

人工智能时代的心理测量革新

工作坊

本次人工智能时代的心理测量革新工作坊汇聚全球 7 位顶尖专家，精心打造连接传统心理测量学与人工智能技术的跨学科对话平台。我们致力于：

1. 理论-技术深度融合：从经典测量理论(项目反应理论、计算机化多阶段测验)到 AI 前沿技术(生成式 AI 测评、监督式机器学习建模)；
2. 测评全流程革新：覆盖题目生成、评分设计、数据分析等关键环节的智能化解决方案；
3. 覆盖多维应用场景：直击教育测评、临床评估、人才选拔等领域的实际痛点。

三大特色体验：

权威领航：剑桥大学、ETS 考试中心、瓦伦西亚理工大学等国际顶尖机构专家倾囊相授；

多元学习：理论精讲配合 R 语言/mlr3 等工具实操，穿插互动式小组案例研讨；

成果转化：提供经过实证检验的 AI+测量学应用框架与实施工具包；

通过系统化的理论解析、真实案例演示(含完整代码实现)和情境化研讨，我们将帮助研究者深化方法学创新、助力实践者提升评估效能，共同推进智能时代心理测量学的范式转型。

欢迎心理测量学研究者、教育评估从业者、AI 技术开发者及相关领域专业人士参与，共同探索测评科学的未来图景！

一、工作坊主题介绍及主讲专家

1. 人工智能评估：范式与技术

课程简介：

要实现本世纪对智能系统的期望，我们亟需融合自然智能与人工智能评估领域中的先进理念与方法。本次课程将从这一宏观视角出发，系统介绍当前人工智能评估的最新进展，涵盖多个主要的评估范式与研究传统，包括：TEVV（测试、评估、验证与确认）体系、基准测试流派、evals 社群、建构导向评估体系、现实影响导向评估路径，以及探索式评估方法等。我们将重点探讨建构导向评估如何更有效地借鉴心理测量学、认知心理学和行为科学的理论，深入理解“能力”本身的含义及其测量方式，而不仅仅是性能表现。同时，我们也将分析为何以人类或 AI 群体为基础的评估可能存在偏差，并介绍可能的替代方案。最后，我们将人工智能评估视

为对系统可预测性的研究：从“整体是否准确或安全”延伸到“在特定使用条件下是否可靠”。当我们将 AI 评估理解为对解释力与预测力的双重追求时，将为该领域带来新的研究挑战与机遇。

主讲人：



何塞·埃尔南德斯奥拉洛，西班牙瓦伦西亚理工大学教授

研究领域：人工智能、机器学习、数据科学和智能测量等人工智能的能力、普遍性、进步、影响和风险分析。

2. 人工智能时代的创造力评估：心理测量与计算方法

课程简介：

本工作坊将介绍创造力评估领域的前沿方法，结合心理测量学与人工智能（AI）技术，探索更加高效、科学的评估路径。工作坊开始将聚焦于评分设计中的核心挑战，介绍如何通过平衡不完全区组设计提升评分效率。随后，讲解如何运用多层面 Rasch 模型（Many-Facet Rasch Model, MFRM），对评分者的严苛程度、偏差及评分一致性进行建模与校正。工作坊后半部分将转向探讨自动化评分，重点展示如何运用 AI，特别是大型语言模型，对创造性作答进行评估。同时，还将演示如何结合 MFRM 模型，对比并分析人类评分与 AI 评分之间的一致性与差异。本工作坊将配套提供相关的 R 语言代码，帮助参与者动手实践所学方法，掌握具备大规模应用潜力、具有良好效度与信度的创造力评估工具。

主讲人：



里卡多·普瑞米博士，巴西坎皮纳斯市圣弗朗西斯科大学副教授

研究领域：心理测量学、数据科学、人格评估、创造力、智力、情商和人工智能在心理评估中的应用。

3. 计算机多阶段测验及其应用

课程简介：

本次工作坊将系统介绍计算机多阶段测验（Multistage Testing, MST）的基本原理、设计理念及其实施流程。课程内容包括 MST 设计的背景与必要性、其与传统线性测验和计算机自适应测验（CAT）的区别、运行机制、建模方法及仿真策略等。我们还将介绍 MST 中的替代评分与估计方法、分类测验设计、路径与评分机制、量表链接策略、测验安全保障措施等，并演示相关软件的操作流程。工作坊内容基于 Yan、von Davier 与 Lewis（2014）主编的专著，结构上涵盖从测验设计到实施及事后分析的完整流程，旨在帮助参与者全面了解 MST 在实际测评中的应用潜力与技术细节。

