

T/ISC

中国互联网络协会团体标准

T/ISC XXXX—XXXX

基础教育体育学业质量 人工智能测评技术 技术规范

Basic education physical education academic quality—Artificial intelligence
evaluation technical specification

（征求意见稿）

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

XXXX – XX – XX 发布

XXXX – XX – XX 实施

中国互联网协会 发布

目 次

前言 II

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 总体原则 1

5 测评体系 2

6 测评内容 3

7 测评方法 4

8 测评流程 5

9 测评报告 6

10 智能测评场景 6

参考文献 8

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国互联网协会提出并归口。

本文件起草单位：

本文件主要起草人：

基础教育体育学业质量 人工智能测评技术规范

1 范围

本文件规定了基础教育（小学、初中、高中）体育学业质量人工智能测评的总体原则、测评体系、测评流程及智能测评场景的要求。

本文件适用于基础教育体育学业质量的人工智能测评指导、使用与管理。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 35273 信息安全技术 个人信息安全规范

GB/T 37229 公共服务效果测评模型和方法指南

GB/T 41867 信息技术 人工智能 术语

3 术语和定义

GB/T 41867界定的以及下列术语和定义适用于本文件。

3.1

基础教育体育学业质量 basic education physical education academic quality

学生在完成基础教育阶段体育与健康课程后的学业成就表现，反映体育核心素养的达成度。

注：基础教育体育学业质量包括知识、能力、行为、健康四个维度。

3.2

人工智能测评 artificial intelligence evaluation

运用计算机视觉、机器学习、可穿戴设备传感等技术，对学生学业质量各维度进行自动化采集、分析与评价的过程。

3.3

动作识别准确率 action recognition accuracy

人工智能系统对人体关键点检测、动作分类、姿态评估的准确程度。

3.4

人机交互测试 human-computer interaction test

通过智能终端向学生推送测评题目、量表或情境任务，由学生通过触摸、语音等方式交互完成，系统自动评分并分析结果的技术手段。

注：人机交互测试主要适用于知识维度及健康维度中认知、心理相关指标的测评。

3.5

过程性测评 process evaluation

在日常教学活动过程中，通过伴随式数据采集对学生学习过程进行持续观察、记录与阶段性反馈的评价方式。

3.6

伴随式数据采集 accompanied data collection

在日常教学、活动过程中，通过非接触式或低介入式设备（如人工智能摄像头、可穿戴设备等）自动持续采集学生运动表现、行为习惯等数据的技术方式。

4 总体原则

4.1 科学性

测评应依据体育学科规律、课程标准和学生身心发展特点，科学设计评价指标体系和测评方法。

4.2 准确性

人工智能测评结果应客观反映体育学业质量真实水平，明确各项测评指标的误差范围和置信度要求，满足硬件与算法的精度指标，匹配各维度测评方法的精度要求。

4.3 过程性

建立过程性与结果性相结合的综合评价机制，覆盖日常教学到终结性考核的全流程，支持体育课、大课间、课外活动、家庭作业等全场景，通过伴随式数据采集对学习过程进行持续评价。

4.4 公平性

人工智能测评系统应对各年龄段学生采用统一的评价标准，测评过程可追溯、可复核，确保测评结果的公平公正。

4.5 安全性

人工智能测评系统和设备应符合GB/T 35273要求。明确人脸、姿态等生物特征数据的采集授权、存储加密、去标识化处理要求，且在系统中设置监护人知情确认界面，确保测评过程中的人身安全和数据隐私安全。

5 测评体系

5.1 概述

人工智能测评系统应由感知层、数据层、分析层、应用层构成，具备数据采集、运动识别、评价分析、数据管理等核心功能，满足基础教育体育学业质量测评需求，覆盖知识、能力、行为、健康四大维度所有指标，适配人机交互测试、AI视觉识别、可穿戴设备、智能量表等多样化测评技术手段。测评系统框架如图1所示。

