

基于多源数据的高校学生学业危机预警指标体系构建研究

杜平 严琳*

广东建设职业技术学院, 中国·广东 广州 510440

摘要: 本研究通过融合学生行为、成绩等多源数据, 基于扎根理论对辅导员访谈资料进行质性分析, 深度萃取学业危机的关键致因维度; 同时, 系统整合各科总成绩、课程排名、不及格门数及获得学分等构成的学习成绩大数据, 作为客观量化依据。在此基础上, 将质性分析形成的核心维度与学习成绩数据指标进行深度融合, 构建了一个涵盖多维度影响因素的综合测度体系, 并据此建立了五级加权学业危机测度模型。实证结果表明, 该模型具有良好的判别效度与应用可行性, 为高校学业危机的精准识别与早期干预提供了可靠的实践工具。

关键词: 学业危机; 预警指标体系; 多源数据融合; 层次分析法; 五级预警模型

Research on the Construction of a Warning Index System for College Students' Academic Crises Based on Multi-source Data

Du Ping, Yan Lin*

Guangdong Polytechnic of Construction, China Guangdong Guangzhou 510440

Abstract: This study integrates multi-source data such as student behavior and academic performance, conducts qualitative analysis of the interview data of counselors based on grounded theory, and deeply extracts the key causal dimensions of academic crises. At the same time, it systematically integrates the big data of academic performance, including total scores of all subjects, course rankings, the number of failed courses, and credits obtained, as objective quantitative basis. On this basis, the core dimensions formed by qualitative analysis are deeply integrated with the academic performance data indicators to construct a comprehensive measurement system covering multiple influencing factors, and a five-level weighted academic crisis measurement model is established accordingly. Empirical results show that the model has good discriminant validity and application feasibility, providing a reliable practical tool for the precise identification and early intervention of academic crises in colleges and universities.

Keywords: Academic crisis; Warning index system; Multi-source data fusion; Analytic hierarchy process; Five-level warning model

0 引言

随着我国高等教育学生规模不断扩大, 高校学业危机现象日益突出。学业危机不仅影响学生个人发展, 也关系到高校人才培养质量与教育资源配置效率。目前多数高校仍主要采取“事后干预”模式, 即在期末成绩公布后对已收到学业警示的学生进行帮扶, 这种方式滞后性强, 学生往往已陷入较深学习困境, 干预难度和成本显著提高。因此, 构建一套科学、前瞻的学业危机识别与预警指标体系, 实现早期发现与精准干预, 已成为高校教育管理的重要课题。

有效的干预始于精准识别, 而精准识别依赖于系统化测度。现有学业危机测度研究主要存在两种路径: 一是基

于管理经验的定性判断, 依托辅导员、教师等一线工作者的观察经验, 归纳学业危机学生的行为特征及成因。其优势在于能深入把握个体情境, 洞察行为背后的复杂动因, 具备较强解释力与人文关怀; 缺点在于主观性强, 缺乏统一量化标准, 难以规模化应用与验证。二是基于数据驱动的定量分析, 利用学生学业表现的结构化数据, 如总成绩、课程排名、不及格门数、已获学分等, 借助统计方法识别学业风险学生。该路径结果客观、便于量化处理, 利于横向比较与趋势判断; 但局限在于多停留于学业表现的表层描述, 难以深入揭示学生实际学习状态、心理变化、行为逻辑及危机形成的深层机制, 难以为精准干预提供充分依据^[1]。

可见，单纯依赖主观经验判断或片面倚重量化数据，均难以实现对学业危机的全面准确测度。为弥补上述不足，本文提出融合“课程考核大数据”与“行为习惯质性大数据”的研究思路。“课程考核大数据”包括各科目总成绩、排名、不及格门数、学分等结构化学业数据；“行为习惯质性大数据”则来源于辅导员通过深度互动、访谈或问卷获取的信息，涵盖课堂行为与学习态度、学业困难应对方式、课外活动参与、个人习惯、心理状态与压力水平、家庭背景及其影响等多个维度。通过将客观学业数据与反映学生实际状态的质性信息相结合，以前者量化学业表现，后者诠释行为逻辑与情境动因，实现数据层面“量”与“质”的协同，为全面精准测度学业危机提供新路径。

