积极自我认知的多维结构与功能预测力

分析思路：

1. 相关分析

1.相关分析（自我报告数据、行为任务指标、自我报告数据与行为任务指标）

2.网络分析（各问卷测量总分与行为任务指标）

3.bifactor模型分析

1. 降维与条目筛选

1.EGA分析（筛选后保留142条目，模型自动提取7个社群）

2.随机数据集共性分析（筛选后保留140条目，提取8个公因子）

1. 预测分析

1.利用随机数据集共性分析保留的条目进行分析➡检验模型拟合能力（样本内部拟合）

2.利用EGA分析保留的条目进行分析➡检验模型泛化能力（划分训练集与验证集，检验对“新个体”的预测能力）

摘要：

本研究旨在探讨自我报告数据与行为任务指标在积极自我认知中的结构特性与心理健康预测效度。研究纳入 135 个自我报告条目（问卷数据）、14 个认知任务指标（行为数据）及四项心理健康变量，通过相关分析、EFA、双因子模型、网络分析及交叉验证预测建模进行系统比较。得到如下结果：

1. 相关分析：自我报告数据内部相关高，说明可能存在良好的聚合结构；行为任务指标内部关联弱，说明内部结构可能较为松散；行为任务指标自我报告与行为任务之间相关性较低，说明二者可能反映了积极自我认知的不同方面。
2. 网络分析：进一步验证了自我报告数据与行为任务指标在结构上各自形成相对独立的子网络，自我报告与行为任务之间在结构层面呈现出明显的分离状态。
3. Bifactor分析：积极自我认知更适合解释为多维度、任务特异性的结构，而非单一潜在因子。
4. EGA分析：通过EGA的冗余检验与网络分区，筛除冗余条目后提取出7个稳定社区，代表了积极自我认知的核心维度。
5. 随机数据集共性分析：同时也采用条目共性与随机基线比较法对原始条目进行筛选，综合结构简洁性与模型拟合指标，最终提取8个因子。EFA分析结果显示大多数条目在某一因子上具有清晰主加载，结构聚合良好。
6. 基于数据集对预测分析（样本内部拟合）：在样本内拟合中，自我报告和行为任务均能预测心理健康指标，且行为任务效果更优（平均 R^2 ：0.77 vs. 0.49）。
7. 十折交叉验证预测（泛化拟合）：在“新个体”预测模拟中，两类数据预测力显著下降（平均 R^2 0.02–0.04），但行为任务仍略优于自我报告。

数据预处理：

一、自我报告数据的预处理

本研究纳入了多种自我报告量表以测量积极自我认知及相关心理特质，涵盖了道德领域与能力领域的多个方面。具体包括：道德同一性量表（Aquino, 2002；中文版修订：万增奎, 2008）、道德自我形象量表（Jordan et al., 2015；中文版翻译：刘青兰等, 2020）、罗森伯格自尊量表（Rosenberg, 1965；中文版修订：孙钦铃, 2007）、核心自我评价量表（Judge et al., 2003；中文版翻译：杜建政等, 2012）、自我概念清晰度量表（Campbell et al., 1996；中文版翻译：牛更枫等, 2016）、乐观量表 LOT-R（Scheier et al., 1994；中文版修订：温娟娟等, 2007）、自欺性拔高与操纵印象子量表（Paulhus, 1988；中文版翻译：汪向东等, 1999）、内在控制源量表（Levenson, 1981；中文版翻译：肖莉 & 陈仲庚, 1989）、显性与过度敏感自恋量表（Ames et al., 2006；Hendin & Cheek, 1997；中文版翻译：王晓燕, 2008）、领域自尊量表（MacDonald et al., 2003）、主观社会经济地位量表（胡牡丽等, 2012）等。上述量表涵盖了积极自我认知的核心成分，包括自尊、乐观、自我清晰、自恋、自我呈现偏差与社会地位感知等。

所有问卷均按照原始计分方法进行处理。例如，道德同一性量表采用 5 点计分法，区分内隐与外显两个维度；罗森伯格自尊量表采用 4 点计分法，其中包含反向条目；乐观量表 LOT-R 包含正向与反向条目，最终通过反向计分后加总得到总分。对于所有量表，首先完成反向条目的统一转换，然后计算总分或分维度分数。

在获得原始得分后，首先对异常值进行了检视：若个别被试在某一量表上的得分超过 ±3 个标准差，则视为极端值并剔除；缺失值比例较低（<5%）的量表条目则采用均值填补，缺失比例较高的个案被排除在分析之外。随后，为保证不同量表间的可比性，本研究对各条目与总分均进行了标准化处理（z 转换）。

信效度方面，所使用的量表在既有研究及本研究样本中均表现出较好的内部一致性。例如，道德同一性量表中文版 Cronbach’s α = 0.85，自尊量表 α = 0.835，核心自我评价量表 α = 0.83，自我概念清晰度量表 α = 0.81，道德自我形象量表 α = 0.88，乐观量表 α = 0.78，自欺性拔高与操纵印象子量表 α = 0.68–0.86。整体上，这些指标表明所用量表具有良好的信度，能较为稳定地反映积极自我认知的相关特质。

二、行为任务指标的预处理

1.内隐联结效应IAT的测量与计算

本研究采用内隐联结测验（Implicit Association Test, IAT）评估个体在两个领域中的积极自我认知偏好，并分别为道德领域与能力领域。IAT 任务旨在捕捉个体对“自我/他人”与“积极/消极”词汇之间的自动联结强度。通过反应时数据，计算标准化的 D 值（D-score）作为内隐偏好的效应量指标。