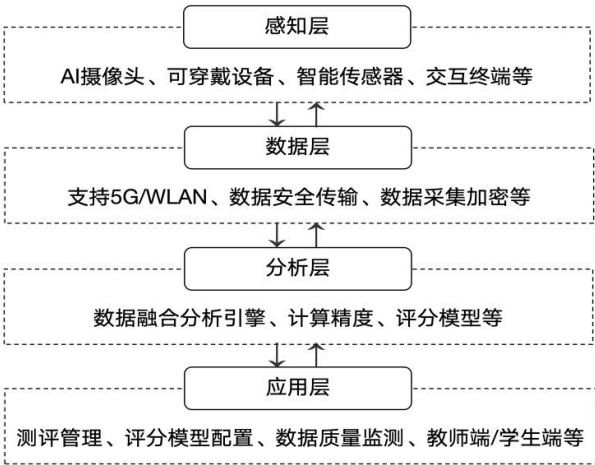


图 1 AI 测评系统的框架

5.2 感知层

感知层包括视频采集设备、传感器、智能穿戴设备、交互终端、智能体质测量设备、运动App/小程序等，用于采集学生运动过程中的图像、视频、姿态、心率、答题结果、量表反馈等原始数据，关键技术指标包括但不限于：

- a) AI 摄像头：分辨率 ≥ 200 万像素、帧率 $\geq 30\text{fps}$ 、支持 H. 265/H. 264 编码、防护等级不低于 IP65；
- b) 边缘计算终端：支持实时视频流处理、动作识别推理延迟 $\leq 100\text{ms}$ ；
- c) 智能可穿戴设备：心率监测精度 $\pm 3\text{bpm}$ 、运动轨迹追踪、数据同步延迟 $\leq 5\text{s}$ ；

- d) 交互终端：触摸屏 ≥ 21 寸、支持语音交互、环境适应性(防尘防水 $\geq \text{IP56}$)，支持人机交互测试题目、智能量表推送；
- e) 视频采集帧率不低于 30fps，分辨率不低于 1920×1080 ，单镜头同时识别人数不少于 6 人；
- f) 智能体质测量设备：支持人脸识别自动检录，身高体重、肺活量等测量误差符合国家标准；
- g) 运动 App/小程序：支持学生查看学生个人测试报告及视频回放，支持生成学生画像，数据同步延迟 $\leq 5\text{s}$ 。

5.3 数据层

数据层负责原始数据的存储、清洗、标注、管理与传输，覆盖四大维度全量测评数据。关键技术要求应包括但不限于：

- a) 传输层支持 5G/WLAN/物联网协议、数据安全传输；
- b) 学生测评数据存储周期不少于学生在校完整学段+2 年，涉及个人生物特征信息的数据应在学生毕业后予以删除或脱敏处理；关键测评数据在学生在校期间保存，心理健康等敏感数据应采取脱敏、加密存储，并严格限制访问权限；
- c) 支持与区域教育数据中心、国家学生体质健康数据平台对接。

5.4 分析层

分析层包括运动姿态识别引擎、数据融合分析引擎、评价模型等，实现四大维度数据的智能分析，算法与模型要求如下：

- a) 人体关键点检测支持 ≥ 17 个关键点、检测帧率 $\geq 25\text{fps}$ ；
- b) 动作识别准确率：基础动作(跑、跳、投) $\geq 95\%$ ，复杂动作(体操技巧) $\geq 90\%$ ；
- c) 计时精度：径赛项目误差 $\leq 0.01\text{s}$ ；距离/计数精度：跳远误差 $\leq 1\text{cm}$ ，跳绳/仰卧起坐计数误差 ≤ 1 次/分钟；
- d) 抗干扰能力：支持不同光照条件($50\text{lux} \sim 10000\text{lux}$)、多目标跟踪、部分遮挡处理；
- e) 智能量表评分模型：涉及心理健康指标的量表应具备信效度验证报告，系统对心理风险信号的识别宜提供置信度标识，预警结果须经教师复核确认，不宜仅凭 AI 结果直接判定学生心理状态。