1 测度指标体系的构建

为构建一套多维度、可量化、可操作的高校学生学业危机测度指标体系，以支持学业危机的早期识别与精准干预，本研究整合了成绩大数据与行为问卷调查数据，通过“质性维度提取—量化指标构建—体系融合确立”的思路，系统完成了测度指标体系的构建工作^[2]。

1.1 基于质性访谈的关键维度提取

通过对辅导员访谈文本的系统梳理，结合《学业情况调查》问卷设计中所涵盖的观测要点，本研究归纳出学业

危机相关的四个关键维度：

学习态度与行为：包括课堂注意力、课后学习时间、作业完成方式、学习求助倾向等；

心理状态与应对方式：涵盖学业压力反应、压力来源认知、情绪状态及心理调适策略等；

个人习惯与时间管理：涉及作息规律、体育锻炼、兼职情况、网络使用习惯及时间分配能力等；

环境支持与影响因素：包括家庭经济与期望压力、专业兴趣与目标明确性、校园社交与支持系统等。

1.2 学业表现量化指标的构建

为弥补质性数据在客观性与规模性上的局限，本研究同步利用某高校 2023 - 2024 学年约 3000 名学生的教务数据进行量化指标构建。原始数据经过缺失值处理、格式统一等清洗步骤后，依据学业表现的关键维度转化为以下指标。

平均学分绩点（GPA）：反映学生整体学业水平；

专业排名相对位置：消除专业间评分差异，增强可比性；

不及格门数与累计不及格学分数：作为学业异常的直接预警指标；

学分完成进度比：通过对比已获学分与应获学分，动

表1 学业危机测度指标体系框架

一级指标	二级指标	三级指标	指标性质	数据来源
A. 学业表现	A1 学业水平	A1-1 平均学分绩点（GPA）	定量	成绩大数据
	A2 学业进展	A2-1 学分完成进度比	定量	成绩大数据
	A3 学业异常	A3-1 不及格门数	定量	成绩大数据
		A3-2 累计不及格学分数	定量	成绩大数据
B. 学习行为与投入	B1 课堂表现	B1-1 出勤率	定量/定性	课堂考勤大数据
		B1-2 课堂参与度	定性	任课教师评价
	B2 课后投入	B2-1 作业完成质量与及时性	定性	任课教师评价
		B2-2 自主学习时间（小时/周）	定量/定性	学生问卷/辅导员访谈
C. 个人心理与习性	C1 心理状态	C1-1 学习压力感知程度	定性	心理问卷/辅导员访谈
		C1-2 考试焦虑水平	定性	心理问卷/辅导员访谈
	C2 行为习惯	C2-1 时间管理与规划能力	定性	学生问卷/辅导员访谈
		C2-2 作息规律性	定性	学生问卷/辅导员访谈
D. 环境支持与影响	D1 家庭环境	D1-1 家庭经济状况	定量/定性	资助系统数据/辅导员访谈
		D1-2 家庭关系与期望压力	定性	辅导员访谈/学生问卷
	D2 校园社交	D2-1 宿舍人际关系融洽度	定性	辅导员观测/学生问卷
		D2-2 同伴学业互助情况	定性	辅导员访谈/学生问卷

