

自我参照的神经成像认知本体论数据集

孙淑婷¹, 王楠², 温佳慧¹, 胡传鹏^{1*}

1. 南京师范大学心理学院, 南京 210024

2. 南京师范大学外国语学院, 南京 210024

* 论文通信作者: 胡传鹏 (hu.chuan-peng@nnu.edu.cn)

摘要: 自我参照 (self-reference, 或自我参照加工) 指人们加工与自己相关的信息时的认知过程。认知神经科学领域对“自我参照”进行诸多研究以理解人类自我认知的神经基础。但是, “自我参照”一词在不同研究中是否表征着相同的心理过程? 这一根本问题鲜受关注。本研究尝试初步建立自我参照的本体论数据集以检验该问题。使用规范化流程系统地检索文献并筛选文章, 随后两位独立的编码者对纳入文章在行为和神经层面对自我参照的操作化定义进行编码并标准化, 形成“自我参照的神经成像认知本体论数据集”。本数据集来自 66 篇神经成像论文, 包括行为与神经层面对自我参照的操作化定义 (以 CSV 文件格式保存)、不同操作化定义下自我参照所激活大脑区域的坐标数据 (以 BrainMap 格式保存)、说明手册。数据编码的一致性分析表明, 编码结果可靠。与自动化元分析数据库 Neurosynth 相比, 本数据集的论文筛选更加精准, 且能够比较自我参照的不同操作化定义所激活脑区的异同, 为理解自我参照的神经基础提供更精确的结果。本数据集为深入理解人类自我认知的神经机制提供基础, 也为其他类似的元研究数据集的创建提供参考, 促进认知本体论的研究。

关键词: 神经成像; 自我参照加工; 元研究; 开放数据; 认知本体论

A cognitive ontological dataset for neuroimaging studies of self-reference

Sun Shu-Ting¹, Wang Nan², Wen Jia-Hui¹, Hu Chuan-Peng^{1*}

1. School of Psychology, Nanjing Normal University, 210024 Nanjing, China
2. School of Foreign Languages and Cultures, Nanjing Normal University, 210024 Nanjing, China

*Email: hu.chuan-peng@nnu.edu.cn

Abstract: Self-reference (or self-referential processing) refers to the cognitive processes underlying self-related information processing. It is widely studied in cognitive neuroscience to understand the neural basis of self-cognition of human beings. However, does the term “self-reference” mean the same psychological processes across studies? This fundamental question is largely neglected. To fill the gap, we built an ontological dataset based on neuroimaging studies of self-reference. We searched the literature and screened the articles following a standard protocol. Then, two independent coders extracted data and standardized operationalizations of self-reference on both behavioral and neural levels, resulting in a cognitive ontological dataset for neuroimaging studies of self-reference. This dataset consists of operationalizations of self-reference (in CSV file format) from 66 neuroimaging articles, coordinates data of brain areas activated by self-reference (in BrainMap format), and corresponding codebooks. The inter-rater reliability analysis suggests that the quality of the coding process is excellent. Compared to automatic meta-analytical platforms, i.e., Neurosynth, the current dataset provides a fine-grained granularity in article selection, which allows the comparison of brain regions activated by different operationalizations of self-reference. This dataset lays a foundation for the understanding of neural mechanisms underlying self-cognition. It may also facilitate the study of cognitive ontology by providing an example for creating similar metascience datasets.

Keywords: Neuroimaging; Self-reference; Metascience; Open data; Cognitive ontology

数据集基本信息简介

数据集（集）名称	自我参照的神经成像认知本体论数据集
数据通信作者	胡传鹏（hu.chuan-peng@nnu.edu.cn）
数据作者	孙淑婷、王楠、温佳慧、胡传鹏
数据时间范围	1990 – 2021年
数据量	21.1MB
数据格式	*.txt, *.csv, *.Rmd, *.xlsx, *.nii
数据服务系统网址	http://doi.org/10.57760/sciencedb.j00001.00469
数据库（集）组成	除说明文件“README.txt”外，数据集包括4个部分：纳入论文的信息及其手册、自我参照的操作化定义及其手册、神经成像的坐标点数据、本论文中方法与结果的补充材料。本数据集纳入66篇文献（包括72个实验、2056个人类被试）。“自我参照_文章信息.csv”（英文版

	<p>“Self_Ref_Article_Info.csv”) 包含文章作者、出版年份、期刊、样本量等，其说明手册文件为“手册_自我参照_文章信息.csv”(对应英文版“Codebook_Self_Ref_Article_Info.csv”)。“自我参照_操作化定义.csv”(英文版“Self_Ref_Operationalization.csv”)包含自我参照操作化定义的信息，如实验刺激、实验设计、实验任务等，其说明手册文件为“手册_自我参照_操作化定义.csv”(英文版“Codebook_Self_Ref_Operationalization.csv”)。文件夹“Self_Ref_Foci_Raw”中包括72个TXT文本，对应66篇文献报告的72个实验中的激活坐标点信息。“Suppl_Materials”文件夹包含本论文的补充材料。</p>
--	---

Dataset Profile

Title	A cognitive ontology dataset for neuroimaging studies of self-reference
Data corresponding author	Hu Chuan-Peng (hu.chuan-peng@nnu.edu.cn)
Data authors	Sun Shu-Ting, Wang Nan, Wen Jia-Hui, Hu Chuan-Peng
Time Range	1990 – 2021
Data volume	21.1MB
Data format	*.txt, *.csv, *Rmd, *xlsx, *nii
Data service system	< http://doi.org/10.57760/sciencedb.j00001.00469 >
dataset composition	<p>Besides the “README.txt” file, the dataset consist of four parts: a CSV file for article information and its codebook, a CSV file for operationalizations of self-reference and its codebook, files for coordinates data of activated brain regions, and supplementary materials for methods and results of the current article. The dataset are from 66 articles (72 experiments and 2056 human participants). “Self_Ref_Article_Info.csv” (Chinese version “自我参照_文章信息.csv”) contains articles’ information, including author, year of publication, journal, sample size etc; “Codebook_Self_Ref_Article_Info.csv” (Chinese version “手册_自我参照_文章信息.csv”) is its codebook. “Self_Ref_Operationalization.csv” (Chinese version “自我参照_文章信息.csv”) contains information of operationalizations of self-reference, including stimuli, experimental designs, experimental task etc; “Codebook_Self_Ref_Operationalization.csv” (Chinese version “手册_自我参照_操作化定义.csv”) is its codebook. The folder “Self_Ref_Foci_Raw” contains 72 TXT files, corresponding to coordinates data from 72 experiments reported by 66 articles. The folder “Suppl_Materials” includes supplementary materials for the current paper.</p>