原始反应时数据中包含四个关键 block（A1, A2, B1, B2），其中 A1 与 A2 表示兼容条件（如“我-积极/他人-消极”配对），B1 与 B2 表示不兼容条件（如“我-消极/他人-积极”配对）。在试次选择方面，我们选择保留 IAT 中关键测量阶段（screen\_id = 3, 4, 6, 7）中的反应数据；在反应式剔除方面，我们剔除所有反应时间超过 10,000 毫秒的异常试次；为消除错误反应对结果的干扰，参照 Greenwald 等人（2003）推荐的方法，对所有错误反应试次进行校正：将错误试次的反应时替换为该 block 内的平均反应时加上 600 毫秒惩罚值。之后，对每位被试在每个 block 下的反应时数据进行汇总，计算每个条件的平均反应时、标准差和试次数。

最终，我们按照标准 D 值计算方法进行效应量提取。分别计算 B1 与 A1、B2 与 A2 之间的平均反应时差，并将其除以对应条件下的合并标准差（pooled SD）进行标准化，获得每位被试在moral领域和ability领域下的 IAT D 值。根据 D 值的绝对大小，将效应划分为三类：|D| ≤ 0.2 为小效应，0.2 < |D| ≤ 0.5 为中等效应，|D| > 0.5 为大效应。

2.ALT实验范式与指标计算方法

在本研究中，我们使用了 ALT（Associative Learning Task）实验范式来测量参与者在道德与能力两个维度下，对“自我”与“朋友”关联积极特质的加工差异。该任务通过呈现一系列“匹配”（如“积极词汇-自我”）与“不匹配”（如“积极词汇-朋友”）的组合，要求被试做出快速反应，从而评估其内隐的积极自我关联倾向。

数据清洗方面，首先筛选了所有来自正式实验阶段的试次并排除反应时小于200ms或大于1200ms 的异常值，以保证数据质量。此外，对于缺失的“正确性”标记（correct），默认按错误处理，防止遗漏。同时，我们提取了两个维度的关键指标：反应时差异（ALT\_SE\_rt） 与信号检测敏感性指数（ALT\_SE\_d′），分别反映个体对匹配刺激的加工速度和辨别能力。

具体计算如下：

（1）ALT\_SE\_rt 指标计算：

首先选取所有“匹配”类型的试次（identit为match），并筛选其中“正性词汇”条件（valence = positive）。对于每位参与者，在“自我”与“朋友”条件下分别计算平均反应时（avg\_rt）。用 friend 条件减去 self 条件的平均反应时，得出 ALT\_SE\_rt 值。该值越大，代表对“自我-积极”匹配的加工越快，表征更强的积极自我加工倾向。

（2）ALT\_SE\_d′ 指标计算：

使用信号检测理论（Signal Detection Theory）框架，对每位参与者在“正性词汇”条件下进行分类统计。将试次按反应正确性与刺激类型（match vs. nonmatch）划分为：击中（hit）、错过（miss）、误报（false alarm） 与 正确拒绝（correct rejection）。计算击中率（hit rate） 与 误报率（false alarm rate），并采用调整策略防止极端值（如 hit rate=1 或 FA rate=0）导致Z值计算错误。根据公式，计算每位被试在“自我”与“朋友”条件下的 d′ 值。最终用 self 条件下的 d′ 减去 friend 条件下的 d′，得出 ALT\_SE\_d 值。该值越大，表示被试对“自我-积极”刺激辨别更敏感，具备更高的积极自我识别能力。

最终得到每位被试在 moral 与 ability 两个维度下的 ALT\_SE\_rt 与 ALT\_SE\_d 指标，分别储存在 moral\_ALT\_rt、ability\_ALT\_rt 及 moral\_ALT\_d、ability\_ALT\_d 四个变量中，用于数据分析。

3. SRET实验范式与指标计算方法

在本研究中，我们使用了 SRET（Self-Referent Encoding Task）范式及其延伸任务，系统测量参与者在“道德”与“能力”两个维度下的积极自我认知水平。任务包括三种子任务：词汇判断任务（SRET-EW 与 SRET-RT）、新旧词判断任务（SRET-RJ1）与来源判断任务（SRET-RJ2）。每个任务分别从评估反应偏好、加工速度与记忆敏感性角度，提取了多个自我相关指标。

数据清洗方面，首先仅保留正式实验阶段的试次，并剔除所有练习或非正式试次。在每类任务中，统一剔除 8 个用于练习或与研究目标不相关的干扰词汇。此外，对于反应时低于200毫秒或高于1200毫秒的试次，亦予以排除，以控制极端值对平均指标的干扰。缺失的判断正确性（correct）变量则默认标为错误，以避免漏判。

本研究共提取了八个 SRET 指标，分别对应三类任务，具体如下：

（1）SRET\_EW 指标计算

SRET\_EW 基于词汇判断任务数据。在每位被试的“自我-积极”与“自我-消极”条件中，分别统计其作出“是”判断的次数，两者之差表示自我条件下选择积极词汇的倾向。该指标在道德（morality）与能力（ability）两个领域分别计算，生成 moral\_SRET\_EW 与 ability\_SRET\_EW 两个变量。数值越大，表明被试更倾向于认同与自我相关的积极特质，反映其积极自我认知水平。

（2）SRET\_rt 指标计算

SRET\_rt 同样来源于词汇判断任务，但关注的是反应时数据。在每位被试“自我-积极”与“自我-消极”条件下，计算反应时均值之差。差值越大表示参与者对积极词汇加工更快，反映更高的积极自我加工效率。该指标同样在两个领域上独立计算，得到 moral\_SRET\_rt 与 ability\_SRET\_rt 两个变量。

1. SRET\_RJ1\_d 指标计算

该指标基于新旧词判断任务。任务要求参与者判断屏幕上呈现的词汇是否曾在之前出现，并作出“新”“旧”或“熟悉”反应。根据信号检测理论，将旧词被正确认定为“旧/熟悉”视为击中（hit），将新词误判为“旧/熟悉”为误报（false alarm）。通过公式 d′计算两者d′值 。以“自我”条件下的 d′ 减去“朋友”条件下的 d′ 值，生成 SRET\_RJ1\_d 指标，分别计算 moral\_SRET\_RJ1\_d 与 ability\_SRET\_RJ1\_d。

