积极自我认知的多维结构与功能预测力

1. 结构层面

相关分析

网络分析

bifactor模型分析

1. 降维与条目筛选

EGA分析（保留142条目，模型自动提取7个社群）

随机数据集共性分析（保留140条目，提取8个公因子）

1. 预测分析

随机数据集共性分析保留的条目➡拟合能力

EGA分析后保留的条目➡泛化能力

摘要

本研究旨在探讨自我报告与认知任务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。结果显示，两类测量在结构层面高度分离：自我报告变量内部一致性高、结构聚合良好；行为任务指标结构松散、内部关联弱。双因子模型与网络分析均表明，积极自我认知呈多维、任务特异性结构，而非单一潜在因子。在控制条目、对噪音较大的条目进行筛选后进行了积极自我认知的功能预测分析，结果显示，在 LASSO 线性模型中，两类测量对心理健康的预测力均有限；在随机森林非线性模型中，预测力相较于线性模型有所提升，随机森林预测结果显示：认知任务表现出较显著的预测优势，但这一结果仅停留在检验数据的“拟合性能”上，不涉及对“新个体”数据预测能力的评估。在这一基础上，我们利用十折交叉验证进一步分析，将数据划分为“训练集”和“验证集”，在独立数据划分下评估模型的“泛化性能”。分析结果显示，认知任务在预测效能上依然略高于自我报告，但整体预测力下降。研究揭示了积极自我认知可能存在多维的结构，同时在其功能预测上，自我报告与认知任务存在差异，并提出“结构—功能悖论”框架，强调在心理测量中需兼顾结构有效性与功能预测价值。

数据分析与结果

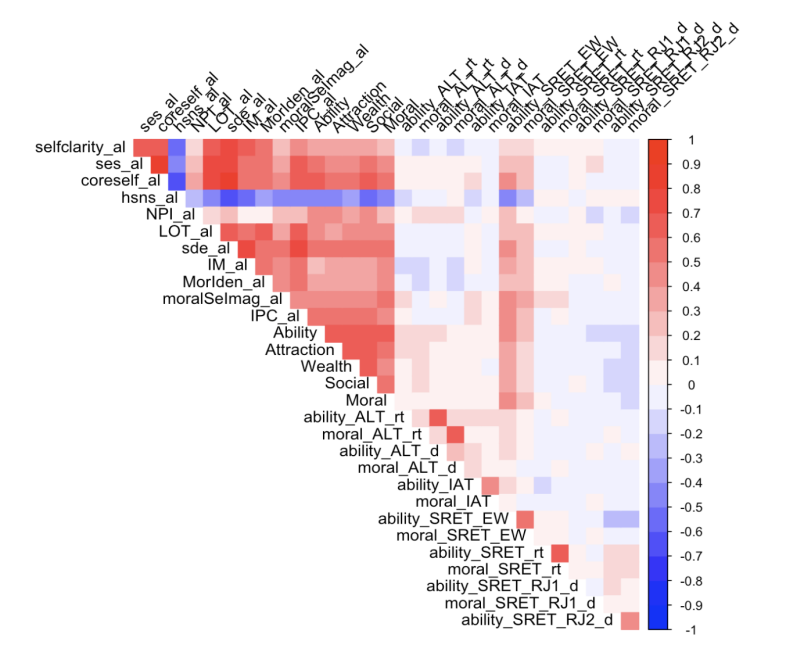
一、结构层面：相关、网络与双因子模型

为检验自我报告与行为任务在结构上的耦合与分离，我们从相关结构、网络结构与潜在因子结构三个层面开展分析。总体上，两类测量在统计结构上高度分离：自我报告条目内部关联强、结构聚合度高；行为任务指标内部关联弱、结构松散，提示二者可能反映不同的心理加工系统。