引 言

自我参照 (self-reference, 或自我参照加工) 是认知科学与认知神经科学中常用的概念, 通常指人们处理与自我相关的信息时的认知过程^[1]。先前研究发现, 人类加工与自我相关信息时, 腹内侧前额叶皮质 (ventromedial prefrontal cortex, vmPFC) 和后扣带皮层 (posterior cingulate cortex, PCC) 等脑区会特异性地激活^[4,5]。自我参照被认为是大脑默认网络 (default model network, DMN) 的重要功能^[6,7]。精神疾病领域的神经成像研究也指出, 自我参照相关脑区的功能异常与多种精神疾病相关, 如抑郁症^[8]、精神分裂症^[9]、自闭症^[10]等。因此, 理解自我参照的神经机制是认知神经科学的重要问题之一。

值得注意的是, 自我参照的神经成像结果受到实验中操作的影响。例如, 将自我与亲密他人相比, 自我相关信息在右侧前额叶皮层 (right lateral prefrontal cortex, rLPFC) 激活更强; 将自我与非亲密他人相比, 则在左侧前额叶皮层 (medial prefrontal cortex, MPFC) 激活更强^[11]。信息的内容也可能会影响到自我参照所激活的脑区: 身体自我信息的处理与右半球的外侧脑区密切相关, 而心理自我信息的处理主要激活皮质中线结构^[12]。自我相关信息呈现的感觉通道也会自我参照加工的激活脑区产生影响: 自我信息使用声音呈现时, 后楔前叶 (posterior precuneus cortex, PCC) 更多地被激活^[13]; 若信息使用视觉方式呈现, 则背外侧前额叶 (dorsolateral prefrontal cortex, DLPFC) 与海马旁回后部 (posterior parahippocampal gyrus) 激活程度更高^[14]。此外, 朱滢等发现, 亲密他人参照与自我参照所激活大脑区域受到被试文化背景的调节^[15]。

上述研究指向“自我参照”这一概念的本体论承诺问题, 即这一概念是否代表某一客观存在的实体 (entity), 如某一特定的心理加工过程或者某一特定的大脑活动模式。从已有的认知本体论研究结果来看, 当前认知科学与认知神经科学中普遍缺乏对概念的严格审视, 同一概念的操作化定义在不同研究中的差异较大^[16]。同一概念在操作化定义上的不一致可能会危害测量的效度 (validity), 也可能是心理科学与认知神经科学中可重复性问题的^[17-19]重要原因之一。在对自我参照的神经机制的探索中, 操作化定义差异意味着研究者难以建立自我参照与脑区/脑网络的稳定关系。

自动化的元分析平台/数据库无法解决本体论承诺的问题。例如, Neurosynth 数据库通过自动摘取神经成像论文中的数据点和关键词, 然后进行自动化的元分析, 建立大脑不同区域和认知任务之间的关系。但该数据库通过算法搜索摘要中出现的术语, 无法提取对认知过程的操作化定义的相关细节^[20]。最近, 研究者使用监督学习建立 Neuroquery 数据库, 通过检索数据库中论文全文, 寻找语义间的相似性, 然后加权组合具有紧密联系的术语所对应的激活坐标, 以预测不同术语可能激活的脑区^[21]。但 Neuroquery 未考虑实验操作化定义, 其建立语义相似性的词汇包括论文的前言和讨论, 这些语义相似性可能更多地反映了研究者的观念, 而非认知过程本身的相似性。因此, 目前尚无数据库能够帮助研究者探索自我参照这一概念的本体论承认问题。

为解决这一问题, 本研究从元研究的视角出发, 严格地检验自我参照的操作化定义, 并进行标准化, 建立自我参照的神经成像本体论数据集。本研究的核心在于整理可能影响自我参照认知过程的操作化定义, 建立标准化的分类索引并使用通用数据格式保存相关信

息。基于本数据集进行的元分析结果将能够揭示自我参照的不同操作化定义带来的差异，推动对人类自我认知的理解与理论构建。

1 数据采集和处理方法

1.1 数据采集

本文按照《开放式荟萃分析的规范化报告》^[21]，对 PubMed 和 Web of Science 进行系统性的文献搜索(最后检索日期 2021 年 12 月 4 日)。具体而言，本文对于自我参照的文献采用关键词“self-referen*”，并使用“AND”将该关键词与“fMRI”和“PET”连接，搜索主题、关键词及摘要包含关键词组合的文献。为尽可能完整地纳入所有涉及自我参照的文献，本文也参考相关的元分析^[23-25]和综述^[7, 26-30]。

对检索得到的全部文献进行筛选时，采用如下标准：

- (1) 研究使用 fMRI 或 PET 扫描；
- (2) 研究必须属于实证研究，而非元分析和文献综述类论文；
- (3) 作者使用英语作为写作语种且文章已正式发表在学术期刊上或存放于预印本平台；
- (4) 实验以健康成年人为研究对象。仅包含神经疾病或其他精神异常以及躯体疾病被试的研究被排除在外；若研究同时包含健康被试与患有疾病的被试，则仅保留健康被试的数据；以年轻和中年成年人为主要被试群体（平均年龄为 18–59 岁），从而避免年龄变化导致的自我相关信息的加工的异常^[31]。
- (5) 使用标准的大脑空间报告空间坐标结果（Talairach 或者 MNI）的研究被纳入数据集；研究需完整报告激活坐标，若研究的坐标数据不完整，研究者将通过邮件向作者询问，排除无回复者。为解决两个标准空间坐标之间的转化问题，本文使用 Lancaster 转换算法^[32]将 Talairach 空间坐标的研究结果转化成为 MNI 的空间坐标结果；
- (6) 研究必须包含全脑分析结果，仅包括部分脑区或使用感兴趣区（region-of-interest, ROI）分析的研究被排除；
- (7) 如果不同文献的结果来自同一组被试的数据，则选择其中一篇文献的数据进入元分析。