（4）SRET\_RJ2\_d 指标计算

该指标来源于来源判断任务。参与者需判断词汇来源于“自我”或“朋友”。以“自我”词汇为信号、朋友词汇为噪音，根据击中（hit）与误报（fa）计算 d′ 值。得出 SRET\_RJ2\_d 指标，分别为 moral\_SRET\_RJ2\_d 与 ability\_SRET\_RJ2\_d。

综上，研究共构建了八个SRET指标，覆盖加工偏好（SRET\_EW）、加工速度（SRET\_rt）与记忆敏感性（SRET\_RJ1\_d 与 SRET\_RJ2\_d）三个层面，分别在道德与能力两个维度上进行计算，反映个体在多个层面上的积极自我认知能力。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **指标名称** | **来源任务** | **计算方式** | **心理学含义** | **领域区分** |
| moral\_IAT | IAT任务 | D值（相容-不相容）反应时差/标准差 | 道德领域中积极自我联结的内隐强度 | Moral |
| ability\_IAT | IAT任务 | D值（相容-不相容）反应时差/标准差 | 能力领域中积极自我联结的内隐强度 | Ability |
| moral\_ALT\_rt | ALT任务 | friend - self 平均反应时差 | 对“自我-积极”匹配加工速度更快 | Moral |
| ability\_ALT\_rt | ALT任务 | friend - self 平均反应时差 | 对“自我-积极”匹配加工速度更快 | Ability |
| moral\_ALT\_d | ALT任务 | d prime = Z(hit) - Z(FA)，self - friend 差值 | 对“自我-积极”匹配的敏感性 | Moral |
| ability\_ALT\_d | ALT任务 | d prime = Z(hit) - Z(FA)，self - friend 差值 | 对“自我-积极”匹配的敏感性 | Ability |
| moral\_SRET\_EW | SRET任务 | positive\_self - negative\_self（yes响应数量差值） | 积极词更倾向归于自我 | Moral |
| ability\_SRET\_EW | SRET任务 | positive\_self - negative\_self（yes响应数量差值） | 积极词更倾向归于自我 | Ability |
| moral\_SRET\_rt | SRET任务 | negative\_self - positive\_self（反应时差值） | 反应更快地接受积极词归于自我 | Moral |
| ability\_SRET\_rt | SRET任务 | negative\_self - positive\_self（反应时差值） | 反应更快地接受积极词归于自我 | Ability |
| moral\_SRET\_RJ1\_d | SRET-新旧词判断任务 | d prime= Z(hit) - Z(FA)，self - friend 差值 | 判断“是否出现过”的敏感性，self为目标 | Moral |
| ability\_SRET\_RJ1\_d | SRET-新旧词判断任务 | d prime= Z(hit) - Z(FA)，self - friend 差值 | 判断“是否出现过”的敏感性，self为目标 | Ability |
| moral\_SRET\_RJ2\_d | SRET-来源记忆任务 | d prime= Z(hit) - Z(FA)，self为信号、friend为噪音 | 判断“来源是否是自己”的能力 | Moral |
| ability\_SRET\_RJ2\_d | SRET-来源记忆任务 | d prime= Z(hit) - Z(FA)，self为信号、friend为噪音 | 判断“来源是否是自己”的能力 | Ability |

