

# 装备制造职业素养的情境化评价

## ——SC-T-P模型的构建与应用

刘江艳<sup>1,2</sup>

<sup>1</sup>浙江工业大学教育学院, 浙江 杭州

<sup>2</sup>湘潭科技职业学院(湘潭技师学院)汽车工程系, 湖南 湘潭

收稿日期: 2025年12月17日; 录用日期: 2026年2月20日; 发布日期: 2026年2月27日

### 摘要

为解决职业素养评价与真实情境脱节的问题, 本研究提出了一种基于SC-T-P模型的情境化评价方法, 并构建了相应的评价体系。该模型整合了情境学习理论、工作过程系统化理论、人格特质理论与绩效评估理论, 形成了由外至内、由情境到个体的动态嵌套关系。研究立足世校赛“育人本位”与装备制造典型情境, 采用反向德尔菲法, 经两轮16位行业专家验证, 确立了包含3个维度、18项可观察行为指标的评价体系。该体系与世校赛官方评分维度无缝映射, 提供了“0/1”判定操作化示例, 贯通了“理论构建、行为观测、赛场评分”全链条。小样本实证检验显示, 工具内容效度(S-CVI/Ave = 0.92)、评分者信度(Krippendorff's  $\alpha = 0.81$ )与结构效度(累计方差解释率71.4%)均表现良好。本研究构建的评价体系兼具理论深度与实践效度, 能够为世校赛评分细则优化与职业院校“精准育人”提供有效的评价工具, 未来可向智能化、数字化方向发展。

### 关键词

世校赛, 装备制造, SC-T-P模型, 职业素养, 反向德尔菲法

# A Contextualized Evaluation of Professional Competency in Equipment Manufacturing

## —The Development and Application of the SC-T-P Model

Jiangyan Liu<sup>1,2</sup>

<sup>1</sup>College of Education, Zhejiang University of Technology, Hangzhou Zhejiang

<sup>2</sup>Department of Automotive Engineering, Xiangtan Vocational College of Science and Technology (Xiangtan Technician College), Xiangtan Hunan

文章引用: 刘江艳. 装备制造职业素养的情境化评价[J]. 职业教育发展, 2026, 15(3): 20-30.

DOI: 10.12677/ve.2026.153121

## Abstract

To address the disconnection between professional competency evaluation and real-world scenarios, this study proposes a contextualized evaluation method based on the SC-T-P model and constructs a corresponding evaluation system. The model integrates situated learning theory, work process systematization theory, personality trait theory, and performance evaluation theory to form a dynamic nested relationship from the outside-in, from situation to individual. Grounded in the “education-oriented” philosophy of the World Vocational College Skills Competition (WVCSC) and typical equipment manufacturing contexts, the research employed the reverse Delphi method. Through two rounds of validation with 16 industry experts, an evaluation system comprising three dimensions and 18 observable behavioral indicators was established. This system maps seamlessly onto the official WVCSC scoring dimensions, provides operational “0/1” judgment examples, and integrates the entire chain of “theoretical construction, behavior observation, competition scoring”. A small-sample empirical test demonstrated good content validity ( $S-CVI/Ave = 0.92$ ), inter-rater reliability (Krippendorff's  $\alpha = 0.81$ ), and construct validity (with a cumulative variance explanation rate of 71.4%). The evaluation system constructed in this study possesses both theoretical depth and practical validity. It can serve as an effective evaluation tool for optimizing the scoring criteria of the WVCSC and for supporting “precise education” in vocational colleges, with potential for future intelligent and digital development.

## Keywords

WVCSC, Equipment Manufacturing, SC-T-P Theoretical Model, Professional Competency, Reverse Delphi Method

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着新一轮科技革命与产业变革纵深推进，装备制造产业对“精密 - 复杂 - 协同”型技术技能人才的需求急剧攀升。教育部《全国职业院校技能大赛执行规划(2023~2027年)》首次将“职业素养”列为与技能同重的核心维度，明确提出培养“具有行业特质、中国情怀、国际视野的综合型技术技能人才、能工巧匠、大国工匠”[1]。世校赛作为全球职教风向标，2024年升级为国际性赛事，吸引74国2.9万名选手，其中装备制造赛道数与赛项数均居19个大类之首；2025年更由3个赛道扩展为7个(含机械设计与制造、数字化生产等)，凸显该领域人才需求旺盛与育人标准升级的双重张力。