主讲人：



颜端丽博士，美国教育考试服务中心数据分析与计算研究主任

研究领域：心理测量、教育评估与人工智能驱动的测评技术（如自适应测试、自动评分、认知建模等）

4. 利用人工智能优化创造力评估

课程简介：

本工作坊将探讨如何运用不同形式的人工智能（AI）提升创造力评估的适用性和实用性。目前，在心理学领域中，具有良好效度与信度的创造力评估工具往往存在使用效率低、人工成本高、费用昂贵等问题。因此，许多研究者转而选择使用较不鲁棒的评估方式（例如自我报告量表）。这些问题也导致创造力评估工具难以为一线用户（如学校教师）所采用。然而，近期已有研究开始探索如何利用 AI 来解决这些难题。本工作坊将介绍多种通过 AI 提升创造力评估

适用性的路径，包括运用机器学习模型评估语言与图形创造力，借助无代码或低代码工具辅助开发机器学习模型，以及利用大型语言模型辅助进行创造力评估。本工作坊将着重探讨如何开发无需具备特殊编程技能即可使用的 AI 创造力评估工具，以提高其实用性与可推广性。

主讲人：



大卫·克罗普利博士，南澳大利亚大学教授

研究领域：变革力，创造力，工程学

5. 心理学中监督式机器学习

课程简介：

近年来，（监督式）机器学习在心理学及其他社会科学领域受到广泛关注。尽管机器学习在预测能力方面展现出巨大潜力，其实际应用仍面临诸多挑战，研究者需对模型进行合理评估，以避免对其性能产生过于乐观的估计。本工作坊将介绍监督式机器学习建模的基本概念、常见误区、核心算法，以及其在心理学评估等领域的实际应用。具体内容包括：多种重采样策略、超参数调优的基本方法、适用于表格数据的基于树的监督学习算法，以及分类与回归任务中常用的性能评估指标。此外，还将简要介绍 mlr3 生态系统——这是一个在 R 语言中广泛应用的机器学习分析工具箱。

主讲人：



大卫·格瑞兹克博士，乌得勒支大学助理教授

研究领域：心理测量学和潜在变量建模与计算统计学的整合，包括数据科学技术、优化和统计学习。

6. 项目反应理论：从核心概念到实际应用

课程简介：

本互动式工作坊将系统介绍项目反应理论（Item Response Theory, IRT）的核心概念，这是一种用于分析测评数据、衡量潜在心理特质（如能力水平）的现代统计方法。工作坊面向初次接触 IRT 的研究人员与实践者，内容涵盖项目反应与潜变量之间的关系，以及该理论所依据的关键假设，如单维性与局部独立性。参与者将学习几种常用的 IRT 模型，包括用于二分类项目的 Rasch 模型、两参数逻辑模型（2PL）与三参数逻辑模型（3PL），以及适用于多分类项目的相关模型。工作坊还将介绍 IRT 的实际应用，包括项目分析、测验信息函数和评分方法，特别强调对模型输出的解读及其在教育与社会科学中的实际运用。本工作坊无需相关背景知识，欢迎对该领域感兴趣的参与者积极参与。

主讲人：



孔瓦尼·穆霍戈博士，剑桥大学心理测量中心心理测量学家

研究领域：心理测量分析，国家参考测试，电子评估，问卷与研究工具评估。

7. 心理测量学中的生成式人工智能

课程简介：

本工作坊探讨生成式人工智能（GenAI）在心理测量学中的当下及未来角色。内容将围绕以下几个核心板块展开：

- 大语言模型的工作原理

简要探究生成式人工智能模型如何学习模式、进行微调并与人类反馈保持一致。

- 头脑风暴：人工智能可在哪些领域革新测试？

小组活动：两人一组，寻找在工作流程中应用人工智能的机会。

- 人工智能生成的测试题目

对比“传统”自动项目编写（模板/规则）与基于人工智能的方法。

小组活动：为一个新特质生成并点评人工智能制作的问题，随后展开讨论。

• 人工智能评分的 essays（注：此处“essays”结合语境或指开放式心理测试文本）
对比“传统”自动评分（简单评分标准/基础统计模型）与大语言模型驱动的工具。初步介绍我们利用大语言模型为心理特质生成评分的最新成果。