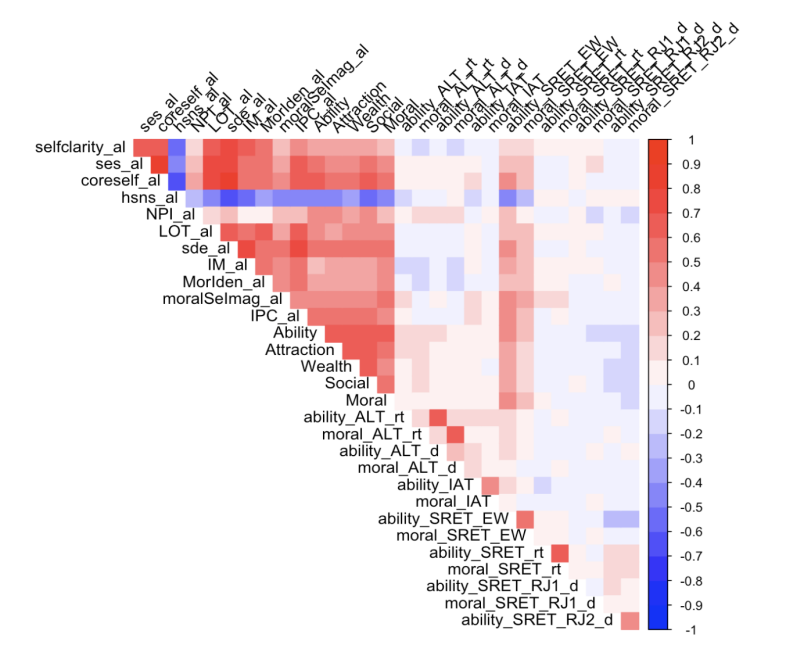
结构层面：

一、基于问卷与行为任务指标的相关分析

1.相关分析（自我报告数据、行为任务指标、自我报告数据与行为任务指标）

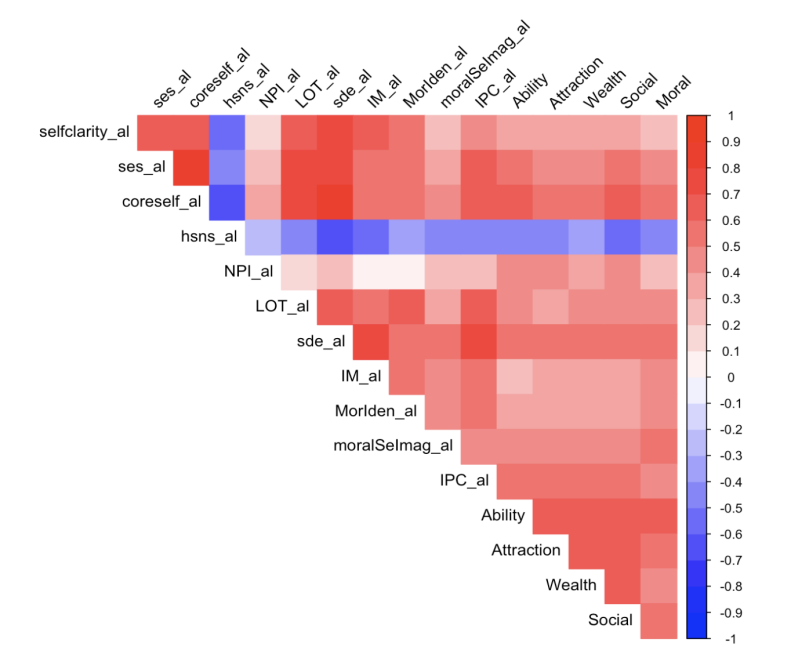
（1）对提取到的所有总分进行相关分析

为探究自评与认知任务指标之间的测量一致性，本研究计算了变量间的皮尔逊相关系数并使用 corrplot 绘制相关热图。结果显示，自我报告变量之间相关较高，而行为任务之间及其与自我报告的相关性普遍较低，提示显性与隐性测量结构存在分离。



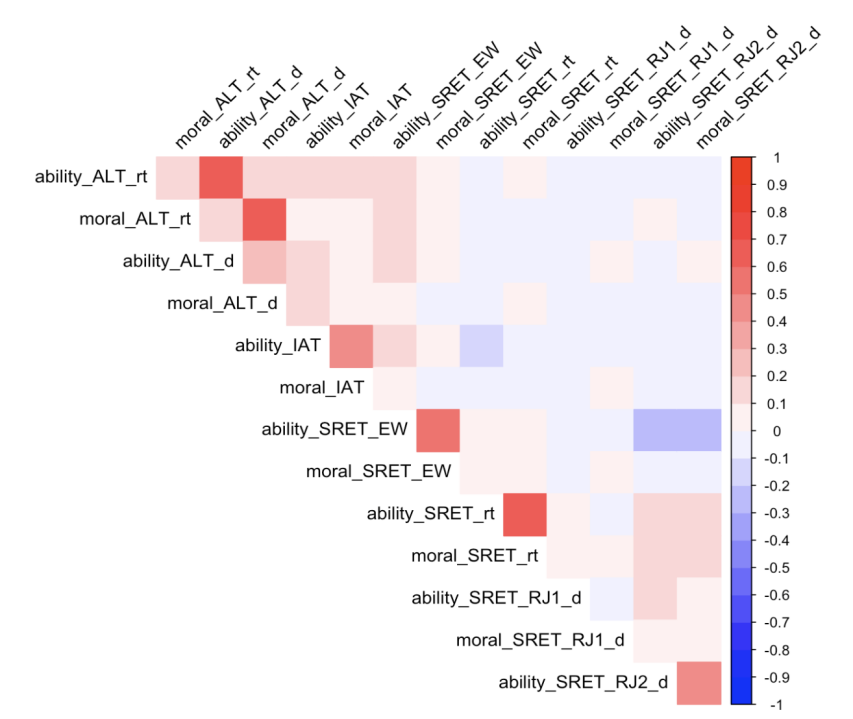
（2）对问卷总分进行相关分析

为检验各自我报告变量之间的关系，计算了前16列变量（自我报告数据）的皮尔逊相关系数并绘制相关热图。结果显示，大部分指标间存在中等到较高的正相关，反映出积极自我相关特质量表之间具有较强的一致性和结构耦合性。



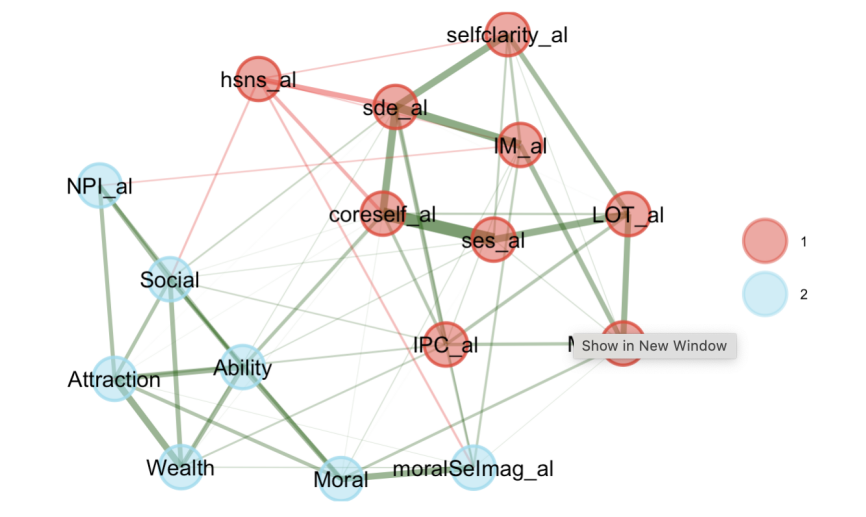
（3）对行为任务进行相关分析

为检验各行为任务指标之间的关系，计算了第 17 至 30 列变量（行为任务指标）的皮尔逊相关系数，并绘制相关热图。结果显示，不同任务间的相关性整体偏低，部分指标甚至呈现无关或负相关，表明行为任务在内部结构上较为分散，可能反映了不同的认知加工机制或任务特异性较强。