5.5 应用层

应用层包括测评管理平台、教师端、学生端、家长端，功能要求如下：

- a) 测评流程管理：支持学生身份识别(人脸/校园卡)、自动检录、成绩实时显示，覆盖知识、能力、行为、健康全维度测评流程；
- b) 评分模型配置：支持按年级、性别、项目设置差异化评分标准，匹配四大维度各指标的评分规则；
- c) 数据质量监控：设备状态监测、异常数据预警、测评过程可追溯；低置信或异常结果转人工复核，不自动计入正式成绩；
- d) 结果输出：自动生成个人报告、班级/年级汇总，支持与国家体质健康网数据对接或输出符合国家体质健康网标准格式的表格；
- e) 支持生成个体学生运动能力画像与群体学情分析报告，支持学生体质健康数据的历次对比与趋势分析，支持针对个体学生推送个性化运动及健康改进建议。

6 测评内容

基础教育体育学业质量人工智能测评内容依据《〈体育与健康〉教学改革指导纲要（试行）》中知识、能力、行为、健康综合评价要求，AI测评需覆盖全部维度实现一体化评价。各维度的测评内容包括但不限于表1。

表 1 测评内容

一级指标	二级指标	三级指标	AI 测评方法
知识	健康知识	健康生活知识	人机交互测评：智能终端推送情境式选择题/判断题，AI 自动评分，分析薄弱点
		生长发育知识	
		安全避险知识	
	体育知识	体育基础知识	
		运动实践知识	
		体育文化知识	
能力	基本运动能力	走、跑、跳、投、攀爬等基本动作	AI 视觉识别+内容结构化评分：摄像头采集测评内容的动作表现，按观测点自动判定完成度
	专项运动能力	球类/体操/武术/游泳等项目	AI 视觉识别+内容结构化评分：对运球、上篮、套路等分解动作逐项评分
行为 ^b	体育锻炼行为	体育课参与度	伴随式采集：课堂 AI 摄像头分析学生个体运动姿态与群体运动数据、注意力集中时长
		大课间/课外活动参与	可穿戴设备传感：记录活动频率、时长、心率强度
		体育家庭作业完成情况	运动 APP/小程序：学生上传视频，AI 辅助确认完成情况
	体育品德行为	规则遵守	AI 辅助行为事件记录+教师综合评定：AI 识别比赛中是否越界、犯规等，教师结合具体情境、规则要求及日常观察进行综合判断
		合作行为	AI 辅助行为事件记录+教师综合评定：AI 识别传球、配合等合作动作频率等，教师结合学生在不同活动中的角色承担、同伴互动、团队协作及日常表现进行综合判断
		坚持与拼搏行为	AI 辅助行为事件记录+教师综合评定：AI 识别学生持续参与、重复练习、应对困难等过程性行为等，教师结合学生个体基础、健康状况、任务难度、进步表现和日常观察进行综合判断
健康 ^c	身体健康	身体形态	智能设备采集：身高体重仪(人脸识别自动检录)/计算BMI
		身体机能	智能设备：肺活量仪、心率监测设备(可穿戴)评估心肺功能
		身体素质	AI 视觉识别+智能设备：包括但不限于《国家学生体质健康标准》身体素质项目，如立定跳远(视觉测距)、仰卧起坐(视觉计数)、坐位体前屈(视觉测距)等
	心理健康	情绪状态	智能量表 ^a ：通过平板推送简短情绪问卷，AI 自动评分并预警
		自我认知	智能量表 ^a ：简短自评问卷(如“我对自己满意吗”等)，AI 辅助分析
		学校适应	人机交互测评+结构化观察：智能终端推送适应量表，系统自动计算课程兴趣、任务适应及规则适应等维度得分，并结合体育课出勤率、任务完成率和课堂参与度分析适应水平
		同伴关系	人机交互测评+结构化观察：智能终端推送同伴关系量表，系统自动计算同伴支持、合作接纳等维度得分，并汇总教师对体育课堂合作互动和冲突事件的结构化观察记录
		注： a 智能量表为经公开发表且具有信效度验证报告的简短量表； b “行为”评价结果由教师结合多次观察、学生说明等多源证据形成； c “健康”测评结果仅作参考，不作为医学临床诊断依据。	

