附录一、研究报告格式(中文)

**对*Rodriguez(2025)*研究结果的计算可复现性检验**

小组成员分工

|  |  |  |  |
| --- | --- | --- | --- |
| 组长 |  | | |
| 组员 | 蔡振辛、祁子陽、赵智康、张华宇 | | |
| 分工 | | | |
| 数据分析 | 蔡振辛（25%）、祁子陽（25%）、赵智康（25%）、张华宇（25%） | PPT 制作 | 蔡振辛（25%）、祁子陽（25%）、赵智康（25%）、张华宇（25%） |
| 文字报告制作 | 蔡振辛（25%）、祁子陽（25%）、赵智康（25%）、张华宇（25%） | PPT 展示 | 蔡振辛（25%）、祁子陽（25%）、赵智康（25%）、张华宇（25%） |

\* 同一名同学可负责多个部分；如同一内容由多位同学负责，可按百分比注明贡献占比

**摘要**：200 ～ 400字，包括背景、方法、结果与讨论/结论几个关键部分

**关键词**：三到五个关键词，需要包括“计算可复现性”

**1 引言**

**1.1 所选文献信息**

**表 1 文献信息表**

|  |  |  |  |
| --- | --- | --- | --- |
| **1 文献基本信息** | | | |
| 所选文献 | Rodriguez, M., Schertz, K. E., & Kross, E. (2025). How people think about being alone shapes their experience of loneliness. Nature Communications, 16(1), 1594. | | |
| 数据来源 | https://doi.org/10.17605/OSF.IO/6GX8R（文章收集的数据）  https://www.cos.io/gfs（Global Flourishing Study,Wave 1, released in 2024需要预注册）  https://doi.org/10.17605/OSF.IO/6GX8R.（Code） | | |
| **2 文献选取** | | | |
| 文献主题是否包含不止一篇研究？ | □是，且包含元分析研究  ☑是，但不包含元分析研究  □否 | 文献此前被其他研究者重复过？ | □是(附上原文链接)  ☑否 |
| **3 研究假设选取** | | | |
| 重复的研究假设 | (指出检验了原文献哪部分的哪些假设结果，如研究一、研究二/ 行为数据、电生理数据/ A任务、B任务等) | | |
| 重复的研究假设是否在其他研究中经过重复？ | □是(附上原文链接)  ☑否 | 文献共几个实验，重复的研究假设是第几个实验中的？ | 文献共5个研究  重复了前四个 |
| 选择该假设的原因 | 按照指南推荐的优先顺序，我们小组重复了前四个研究，第五个研究无法完全复现的原因是，Global Flourishing Study,Wave 1数据库暂未公开 | | |
| **4 数据集选取** | | | |
| 是否采用原始数据？ | ☑是  □否 | 是否对样本量进行修改？ | □是(说明原因)  ☑否 |
| 若修改，报告原文样本量大小和修改后的样本量大小 | 无 | 若修改，报告使用 G-power 计算的修改后的样本量对应的效应量 | 无 |

**1.2 文献介绍**

孤独感是媒体和公共卫生运动经常关注的普遍问题。我们认为，这些善意的努力可能通过消极影响孤独感的潜在驱动因素——人们对独处状态的认知——反而加剧这一问题。支持这一论点的证据显示：当代美国新闻文章将独处描述为有害而非有益的可能性高出十倍（研究1，样本量N=144篇文章）；接触此类文章会直接影响人们对独处的认知（研究2，N=456）。此外，一项为期两周的经验抽样研究（研究3，N=161）表明，对独处持消极认知的人在日常生活中独处后孤独感会急剧上升，而持积极认知者独处后孤独感反而降低。最终，多项跨国研究证实这些结论在跨越六大洲的九个国家中具有普适性（研究4和5，N=2930）。这些发现共同揭示了人们对独处的认知在塑造孤独感过程中起到的关键作用。

**2 方法**

**2.1 样本**

研究一：第一项研究采用档案数据来考察当代美国新闻媒体如何描绘独处现象。数据收集：我们从ProQuest数据库（https://www.proquest.com/）检索了2020年1月1日至2022年12月31日期间，美国印刷发行量排名前十的日报媒体（即《纽约时报》《华尔街日报》《华盛顿邮报》《芝加哥论坛报》《今日美国》《坦帕湾时报》《纽约邮报》《洛杉矶时报》《新闻日报》和《明星论坛报》）发布的印刷版和网络版新闻记录。搜索包含以下术语的文章："being alone"（独处）、"time alone"（独处时光）、"solitude"（孤独）、"solitary"（独居）、"aloneness"（孤独感）或"isolation"（隔离）。我们确认遵守了ProQuest数据库的服务条款。

研究二：本实验检验接触关于独处信息是否会影响人们对独处的信念。我们已在AsPredicted平台预注册研究假设与分析方案（编号#148930，2023年10月28日，https://aspredicted.org/5xt83.pdf）。参与者 通过G\*Power软件进行的先验功效分析表明，要检测小到中等效应量（Cohen's d = 0.30）并在0.05显著性水平达到90%统计功效，需要417名参与者。考虑到可能的样本排除，我们通过学术研究专用众包平台Prolific招募了456名英语流利的成年人。最终分析纳入439人（排除情况见下文；平均年龄=44.75岁，年龄标准差=14.24；女性47.4%，男性50.1%，非二元性别/其他2.5%；白人76.9%，黑人10.0%，拉丁裔6.9%，亚裔5.6%，其他种族0.7%；高中以下学历0.2%，高中学历12.1%，部分大学教育20.3%，副学士学位10.3%，学士学位38.3%，硕士学位15.5%，专业学位2.3%，博士学位1.1%）。入选标准为年满18周岁且英语流利。每位参与者获得3美元报酬。