1. 相关分析

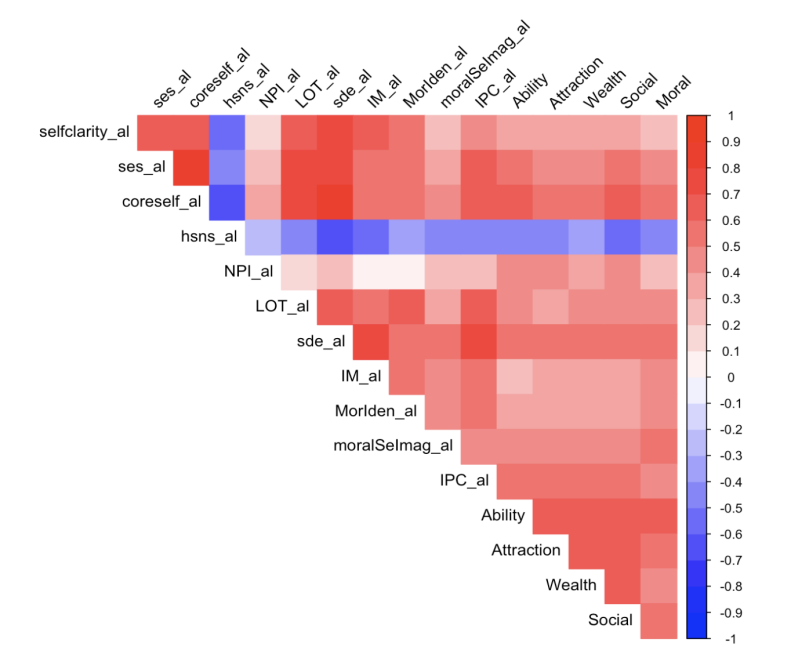
（1）对提取到的所有总分进行相关分析

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。



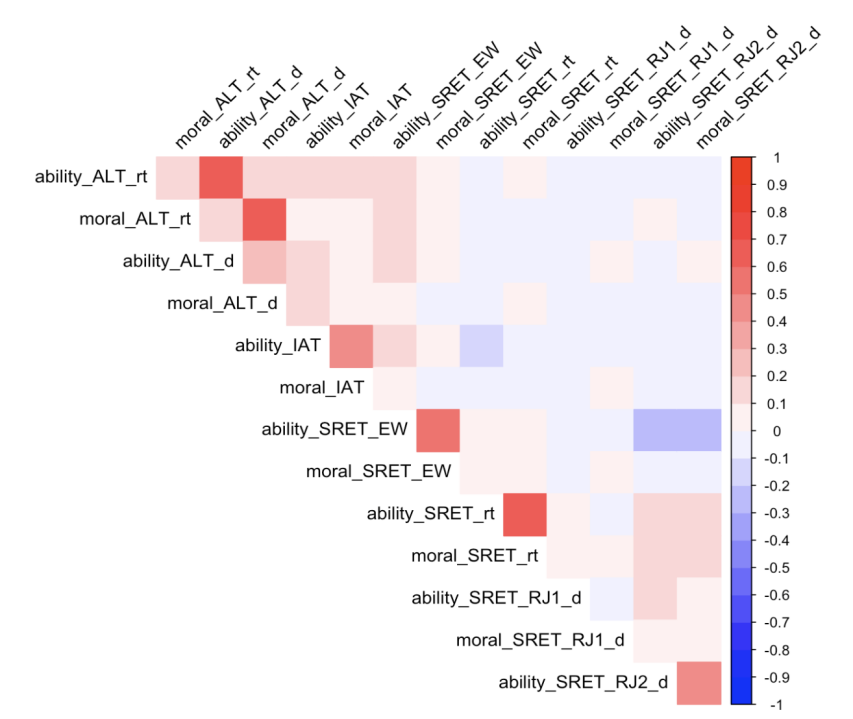
（2）对问卷总分进行相关分析

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）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。



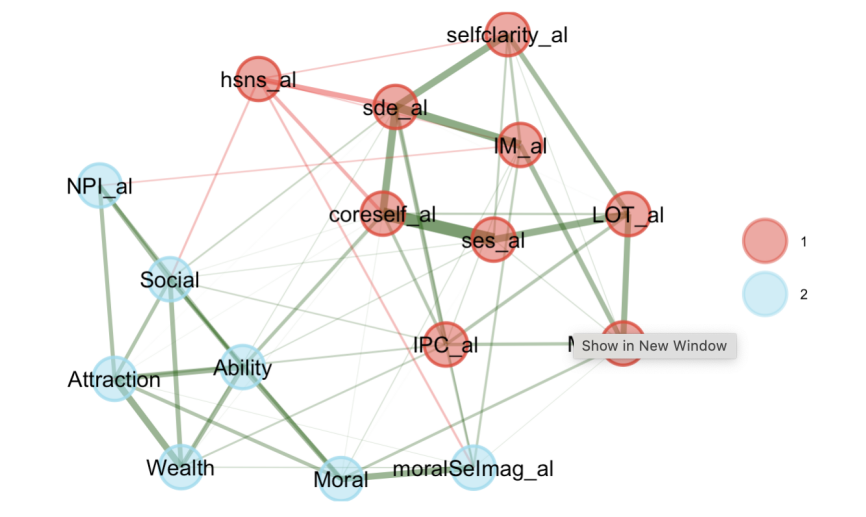
（3）对行为任务进行相关分析

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）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。