表2 学业危机测度指标体系及权重分配

一级指标	二级指标	三级指标
A. 学业表现 (0.45)	A1 学业水平 (0.60)	A1-1 平均学分绩点 (GPA) (1.00)
	A2 学业进展 (0.15)	A2-1 学分完成进度比 (1.00)
	A3 学业异常 (0.25)	A3-1 不及格门数 (0.60)
B. 学习行为与投入 (0.25)	B1 课堂表现 (0.60)	A3-2 累计不及格学分数 (0.40)
		B1-1 出勤率 (0.60)
	B2 课后投入 (0.40)	B1-2 课堂参与度 (0.40)
		B2-1 作业完成质量与及时性 (0.50)
C. 个人心理与习性 (0.20)	C1 心理状态 (0.60)	B2-2 自主学习时间 (0.50)
		C1-1 学习压力感知程度 (0.50)
	C2 行为习惯 (0.40)	C1-2 考试焦虑水平 (0.50)
		C2-1 时间管理与规划能力 (0.60)
D. 环境支持与影响 (0.10)	D1 家庭环境 (0.70)	C2-2 作息规律性 (0.40)
		D1-1 家庭经济状况 (0.40)
	D2 校园社交 (0.30)	D1-2 家庭关系与期望压力 (0.60)
		D2-1 宿舍人际关系融洽度 (0.60)
		D2-2 同伴学业互助情况 (0.40)

态监测学业进展。

1.3 质性与量化指标的融合与体系确立

在明确质性维度与量化指标的基础上，本研究遵循“维度 - 观测点 - 量化标定”的逻辑，将二者进行系统融合，形成如表 1 所示的多层测度指标体系。该体系共包含 4 个一级指标、9 个二级指标与 16 个三级指标^[3]。

结构互补：一、二级指标源于质性分析，阐释危机成因；三级指标明确测量内容与方法，实现机理洞察与客观测量的统一。

数据互补：学业表现以客观数据为主，确保评价公正；其余维度融入问卷、访谈等质性数据，捕捉成绩之外的致因信息。

全景画像：体系兼顾学业成果、行为模式、心理状态与环境因素，为精准干预提供立体依据。

通过以上步骤，我们成功构建了一个包含 4 个一级指标、9 个二级指标和 16 个三级指标的多维、多层学业危机测度指标体系，为后续构建五级加权测度模型奠定了坚实的理论与数据基础。

2 学业危机五级加权测度模型的建立

在构建综合测度指标体系的基础上，为将其有效应用于学业危机的识别与预警，需将多维指标整合为可量化比较的综合评价价值，并据此对学生学业危机程度进行科学分级。本研究通过层次分析法确定各指标权重，基于百分位数法划定危机等级阈值，最终构建了一个兼具科学性与操作性的五级加权测度模型^[4,5]。

2.1 基于层次分析法的指标权重确定

为客观反映不同指标在学业危机评估中的相对重要性，本研究采用层次分析法进行权重确定。该方法可将专

家经验判断量化，并通过一致性检验确保逻辑合理性，适用于本研究多维度、多层次的综合评价问题。

（1）以“学业危机综合测度”为目标层，以 4 个一级指标（学业表现、学习行为与投入、个人心理与习性、环境支持与影响）为准则层，其下的 9 个二级指标和 16 个三级指标构成方案层，层次结构与表 1 一致。

（2）邀请来自高校学生处、教务处及具有丰富一线工作经验的辅导员组成专家组，采用 1 - 9 标度法对同层次指标进行两两比较。例如，专家普遍认为“学业表现”因客观性强、直接影响显著，其重要性高于“环境支持与影响”；在“学业表现”内部，“学业水平”与“学业异常”的权重又高于“学业进展”。

（3）通过计算与检验，最终确定了各层级指标的权重。其中，一级指标的权重分别为：学业表现（0.45）、学习行为与投入（0.25）、个人心理与习性（0.20）、环境支持与影响（0.10）。该权重分配既凸显了学业成果的核心地位，也肯定了学习过程中心理与行为因素的重要预警作用。完整的指标体系权重分配如表 2 所示。

2.2 危机等级划分与综合测度模型构建

确定权重后，需将学生多维信息汇总为综合测度分，并据此划分危机等级，以支持分级干预。

2.2.1 指标标准化与综合测度分计算

由于各三级指标量纲与指向不同，首先进行标准化处理，将其统一转换至[0, 100] 区间。定量指标（如 GPA、不及格门数）采用最小 - 最大归一化；定性指标（如课堂参与度、心理压力）则依据评估等级量化（如“好、中、差”对应 100、60、20 分）。