7 测评方法

7.1 概述

基础教育体育学业质量人工智能测评方法应综合运用AI视觉识别、可穿戴设备传感、量表测评、行为事件记录等多种技术手段，针对不同指标特点选择适配的测评方式。

7.2 测评方法分类

7.2.1 AI 视觉识别

AI视觉识别法是指通过部署在运动场景中的摄像头采集学生运动视频，利用人体关键点检测、动作识别、多目标跟踪等计算机视觉技术，对学生的运动能力表现进行自动化分析与评分的方法。

7.2.2 智能设备测量

智能设备测量法是指通过可穿戴设备、智能体质测试设备及其他传感装置，采集学生身体形态、身体机能、身体素质、生理反应和运动过程数据，并依据规定的测量程序和计算规则形成测评结果的方法。

7.2.3 人机交互测评

人机交互测评法是指通过智能终端或移动应用程序向学生呈现结构化试题、情境任务、问卷或量表，由学生采用触摸、语音等方式完成作答，系统记录作答数据，并依据预先确定的评分规则形成测评结果的方法。

7.2.4 结构化观察评定

结构化观察评定法是指依据预先设定的观察指标和评价规则，在体育活动情境中，对学生可观察的行为表现进行系统观察、记录和评价的方法。

8 测评流程

8.1 概述

基础教育体育学业质量人工智能测评流程应适配知识、能力、行为、健康四大维度测评要求，并按下列步骤开展：

- a) 明确测评目的；
- b) 确定测评内容；
- c) 设计测评工具；
- d) 处理测评数据；
- e) 分析测评结果；
- f) 完成测评报告。

8.2 明确测评目的

开展测评前应明确测评目的，包括但不限于：

- a) 学校日常体育教学评价；
- b) 学生体育学业质量测评；
- c) 区域体育教育质量监测。

8.3 设计测评工具

测评工具适配各维度测评方法，包括但不限于：

- a) 知识维度：采用人机交互测试工具，通过智能终端推送情境式题目，实现自动评分与薄弱点分析；
- b) 能力维度：采用 AI 视觉识别工具，通过摄像头采集动作表现，按观测点自动判定完成度；
- c) 行为维度：采用学生自评量表、伴随式数据捕捉设备（课堂 AI 摄像头）、可穿戴设备、校园一卡通、运动 APP/小程序、教师观察记录表，实现全场景行为数据采集；
- d) 健康维度：采用智能体质测量设备、可穿戴监测设备、智能心理测评工具（平板推送量表），实现体质与心理健康数据采集。测评形式包括现场表现性测试、人机交互测试及伴随式数据采集三类。

8.4 处理测评数据

依据测评指标和测评工具，选择适配的数据采集渠道：知识维度采用线上答题，能力维度采用动作捕捉，行为维度采用轨迹记录，健康维度采用体质测试和心理测评。数据采集和处理技术应符合 GB/T 37229 的要求。

8.5 分析测评结果

8.5.1 测评等级与分数区间

依据测评数据，从知识、能力、行为、健康四个维度进行综合分析。测评等级与分数区间见表2。

表 2 测评等级与分数区间

测评等级	分数区间
优秀（A）	90.0 分~100 分
良好（B）	80.0 分~89.9 分
合格（C）	60.0 分~79.9 分
待提高（D）	59.9 分及以下