2. 网络分析（EGA）

为检验自我报告与行为任务变量的结构联结模式，分别对两类变量进行了基于图模型的网络分析（EGA, model = “glasso”），并构建了合并的邻接矩阵进行联合可视化。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，跨组边稀疏，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。

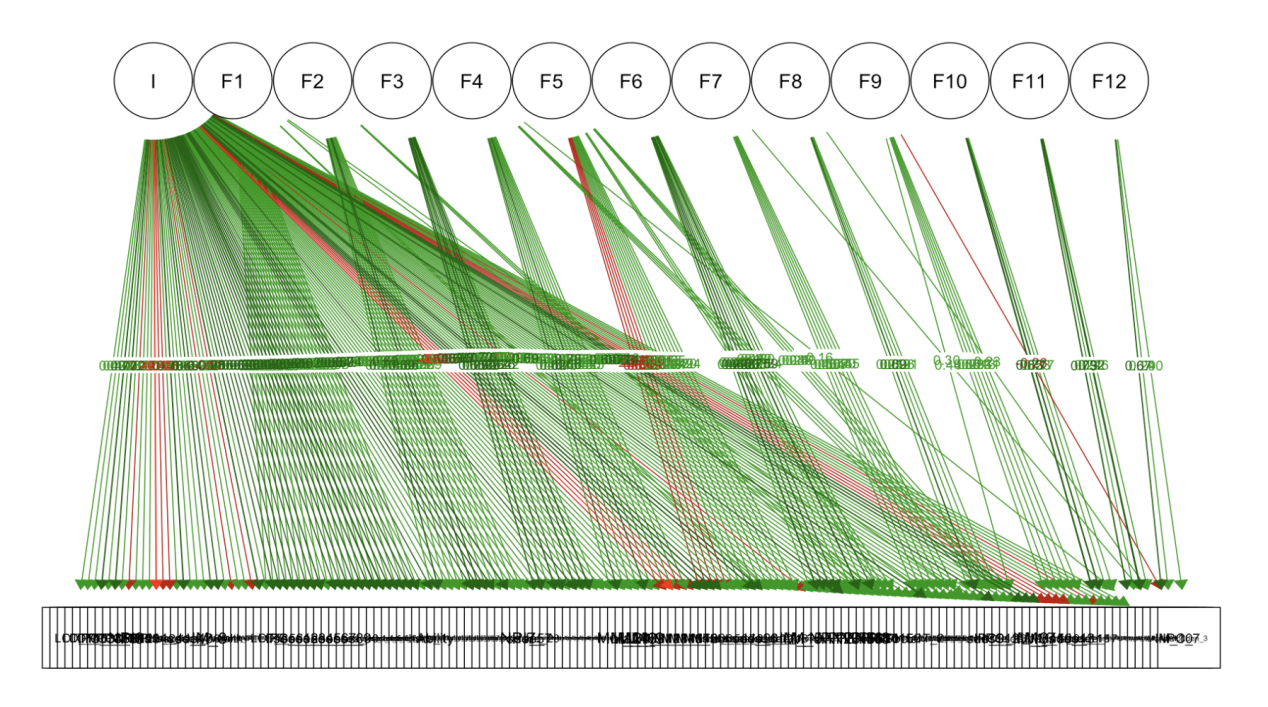


3. 双因子（bifactor）模型

在进一步检验积极自我认知结构的潜在共性时，我们基于标准化数据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。

结合模型结构与加载结果可见，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。这一结果暗示，在积极自我认知的测量中，尽管不同任务（如IAT、ALT、SRET）均意在揭示个体对自我持有的积极加工倾向，但它们可能反映了不同的心理机制——如反应偏好、加工速度、信号辨别能力等。也就是说，积极自我认知可能并非源于统一的潜在构念，而更可能是一种多维且任务特异性较强的结构。

因此，本研究结果不支持将积极自我认知简化为单一共性因子所主导的结构模型，而支持其呈现出多种潜在加工成分的可能性。这也为未来在建构积极自我认知的理论模型时提供了证据支持：应充分考虑其心理加工的多样性与测量工具的异质性，而非假设其具有单一、内隐一致的潜在因子结构。