在此基础上，采用线性加权法计算每位学生的学业危

表3 学业危机五级预警等级与干预措施

危机等级	等级名称	分数区间（S）	学业状态描述	干预措施
I 级	正常（安全区）	$S \geq 80$	学业状态良好，无显著风险。	保持常态关注。
II 级	轻度危机（关注区）	$70 \leq S < 80$	出现潜在风险点，如：偶尔缺勤、单科成绩下滑。	需辅导员关注。
III 级	中度危机（警示区）	$55 \leq S < 70$	出现多门不及格、学习投入明显不足等情况。	需启动学业警示并制定帮扶计划。
IV 级	重度危机（高危区）	$40 \leq S < 55$	学业濒临崩溃，可能伴有严重心理或家庭问题。	需学院介入并启动家校联动与心理辅导。
V 级	极端危机（极端区）	$S < 40$	面临退学风险。	需学校成立专项工作组进行危机处置。

机综合测度分（S），公式如下：

$$S = \sum_{i=1}^n (W_i \cdot \sum_{j=1}^m (W_{ij} \cdot \sum_{k=1}^p (W_{ijk} \cdot X_{ijk})))$$

其中 W_i ， W_{ij} ， W_{ijk} ，分别表示第一个一级指标、其下第个二级指标及第个三级指标的权重； W_{ijk} 为对应三级指标标准化后的值。S 值越高，表明学业危机风险越低。

2.2.2 五级危机等级阈值划定

为直观标识风险程度并匹配干预策略，依据综合测度分 S 的分布并结合高校学业管理实际，将学业危机划分为以下五个等级，如表 3 所示。

该五级加权测度模型通过指标集成与等级划分，将多源异构数据转化为直观、可操作的危机信号，实现了从“数据”到“洞察”的跨越，为高校开展精准、高效的学业预警与干预提供了核心决策支持^[6]。

3 结语

为应对高校学业危机预警滞后、干预被动的问题，本研究提出了一条融合课程考核数据与行为习惯质性数据的综合路径。通过质性分析提取关键致因维度，结合教务数据构建量化指标，建立了涵盖学业表现、学习行为与投入、个人心理与习性、环境支持与影响四个维度的测度指标体系。在此基础上，运用层次分析法确定指标权重，通过线性加权与阈值划分，构建了五级学业危机加权测度模型。

本研究的主要贡献在于：在理论层面，整合了质性洞察与量化描述，构建了多维度、多层级的学业危机测度框架；在方法层面，形成了从指标构建、权重赋权到等级划分的完整方法链；在实践层面，提供了可实现早期预警、

分级干预的量化工具，有助于推动学业支持从“事后处理”向“精准预防”转变。

参考文献：

[1] 徐小婷. 大数据背景下高职院校学生学业预警研究[J]. 湖北开放职业学院学报, 2024, 37(5): 164-166.

[2] 邹挺. 基于随机森林算法的学生学业预警分析[D]. 南昌: 南昌大学, 2024.

[3] 吴修国, 孙涛. 基于教育大数据挖掘的大学生学业预警研究[J]. 中国教育信息化, 2020(7): 55-57, 62.

[4] 何锴. 基于学习者画像的学业预警模型研究[D]. 武汉: 武汉工程大学, 2024.

[5] 王春洁. 基于机器学习的学习者画像构建与教学预警研究[J]. 电脑知识与技术, 2024, 20(22): 107-109, 113.

[6] 熊德兰. 基于大数据学生画像平台的高校学生学业预警与帮扶研究——以许昌学院为例[J]. 许昌学院学报, 2023, 42(5): 104-107.

基金项目：2023 年广东省高等职业教育教学质量与教学改革工程项目（课题编号：2023JG067）研究成果；2024 年广东省教育科学规划课题（高等教育专项）（课题编号：2024GXJK1121）研究成果。

作者简介：杜平（1982-），女，汉族，工程师、信息系统项目管理师，工学硕士，研究方向：大数据技术、人工智能技术、高职教育。

通讯作者：严琳（1988-），女，汉族，江西赣州，硕士研究生，讲师，研究方向：主要从事智能电子产品研发和职业教育的研究工作。