研究三：这项经验取样研究探讨了人们对独处状态的信念是否会影响其在日常生活中独处后孤独感及其他情绪状态的变化程度。本研究未进行预注册。参与者通过Prolific平台招募了161名受试者参与关于"日常情绪"的研究，未告知研究团队对独处时间及孤独感的特别关注。入选标准为：年龄18-30岁、居住在美国、英语流利、持有安卓智能手机。样本量参考近期评估日常生活独处时间的经验取样研究确定68,69。经筛选后（见流程），最终纳入138名参与者（总观察量9570次）（平均年龄=25.04岁，标准差=3.37；女性42.8%，男性50.7%，非二元性别/其他6.5%；白人55.1%，亚裔13.0%，黑人13.8%，拉丁裔13.0%，其他种族5.1%；1%高中在读，21.7%高中学历，30.4%大学在读，37.7%本科学历，8.0%硕士学历，1.4%博士学历）。参与者分布美国本土各时区（东部时间42.0%，中部时间34.8%，山地时间2.2%，太平洋时间21.0%）。

研究四：本研究者招募了四个跨国样本，分别来自美国和日本。美国样本1共有251名参与者通过Prolific平台招募，最终244人纳入分析，平均年龄为37.32岁（SD = 12.73），其中46.3%为女性，50.4%为男性，3.3%为非二元或其他性别；种族构成以白人为主（67.2%），其余为亚裔（10.2%）、黑人（9.8%）、拉美裔（8.2%）及其他（4.5%）；教育程度涵盖从未完成高中（1.6%）到博士（0.004%）。参与者需年满18岁，在美国长大并现居美国，报酬为2美元。日本样本1共有269人通过Lancers平台招募，最终255人纳入分析，平均年龄为41.83岁（SD = 9.25），其中38.4%为女性，61.2%为男性，0.004%为非二元或其他性别，全部为日本人；教育背景从未完成高中（2.4%）到博士（2.0%），报酬为200日元（约1.5美元），参与者需在日本长大并现居日本。为验证可重复性，研究者开展了样本量更大的美国样本2和日本样本2。美国样本2共招募604人，576人纳入分析，平均年龄为41.27岁（SD = 13.58），49.7%为女性，48.8%为男性，1.6%为非二元或其他性别；种族构成主要为白人（74.0%），其他包括亚裔（7.5%）、黑人（7.8%）、拉美裔（7.3%）及其他（3.5%）；教育程度从未完成高中（1.0%）至博士（3.0%），报酬同为2美元。日本样本2共招募546人，最终511人纳入分析，平均年龄为43.16岁（SD = 9.88），39.9%为女性，59.1%为男性，1.0%为非二元或其他性别，所有参与者皆为日本人；教育程度从未完成高中（2.2%）至博士（1.0%），报酬为200日元（约1.5美元）。由于数据收集在17小时后遇到瓶颈，日本样本2未达到原定的600人目标，但数据收集终止前未查看任何数据，符合预注册计划。

研究五：通过众包平台Prolific（南非、英国、墨西哥、西班牙、澳大利亚、波兰）和Besample（巴西）在线招募参与者。选择Besample而非Prolific招募巴西参与者是因为截至2024年7月，Prolific平台上活跃的巴西用户不足25人。参与者需满足以下条件：(a)年满18周岁；(b)在对应国家成长；(c)当前居住在该国。每个国家有180名参与者完成研究（总样本量N=1260）。排除标准详见后文"排除标准"部分。纳入分析的参与者人口统计学特征详见补充信息。

本研究选取了作者在 OSF（https://doi.org/10.17605/OSF.IO/6GX8R）上公开的数据集进行复现，重点使用研究一至研究四的数据。所用样本与原始研究一致，无需进行额外的数据替换或清洗，并报告原本的样本量和效应量的结果。

**2.2 原研究方法简介**

研究一：研究1采用内容分析法，基于ProQuest数据库收集2020年至2022年间美国前十大发行量报纸中与“独处”相关的新闻文章。研究分为两个阶段：首先通过关键词检索并筛选标题与全文，以人工双盲编码的方式剔除与人际独处无关的内容；其次进行系统的标题和全文内容分析。标题分析由四名编码员对每篇文章标题进行情感倾向分类，并计算Fleiss’s Kappa检验一致性。全文分析部分由八名编码员组成四组，使用二元变量编码判断是否提及独处的具体益处与风险，并通过Cohen’s Kappa检验组内一致性。研究未使用自动化文本分析工具，而是依赖人工内容判定与统计分析方法，编码一致性数据表明分析过程具有高度信度。相关统计分析借助R完成，涉及的统计方法包括Cohen’s Kappa与Fleiss’s Kappa等一致性系数计算。

研究二：研究2采用实验设计，检验媒体中关于独处信息的呈现是否会影响人们对独处的信念。研究预先在 AsPredicted 注册假设与分析计划（注册号 #148930），通过 Prolific 平台招募456名符合条件的成年参与者，并最终纳入439人用于分析。实验采用阅读理解掩饰任务，将参与者随机分配至三组：独处益处组、独处风险组与中性控制组，每组随机顺序阅读四篇文章（每篇150词），三篇填充文章相同，都由ChatGPT草拟、由第一作者修订。目标文章分别包含独处的益处、独处的危害或与独处无关的内容，规范了独处益处和独处风险文章的内容;它们包含相同的短语，只有关键词被改变了。为了检测参与者对材料的理解与注意力，研究设计了高亮关键语句任务与开放式的总结文章大意，统计完成情况并对两位编码员判断的完成情况进行一致性评估（Gwet's AC1 系数均在0.993以上），剔除未通过者。随后，参与者完成四份简短问卷以评估其信念，其中仅一份涉及独处相关信念（后续研究独处信念量表的简化版）项目一致性高（Cronbachα = 0.93），采用1到7的李克特量表测量，并计算信念总分。为避免猜测研究目的对结果的干扰，研究还设置了漏斗式报告，根据一般性逐渐下降的三个问题判断参与者是否准确识别研究假设，进行相应剔除。最终样本为440人数据分析的方法：