⸻

⸻

二、条目筛选与降维

为缓解维度高、冗余大等问题，我们实施了两条并行路径：

1. EGA 驱动的筛选：

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 133 个自报条目与 7 个行为指标。

本研究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），最终划分出七个稳定的社区。

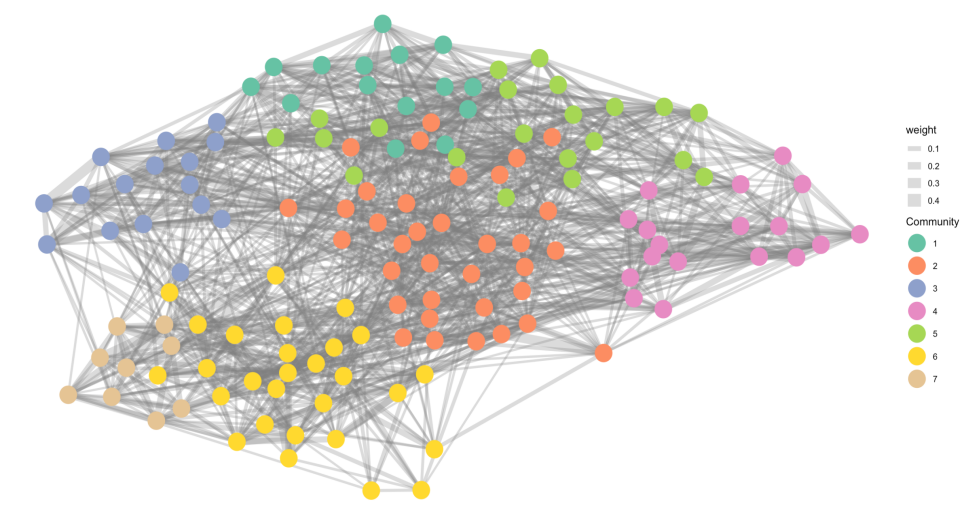
具体来看，七个社区中指标的分布及心理学意义如下：

社区1 主要由selfclarity\_1至selfclarity\_12等12个“自我清晰度”相关指标组成，代表个体自我认知的稳定性和明确度；社区2 包含LOT\_1、LOT\_4、LOT\_6及IPC\_1至IPC\_3、IPC\_5、IPC\_6、IPC\_7、IPC\_8，以及sess\_1至sess\_5等多个社会支持和认知风格指标，反映个体的乐观倾向、社会支持感知及信息加工特点；社区3 由LOT\_3、ses\_6至ses\_10、coreself\_2、coreself\_5、coreself\_6等组成，代表社会自我认同及情绪调节相关维度；社区4 以IPC\_4及多项自我内省测量（如IM系列）为主，反映个体内部动机及自我监控能力；社区5 由多个coreself项目和sde系列指标构成，体现自我核心概念及其稳定性；社区6 主要聚集了以NPI（自恋人格量表）、ability、attraction、wealth、social和moral等行为任务相关指标，代表个体能力、自我吸引力及社会-道德认知表现；社区7 包含moralSeImag\_1至moralSeImag\_9等9个道德自我意象相关指标，反映个体在道德领域的自我表征。

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，其中第1社区（自我清晰度）和第7社区（道德自我意象）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。第2、3、5、6社区的结构一致性中等（0.484至0.722之间），而第4社区稳定性较低（0.278），提示该维度可能存在结构不确定性。

条目稳定性分析结果显示，大部分指标在500次重采样中保持高度稳定（条目重复比例大多数接近或等于1.0），少数条目如IPC\_4（0.316）、coreself\_4（0.264）等稳定性较低，可能需进一步考察其测量特征。