本数据集依据 PRISMA^[33]进行文献筛选，具体筛选过程如图 1 所示。

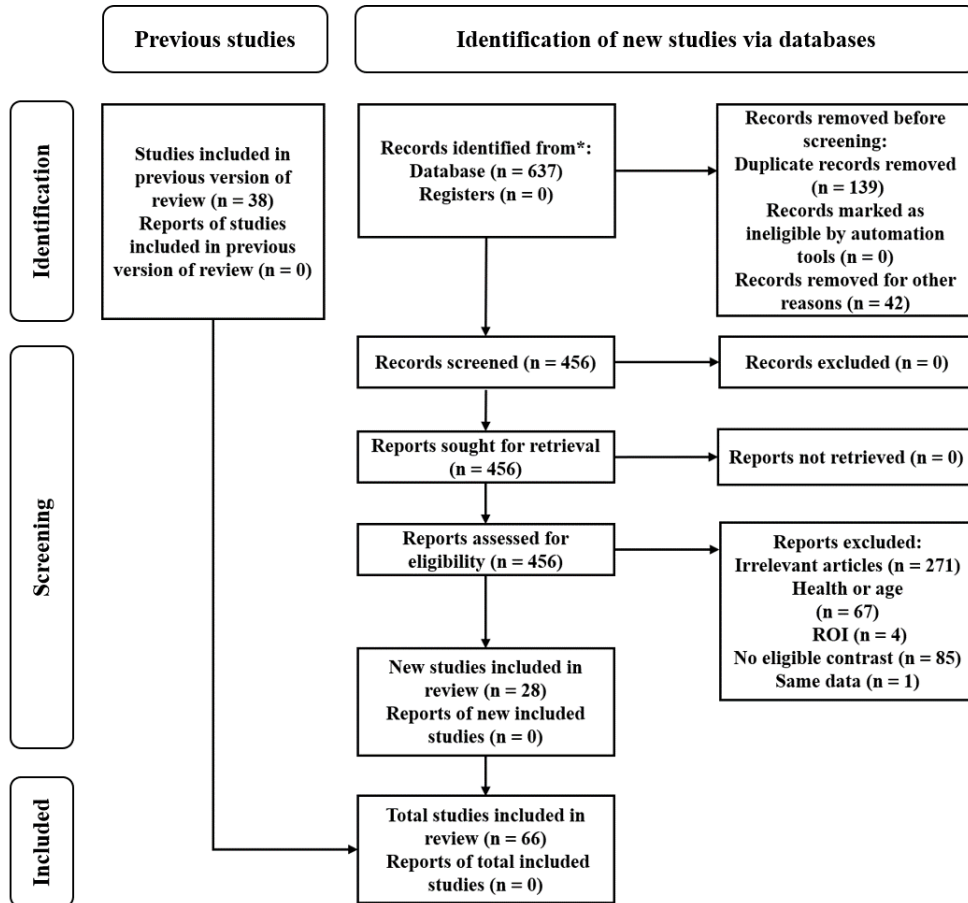


图1 对自我参照的 fMRI 与 PET 文献搜索及筛选的流程图^[22,33]。注：“Health or age”表示“被试是非健康群体，或年龄小于 18 岁或大于 59 岁”；“ROI”表示 regions of interest，即“研究仅报告感兴趣区分析的脑成像坐标”；“No eligible contrast”表示研究不包括自我参照相关的脑成像结果；“Same data”表示有两个研究使用同一批被试的数据，仅一篇文章被纳入。

Figure 1 Flowchart of fMRI and PET literature search and screening for self-reference ^[22, 33]. Note : “Health or age” indicates that studies are excluded when participants were unhealthy, younger than 18 years old or older than 59 years old; “ROI” means region of interest, which indicates that studies are excluded when results from region of interest (ROI) analyses were reported only; “No eligible contrast” indicates that studies are excluded if they did not contain brain imaging results for self-reference; and “Same data” indicates that when two studies used data from the same group of participants, only one study is included.

1.2 数据处理

对纳入文章的编码包括四个阶段。第一阶段，编码手册的编制，两位编码者和通讯作者共同编制初步的编码手册；随后两位编码者依据初步的编码手册对少量（3~5 篇）文献进行独立的预编码，比较预编码结果并与通讯作者讨论，改进编码手册；本阶段经过多次迭代，形成最终编码手册。第二阶段，两位编码者根据最终编码手册进行独立的数据摘录。第三阶段，两位编码者对摘录的数据分别进行编码一致性的评分，并对摘录内容进行汇总和统一。第四阶段，对所有纳入文章按照最终编码手册进行再次检查和汇总。为确保脑成

像数据摘录的准确性，对大脑激活坐标点数据的收集也经过三个阶段：第一阶段两位编码者独立摘录，第二阶段两者独立核对，核对后讨论不同，第三阶段整合汇总所有坐标，并进行分类。

2 数据样本描述

2.1 命名格式

本数据集包括四个部分（除“README”说明文件外）。第一部分为文献信息记录，文件名为“自我参照_文章信息.csv”（英文版“Self_Ref_Article_Info.csv”），格式为 csv，包括入选文献的基本信息，如作者、期刊、样本量等，其说明手册为“手册_自我参照_文章信息.csv（英文版为“Codebook_Self_Ref_Article_Info.csv”）”。第二部分为整理后的自我参照操作化定义信息，格式为 csv，记录入选文献中对自我参照效应的操作化定义，其命名形式为“自我参照_操作化定义.csv”（英文版“Self_Ref_Operationalization.csv”），其说明手册为“手册_自我参照_操作化定义.csv”（英文版为“Codebook_Self_Ref_Operationalization.csv”）。第三部分为 fMRI 激活坐标点数据，统一保存于“Self_Ref_Foci_Raw”文件夹之中，使用与 BrainMap 数据库^[34]相同的文件格式，以 TXT 格式文件保存，其命名形式为“FirstAuthor_Year_JournalAbbreviation.txt”，其中 FirstAuthor 代表文章的第一作者的姓氏，Year 代表文章的发表年份，JournalAbbreviation 代表文献发表期刊的简写。例如，“Hornung_2019_FrontBehavNenurosci.txt”代表该文章发表于 2019 年，第一作者的姓氏为 Hornung，发表期刊为 Frontiers in Behavioral Neuroscience。第四部分为本论文的方法与结果的补充材料，保存于“Suppl_Materials”文件夹。

2.2 数据样本

本数据集的主体数据来自 66 篇文献，共包括 72 个实验，2056 个被试。数据文件包括 8 个 csv 文件和 72 个 TXT 文件。其中 8 个 CSV 文件分别是文献信息和自我参照的操作化定义的中英文版本及其对应的说明手册；TXT 文件为激活坐标点数据。这两部分数据中，文献编号采用统一编号。此外，本数据集的在线版本中，还包括对本数据集的文件夹结构进行说明的 README.txt（内容见图 2A）。