然而，WVCSC现行评分细则对“职业素养”仅给出“职业道德、工匠精神、安全意识”3个观测点，缺乏可观察、可量化、与装备制造典型任务深度耦合的行为指标，赛场评价与日常教学难以同频共振。因此，本研究提出SC-T-P四重嵌套模型，旨在构建一套装备制造职业素养的情境化评价体系。本文首先阐述SC-T-P模型的理论基础与建构逻辑；其次，介绍基于该模型的指标体系精炼与反向德尔菲法验证过程；再次，通过实证检验分析评价体系的信效度与权重；最后，总结该模型与评价体系的理论价值与实践意义，以期职业院校的精准育人提供参照。

## 2. 文献综述

### 2.1. 职业素养评价方法的演进

现有研究对“职业素养是什么”的讨论很充分，但对“如何评”的关注不足，尤其是缺乏针对技能大赛特定情境的系统性综述。职业素养评价方法历经从抽象量表向行为锚定模型的演进，其核心矛盾在于信度与效度的平衡。现有研究可分为三类(如表 1):

Table 1. Evolution table of professional competency evaluation methods

表 1. 职业素养评价方法演进表

方法类型	核心特征	优势	局限性	代表性研究
通用量表法	李克特量表测量抽象维度(如责任感)	结构简洁、信度高( $\alpha > 0.90$ )	脱离实操情境,无法捕捉具体行为	王寿斌等构建的高职生职业素养量表[2]
行为观察法	现场记录行为频次或质量(如工具归位)	贴近真实情境,可直接观察	缺乏标准化锚点,主观性强(记录差异 $> 30%$ )	徐国庆等的工具归为频次观察[3]
行为锚定量表	将等级与具体行为描述绑定	明确评分标准,提升评分一致性( $\kappa = 0.72$ )	设计成本较高,需大量行为样本	European Training Foundation (2024)对 8 国技能大赛的实证研究[4]

上述演进表明,单纯的量表或观察法已难以满足装备制造赛道对“可观测、可追溯、可量化”职业素养的需求,行为锚定量表结合虚拟仿真技术,已成为提升评价信效度的关键路径。然而,针对装备制造赛道的行为锚定量表模型仍属空白。

### 2.2. 装备制造行业职业素养的核心维度

当前,装备制造业正逐渐成为科技创新与产业创新深度融合的核心领域,是抢占未来经济和科技发展制高点的关键所在。同时,面对全球贸易保护主义的不断升级,装备制造业已成为实现科技自立自强、构建安全可靠产业链和供应链的重要支撑[5],2023年,装备制造业增加值占规模以上工业增加值比重为 33.6% [6],产业升级对技术技能人才提出“数字化 + 绿色化 + 国际化”的复合素养要求,倒逼人才培养标准升级。

《中国制造 2025》技术路线图指出,装备制造对“安全、精密、协作”要求极高[7]。在装备制造领域,职业素养的内涵已由传统的“通用素质”细化为与智能制造工艺深度关联的三项关键特征:安全规范、精度意识、团队协作。这些特征不仅在《职业教育专业教学标准-2025 年修(制)订》等教育政策中明确规定[8],也是世校赛竞赛规程装备制造赛道的评分要点,成为衡量中职装备制造专业学生职业素养的权威标尺。但现有评价体系未嵌入上述情境,导致与岗位实际脱节。

### 2.3. 技能大赛评分体系的范式差异

世赛与世校赛在职业素养评价上存在根本性范式差异,现有研究多聚焦技能水平测试,却忽视了两种赛事在定位、结构与评价取向的范式差异,对职业素养关注不足,缺乏针对细分赛道的模型验证。为厘清本研究“面向世校赛的职业素养模型”之边界与权重来源,有必要对两项赛事的官方评分体系进行系统比较。依据 WorldSkills International (2023)《WS Competition Rules》与中华人民共和国教育部《世界职业院校技能大赛竞赛规程》的原始文本[9],可发现二者在“技术技能 - 职业素养”双维度上的权重分配与指标内涵存在显著差异,具体表现为以下三点:

(1) 定位差异:世赛以遴选全球顶尖技能人才为导向,强调在极端工况下的“零缺陷”表现;世校赛

则以“以赛促教、以赛促学”为核心功能，更关注学生的基础职业规范是否达标。

(2) 结构差异：世赛将安全、可持续、职业行为等要素统一归入“职业素养”维度，形成“隐性-显性”连续的测量逻辑；世校赛则将“团队合作”“创新创业”等原属世赛职业素养范畴的指标独立为单列维度，致使“职业素养”在统计意义上被窄化为“职业道德-工匠精神-安全意识”三因子，权重进一步稀释。

(3) 评价取向差异：世赛以结果为导向，把职业素养融入专业技能之中，评价的是“在顶尖技能中的职业素养”，职业素养被量化为可观察的绩效指标；世校赛以过程为导向，体现了“育人本位”的发展主义取向，职业素养更关注“学生是否养成了良好职业习惯”，承担“价值引领”与“基础规范”双重功能。

直接移植世赛框架将造成“高权重-高要求”与“低权重-基础要求”之间的结构性错位，导致指标过高、情境缺失、判定模糊。

## 2.4. 研究缺口与本研究的创新点

### 2.4.1. 现有研究缺口

通过对国内外职业素养评价相关文献的系统梳理与实践反思，本研究识别出现有研究存在三个显著缺口，制约了评价体系对装备制造类专业人才培养的支撑效能。

方法论滞后与定位偏差：现有评价范式多源于对世界技能大赛(WorldSkills)技术标准的直接移植与简化应用，忽视“世校赛”在“育人本位”核心目标上的根本差异。这种移植导致评价标准与日常教学情境适配性不足，难以实现对教学过程的逆向引导与改进功能。

核心职业情境嵌入缺失：装备制造领域对“安全规范(Safety)、精度意识(Precision)、团队协作(Collaboration)”的素养要求具有极端情境依赖性，现有通用评价模型未能将上述核心素养嵌入具体工作流程的可观测行为指标，导致评价悬浮于真实岗位情境之上。

模型实证验证不足与推广性存疑：多数研究集中于构建宏观评价框架，缺乏针对“装备制造”这一细分赛道的测评工具，所构建模型未能基于大赛与日常教学双场景。

### 2.4.2. 本研究创新点

针对上述研究缺口，本研究的创新性工作主要体现在以下三个层面：

理论创新：提出“世校赛本位”的职业素养评价新范式。本研究突破“世赛标准降维移植”的路径依赖，旨在将竞赛的先进标准转化为驱动教学改革与全员素养提升的内生变量，实现“以赛促教、以评促学”的有机统一。

方法创新：开发基于行为锚定量表的“安全-精密-协作”情境化测评工具体系。将装备制造的核心素养要求，解构并转化为一套基于二元行为锚点(0/1)的、可观测、可记录的情境化行为指标集。例如，将“精度意识”操作化为“加工前执行二次图样确认”、“使用规定量具执行首件检验并记录”等具体行为，彻底取代依赖主观判断的抽象评分，实现了评价的客观化、精准化与情境化。

实践创新：实现“赛场-日常”一体化应用与实证验证。本研究将所构建的SC-T-P模型与评价工具同步应用于大赛赛场与常规实训教学双场景，旨在打通竞赛与教学之间的应用壁垒，形成一套既服务于竞赛选拔、又适用于全员素养诊断与培养的、标准化、可推广的评价解决方案，切实弥合人才培养与岗位需求之间的断层。

## 3. SC-T-P 模型：情境化评价的理论框架

### 3.1. 理论基础

本研究构建的SC-T-P四重嵌套模型，其理论基础整合了情境学习理论、工作过程系统化理论、人格

特质理论与绩效评估理论，以系统回应装备制造领域职业素养评价的多维性、情境性与可测性需求。四项理论以“嵌套”方式分别支撑模型的四个逻辑层次，形成一个有机整合的理论框架：

情境层(Situation)以情境学习理论为主要依据[10]。该理论深刻指出，职业素养的有效习得必须发生在真实的“实践共同体”中，通过“合法的边缘性参与”来实现。这为本模型将评价嵌入“精密-危险-协同”的装备制造典型情境提供了核心哲学基础，明确了脱离具体工作情境的抽象评价难以迁移至实际工作环境。