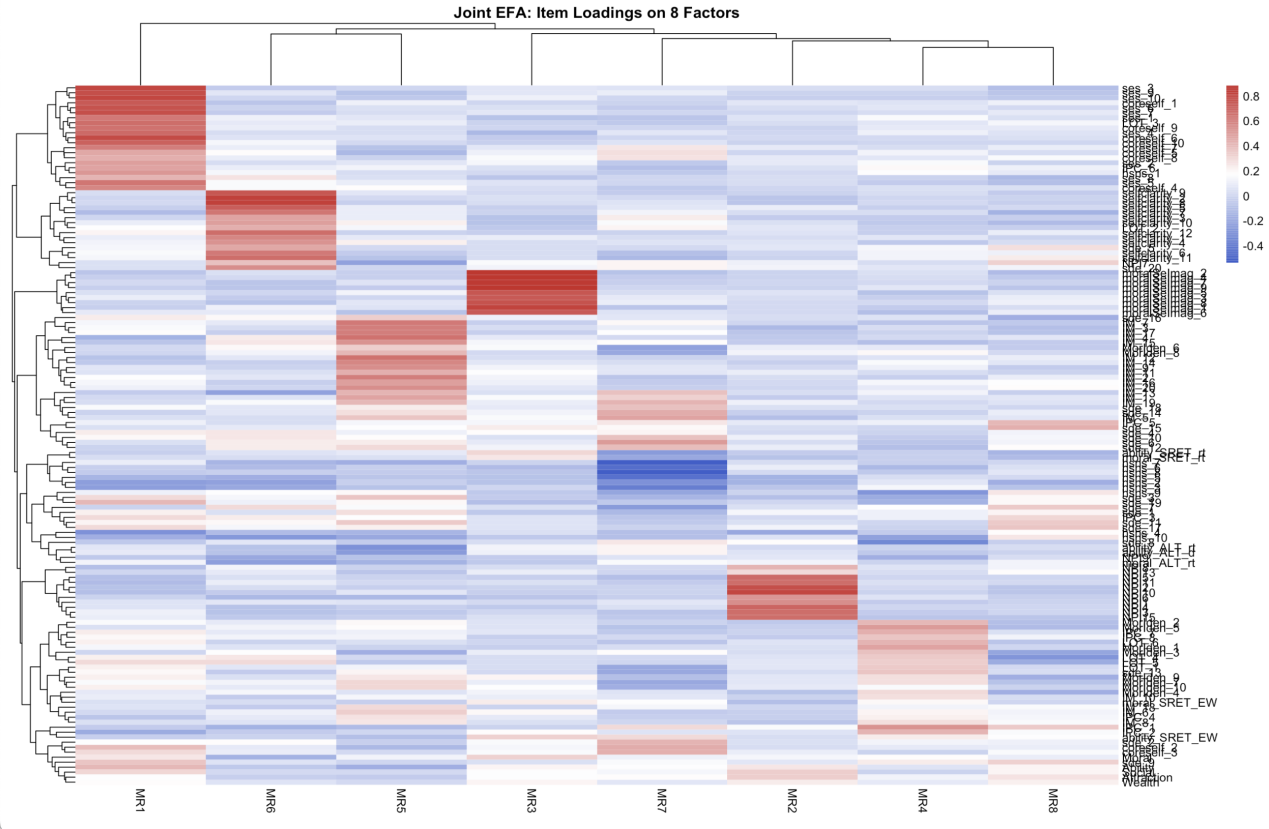
总体而言，EGA结果揭示了样本数据中七个具有较好解释力和稳定性的潜在维度，涵盖了自我清晰度、认知风格、社会支持、自我监控、核心自我概念、行为任务表现及道德自我意象等多个心理学领域。在问卷与行为数据方面，EGA结果显示，积极自我认知指标在网络中明显分为多个社区，反映出测量变量间的聚合结构。具体来看，自我报告条目高度聚合于同一社区，表明其在结构上表现出较强的内部一致性；而行为任务指标则分布较为分散，社区划分不够明确，反映其内部关联较弱，结构松散。社区划分结果与理论预期基本吻合，支持显性测量与隐性测量在结构上呈现显著分离。本部分分析揭示了积极自我认知测量变量的多维结构特性，并凸显显性测量与隐性测量在统计结构上的差异性。这一发现为后续基于网络结构的心理机制分析和功能预测提供了重要基础。



1. “公平对比”的等量抽样策略：

考虑显性条目数量远多于隐性指标，采用等量抽样 + 多次重复建模：从 133 个显性条目中随机抽取与隐性指标等量的 7 项，分别与 7 个隐性指标建模，重复 100 次，对比平均 R²/ RMSE 与变量重要性分布，以避免“数量优势”带来的偏倚。

为提升因子结构的清晰度与条目的结构贡献度，这部分研究首先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）。此方法有效排除了结构不明确或解释力较弱的变量，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，仅少数条目因载荷不足被进一步剔除。最后，通过因子–条目聚类热图直观展示了条目在因子空间中的结构分布，验证了条目归属的合理性与因子结构的可解释性。