1. 网络分析（各问卷测量总分与行为任务指标）

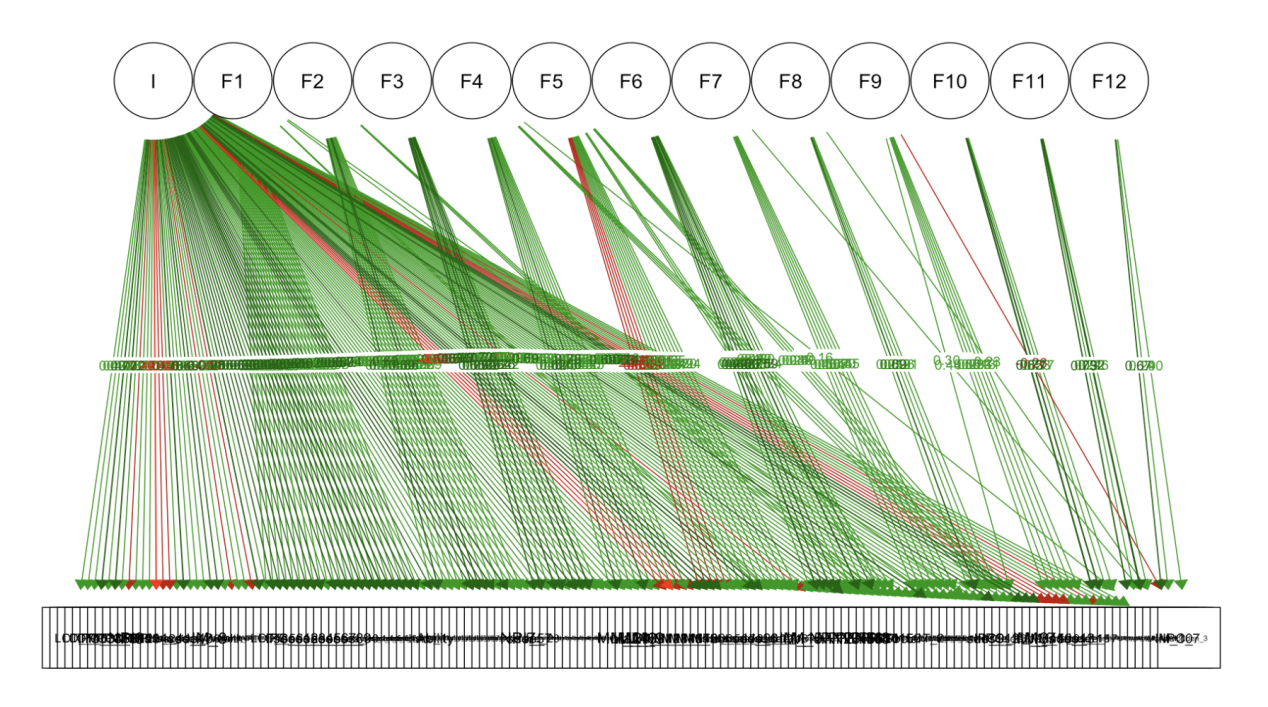
为检验自我报告数据与行为任务指标变量的结构联结模式，分别对两类变量进行了网络分析（EGA, model = “glasso”）。结果显示，两组变量在图结构上各自形成相对独立的子网络，节点之间的连接主要集中在各自组内，表明自我报告与行为任务之间在结构层面呈现出明显的分离状态。



3.Bifactor模型分析

在进一步检验积极自我认知结构的潜在共性时，我们构建了包含通用因子（general factor）与多个特异因子（specific factors）的双因子（bifactor）模型。拟合结果显示，模型的 RMSEA = 0.049 与 SRMR = 0.069，均处于可接受甚至理想范围，表明残差结构尚可；但 CFI = 0.735 与 TLI = 0.728 明显低于理想拟合标准（通常要求 > 0.90），提示该模型在整体结构拟合方面存在偏差。

模型结果显示，尽管我们设定了一个统摄全部测量指标的共同因子（I），但其解释力相对有限，模型更倾向于表现为由多个特异因子主导的多维结构。



二、条目筛选与降维

1.EGA分析（筛选后保留142条目，模型自动提取7个社群）

软件包：

步骤：

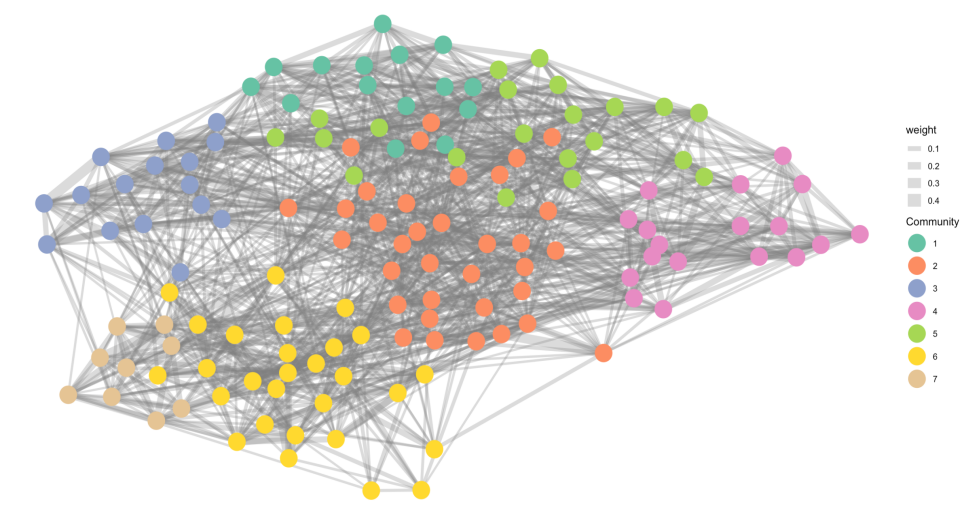
相关参考文献：

EGA分析先做冗余剔除与网络分群，再在结构框架内保留稳定条目。最终保留约 132 个自报条目与 10 个行为指标。

基于经冗余剔除后的标准化数据（n = 142项指标）采用图形估计法（EGA）进行维度结构探索，使用GLASSO模型结合Walktrap算法识别潜在社区结构。结果显示，指标间存在丰富的连接（142条节点，1427条边，边密度为0.143），根据条目的聚合，可以初步划分为自我清晰度、认知风格、社会支持、自我监控、核心自我概等七个稳定的社区。