文章信息数据包括文章编号、第一作者姓氏、出版年份、期刊等文献信息以及样本量、性别、年龄等被试相关的信息（见图 2B）。

操作化定义数据包括文章编号、实验设计、实验任务等（见图 2C）。本部分信息编码了各个文章对自我参照加工的操作化定义细节。激活坐标点数据采用文本文件（.txt），使用 BrainMap 格式（又称“Sleuth 格式”）摘录实验简要信息与脑成像结果的坐标点信息。单个文本文件仅记录来自同一篇文章的坐标点信息，按照全脑分析的结果进行分组。Sleuth 格式使用“//”来作为行的分割符。通常第一行记录文章采用的脑成像坐标模板，“// Reference=MNI”代表 MNI 坐标模板；第二行记录实验相关信息；第三行记录样本量，“// Subjects =14”表示参与脑成像扫描的样本量为 14；第四行为额外的标记信息，本数据集在这一行记录了坐标点信息对应的标准化比较(contrast)的名称，用于保存自我参照在神经成像层面的操作化定义；从第五行开始记录坐标点信息，每行为一个坐标点，从左往右

依次为 x、y、z，每行单个数值用通配符隔开。通常一个实验中会报告多个全脑分析的结果，通过空行分割不同的结果（见图 2D）。

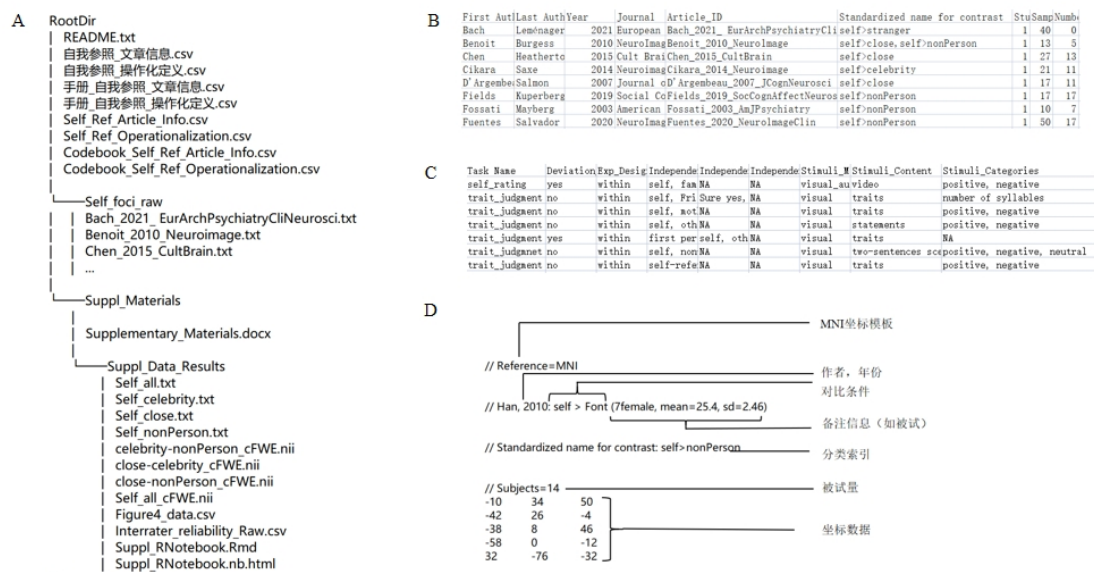


图 2 数据样例截图: (A) “README.txt”中关于本数据集的文件夹结构的截图; (B) “Self_Ref_Article_Info.csv”的部分截图; (C) “Self_Ref_Operationalization.csv”的部分截图; (D) 单个研究的激活坐标点数据的 BrainMap 格式

Figure 2 Screenshot of data sample : (A) The folder structure of the dataset in “README.txt”; (B) Screenshot of “Self_Ref_Article_Info.csv”; (C) Screenshot of “Self_Ref_Operationalization.csv”; (D) BrainMap format for coordinates data of each study

数据说明手册涵盖记录文献信息和操作化定义时所采用的变量，具体包括变量的中英文名称、变量值、变量分类，以及变量具体含义的说明（见图 3A 和图 3B）。

A	Variable_Name	Variable_description	Variable_value	Variable_category
	First Author	the first named author of the	NA	categorical
	Last Author	the last named author of the	NA	categorical
	Year	the year when the article is	NA	numeric
	Journal	the journal the article is published in	NA	categorical
	Article_ID	the ID of the article	Author_Year_Journal	categorical
	Standardized name	the contrasts are standardized	NA	categorical
	Study	order of independent studies	NA	numeric
	Sample_Size	number of samples	NA	numeric
	Number_of_Female	number of female participants	NA	numeric

B	Variable_Name	Variable_description	Variable_value	Variable_category
	First Author	the first named author of the	NA	categorical
	Last Author	the last named author of the	NA	categorical
	Year	the year when the article is	NA	numeric
	Journal	the journal the article is published in	NA	categorical
	Study	order of independent studies	NA	numeric
	Article_ID	the ID of the article	Author_Year_Journal	categorical
	Definition of self-reference	author's definition of SRE in the article	NA	categorical
	Self-reference effect	how self-reference effect is defined	NA	categorical
	Standardized name	the contrasts are standardized	NA	categorical

图 3 数据编码手册截图：（A）“Codebook_Article_info.csv”的部分截图；（B）“Codebook_Operationalization.csv”的部分截图

Figure 3 Screenshot of codebooks: （A）Screenshot of “Codebook_Article_info.csv”；（B）Screenshot of “Codebook_Operationalization.csv”

3 数据质量控制和评估

根据刘宇等^[22]的建议，本数据集的编码中由两位独立的研究者完成，以减少编码的主观性。独立完成编码后，两位研究者共同核对编码结果并通过 0/1 的方式对编码是否相同进行打分，最后对打分结果进行评分者一致性分析（一致性打分及评分者一致性系数的计算代码见在线补充材料）。

作为基于元研究的数据集，本数据集质量的核心是编码过程的准确性与一致性。本文采用 Gwet^[35]提出的 AC_1 系数量化评分者一致性。 AC_1 系数比经典的评分者一致性指标 Kappa 系数更稳健^[36]。其具体计算公式如下：

$$AC_1 = \frac{p_a - p_{ey}}{1 - p_{ey}} \quad (1)$$

其中 p_a 是包括偶然或非偶然的总体一致概率， p_{ey} 是偶然一致性（chance agreement）概率。通过 R 包 irrCAC^[37] 计算 AC_1 系数。结果表明，数据编码的 AC_1 系数为 0.871，高于 0.8，表明本数据集的文献编码具有高度一致性^[36]。