⸻

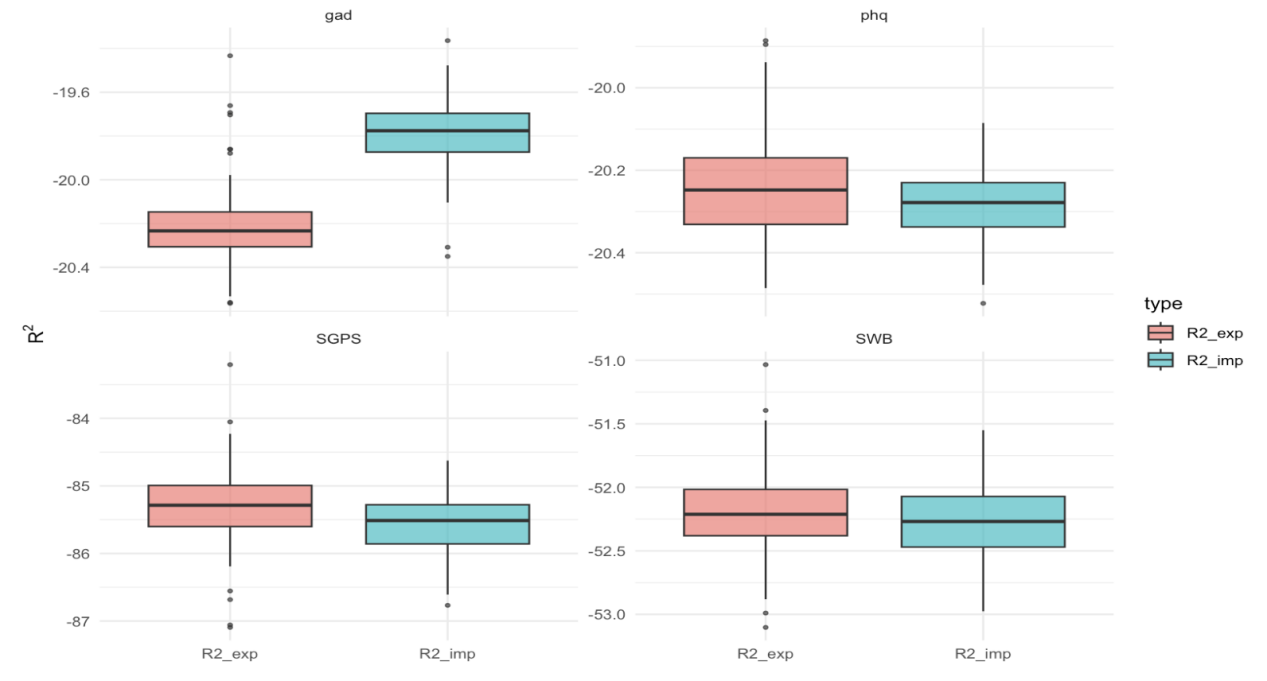
结合前述相关分析与因子结构提取的结果可以看出，积极自我认知可能并非单一维度构念。在不同测量方式之间相关性较低的同时，结构性分析进一步表明，在共性筛选后的条目中存在多个聚合良好的因子维度。该结果提示，个体的积极自我可能由多个侧面构成，具有内在结构复杂性。