主讲人：



乔·沃森博士，剑桥大学心理测量中心心理测量学家

研究领域：教育与社会影响领域的量化研究与心理测量；运用数据挖掘、自然语言处理与数据可视化技术，开展法律分析、教育政策研究及社交媒体研究。

二、工作坊安排

报名学员可根据自己的需求和兴趣，以及具体的课程时间安排，从以下 7 个工作坊中选择 5 个工作坊进行参与和学习。（注：时间重叠的课程只能二选一）

1. 课程安排

时间	课程名称	学时	主讲人
7 月 10 日 星期四 18:00-21:00	人工智能评估：范式与技术	4	何塞·埃尔南德斯奥拉洛
7 月 12 日 星期六 18:00-21:00	人工智能时代的创造力评估：心理测量与计算方法	4	里卡多·普瑞米
7 月 13 日 星期日 9:00-12:00	计算机多阶段测验及其应用 或 项目反应理论：从核心概念到实际应用 (二选一)	4	颜端丽 孔瓦尼·穆霍戈
7 月 13 日 星期日 14:00-17:00	利用人工智能优化创造力评估 或 心理测量学中的生成式人工智能 (二选一)	4	大卫·克罗普利 乔·沃森

7月13日 星期日 18:00-21:00	心理学中监督式机器学习	4	大卫·格瑞兹克
-----------------------------	-------------	---	---------

2. 课程语言：英语，不提供实时人工翻译。

3. 课程地点：北京师范大学心理学部 EDP 中心培训教室：北京市西城区新街口外大街 8 号金丰和创业园 A 座（北师大东门对面）。具体教室另行通知。

三、结业证书

课程共计 20 学时。修满学时、考核合格的学员将获得“人工智能时代的心理测量革新工作坊”培训结业证书，可在北京师范大学培训管理信息平台

(<http://jpc.bnu.edu.cn/cms/zscx/index.htm>) 进行证书查验。



四、招生人数

60 人

五、培训费

1. 社会人士：2980 元/人。

2. 三人及以上团报、北师大校友、参加过北师大心理学部往期培训学员：1980 元/人。

备注：培训费用包括：培训费、资料费等。

培训食宿、交通费用自理。

开课后不再受理退费事宜。

六、报名方式、时间

1. 预报名方式:

扫描二维码，进入快速预报名通道



2. 报名截止时间: 2025 年 7 月 5 日 24: 00

七、缴费及发票

1. 登录缴费平台: <http://wszf.bnu.edu.cn/publish/>，登录后查找【人工智能时代的心理测量革新工作坊】进行缴费。（需开发票的学员推荐使用此平台进行缴费）

2. 银行对公转账账号【可以通过手机银行、支付宝、微信等转入如下账户】:

开 户 行: 中国银行北京文慧园支行

账 户 名: 北京师范大学

账 号: 340256015272

银行名称: 中国银行

纳税人识别号: 12100000400010056C

地 址: 北京市海淀区新街口外大街 19 号

备注: 请在汇款单上备注“心理测量革新工作坊+姓名+电话”。

缴费截止时间: 2025 年 7 月 5 日 24: 00

3. 发票:

发票内容: 函大电大夜大及短期培训费

发票类型: 中央财政中央非税收入统一票据

发票形式: 电子发票

发放方式: 发送至报名者提供的邮箱

八、联系我们

1. 咨询地点: 北京师范大学心理学部 EDP 中心

(北师大东门对面, 北京市海淀区新街口外大街 8 号金丰和创业园 A 座 6 层 606 室)

2. 报名邮箱: edpbnu@bnu.edu.cn

3. 联系人: 原老师、杨老师、赵老师、乔老师、谭老师、张老师

咨询电话: 010-62204669; 010-62204009; 010-82240811; 13260061101

4. 咨询微信: 13260061101 (北师大心理学部 EDP 中心)