利用Bootstrap EGA进行500次重采样验证社区划分的稳定性。维度稳定性分析显示，七个社区的结构一致性差异明显，其中2个社区的结构稳定性较高（结构一致性分别为0.904和0.952），表明这两个社区的维度划分高度稳健。4个社区的结构一致性中等（0.484至0.722之间），1个社区稳定性较低（0.278），提示该维度可能存在结构不确定性。

总的来说，分析揭示了积极自我认知测量变量的多维结构特性，以及两类测量在统计结构上的差异性。具体来看，自我报告条目高度聚合于同一社区，而行为任务指标则分布较为分散，社区划分不够明确。社区划分结果与理论预期基本吻合，支持显性测量与隐性测量在结构上呈现显著分离。



2.随机数据集共性分析（筛选后保留140条目，提取8个公因子）

软件包：

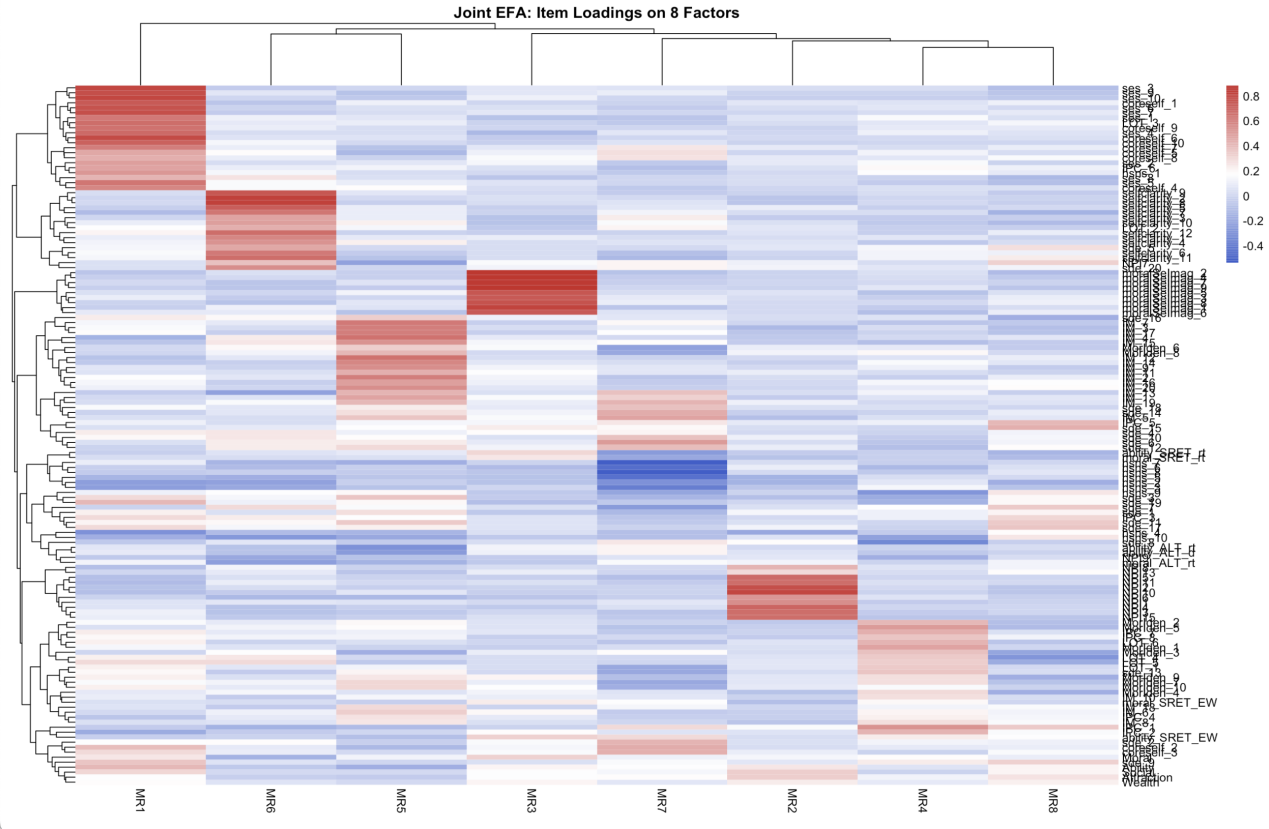
步骤：

相关参考文献：

为排除结构不明确或解释力较弱的变量，采用条目共性与随机基线比较法对原始条目进行筛选。具体地，通过计算每个条目在真实数据下的共性，并与相同结构的随机数据中该条目的共性分布进行比较，仅保留共性高于随机分布95%分位的条目（即在结构贡献上显著强于随机性的条目）。

最终保留了 140 个结构贡献显著的条目。随后，使用平行分析、MAP检验与BIC等方法综合判断因子数，结果显示建议因子数在8～11之间。综合结构简洁性与模型拟合指标，最终提取8个因子，采用最小残差法（minres）与oblimin斜交旋转进行EFA分析。结果显示大多数条目在某一因子上具有清晰主加载，结构聚合良好。