三、功能层面：预测分析

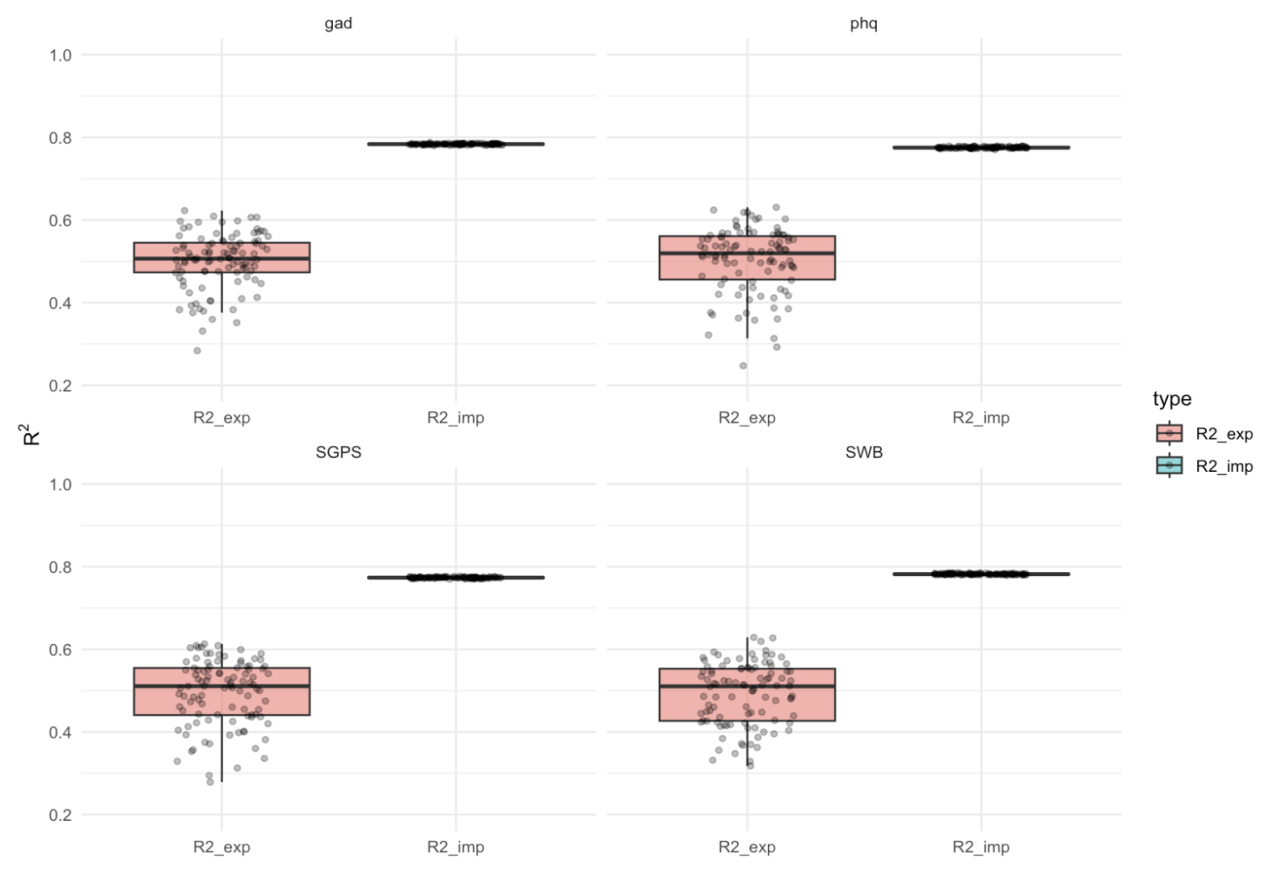
我们依次开展了样本内拟合与泛化评估两步，以避免只看“好看的”样本内指标。

1.样本内拟合（全部样本上训练与评估）

线性模型（LASSO）：无论显性或隐性，R² 皆为负，表明线性关系不足以解释心理健康变量的方差。



随机森林（RF）：预测效度在样本内显著提升，且隐性测量显示出明显优势（文档报告隐性 R² 可达约 0.78；显性 R² 约 0.5–0.6），提示非线性与高阶交互可能承载了额外的预测信息。