数据分析使用 R 语言（版本1.4.1717）进行，按照预注册方案执行对比检验，重点比较“独处益处组”与“控制组”、“独处益处组”与“独处风险组”之间在独处信念评分上的差异。分析中遵循正态性和方差齐性的统计假设，并报告了由设定参考组所导致的负向回归系数（以绝对值呈现）。整个分析流程严格遵循实验控制与统计推断标准，确保了研究的信度与解释效度。

研究三：研究三采用经验抽样法（Experience Sampling Method, ESM）考察人们对独处的信念是否会影响其在日常生活中独处后的孤独感变化。研究共招募161名18至30岁的美国参与者（最终样本为138人），要求为Android手机用户并能流利使用英语。研究共分两个阶段：第一阶段，参与者填写背景问卷，测量包括“对独处的信念”在内的多项心理变量；第二阶段，在为期14天的追踪期间，参与者每天接收5次短信（9:00、12:00、15:00、18:00、21:00），每次填写关于当前孤独感、过去数小时内的独处时间及多种情绪状态（如压力、无聊、正负性情绪等）的简短调查问卷。数据分析使用R软件（版本1.4.1717）中的lme4（版本1.1-26）和lmerTest（版本3.1-3）包，建立线性混合效应模型（Linear Mixed-Effects Models），处理时间点嵌套于个体的数据结构。模型中将独处时间预测相邻时间点之间的孤独感变化，控制前一时间点的孤独感，并检验“对独处的信念”是否调节上述关系。模型允许随机截距和斜率，并对一级变量进行个体均值中心化、二级变量进行总均值中心化，同时排除了跨日的滞后效应。为检验稳健性，研究还将抑郁、焦虑、孤独、社会支持与外向性作为协变量纳入模型，所有统计分析均符合常规假设检验前提。

研究四：研究4采用跨国横断面设计，包含四个样本（美国两个，日本两个），旨在探讨对独处的信念是否调节了过去一个月独处时间与孤独感之间的关系。美国样本1通过Prolific招募251名参与者，分析纳入244人；日本样本1通过Lancers招募269人，分析纳入255人。美国样本2招募604人，分析纳入576人；日本样本2招募546人，分析纳入511人。样本量通过G\*Power进行事前功率分析确定，样本1目标为95%功率检测Cohen’s d = 0.35效应（每国178人，计划招募250人），样本2目标为检测d = 0.20效应（每国542人，计划招募600人）。参与者完成问卷，评估独处信念（6项量表）、孤独感（UCLA孤独感量表，8项）、独处时间（MIDUS/MIDJA社交接触项反向评分）及文化维度（关系流动性、紧密-松散、个人/集体主义量表）。数据分析包括独立样本t检验，比较两国独处信念差异；使用R（版本1.4.1717）中的lme4包（版本1.1-26）构建线性混合效应模型，检验独处时间与独处信念的交互作用对孤独感的影响，样本作为随机截距，控制文化和人口变量。数据满足残差独立性、正态性和同方差性假设。双变量相关和媒体暴露分析按预注册计划执行，结果见补充信息。

研究五：研究5采用与研究4相同的横断面设计，在七个国家（巴西、南非、英国、墨西哥、西班牙、澳大利亚、波兰）重新测试，以检验结果在不同文化背景下的普遍性。国家选择基于2024年全球繁荣研究（Global Flourishing Study）中的孤独感均值，通过k均值聚类分为低、中、高孤独感组别，确保地理代表性。样本量通过R中simr包的powerCurve函数进行事前功率分析，设定交互效应（独处时间与独处信念）的未标准化β为-0.3057（基于研究4最小效应），随机截距方差为1.05，残差标准差为4.465，分析显示每国150名参与者可提供97.10%功率（95% CI [95.86, 98.05]），计划每国招募180人以应对排除情况。参与者通过Prolific（南非、英国、墨西哥、西班牙、澳大利亚、波兰）和Besample（巴西）在线招募，总计1260人（每国180人），需年满18岁、在研究国家成长并居住。程序与研究4一致，但移除媒体暴露问题，支付统一为2美元。非英语国家（巴西、墨西哥、西班牙、波兰）的问卷翻译成葡萄牙语、西班牙语和波兰语，先由Google Translate初步翻译，再由母语者与第一作者（M.R.）校对并回译验证。数据分析包括各国的双变量相关（Pearson’s r）检验独处信念与孤独感的关系，以及混合效应线性回归模型，合并所有国家数据，测试独处时间、独处信念及其交互对孤独感的影响，以国家为随机截距，满足残差独立性、正态性和同方差性假设。所有分析使用R软件（版本未指定），涉及simr包进行功率分析，探索性分析（如文化差异、交互效应变异）见补充信息。