能力层(Competency)的构建主要基于工作过程系统化理论[11]。该理论强调，职业能力是由嵌入在工作流程中的隐性、情境依赖与协同性知识构成的。它支持本研究将岗位关键技能(如能量锁定、零缺陷质量控制与节拍协同)转化为一系列可观察、可测量的具体行为指标。

特质层(Trait)引入人格特质理论的视角[12]。该理论认为，“责任心”与“谨慎性”等稳定心理特质是预测个体行为跨情境一致性的关键变量。本研究创新性地将其操作化为“安全人格”与“责任特质”作为调节变量，并通过外显行为(如是否主动进行安全互检、是否持续遵守操作规范等)进行间接推断，此举避免了引入复杂心理测量导致的学科偏离，确保了模型在教育评价场域中的适用性。

绩效层(Performance)以绩效评估理论为核心方法论[13]。该理论依托 Kirkpatrick 四级评估模型，尤其关注“行为层”与“结果层”的显性输出。在本模型中，以实际行为达成度(如操作合规性)与任务完成质量(如加工精度)作为素养的终极表征和评价依据，确保评价结果客观、可量化。

上述理论之间形成了由外至内、由情境到个体的动态嵌套关系：外层情境激活中层能力表现，能力层受内层稳定特质调节，最终共同外化为可观测的绩效。该框架既保持了教育评价的实践导向，也具备了多理论协同的深度解释力。

### 3.2. SC-T-P 四重嵌套模型的建构逻辑

基于上述理论的交叉融合，本研究建构了 SC-T-P 四重嵌套模型(见图 1)，其核心设计逻辑在于通过由外至内的层级嵌套与动态传导，将隐性的综合素养转化为可观测、可测量的显性绩效，从而实现操作化评价。该模型的层次结构与内在逻辑如下：

情境层(Situation)：将装备制造的典型工作情境抽象并定义为“精密操作”、“危险规避”与“协同作业”三类耦合条件，作为素养表现的宏观环境、驱动力与评价场域，限定并激活了后续所有层次的表现(此层主要依托情境学习理论)。

能力层(Competency)：依据职业能力发展模型(如 COMET)的理论框架[14][15]，结合工作过程分析，提炼出“能量锁定完整性”、“零缺陷首检意识”、“节拍协同响应能力”等关键可观测技能点。这些技能点是响应情境层要求的具体表现，也是通往最终绩效产出的必要桥梁(此层主要依托工作过程系统化理论)。

特质层(Trait)：引入“责任心”与“安全人格”作为内在调节变量，其效应通过“行为一致性”、“规则主动遵守度”等外显行为指标间接反映与测量。该设计避免了引入复杂心理测评工具导致模型过度心理学化，确保了其在教育评价框架中的适用性(此层借鉴了人格特质理论的解释视角)。

绩效层(Performance)：以行为输出(操作过程合规性)与任务结果(产品加工质量)为核心评价对象，参照 Kirkpatrick 评估模型的行为与结果层级，进行基于证据的达成性判定(通过/未通过)，形成闭环评价。此层是所有内外部因素相互作用的最终凝结与客观呈现(此层严格遵循绩效评估理论的方法论)。

模型各层次间呈现动态嵌套与传导关系：外层情境激活中层能力要求，能力表现受内层稳定特质调节，最终所有因素共同投射并凝结为可量化评估的绩效结果。该设计不仅在纵向上实现了多理论的有机融合，更在横向上保证了模型在实操评价中的可行性、客观性与适用性。

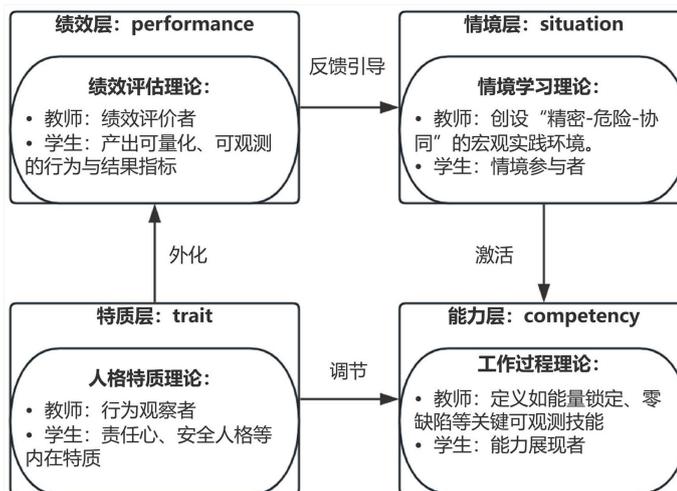


Figure 1. Schematic diagram of the SC-T-P four-layer nested model  
图 1. SC-T-P 四层嵌套模型示意图