通过对各文章实验细节的编码，发现在实验设计、材料和数据分析等步骤上能够体现对自我参照的操作化定义上的变异性（见图 4）。这些结果初步表明，不同研究间对自我参照的操作化定义上存在较大变异性：在感觉通道上，存在视觉、听觉以及两种感觉混合的情况；刺激种类包括特质词、句子与图片；被试所需做出的反应既包括判断特质词是否适合描述自己（或他人），也包括不做任何反应；在统计分析中为得到自我参照的特异性激活脑区而设置的控制条件也可标准化为亲密他人（close other）、名人（celebrities）、非人称（non-person）和陌生人（stranger）四种条件。这些初步的分析表明，自我参照的本体论承诺的确是一个被忽略的问题，其对自我参照神经基础的影响需要进一步的研究。

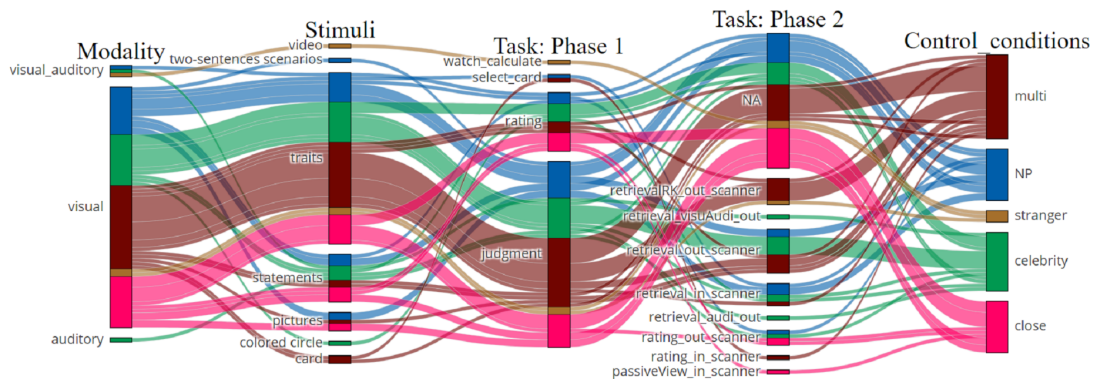


图 4 自我参照的操作化定义在五个维度上（刺激呈现的感觉通道、刺激内容、任务的阶段 1、阶段 2 及控制条件）上的变异性。每一列代表一个维度，列之间线条的粗细代表研究的数量，线条颜色由第五列（控制条件维度）进行区分。注：第四列“Task: Phase 2”中的“NA”表示这些研究没有第二阶段任务；第

五列“Control condition”中的“NP”表示“non-person”。本图绘制的数据与代码见补充材料。

Figure 4 The variability of operationalizations of self-reference across five dimensions (sensory modalities of stimuli presentation, content of stimuli, stage 1 of the task, stage 2 of the task, and control conditions). Each column represents a dimension, the thickness of the lines, which flow through the columns, represents the number of studies, and the colors of lines are defined by the fifth column (control conditions). Note: In the forth column, “Task: Phase 2”, “NA” means tasks in these studies did not have a second phase; the “NP” in the fifth column, “Control conditions”, means “non-person”. See online supplementary material for data and code to reproduce this figure.

本数据集相比已有的自动化元分析数据库具有如下优势：第一，在文献的选择上更加精细和准确，弥补自动元分析数据库的局限性^[20,21]。Neurosynth 数据库中使用与自我参照最相近的英文字段“self-referential”进行搜索时得到 166 篇文献，与本数据集纳入的文章有重合的仅 14 篇。类似地，在 NeuroQuery 中搜索“self-referential”得到 72 个相关联的术语，30 篇相关文献，其中与本数据集所纳入的文章有重合的仅 11 篇。文献选择上的差异导致 ALE 元分析结果的差异。比较 Neurosynth 的结果（图 5A）和本数据集中数据 ALE 元分析的结果（图 5B，ALE 元分析的方法细节见本数据集中的补充材料）可以发现，两者的模式大致相同，但本数据集的 ALE 元分析结果中激活脑区更加集中，且不包括颞叶部，与先前手动元分析的结果相同^[12, 23]。

第二，本数据集能够揭示操作化定义的变异性带来的效应。作为展示，本文根据自我参照操作化定义中的一个维度——统计分析中的控制条件——对文献进行分类和 ALE 元分析。如前所述，在当前文献中，依据统计分析中控制条件的不同可分为四类操作化定义：自我 vs 亲密他人（如家人或朋友）、自我 vs 名人（如政治人物或娱乐明星）、自我 vs 非人称条件（如字体判断或语义判断）和自我 vs 陌生人。由于最后一类的文献数量较少（ $n=5$ ），不满足 ALE 元分析的需求，因此仅对其他三类自我参照加工的结果进行元分析并对元分析结果进行差异分析。结果表明，控制条件的差异对自我参照效应的脑网络有显著的影响，见图 5 中的 C、D、E 和表 1。