**2.3 重复思路说明**

对选用的文章进行复现的思路，包括选取了哪些部分重复及其原因，在使用原文献方法进行重复时是否有需要改动的情况、具体如何改动的。

*此外，如果使用了创新方法，需要具体介绍，并结合相应理论对该方法的可行性、及其相比原方法的优势进行说明。*

**3 结果**

**3.1 描述性统计**

**表 2.1 研究二描述性统计结果的比较**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | 原研究报告结果 | 本研究 | δ | 评级 |
| 人口统计学变量：年龄 | *N* | 440 | 439 | 0.23% | 偏差较小 |
| Mean | 44.75 | 44.75 | 0% | 完全一致 |
| *SD* | 14.24 | 14.24 | 0% | 完全一致 |
| 人口统计学变量：性别 | Frequency | 47.4% Female,  50.1% Male,  2.5%Non-binary/Other. | 47.4% Female,  50.1% Male,  2.5%Non-binary/Other. | 0% | 完全一致 |
| 人口统计学变量：种族(*N=432)* | Frequency | 76.9% White,  10.0% Black,  6.9% Latino,  5.6% Asian,  0.7% Other Race. | 76.9% White,  10.0% Black,  6.9% Latino,  5.6% Asian,  0.7% Other Race. | 0% | 完全一致 |
| 人口统计学变量：教育程度 | Frequency | 0.2%Less Than High School,  12.1%High School,  20.3% Some College,  10.3% Associate’s Degree,  38.3% Bachelor’s Degree,  15.5% Master’s Degree,  2.3% Professional Degree,  1.1% Doctorate Degree). | 0.2%Less Than High School,  12.1%High School,  20.3% Some College,  10.3% Associate’s Degree,  38.3% Bachelor’s Degree, 15.5% Master’s Degree,  2.3% Professional Degree, 1.1% Doctorate Degree). | 0% | 完全一致 |
| 排除前编码一致性检查 | *N* | 456 | 456 | 0% | 完全一致 |
| 两位评分者对参与者摘要的一致性检查（*N*=456） | AC1（Gwet’s） | TikTok article: 0.996;  Red Wine article: 1.00;  Cold Plunge article: 0.998;  Alone Risks article: 1.00;  Alone Benefits article: 0.993;  Control article: 1.00. | TikTok article: 0.996;  Red Wine article: 1.00;  Cold Plunge article: 0.998;  Alone Risks article: 1.00;  Alone Benefits article: 0.993;  Control article: 1.00. | 0% | 完全一致 |
| 两位评分者对漏斗式报告的一致性检查（*N*=456） | AC1（Gwet’s） | Q1: 0.993,  Q2:0.993,  Q3: 0.998. | Q1: 0.993,  Q2:0.993,  Q3: 0.998. | 0% | 完全一致 |
| 独处信念测量项目 | Mean（**beliefs）** | 5.18 | 5.18 | 0% | 完全一致 |
| *SD* | 1.38 | 1.38 | 0% | 完全一致 |
| **Cronbach's α** | 0.93 | 0.93 | 0% | 完全一致 |

**\***主要变量。

**表 2.2 研究三描述性统计结果的比较**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | 原研究报告结果 | 本研究 | δ | 评级 |
| 人口统计学变量：年龄 | *N* | 138 | 138 | 0% | 完全一致 |
| Mean | 25.04 | 25.04 | 0% | 完全一致 |
| *SD* | 3.37 | 3.37 | 0% | 完全一致 |
| 背景问卷 | Mean | 4.82(belief)  19.40(UCLA)  57.04(pss)  4.00(PHQ)  13.96(GAD)  18.72(extra) | 4.82(belief)  19.40(UCLA)  57.04(pss)  4.00(PHQ)  13.96(GAD)  18.72(extra) | 0% | 完全一致 |
| *SD* | 1.13(belief)  5.97(UCLA)  16.56(pss)  1.67(PHQ)  5.16(GAD)  7.48(extra) | 1.13(belief)  5.97(UCLA)  16.56(pss)  1.67(PHQ)  5.16(GAD)  7.48(extra) | 0% | 完全一致 |
| Cronbach *α* | 0.88(belief)  0.90(UCLA)  0.91(pss)  0.85(PHQ)  0.89(GAD)  0.89(extra) | 0.88(belief)  0.90(UCLA)  0.91(pss)  0.85(PHQ)  0.89(GAD)  0.89(extra) | 0% | 完全一致 |
| 情绪追踪取样 | Mean | 23.43(lonely) 2.99(time)  6.10(pos)  3.02(neg)  31.88(bored)  35.38(stressed)  57.41(content)  53.18(grateful) | 23.43(lonely) 2.99(time)  6.10(pos)  3.02(neg)  31.88(bored)  35.38(stressed)  57.41(content)  53.18(grateful) | 0% | 完全一致 |
| *SD* | 27.19(lonely) 1.50(time)  2.41(pos)  2.60(neg)  28.20(bored)  29.79(stressed)  27.09(content)  29.08(grateful) | 27.19(lonely) 1.50(time)  2.41(pos)  2.60(neg)  28.20(bored)  29.79(stressed)  27.09(content)  29.08(grateful) | 0% | 完全一致 |