**（同样作为条目筛选的方法，EGA自动将输入的数据进行冗余条目的剔除，剔除后自动将剩余的条目与指标进行“社群”的划分，并且就筛选结果来看，最大程度上的保留了行为任务指标，同时可以进行社区稳定性和条目稳定性的检验，并对生成的社群进行命名；而另一种方法则是通过生成的随机数据的共性作为引入的“黄金标准”来筛选，再将筛选后的数据通过EFA来提取公因子。两者作用相似，保留其一即可。）**



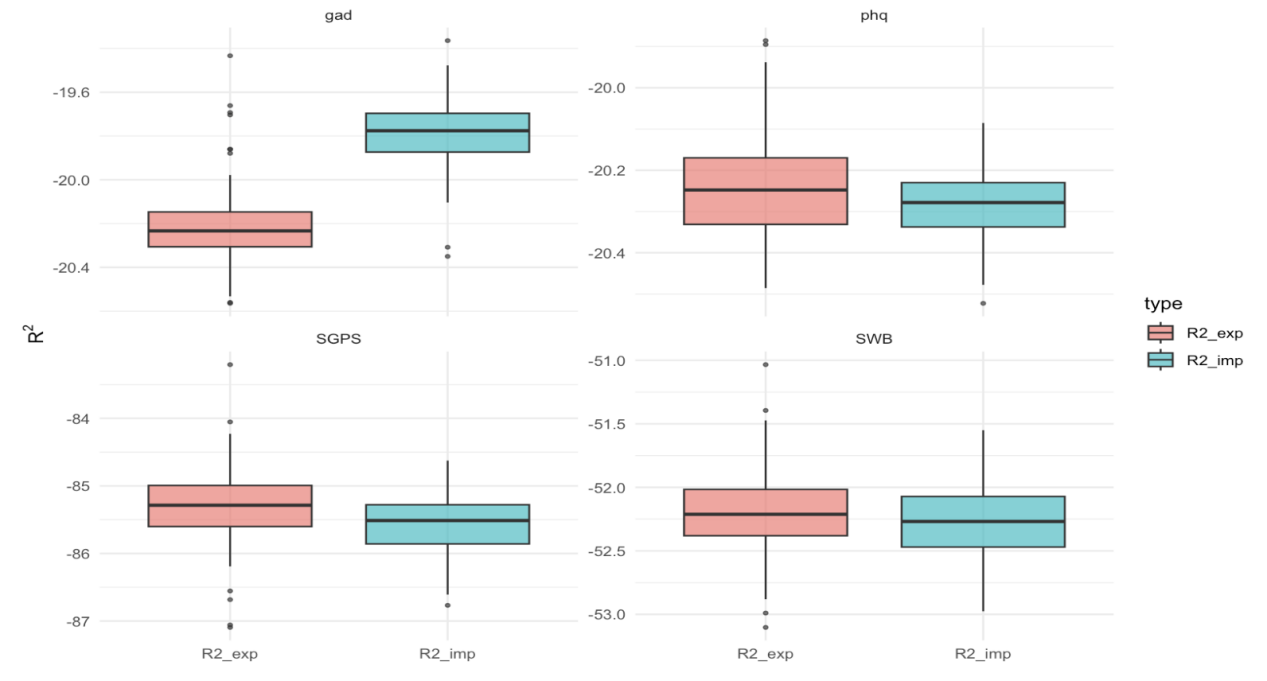
功能层面：

三、预测分析（**可以选择上一步骤中一个筛选方法的结果作为预测分析的数据**）

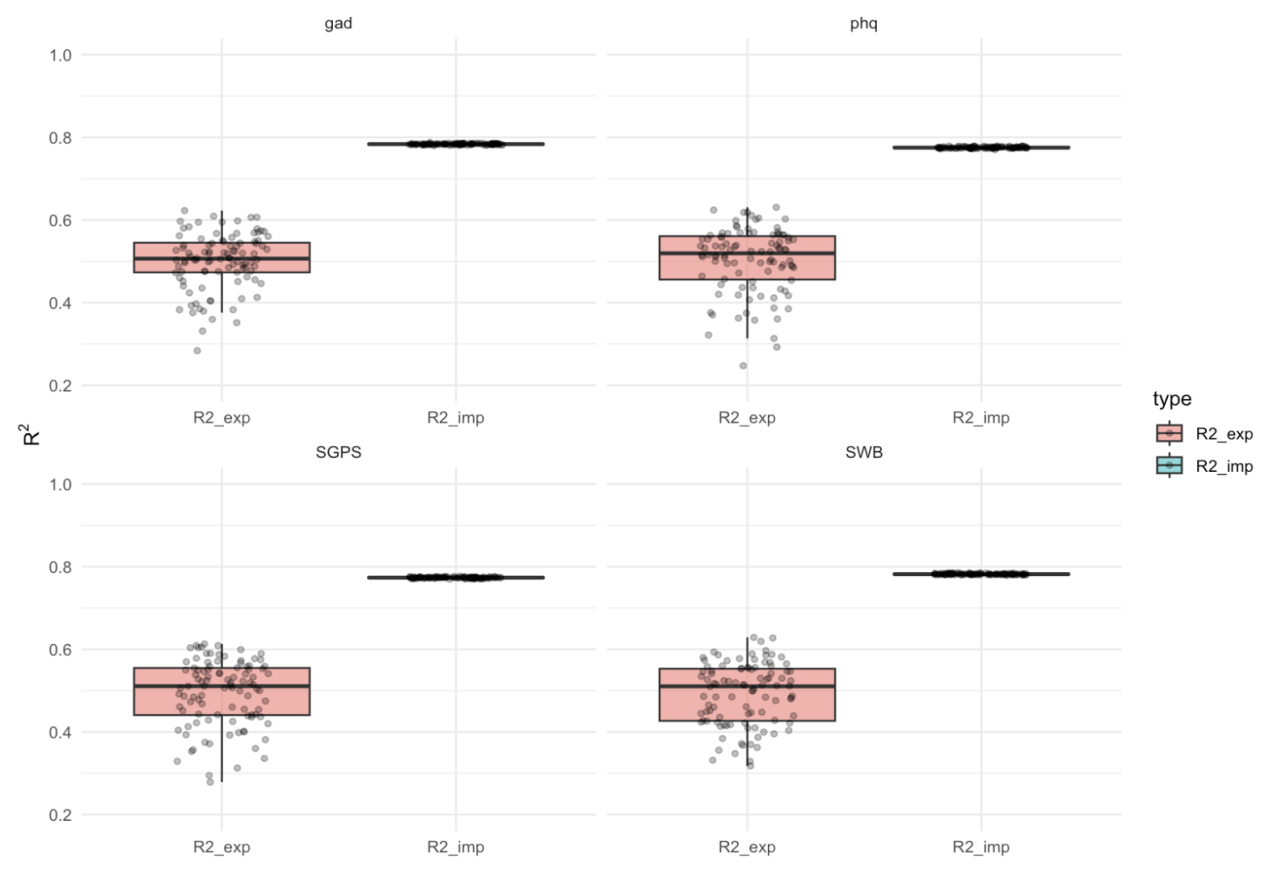
1.利用随机数据集共性分析保留的条目进行分析➡检验模型拟合能力（样本内部拟合）

在这一部分，我们利用筛选后的全部的数据进行初步的预测分析。为了比较自我报告数据与行为任务指标在健康变量预测中的效力，我们采用了两种建模策略：稀疏线性模型（LASSO 回归） 与 非线性集成模型（随机森林）。考虑自我报告数据数量远多于行为任务指标，所以我们采用“等量抽样 + 多次重复建模”的方法，具体操作如下：

1. 线性模型：我们利用 glmnet 包中的 cv.glmnet 函数，基于 LASSO 回归（α = 1） 进行建模。该方法通过引入 L1 正则化项对预测变量进行收缩与选择，从而抑制多重共线性并筛选出最具预测价值的特征。为了防止过拟合并确定最优的正则化参数 λ，模型在训练过程中嵌入了 10 折交叉验证，即将训练集划分为十个子集，轮流作为验证集，计算各 λ 下的预测误差，并选取最优 λ 拟合最终模型。线性模型（LASSO）结果显示，无论自我报告数据还是行为任务指标，R² 皆为负，表明线性关系不足以解释心理健康变量的方差。



（2）非线性模型：采用随机森林回归模型（randomForest）。在每一次重复迭代中，对于自我报告数据，我们从 133 个条目中 随机抽取 7 个条目进入模型，以模拟有限指标下的预测效力；对于行为任务数据，则使用全部 7 个指标。每次模型包含 500 棵决策树，预测性能以 均方根误差（RMSE） 与 决定系数（R²） 衡量。为了获得稳健结果，我们进行了 100 次重复迭代，并在每次迭代后记录预测性能，最终对结果取平均。随机森林（RF）结果显示，预测效度在样本内显著提升，且行为任务测量显示出明显优势（行为任务数据R² 约 0.78；自我报告数据R² 约 0.5–0.6）。