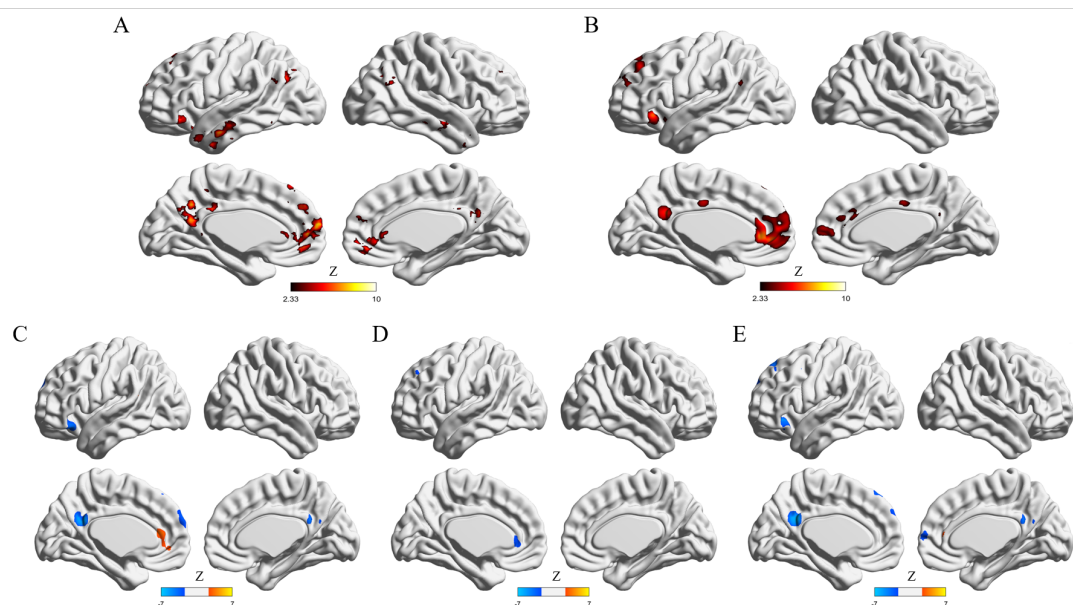


图 5 比较不同数据来源的自我参照神经成像元分析结果。（A）Neurosynth 中“self-referential”字段进行自动元分析结果图；（B）本数据集中纳入研究的元分析结果图；（C）本数据集中“自我 vs 名人”与“自我 vs 非人称”两种不同操作化定义的 ALE 元分析结果的差异分析结果图；（D）“自我 vs 亲密他人”与“自我 vs 名人”的 ALE 元分析结果的差异分析结果图；（E）“自我 vs 亲密他人”与“自我 vs 非人称”的 ALE 元分析结果的差异分析结果图。

Figure 5 Comparisons of meta-analytical results from different sources. (A) Results of automatic meta-analysis of term “self-referential” in Neurosynth; (B) Meta-analytical results for all studies included in the current dataset; (C) Results of contrast analysis between meta-analytical results of “self vs celebrity” and that of “self vs non-personal”; (D) Results of contrast analysis between meta-analytical results of “self vs close” and that of “self vs celebrity”; (E) Results of contrast analysis between meta-analytical results of “self vs close” and that of “self vs non-personal”.

表 1 元分析结果的差异分析

Table 1. Contrast analyses of meta-analytical results

脑区	体积 (voxel)	Z-值峰值坐标			解剖位置
		x	y	z	
(Self - Close) > (Self > non-Person)					
1	72	4	40	10	Right Cingulate Gyrus, anterior division
(Self - non-Person) > (Self - Close)					
1	473	-6	-52	28	Left Cingulate Gyrus, posterior division
2	431	-42	28	-8	Left Frontal Orbital Cortex
3	430	10	64	8	Right Frontal Pole
4	154	-48	-58	20	Left Angular Gyrus
5	143	-8	44	52	Left Frontal Pole
6	86	-38	10	50	Left Middle Frontal Gyrus

脑区	体积 (voxel)	Z-值峰值坐标			解剖位置
		x	y	z	
(Self - Celebrity) > (Self - non-Person)					
1	504	-2	34	12	Left Anterior Cingulum
2	97	-60	-46	16	Left Supramarginal Gyrus, posterior division
3	66	-22	40	36	Left Frontal Pole
(Self - non-Person) > (Self - Celebrity)					
1	437	-6	-52	28	Left Cingulate Gyrus, posterior division
2	374	-46	24	-18	Left Temporal Pole
3	350	-4	62	28	Left Frontal Pole
4	99	-6	38	48	Left Superior Frontal Gyrus
5	78	-48	-62	32	Left Lateral Occipital Cortex, superior division
6	70	-36	4	50	Left Middle Frontal Gyrus
(Self - Close) > (Self - Celebrity)					
NA	NA	NA	NA	NA	NA
(Self - Celebrity) > (Self - Close)					
1	213	-4	32	0	Left Anterior Cingulum
2	127	-18	40	44	Left Frontal Pole
3	106	-54	-50	20	Left Angular Gyrus
4	12	-26	56	26	Left Frontal Pole

4 数据价值

本数据集对自我参照的 fMRI 和 PET 研究进行梳理，详细地分析现有文献中对自我参照的操作化定义。在此基础上，统一描述神经成像研究中关于自我参照操作化定义的细节，区分几类可能会影响自我参照的操作化定义，有助于研究者意识到操作化定义的重要性以及“自我参照”这一概念的本体论承诺问题，促进“自我参照”这一概念使用的规范性。

其次，本数据集为自我参照提供更精细的元分析数据，将利于比较不同操作化定义对自我参照的影响，提升对心理概念与操作化定义之间的关系的理解。如何操作化定义一个心理过程本身包含研究者对该心理过程的理论预设。比较不同操作化定义将有助于研究者更加清晰地将自我参照的理论预设进行表述，促进关于自我参照加工的理论建构。精细的区分不同操作化定义也将有助于提高对自我参照激活脑区的精细认识，增进对自我参照的认知神经机制的理解，为跨精神疾病诊疗提供依据。最后，本数据集的建立为后续构建类似的认知本体论数据集提供一定参考依据。

本数据集作为首个针对单个心理构念的神经成像认知本体论数据集，在数据量和数据格式上仍然具有较大的提升空间。在数量上，未来需要加入更多的自我信息加工相关的神经成像研究，例如关于自我面孔识别的神经成像研究^[12]、自传体记忆的神经成像研究等^[24]。在数据格式上，未来可能整合更加方便机读和自动化元分析的技术，例如，与 Datalad^[38]进行整合。

5 数据使用方法和建议

本数据集包括自我参照操作化定义及其 fMRI 和 PET 激活坐标点的结果，并使用 ALE 神经成像元分析中常见的格式。研究者可采用 GingerALE、Matlab 和 Python 等软件进行数据的读取与分析。未来研究可以根据操作化定义对感兴趣的自我参照加工过程进行分类并进行基于坐标的神经成像元分析或激活网络图谱（Activation Network Mapping）^[39]分析，形成新的研究假设或与其他认知过程的元分析结果进行比较分析。具体而言，研究者可将自己的研究兴趣与数据集中关于自我参照的操作化定义信息相结合，对包含在本数据集中的研究进行二次分类和筛选，并提取各个实验中的激活坐标点的数据进行神经成像的元分析^[12, 40-42]。

6 数据可用性声明

本数据集由已发表期刊论文中的公开数据构成，可在科学数据银行(Science Data Bank)中获得本数据集的原始数据，具体而言，本数据集可通过访问 <http://doi.org/10.57760/sciencedb.j00001.00469> 获得。如果您在研究中使用本数据集，请在参考文献中标注引用。本数据集禁止用于商业用途。

数据作者分工职责

胡传鹏（1987—），男，湖北省荆州市人，博士，教授，研究方向为元科学、社会认知与计算认知神经科学。主要承担工作：总体方案设计、数据核对、组织实施及论文撰写与修改。