**\***主要变量即belief、lonely、time。

**表 2.3 研究四描述性统计结果的比较**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | 原研究报告结果 | 本研究 | δ | 评级 |
| 人口统计学变量：年龄及性别 | USSample1 | Mage = 37.32, SDage = 12.73; 46.3%女, 50.4%男, 3.3% 非两元性别 | Mage = 37.32, SDage = 12.73; 46.3%女, 50.4%男, 3.3% 非两元性别 | 0% | 完全一致 |
| USSample2 | Mage = 41.83, SDage = 9.25; 38.4%女, 61.2% 男, 0.004% 非二元性别 | Mage = 41.83, SDage = 9.25; 38.4%女, 61.2% 男, 0.004% 非二元性别 | 0% | 完全一致 |
| JASample1 | Mage = 41.27, SDage = 13.58; 49.7% 女, 48.8% 男, 1.6% 非二元性别 | Mage = 41.27, SDage = 13.58; 49.7% 女, 48.8% 男, 1.6% 非二元性别 | 0% | 完全一致 |
| JASample2 | Mage = 43.16, SDage = 9.88; 39.9%女, 59.1%男, 1.0%非二元性别 | Mage = 43.16, SDage = 9.88; 39.9%女, 59.1%男, 1.0%非二元性别 | 0% | 完全一致 |
| 人口统计学变量：种族及受教育程度 | USSample1 | 67.2%白人，10.2%亚裔，9.8%黑人，8.2%拉丁裔，4.5%其他种族；1.6% 为高中未毕业，13.9% 为高中毕业，26.2% 有部分大学经历，40.2% 拥有学士学位，17.6% 拥有硕士学位，0.004% 拥有博士学位。 | 67.2%白人，10.2%亚裔，9.8%黑人，8.2%拉丁裔，5.7%其他种族；1.6% 为高中未毕业，13.9% 为高中毕业，26.2% 有部分大学经历，40.2% 拥有学士学位，17.6% 拥有硕士学位，0.004% 拥有博士学位。 | 21% | 主要偏差 |
| USSample2 | 74.0%白人，7.5%亚裔，7.8%黑人，7.3%拉丁裔，3.5% 其他种族；1.0% 高中未毕业，14.1% 高中毕业，28.6% 有部分大学经历，41.8% 拥有学士学位，11.5% 拥有硕士学位，3.0% 拥有博士学位 | 74.0%白人，7.5%亚裔，7.8%黑人，7.3%拉丁裔，2.4% 其他种族；1.0% 高中未毕业，14.1% 高中毕业，28.6% 有部分大学经历，41.8% 拥有学士学位，11.5% 拥有硕士学位，3.0% 拥有博士学位 | 46% | 主要偏差 |
| JASample1 | 100% 日本人；2.4% 高中未毕业，29.4% 高中毕业，22.4% 有部分大学经历，39.2% 拥有学士学位，4.7% 拥有硕士学位，2.0% 拥有博士学位。 | 100% 日本人；2.4% 高中未毕业，29.4% 高中毕业，22.4% 有部分大学经历，39.2% 拥有学士学位，4.7% 拥有硕士学位，2.0% 拥有博士学位。 | 0% | 完全一致 |
| JASample2 | 100% 日本人；2.2% 高中未毕业，30.3% 高中毕业，22.3% 有部分大学经历，38.4% 拥有学士学位，5.9% 拥有硕士学位，1.0% 拥有博士学位。 | 100% 日本人；2.2% 高中未毕业，30.3% 高中毕业，22.3% 有部分大学经历，38.4% 拥有学士学位，5.9% 拥有硕士学位，1.0% 拥有博士学位。 | 0% | 完全一致 |
| 各量表之间的描述性统计及Cronbach α比较 | USSample1 | *M* = 5.40, *SD* = 1.17, Cronbach α = .91(独处信念) *M* = 18.09, *SD* = 5.59, Cronbach α = .87(孤独感)  *M* = 3.84, *SD* = 1.46(社交频率) | *M* = 5.40, *SD* = 1.17,  Cronbach α = .91(独处信念) *M* = 18.09, *SD* = 5.59, Cronbach α = .87(孤独感)  *M* = 3.84, *SD* = 1.46(社交频率) | 0% | 完全一致 |
| USSample2 | *M =* 5.02*, SD =* 1.23, Cronbach α = .93(独处信念) *M* = 16.59, *SD* = 5.47; Cronbach α = .88(孤独感)  *M* = 3.47, *SD* = 1.41(社交频率) | *M =* 5.02*, SD =* 1.23, Cronbach α = .93(独处信念) *M* = 16.59, *SD* = 5.47; Cronbach α = .88(孤独感)  *M* = 3.47, *SD* = 1.41(社交频率) | 0% | 完全一致 |
| JASample1 | *M =* 5.58*, SD =* 1.07, Cronbach α = .91(独处信念) *M* = 18.58, *SD* = 5.28; Cronbach α = .87(孤独感)  *M* = 4.13, *SD* = 1.59(社交频率) | *M =* 5.58*, SD =* 1.07, Cronbach α = .91(独处信念) *M* = 18.58, *SD* = 5.28; Cronbach α = .87(孤独感)  *M* = 4.13, *SD* = 1.59(社交频率) | 0% | 完全一致 |
| JASample2 | *M =* 5.56*, SD =* 1.02, Cronbach α = .90(独处信念) *M* = 18.49, *SD* = 4.70; Cronbach α = .83(孤独感)  *M* = 3.80, *SD* = 1.58(社交频率) | *M =* 5.56*, SD =* 1.02, Cronbach α = .90(独处信念) *M* = 18.49, *SD* = 4.70; Cronbach α = .83(孤独感)  *M* = 3.80, *SD* = 1.58(社交频率) | 0% | 完全一致 |