8.5.2 结果分析

- 测评结果分析应包括但不限于：
- a) 知识维度分析：自动定位学生健康生活、生长发育、安全避险、体育基础、运动实践、体育文化等知识的薄弱点，推送针对性学习内容，生成知识掌握情况评估报告；
 - b) 能力维度分析：自动生成学生基本运动能力、专项运动能力等级，识别优势项目与薄弱环节，输出能力提升建议；
 - c) 行为维度分析：综合多源数据判断锻炼行为各维度水平，对体育精神、体育品格、体育道德等品德行为以等级描述为主，生成行为习惯改进建议；
 - d) 健康维度分析：综合智能设备采集、体质测试、心理测评等多源数据，判断体质健康与心理健康各维度水平，自动生成学生体质健康档案与心理健康评估报告。要求心理健康数据使用独立数据域或等效隔离措施、单独密钥和最小权限；高风险或主管部门明确要求时采用物理隔离。未经授权不得访问，并保留访问审计。
 - e) 人工复核：AI 系统生成的测评等级、分数和改进建议，须经由教师审核确认后方可作为正式学业质量评价结论，且教师应在复核记录中签名确认。

9 测评报告

基础教育体育学业质量人工智能测评报告是基于测评数据分析以差异化报告形式提供，具体包括但不限于：

- a) 学生报告：包含运动能力等级评定、健康行为水平判断、体育品德表现描述、综合评价等级、个性化运动及健康改进建议；
- b) 教师报告：包含班级及年级测评结果分析、群体优势与薄弱环节诊断、分层教学建议；
- c) 学校报告：学校各年级水平分布、男女生差异分析、年度纵向对比、优势与薄弱维度诊断、学校体育工作改进建议。
- d) 教育管理部门报告：包含区域总体状况、城乡及校际差异分析、突出问题预警、政策改进建议。
- e) 家长报告：包含学生个人运动能力等级、健康行为表现、体育品德描述，以及适合家庭场景的运动建议。

10 智能测评场景

10.1 概述

智能测评场景根据测评目的和使用环境选择确定，对应三类测评层级：课堂过程性测评、学期终结性测评、国家/区域质量监测，适配知识、能力、行为、健康四大维度全场景测评。

采用人机交互测试、AI视觉识别、智能量表等技术手段实现自动化测评与风险预警。

10.2 课堂过程性测评

在日常体育课堂教学过程中，运用人工智能技术对学生运动表现、课堂参与、技能掌握情况进行即时采集与反馈。系统应支持教师实时查看学生运动负荷、动作规范性、课堂参与度等数据，辅助教师及时调整教学策略。

10.3 学业质量测评

在学期末或单元教学结束后，运用人工智能技术对学生体育学业质量进行终结性评价。系统依据课程标准要求，对学生的知识、能力、行为、健康进行全面测评，生成学生学业质量等级评定报告。

测评结果作为学生体育课程学业成绩的参考依据之一，适用场景为期末体育考试，重点覆盖能力维度的专项运动能力测评、健康维度的身体素质测评、知识维度的期末线上答题、行为维度的学期体育活动参与情况汇总等。

10.4 区域质量监测

在区域或学校层面，运用人工智能技术定期开展体育教育质量监测。系统具备跨校、跨区域数据汇聚与分析能力，支持多维度交叉分析与趋势预测，生成区域学生体育学业质量数字图，测评结果用于诊断区域体育教育质量问题、评估政策实施效果、指导体育教育改革，适用场景为国家义务教育质量监测、区域学业质量抽测，重点覆盖四大维度的区域整体数据对比、城乡校际差异分析、学生体质健康与心理健康风险预警。

参 考 文 献

- [1] 教育部办公厅关于印发《〈体育与健康〉教学改革指导纲要（试行）》的通知 教体艺厅函〔2021〕28号
- [2] 中共中央 国务院印发《深化新时代教育评价改革总体方案》
- [3] 教育部等五部门关于印发《“人工智能+教育”行动计划》的通知 教科信〔2026〕1号
- [4] 中华人民共和国教育部 义务教育体育与健康课程标准（2022年版）
- [5] 中华人民共和国教育部 普通高中体育与健康课程标准（2017年版2020年修订）
-