孙淑婷（1998—），女，福建省福州市人，硕士研究生在读，研究方向为元科学、社会认知和心理健康。主要承担工作：数据采集、数据校对、汇总整理、数据分析及论文撰写与修改。

王楠（2000—），女，江苏省盐城市人，本科生在读，研究方向为元科学和心理语言学。主要承担工作：数据采集、数据校对、汇总整理及论文撰写与修改。

温佳慧（2000—），女，山西省太原市人，硕士研究生在读，研究方向为元科学和社会认知。主要承担工作：数据的采集。

参考文献

- [1] BENOIT R G, GILBERT S J, VOLLE E, et al. When I think about me and simulate you: Medial rostral prefrontal cortex and self-referential processes[J/OL]. NeuroImage, 2010, 50(3): 1340-1349. DOI:10.1016/j.neuroimage.2009.12.091.
- [2] MORAY N. Attention in Dichotic Listening: Affective Cues and the Influence of Instructions[J/OL]. Quarterly Journal of Experimental Psychology, 1959, 11(1): 56-60. DOI:10.1080/17470215908416289.
- [3] ROGERS T B, KUIPER N A, KIRKER W S. Self-reference and the encoding of personal information[J/OL]. Journal of Personality and Social Psychology, 1977, 35(9): 677-688. DOI:10.1037/0022-3514.35.9.677.

- [4] FINK G R, MARKOWITSCH H J, REINKEMEIER M, et al. Cerebral Representation of One's Own Past: Neural Networks Involved in Autobiographical Memory[J/OL]. *Journal of Neuroscience*, 1996, 16(13): 4275-4282. DOI:10.1523/JNEUROSCI.16-13-04275.1996.
- [5] KIRCHER T T J, SENIOR C, PHILLIPS M L, et al. Towards a functional neuroanatomy of self processing: effects of faces and words[J/OL]. *Cognitive Brain Research*, 2000, 10(1): 133-144. DOI:10.1016/S0926-6410(00)00036-7.
- [6] ANDREWS-HANNA J R, REIDLER J S, SEPULCRE J, et al. Functional-Anatomic Fractionation of the Brain's Default Network[J/OL]. *Neuron*, 2010, 65(4): 550-562. DOI:10.1016/j.neuron.2010.02.005.
- [7] QIN P, NORTHOFF G. How is our self related to midline regions and the default-mode network?[J/OL]. *NeuroImage*, 2011, 57(3): 1221-1233. DOI:10.1016/j.neuroimage.2011.05.028.
- [8] SHELINE Y I, BARCH D M, PRICE J L, et al. The default mode network and self-referential processes in depression[J/OL]. *Proceedings of the National Academy of Sciences*, 2009, 106(6): 1942-1947. DOI:10.1073/pnas.0812686106.
- [9] VAN DER MEER L, COSTAFREDA S, ALEMAN A, et al. Self-reflection and the brain: A theoretical review and meta-analysis of neuroimaging studies with implications for schizophrenia[J/OL]. *Neuroscience & Biobehavioral Reviews*, 2010, 34(6): 935-946. DOI:10.1016/j.neubiorev.2009.12.004.
- [10] PADMANABHAN A, LYNCH C J, SCHAER M, et al. The Default Mode Network in Autism[J/OL]. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 2017, 2(6): 476-486. DOI:10.1016/j.bpsc.2017.04.004.
- [11] OCHSNER K N, BEER J S, ROBERTSON E R, 等. The neural correlates of direct and reflected self-knowledge[J/OL]. 2005: 18. DOI:10.1016/j.neuroimage.2005.06.069.
- [12] HU C, DI X, EICKHOFF S B, et al. Distinct and common aspects of physical and psychological self-representation in the brain: A meta-analysis of self-bias in facial and self-referential judgements[J/OL]. *Neuroscience & Biobehavioral Reviews*, 2016, 61: 197-207. DOI:10.1016/j.neubiorev.2015.12.003.
- [13] Pfeifer J H, Lieberman M D, Dapretto M. "I know you are but what am I?": neural bases of self-and social knowledge retrieval in children and adults[J]. *Journal of Cognitive Neuroscience*, 2007, 19(8): 1323-1337. DOI:10.1162/jocn.2007.19.8.1323
- [14] Schmitz T W, Kawahara-Baccus T N, Johnson S C. Metacognitive evaluation, self-relevance, and the right prefrontal cortex[J]. *Neuroimage*, 2004, 22(2): 941-947. DOI:10.1016/j.neuroimage.2004.02.018
- [15] Zhu, Y., Zhang, L., Fan, J., & Han, S. (2007). Neural basis of cultural influence on self-representation. *Neuroimage*, 34(3), 1310-1316. DOI:10.1016/j.neuroimage.2006.08.047.
- [16] EISENBERG I W, BISSETT P G, ZEYNEP ENKAVI A, et al. Uncovering the structure of self-regulation through data-driven ontology discovery[J/OL]. *Nature Communications*, 2019, 10(1): 2319. DOI:10.1038/s41467-019-10301-1.