**3.2 推断性统计**

**3.2.1 使用与原文献相同方法的推断性统计**

**表3.1 研究一推断性统计结果的比较(原文献方法)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | 原文献 | 本研究 | δ | 评级 |
| 假设1 | 样本量(*N*)\* | 675 | 675 | 0% | 完全一致 |
| 统计量 | - | - | - | - |
| 效应量  *Cohen’s Kappa κ* | 0.65 | 0.647 | 0% | 因舍入导致的偏差 |
| 显著性指标(*p*) | - | 0 | - | - |
|  |  |  |  |  |  |
| 假设2 | 样本量(*N*)\* | 377 | 372 | 1.33% | 偏差较小 |
| 统计量 | - | - | - | - |
| 效应量  *Cohen’s Kappa κ* | 0.78 | 0.772 | 1.03% | 偏差较小 |
| 显著性指标(*p*) | - | 0 | - | - |
|  |  |  |  |  |  |
| 假设3 | 样本量(*N*)\* | 144 | 144 | 0% | 完全一致 |
| 统计量 | - | - | - | - |
| 效应量  *Cohen’s Kappa κ* | 0.82 | 0.827957 | 0% | 因舍入导致的偏差 |
| 显著性指标(*p*) | - | 0 | - | - |
|  |  |  |  |  |  |
| 假设4 | 样本量(*N*)\* | 144 | 144 | 0% | 完全一致 |
| 统计量 *χ2* | 63.75 | 63.753 | 0% | 因舍入导致的偏差 |
| 效应量  Cramér’s *V* | 0.83 | 0.827957 | 0% | 因舍入导致的偏差 |
| 显著性指标(*p*) | p < .001 | 0 | 1.411e-15 | 完全一致 |
|  |  |  |  |  |  |
| 假设5 | 样本量(*N*)\* | 144 | 144 | 0% | 完全一致 |
| 统计量 *χ2* | 13.03 | 13.031 | 0% | 因舍入导致的偏差 |
| 效应量  Cramér’s *V* | 0.32 | 0.3178295 | 0% | 因舍入导致的偏差 |
| 显著性指标(*p*) | p < .001 | p = 0.0003064 | 0% | 完全一致 |
|  |  |  |  |  |  |
| 假设6 | 样本量(*N*)\* | 144 | 144 | 0% | 完全一致 |
| 统计量 *χ2* | 41.78 | 41.779 | 0% | 因舍入导致的偏差 |
| 效应量  Cramér’s *V* | 0.66 | 0.6631579 | 0% | 因舍入导致的偏差 |
| 显著性指标(*p*) | p < .001 | p = 1.022e-10 | 0% | 完全一致 |
|  |  |  |  |  |  |
| 假设7 | 样本量(*N*)\* | 144 | 144 | 0% | 完全一致 |
| 统计量 *χ2* | 49.33 | 49.33 | 0% | 完全一致 |
| 效应量  Cramér’s *V* | 0.74 | 0.7362637 | 0% | 因舍入导致的偏差 |
| 显著性指标(*p*) | p < .001 | p = 0.0003064 | 0% | 完全一致 |
|  |  |  |  |  |  |
| 假设8 | 样本量(*N*)\* | 144 | 144 | 0% | 完全一致 |
| 统计量 *χ2* | 0.33 | 0.32705 | 0% | 因舍入导致的偏差 |
| 效应量  Cramér’s *V* | - | - | - | - |
| 显著性指标(*p*) | p=0.567 | p=0.5674 | 0% | 因舍入导致的偏差 |
|  |  |  |  |  |  |
| 假设9 | 样本量(*N*)\* | 144 | 144 | 0% | 完全一致 |
| 统计量 *χ2* | 0.54 | 0.5401 | 0% | 因舍入导致的偏差 |
| 效应量  Cramér’s *V* | - | - | - | - |
| 显著性指标(*p*) | p=0.462 | p=0.4624 | 0% | 因舍入导致的偏差 |

**表 3.2 研究二推断性统计结果的比较(原文献方法)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 原文献结果 | 本研究结果 | δ | 评级 |
| 样本量(*N*)\* | 439 | 439 | 0% | 完全一致 |
| 统计量b | 0.73（(Benefits vs. Risks)  0.42（Benefits vs. Control)  -0.31(Risks vs. Control) | 0.73（(Benefits vs. Risks)  0.42（Benefits vs. Control)  -0.31(Risks vs. Control) | 0% | 完全一致 |
| *SE* | 0.16(Benefits vs. Risks)  0.16(Benefits vs. Control)  0.16(Risks vs. Control) | 0.16(Benefits vs. Risks)  0.16(Benefits vs. Control)  0.16(Risks vs. Control) | 0% | 完全一致 |
| 显著性指标(*p*) | <0.001(Benefits vs. Risks)  0.008(Benefits vs. Control)  0.052(Risks vs. Control) | <0.001(Benefits vs. Risks)  0.008(Benefits vs. Control)  0.052(Risks vs. Control) | 0% | 完全一致 |
| 95%CI | [0.41, 1.04]  [0.11, 0.73]  [-0.62, 0.003] | [0.41, 1.04]  [0.11, 0.73]  [-0.62, 0.003] | 0% | 完全一致 |
| *Cohen's d* | 0.54(Benefits vs. Risks)  0.31(Benefits vs. Control)  -0.22(Risks vs. Control) | 0.54(Benefits vs. Risks)  0.31(Benefits vs. Control)  -0.22(Risks vs. Control) | 0% | 完全一致 |

**表 3.3 研究三推断性统计结果的比较(原文献方法)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | | 原文献结果 | 本研究结果 | δ | 评级 |
| 假设一 | 样本量(*N*)\* | 139 | 139 | 0% | 完全一致 |
| 统计量b | -1.19 | -1.19 | 0% | 完全一致 |
| *SE* | 0.27 | 0.27 | 0% | 完全一致 |
| 显著性指标(*p*) | ＜0.001 | ＜0.001 | 0% | 完全一致 |
| 95%CI | -1.73,-0.65 | -1.73,-0.65 | 0% | 完全一致 |
| 假设二 | simple slope | 3.64 | 3.64 | 0% | 完全一致 |
| *SE* | 0.43 | 0.43 | 0% | 完全一致 |
| 显著性指标(*p*) | ＜0.001 | ＜0.001 | 0% | 完全一致 |
| 研究三 | simple slope | 0.96 | 0.96 | 0% | 完全一致 |
| *SE* | 0.43 | 0.43 | 0% | 完全一致 |
| 显著性指标(*p*) | 0.030 | 0.030 | 0% | 完全一致 |