2.利用EGA分析保留的条目进行分析➡检验模型泛化能力（划分训练集与验证集，检验对“新个体”的预测能力）

为评估积极自我认知的两类测量在健康指标预测中的泛化效力，我们分别基于自我报告数据（132个问卷条目）与行为任务数据（10个认知任务指标）对四个连续健康结局（SGPS、SWB、GAD、PHQ）进行建模，依然分别建立了线性模型与非线性模型：

线性模型：首先采用了基于 LASSO/Ridge/Elastic Net 正则化的线性模型（glmnet）。该模型在十折交叉验证（10-fold cross-validation）的框架下运行：将全部样本随机划分为十个子集，每次用其中九个子集训练模型，剩余一个子集用于验证，循环十次并取平均性能指标，从而估计模型在新数据上的泛化能力。同时，glmnet 在每一折训练集内部还进一步使用了内嵌交叉验证，以选择最优的正则化参数 λ，从而在偏差与方差之间取得平衡，避免过拟合并确保模型稀疏性与稳健性。因此在外层交叉验证中估计泛化性能的同时，也通过内层交叉验证实现了参数调优。

非线性模型：采用随机森林（method="rf"），在十折交叉验证的框架下进行性能评估。在此过程中，数据同样被划分为十个子集，依次进行训练与验证，以获得模型在未见数据上的预测表现。不同于线性模型的双层交叉验证，随机森林的调参过程直接嵌入在外层十折交叉验证中：通过设定 tuneLength=5，确定最优超参数（tuneLength=5，importance=TRUE），并基于该参数选择的结果计算预测性能，捕捉特征之间的复杂非线性关系。两类模型均以均方根误差（RMSE）与决定系数（R²）作为性能指标，度量在未见数据上的预测误差与解释度。对每个健康指标，分别基于自我报告数据与行为任务数据独立建模并评估性能；最终输出各模型在十折交叉验证下的平均 RMSE 与 R² 作为泛化性能估计。

结果显示：

1.总体预测效力：在十折交叉验证框架下，两类数据源的总体 R² 较低（通常接近 0–0.04）。这表明以当前特征集和样本规模，对健康指标的可泛化预测效力有限。

2.自我报告数据 vs 行为任务数据：行为任务数据在多数结局上表现出略高的交叉验证 R²，同时 RMSE 相当或略低，而自我报告数据整体预测力稍弱，这一模式在随机森林模型中更为一致。这说明，行为任务所捕捉的过程性特征可能与健康状态存在更直接（或更非线性）的关联，自我报告可能更多反映主观体验，与健康结果的关联在预测层面较弱。