- [17] 胡传鹏, 王非, 过继成思, 等. 心理学研究中的可重复性问题: 从危机到契机[J/OL]. 心理科学进展, 2016, 24(9): 1504. DOI:10.3724/SP.J.1042.2016.01504.[HU Chuanpeng; DI Xin; LI Jiawei; SUI Jie; PENG Kaiping. (2015). Meta-analysis of Neuroimaging Studies. *Advances in Psychological Science*, 23(7), 1118-1129.]
- [18] BRACHEM J, FRANK M, KVETNAYA T, et al. Replication Crisis, p-Hacking, and Open Science. An Inquiry into Questionable Research Practices in Student Projects and Impulses for the Teaching Environment[J/OL]. *PSYCHOLOGISCHE RUNDSCHAU*, 2022, 73(1): 1-17. DOI:10.1026/0033-3042/a000562.
- [19] EARP B D, TRAFIMOW D. Replication, falsification, and the crisis of confidence in social psychology[J/OL]. *FRONTIERS IN PSYCHOLOGY*, 2015, 6. DOI:10.3389/fpsyg.2015.00621.
- [20] YARKONI T, POLDRACK R A, NICHOLS T E, et al. Large-scale automated synthesis of human functional neuroimaging data[J/OL]. *Nature Methods*, 2011, 8(8): 665-670. DOI:10.1038/nmeth.1635.
- [21] DOCKÈS J, POLDRACK R A, PRIMET R, 等. NeuroQuery, comprehensive meta-analysis of human brain mapping[J/OL]. *eLife*, 2020, 9: e53385. DOI:10.7554/eLife.53385.
- [22] 刘宇, 陈树铨, 樊富珉等. 开放式荟萃分析的规范化报告[J]. 中国科学 (生命科学), 2021(6): 764-778.[Liu Y, Chen S Q, Fan F M, et al. A standardized checklist on meta-analysis reporting in the open science era. *Sci Sin Vitae*, 2021, 51: 764 – 778, DOI:10.1360/SSV-2021-0009]
- [23] NORTHOFF G, HEINZEL A, DE GRECK M, et al. Self-referential processing in our brain--a meta-analysis of imaging studies on the self[J/OL]. *NeuroImage*, 2006, 31(1): 440-457. DOI:10.1016/j.neuroimage.2005.12.002.
- [24] ARAUJO H, KAPLAN J, DAMASIO A. Cortical midline structures and autobiographical-self processes: An activation-likelihood estimation meta-analysis[J/OL]. *Frontiers in Human Neuroscience*, 2013, 7. <https://www.frontiersin.org/article/10.3389/fnhum.2013.00548>. DOI:10.3389/fnhum.2013.00548.
- [25] QIN P, WANG M, NORTHOFF G. Linking bodily, environmental and mental states in the self-A three-level model based on a meta-analysis[J/OL]. *Neuroscience and Biobehavioral Reviews*, 2020, 115: 77-95. DOI:10.1016/j.neubiorev.2020.05.004.
- [26] FREWEN P, SCHROETER M L, RIVA G, et al. Neuroimaging the consciousness of self: Review, and conceptual-methodological framework[J/OL]. *Neuroscience & Biobehavioral Reviews*, 2020, 112: 164-212. DOI:10.1016/j.neubiorev.2020.01.023.
- [27] NORTHOFF G. Is the self a higher-order or fundamental function of the brain? The “basis model of self-specificity” and its encoding by the brain’s spontaneous activity[J/OL]. *Cognitive Neuroscience*, 2016, 7(1-4): 203-222. DOI:10.1080/17588928.2015.1111868.
- [28] Northoff, G. (2021). Brain networks and the emergence of the self: A neurophenomenal perspective. V. A. Diwadkar & S. B. Eickhoff (Ed.), *Brain network dysfunction in neuropsychiatric illness: Methods, applications, and implications* (page 433–453). Springer International Publishing. DOI:10.1007/978-3-030-59797-9_21
- [29] SCALABRINI A, SCHIMMENTI A, DE AMICIS M, et al. The self and its internal thought:

- In search for a psychological baseline[J/OL]. *Consciousness and Cognition*, 2022, 97: 103244. DOI:10.1016/j.concog.2021.103244.
- [30] WALLA P, NORTHOFF G, HERBERT C. The Human Self Has Two Serial Aspects and Is Dynamic: A Concept Based on Neurophysiological Evidence Supporting a Multiple Aspects Self Theory (MAST)[J/OL]. *Life-Basel*, 2021, 11(7): 611. DOI:10.3390/life11070611.
- [31] LESHIKAR E D, DUARTE A. Medial prefrontal cortex supports source memory for self-referenced materials in young and older adults[J/OL]. *Cognitive, Affective, & Behavioral Neuroscience*, 2014, 14(1): 236-252. DOI:10.3758/s13415-013-0198-y.
- [32] LANCASTER J L, TORDESILLAS - GUTIÉRREZ D, MARTINEZ M, 等. Bias between MNI and Talairach coordinates analyzed using the ICBM - 152 brain template[J/OL]. *Human Brain Mapping*, 2007, 28(11): 1194-1205. DOI:10.1002/hbm.20345.
- [33] PAGE M J, MCKENZIE J E, BOSSUYT P M, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews[J/OL]. *BMJ*, 2021: n71. DOI:10.1136/bmj.n71.
- [34] FOX P T, LANCASTER J L. Opinion: Mapping context and content: the BrainMap model[J/OL]. *Nature Reviews. Neuroscience*, 2002, 3(4): 319-321. DOI:10.1038/nrn789.
- [35] Gwet, K. L. 2008. Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology* 61 (1):29–48. DOI:10.1348/000711006X126600.
- [36] Wongpakaran, N., Wongpakaran, T., Wedding, D., & Gwet, K. L. (2013). A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. *BMC medical research methodology*, 13(1), 1-7. DOI:10.1186/1471-2288-13-61
- [37] Gwet K L. irrCAC: computing chance-corrected agreement coefficients (CAC)[J]. R package version, 2019, 1: 2019.
- [38] HALCHENKO Y O, MEYER K, POLDRACK B, et al. DataLad: distributed system for joint management of code, data, and their relationship[J/OL]. *Journal of Open Source Software*, 2021, 6(63): 3262. DOI:10.21105/joss.03262.
- [39] Peng, S., Xu, P., Jiang, Y., & Gong, G. (2022). Activation network mapping for integration of heterogeneous fMRI findings. *Nature Human Behaviour*, 1–13. DOI:[10.1038/s41562-022-01371-1](https://doi.org/10.1038/s41562-022-01371-1)
- [40] EICKHOFF S B, LAIRD A R, GREFKES C, et al. Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty[J/OL]. *Human Brain Mapping*, 2009, 30(9): 2907-2926. DOI:10.1002/hbm.20718.
- [41] EICKHOFF S B, BZDOK D, LAIRD A R, et al. Activation likelihood estimation meta-analysis revisited[J/OL]. *NeuroImage*, 2012, 59(3): 2349-2361. DOI:10.1016/j.neuroimage.2011.09.017.
- [42] TURKELTAUB P E, EICKHOFF S B, LAIRD A R, et al. Minimizing within-experiment and within-group effects in Activation Likelihood Estimation meta-analyses[J/OL]. *Human Brain Mapping*, 2012, 33(1): 1-13. DOI:10.1002/hbm.21186.

论文引用格式

孙淑婷, 王楠, 温佳慧等. 自我参照的神经成像认知本体论数据集[J/OL]. 中国科学数据, 2022. (2022-07-01). DOI: 10.11922/11-6035.csd.2022.0047.zh.

数据引用格式

孙淑婷, 王楠, 温佳慧等. 自我参照的神经成像认知本体论数据集[DS/OL]. Science Data Bank, 2022. (2022-07-25). DOI: 10.57760/sciencedb.j00001.00469.