\* 此前描述性统计已报告，因此此处仅需进行报告，不需作为一个额外的结果

**表3.4 研究四推断性统计结果的比较(原文献方法)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | | 原文献 | 本研究 | δ | 评级 |
| 跨文化差异分析 | 样本量(*N*)\* | 1670 | 1670 | 0% | 完全一致 |
| *MDiff* | 0.43 | -1.19 | 0% | 完全一致 |
| 统计量*t* | 760 | 760 | 0% | 完全一致 |
| *df* | 1568.10 | 1568.10 | 0% | 完全一致 |
| 显著性指标(*p*) | ＜0.001 | ＜0.001 | 0% | 完全一致 |
| *Cohen’d* | 0.38 | 0.38 | 0% | 完全一致 |
| 95%CI | [0.32,0.54] | [0.32,0.54] | 0% | 完全一致 |
| 相关分析 | *r* | -.29(US)  -.21(JA) | -.29(US)  -.21(JA) | 0% | 完全一致 |
| *df* | 818(US)  764(JA) | 818(US)  764(JA) | 0% | 完全一致 |
| 显著性指标(*p*) | ＜0.001 | ＜0.001 | 0% | 完全一致 |
| 95%CI | [-0.35,-0.22](US)  [-0.27,-0.14](JA) | [-0.35,-0.22](US)  [-0.27,-0.14](JA) | 0% | 完全一致 |
| 交互作用 | b | -0.33 | 0.96 | 0% | 完全一致 |
| *SE* | 0.06 | 0.43 | 0% | 完全一致 |
| 显著性指标(*p*) | ＜0.001 | ＜0.001 | 0% | 完全一致 |
| 效应量(*f2*) | 0.013 | 0.013 | 0% | 完全一致 |
| 95%CI | [-0.45,-0.20] | [-0.45,-0.20] | 0% | 完全一致 |
| 斜率分析 | simple slope（b） | 1.16（1SD）  1.91（-1SD） | 1.13（1SD）  1.80（-1SD） | 3%和6% | 次要偏差 |
| *SE* | 0.10（1SD）  0.11（-1SD） | 0.09（1SD）  0.11（-1SD） | 11.1% | 主要偏差 |
| 显著性指标(*p*) | 均＜0.001 | 均＜0.001 | 0% | 完全一致 |
| 三重交互作用 | b | 0.02 | 0.96 | 0% | 完全一致 |
| *SE* | 0.13 | 0.43 | 0% | 完全一致 |
| 显著性指标(*p*) | 0.861 | 0.861 | 0% | 完全一致 |
| 95%CI | [-0.24,0.28] | [-0.24,0.28] | 0% | 完全一致 |

**3.3 对原文计算可复现性进行评估**

**3.3.1 使用与原文献相同方法**

报告原文献的值的评级分布、推论的一致情况，整理成表格，如下表所示：

**表 4 结果可复现性的评估表**

|  |  |  |  |
| --- | --- | --- | --- |
| 研究 | 可复现性情况 | 数量及占比 | |
| *N* | *%* |
| 研究一 | 完全一致(*δ* = 0%) | 13 | 46.42 |
| 偏差较小(0% < *δ* < 10%) | 2 | 7.14 |
| 偏差较大(*δ* ＞ 10%) | 0 | 0 |
| 因舍入导致的偏差 | 13 | 46.42 |
| 研究二 | 完全一致(*δ* = 0%) | 16 | 94 |
| 偏差较小(0% < *δ* < 10%) | 0 | 0 |
| 偏差较小(*δ* ＞ 10%) | 0 | 0 |
| 因舍入导致的偏差 | 0 | 0 |
| 研究三 | 完全一致(*δ* = 0%) | 10 | 100 |
| 偏差较小(0% < *δ* < 10%) | 0 | 0 |
| 偏差较小(*δ* ＞ 10%) | 0 | 0 |
| 因舍入导致的偏差 | 0 | 0 |
| 研究四 | 完全一致(*δ* = 0%) | 25 | 83.3 |
| 偏差较小(0% < *δ* < 10%) | 2 | 6.7 |
| 偏差较小(*δ* ＞ 10%) | 3 | 10 |
| 因舍入导致的偏差 | 0 | 0 |

\* 结果数量*N*指在重复分析中，对重复分析结果与原结果进行配对比较的次数。对于每个目标效应，结果包括一组数值，如汇总效应估计(summary estimate，如*t*值/ F值)、置信区间界限(confidence interval bound)、效应量(effect size)样本大小(size effect)等，应将原文中报告的每个数值与重复结果进行比较。例如，在一个*t*检验中，原文献报告了*t*值、95%置信区间、cohen’s d和样本大小，则这个效应中*N*＝4。将各效应的*N*求和即为全体数量。

**表 5 推论的一致性的评估表(原分析方法)**

|  |  |  |  |
| --- | --- | --- | --- |
| 实验 | 推论的一致性 | 数量及占比 | |
| *N \** | *%* |
| 实验一 | 一致 | 28 | 100 |
| 不一致 | 0 | 0 |
| 实验二 | 一致 | 6 | 100 |
| 不一致 | 1 | 1 |
| 实验三 | 一致 | 3 | 100 |
| 不一致 | 0 | 0 |
| 实验四 | 一致 | 5 | 83.3 |
| 不一致 | 1 | 16.7 |

\* 推论数量*N*指在重复分析中，对效应做出统计推断的次数。例如，仅进行了一个*t*检验，则*N*＝1；如果进行了一个2\*2的方差分析，并进行简单效应分析，则有可能有7个统计推断：两个主效应的推论，一个交互作用的推论，四个可能的简单效应分析的推论，因此*N*＝7。如果报告的*p*值相对于重复的*p*值落在显著性水平边界的另一侧，则被归类为推论不一致；反之为推论一致。