这些样本内结果呈现出典型的“结构—功能悖论”：结构更清晰的显性测量未必在功能预测上更强，而结构松散的隐性指标在非线性模型中可能更有用。

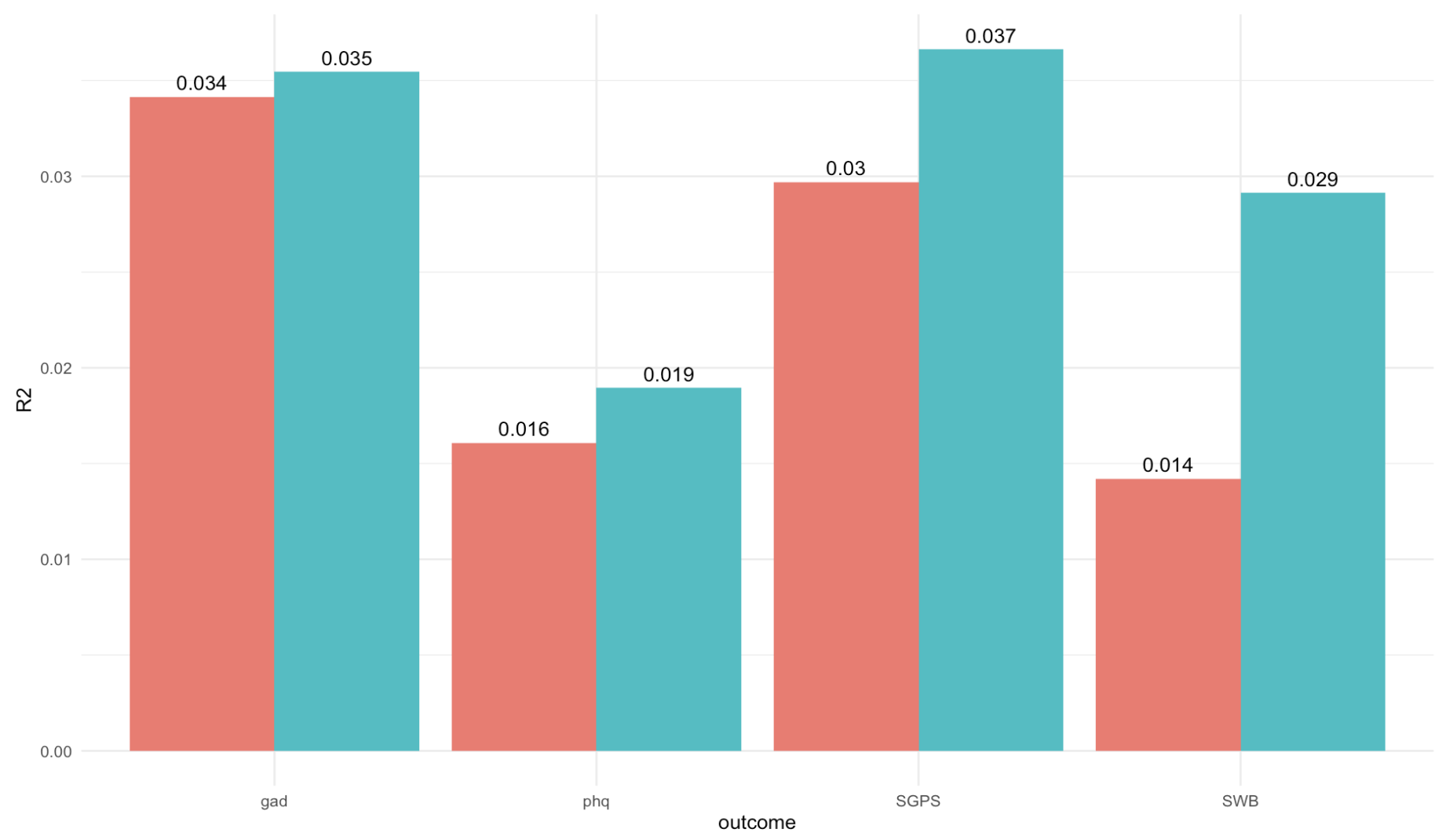
2.泛化评估（十折交叉验证）

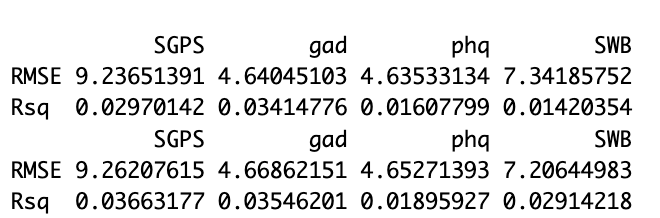
为评估对新个体的可推广性，我们对显性与隐性分别进行十折交叉验证的预测。

为了进一步探索积极自我认知对心理健康对预测效果，同时验证在行为数据与问卷数据在预测效果方面发现对差异，比较行为任务与自我报告在结果变量预测中的相对效能，我们基于随机森林模型分别构建两类预测模型。在初步分析中，我们使用全部数据拟合模型，计算各模型的决定系数（R^2）以评估拟合性能。随后，为检验预测效能的稳健性并降低过拟合风险，我们采用十折交叉验证（10-fold cross-validation），在独立数据划分下评估模型的泛化性能。

分析结果表明，行为任务测量指标在所有结果变量上的拟合性能（R^2）均高于问卷测量指标，显示出行为数据在预测模型中的相对优势。在十折交叉验证条件下，两类模型的预测效能均有所下降，符合独立样本预测趋于保守的预期。然而，行为测量的预测效能在各结果变量上依然略高于问卷测量，且该差异虽较原分析有所减小，但仍保持一致的方向。这一结果提示，行为测量在预测上的优势在控制过拟合风险后依然稳健存在。

在整体预测力方面，无论是基于问卷的显性积极自我指标，还是基于行为任务的隐性积极自我指标，对心理健康相关变量（SGPS、GAD、PHQ、SWB）的预测效应均较为有限。在十折交叉验证中，两类指标的决定系数（R²）均处于较低水平（约 0.014–0.036），其中隐性指标在多数心理健康变量上表现出略高的预测力，但优势幅度极小（差异约 0.002–0.007）。与原始拟合结果相比，交叉验证显著降低了模型的预测效应，提示原模型可能存在一定的过拟合现象，而交叉验证结果更能反映其真实的泛化能力。

⸻



⸻