study. In that case, the pre-test study is coded as 1, and the subsequent studies are numbered sequentially; 3. If multiple studies in an article use the same batch of participants, only the first study must be recorded, and the remaining studies only need to be recorded with the study number and a note in the remark section that the samples were used repeatedly.*

* **Participants’ Group**

*Note: If multiple groups of samples were reported, we will code them separately, and all demographic information of each group will be recorded in detail.*

* **Study Type**
  + - 1. Experimental/quasi-experimental; 2. Questionnaire; 3. Qualitative research; 4. Others
* **Sample Type**
  + - 1. University students (including graduate students); 2. Students but not university students; 3. Infants and toddlers; 4. Preschool children; 5. Adults who are not students; 6. Others
    - Extract sentences or words of papers that associated with sample type.

*Note: 1. The age range for infants is 0-1 years, the age range for toddlers is 1-3 years, and the age range for preschoolers is 3-6 years. If there is a study that includes all types of groups or is not clearly stated, it is recorded as other. 3. If it is a tracking study, the identity of the subjects has changed at different tracking times and the sample type is mainly the main target population stated in the study. If this cannot be determined, it is the subject's last type of participation in the study, and other demographic information is also based on the last participation. 4. Master's and Ph.D. students should be coded as "4. Non-student adults". When extracting information from the original text, please ensure that it is a complete sentence and that the sentence reflects the type of subject.*

* **Sample Size**

*Note: 1. Sample size refers to the number of valid samples involved in the study; 2. Textual descriptions take priority over tabular descriptions when coding articles with both types of descriptions.*

* **Gender**
  + - 0. Not reported in the article or supplementary materials, and the original data cannot be obtained; 1. Reported in the article or supplementary materials, but the original data cannot be obtained; 2. Raw data can be obtained.
    - If reported in the article or supplementary materials, but the original data are not available, record: Number of men (%) / Number of women (%) / Number of other gender(%)

*Note: 1. If the article describes the number of males, females, and other genders, as well as the proportion of each gender, please record the number of people. 2. If the article gives only the proportions of males, females, and other genders, please record the proportions separately.*

* **Age**
  + - 0. Not reported in the article or supplementary materials, and the original data cannot be obtained; 1. Reported in the article or supplementary materials, but the original data cannot be obtained; 2. Raw data can be obtained.
    - If reported in the article or supplementary materials, but the original data are not available, please extract details of age.

*Note: Please extract the specific description of age in the article if neither the M±SD nor the range of the sample’ age is mentioned.*

* **Socioeconomic Status**
  + - 0. Not reported in the article or supplementary materials, and the original data cannot be obtained; 1. Reported in the article or supplementary materials, but the original data cannot be obtained; 2. Raw data can be obtained.
    - If reported in the article or supplementary materials, but the original data are not available, extract the type of SES information

*Note: The types of socioeconomic information include: 1. Income or other property; 2. Subjective social class. When coding, simply record the corresponding number. Record as 12 if the article contains both types of participants’ information.*

* **Educational Attainment**
  + - 0. Not reported in the article or supplementary materials, and the original data cannot be obtained; 1. Reported in the article or supplementary materials, but the original data cannot be obtained; 2. Raw data can be obtained.
    - If reported in the article or supplementary materials, but the original data are not available, extract the type of reported information

*Note: 1. Enter this code only if the sample type is 4 or 5; 2.the types of information on education level include: 1. Lower than college; 2. College and higher. 3. If studies involve participants with both educational backgrounds, record as 12.*

* **Ethnicity**
  + - 0. Not reported in the article or supplementary materials, and the original data cannot be obtained; 1. Reported in the article or supplementary materials, but the original data cannot be obtained; 2. Raw data can be obtained.
    - If reported in the article or supplementary materials, but the original data are not available, extract the ethnicity information (e.g., white, black)
* **Occupation**
  + - 0. Not reported in the article or supplementary materials, and the original data cannot be obtained; 1. Reported in the article or supplementary materials, but the original data cannot be obtained; 2. Raw data can be obtained.

*Note: 1. Enter this code only if the sample type is 4 or 5.*

* **Religious**
  + - 0. Not reported in the article or supplementary materials, and the original data cannot be obtained; 1. Reported in the article or supplementary materials, but the original data cannot be obtained; 2. Raw data can be obtained.
    - If reported in the article or supplementary materials, but the original data are not available, extract the detailed information

*Note: Religious information is classified as follows: 1. Buddhism; 2. Christianity; 3. Islam; and 4. Other. Only the corresponding numbers need to be recorded.*

* **National** 
  + - 0. Not reported in the article or supplementary materials, and the original data cannot be obtained; 1. Reported in the article or supplementary materials, but the original data cannot be obtained; 2. Raw data can be obtained.
    - If reported in the article or supplementary materials, but the original data are not available, extract the type of national information.

*Note: 1. The coding record for a country or region is a two-letter ISO 3166-1 code, for example: China is coded as CN; 2. When a study includes multiple countries or regions, each country or region coded with a specific country/region is recorded separately. 3. If a study does not identify a specific country or region, but only a larger area, the region is recorded directly, for example, "Europe".*

* **Sampling Method**
  + - 0. Unreported; 1. Convenience sampling; 2. Random sampling; 3. Others
* **Participant Recruitment Method**
  + - 0. Unreported; 1. reported.
    - If reported, record specific methods. For online data collection, record the name of the online platform. If the platform is not available, record it as "crowdsourcing". For offline data collection, record it as "offline". The different methods are separated by commas.
* **Abstract**
  + - Whether articles mentioned the participants’ demographic information (0. Not mentioned; 1. Mentioned)
* **Generalization Statement**
  + - Target population: 1. stated specific population; 2. stated general population; 3. inferred general population.
    - Target population coding basis: exacting sentences/words excerpted from the full text of the paper that are associated with the statement about the target population
    - Restricted generalizability: 0. No relevant description; 1. Relevant description exists
    - Restricting generalizability coding basis: Record the basis on which the coder restricted the generalizability coding, i.e., extracted the original text.

*Note: 1. For the purpose of coding the target population and describing further restrictions on generalizability, the discussion and conclusion sections of the article are primarily read, but sometimes obvious judgments can be made in the title, abstract, and other parts of the article. The coder can make judgments quickly by integrating multiple pieces of information. 2. For coding the target population, a decision tree is provided below for everyone to understand.*

yes

stated specific population

Is it a specific population?

no

inferred general population

no

Are generic terms such as "human" used directly to describe?

stated general population

yes

Figure 2. Target population decision tree

* **Remarks**

*Note: 1. Please note large data sets, secondary data, case studies, animal studies, and mining data studies separately as database, secondary data, case, animal, and mining data; 2. For multiple notes, please add the serial number before each note; 3. In addition to the mandatory notes, coders can note what they want to record when they coding.*

1. The codes or information are separated by a semicolon in the English state if more than one need be recorded in

   the same section. [↑](#footnote-ref-1)