**4 讨论**

**4.1 计算可复现性检验结果分析**

结合下表，对原文献进行分析，推测可能导致可复现性检验结果差异的原因。对于重要的原因，逐段进行展开说明。

**表 8 计算上（不）可重复的原因分析表**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **可能原因** | | | **研究一** | **研究二** | **研究三** | **研究四** |
| **原文献开放性问题** | **一般性开放获取问题** | 几个结果的微小差异，可能是由于分析中使用了没有设置固定种子的随机数； |  |  |  |  |
| 个别结果的微小差异，可能是由于印刷或复制粘贴错误； |  | 是，排除后被试总量不对。 |  |  |
| 文章文本中程序报告不明确，包括纳入亚组的标准、缺乏或不正确报告用于回归模型的变量、以及未报告的单侧分析； |  |  |  |  |
| 在文章的开放实践声明中对研究的模糊标记。 |  |  |  |  |
| **OSF 开放获取特定问题** | OSF 中缺乏对数据和/或代码内容进行说明的文档(readme文档)； |  |  |  |  |
| OSF 上的数据与代码文件不一致，如代码中对部分数据进行了操作，但这部分数据在数据文件中无对应； |  |  |  |  |
| OSF上的数据存储问题，包括文件损坏或无法下载。 |  |  |  |  |
| **数据开放获取特定问题** | 没有提供原始数据； |  |  |  |  |
| 没有提供处理后的数据； |  |  |  |  |
| 没有提供数据处理过程的描述或代码。 |  |  |  |  |
| **代码开放获取特定问题** | 缺乏共享的分析代码或建模代码； |  |  |  |  |
| 软件包或软件版本的问题。 |  |  | 是 | 是 |
| **重复过程的原因** | **重复研究与院研究的区别** | 是否使用同样的数据集； | 是 | 是 | 是 | 是 |
| 是否使用同样的数据分析软件及软件包； | 是 | 是 | 是 | 是 |
| 是否使用同样的数据分析方法。 | 是 | 是 | 是 | 是 |
| **重复者相关因素** | 重复者此前是否有过 R 使用经验； | 是 | 是 | 是 | 是 |
| 重复者对关于 R 的知识或操作上存在漏洞，较难理解原文章中的部分操作(可做简单说明)。 | 否 |  | 是，除文章中报告的部分数据的相关代码外，还存在着大量其他代码，不明确具体作用 | 是，除文章中报告的部分数据的相关代码外，还存在着大量其他代码，不明确具体作用 |
| **其他影响因素** | **文献年份** | 文献发表年份是否较为久远，是否在开放科学运动之前； | 否 | 否 | 否 | 否 |
| **文献质量** | 文献引用量大小； | 否 | 否 | 否 | 否 |
| 是否有其他研究支持本文献结果； |  |  |  |  |
| 是否有其他研究对本文献结果进行了重复，重复结果如何(可做简单说明)。 | 否 | 否 | 否 | 否 |

**参考文献(APA格式)**

Rodriguez, M., Schertz, K. E., & Kross, E. (2025). How people think about being alone shapes their experience of loneliness. Nature Communications, 16(1), 1594.

Bauer, G., Breznau, N., Gereke, J., Höffler, J. H., Janz, N., Rahal, R. M., Rennstich, J. K., & Soiné, H. (2023). Constructive replication in the social sciences: Teaching companion and sample syllabus “TC”. Open Science Framework. <https://osf.io/g3k5t/>

Bauer, G., Breznau, N., Gereke, J., Höffler, J. H., Janz, N., Rahal, R.-M., Rennstich, J. K., & Soiné, H. (2025). Teaching Constructive Replications in the Behavioral and Social Sciences Using Quantitative Data. *Teaching of Psychology*, *52*(1), 117-123.

Crüwell, S., Apthorp, D., Baker, B. J., Colling, L., Elson, M., Geiger, S. J., Lobentanzer, S., Monéger, J., Patterson, A., Schwarzkopf, D. S., Zaneva, M., & Brown, N. J. L. (2023). What’s in a badge? A computational reproducibility investigation of the open data badge policy in one issue of Psychological Science. *Psychological Science*, *34*(4), 513–522.

Goodman, S. N., Fanelli, D., & Ioannidis, J. P. (2016). What does research reproducibility mean? *Science translational medicine*, *8*(341).

Janz, N. (2016). Bringing the Gold Standard into the Classroom: Replication in University Teaching. *International Studies Perspectives*, *17*(4), 392–407,

Hardwicke, T. E., Mathur, M. B., MacDonald, K., Nilsonne, G., Banks, G. C., Kidwell, M. C., ... & Frank, M. C. (2018). Data availability, reusability, and analytic reproducibility: Evaluating the impact of a mandatory open data policy at the journal Cognition. *Royal Society Open Science*, *5*(8), 180448.

Hardwicke, T. E., Bohn, M., MacDonald, K., Hembacher, E., Nuijten, M. B., Peloquin, B. N., deMayo, B. E., Long, B., Yoon, E. J., & Frank, M. C. (2021). Analytic reproducibility in articles receiving open data badges at the journal Psychological Science: An observational study. *Royal Society Open Science*, *8*(1), 201494.

Kambouris, S., Wilkinson, D. P., Smith, E. T., & Fidler, F. (2024). Computationally reproducing results from meta-analyses in ecology and evolutionary biology using shared code and data. *PLOS ONE, 19*(3), e0300333. https://doi.org/10.1371/journal.pone.0300333

Karathanasis, N., Hwang, D., Heng, V., Abhimannyu, R., Slogoff-Sevilla, P., Buchel, G., Frisbie, V., Li, P., Kryoneriti, D., & Rigoutsos, I. (2022). Reproducibility efforts as a teaching tool: A pilot study. *PLOS Computational Biology, 18*(11), e1010615.

Kitzes, J., Turek, D., & Deniz, F. (2017). *The practice of reproducible research: Case studies and lessons from the dataintensive sciences*. Oakland: University of California Press.

LeBel, E. P., McCarthy, R. J., Earp, B. D., Elson, M., & Vanpaemel, W. (2018). A unified framework to quantify the credibility of scientific findings. *Advances in Methods and Practices in Psychological Science*, *1*(3), 389–402.

Marshall, E., C., & Underwood, A. (2019). Writing in the discipline and reproducible methods: A process-oriented approach to teaching empirical undergraduate economics research. *The Journal of Economic Education, 50*(1), 17–32.

Open Science Collaboration. (2014). The Reproducibility Project: A Model of Large-Scale Collaboration for Empirical Research on Reproducibility. in I*mplementing Reproducible Computational Research (A Volume in The R Series)*, V. Stodden, F. Leisch, R. Peng, Eds. (Taylor & Francis, New York, 2014), pp. 299–323

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349,* aac4716.

Standing, L.G., Grenier, M.L., Lane, E.A., Roberts, M.S., & Sykes, S. (2014). Using replication projects in teaching research methods. *Psychology Teaching Review*, 20(1), 96–104.

Stojmenovska, D., Bol, T., & Leopold, T. (2019). Teaching Replication to Graduate Students. *Teaching Sociology, 47*(4), 303-313.