综上所述, SC-T-P 四重嵌套模型通过“情境层”激活能力、“特质层”调节表现、“绩效层”客观评价的动态传导机制, 为将抽象的职业素养转化为可观测的行为指标提供了坚实的理论框架。后续的指标池精炼与专家验证, 均是在此模型指导下展开的具体实践。

### 3.3. 任务池精炼机制

为将 SC-T-P 抽象框架转化为可直接落地的行为量表, 本研究以近五年全国职业院校技能大赛“现代加工技术”等赛项的技术文件、世界职业院校技能大赛装备制造赛道竞赛规程、美国 O\*NET 职业数据库中的典型岗位(如 51-4041.00 机械师、51-9162.00 数控程序员、51-9161.00 数控操作员)任务描述, 经去重合并, 剔除表述重复及仅语序调整的条目, 最终形成包含 138 项任务的装备制造岗位任务集(附件 A), 其内容覆盖精密加工、设备调试、协同作业等典型工作场景。

将以上 138 条任务按照 COMET (Competency Model) 聚类, 将装备制造岗位任务划分为三个维度: 概念性任务(涉及理论知识、原理解释)、程序性任务(涉及操作流程、技能执行)和情境性任务(涉及实际工作场景中的问题解决与应变)。该划分旨在全面覆盖职业能力的认知、技能与情境应用层面, 确保评价指标体系的系统性与科学性。去除重复或意思相近的内容, 保留职业院校学生在实训场或世校赛现场能完成的任务, 并将任务名称结合国标映射、世校赛映射, 压缩为可观察、可量化、合法合规、可溯源的口语化、赛场化名称, 形成指标名称, 初步构建的装备制造岗位核心职业素养指标库包含 30 项指标, 其中情境性(S) 15 项, 程序性(P) 10 项, 概念性(C) 5 项, 完整指标库见附件 B。该流程既保障了指标的国际岗位覆盖度, 又确保了其在职业院校本土化实践中的可行性。

### 3.4. 反向德尔菲过程

为避免传统德尔菲法因“任务过载”引发专家疲劳、共识虚高等问题, 本研究采用“反向德尔菲”(Validation-Consensus Delphi)范式, 研究者先行构建情境化指标池, 专家仅承担“验证-修订”职能, 从而更高效、更聚焦专家共识、避免初始发散, 提升决策效率与内容效度。该方法在职业教育评价领域已被证实可在控制专家规模的同时, 快速收敛共识并保留情境适切性[16]。

首轮反向德尔菲专家邀请 16 名世校赛装备制造赛道竞赛指导教师(广东 3 人、浙江 3 人、湖南 4 人、四川 3 人、河南 3 人), 均具备技师以上职业资格, 平均竞赛指导经验 4.3 年, 专业领域覆盖数控技术(5

人)、机电一体化(7人)、工业机器人(4人),确保对“精密-危险-协同”情境具备深度解读力。通过问卷星平台进行30项条目的三级量表评估(1=完全同意,2=微调意见,3=删除并说明理由),回收率100%。依据共识阈值(同意率 $\geq 70\%$ 保留,删除率 $\geq 30\%$ 剔除),首轮结果:15项直接保留,5项剔除,10项需修订。

第二轮采用现场结构化研讨,邀请7名湖南省职业院校技能大赛裁判(含5名世校赛铜奖指导教师)对修订条目进行深度论证。修订逻辑包括:①语义明确性(如“节拍协同能力”调整为“节拍协同响应能力”);②行为可观测性(如“安全人格”转化为“安全互保意识”);③评分可操作性(如“环保合规意识”增加废弃物分类处置范例),修订逻辑与结果见表2。

**Table 2.** Revision logic and results

**表 2.** 修订逻辑与结果

原指标名称	处置方式	修订或删除依据	最终归属指标
首件标识管理	合并	行为域被“零缺陷首检意识”完全覆盖	零缺陷首检意识
个人防护装备 冷却液管理 现场6S维持 安全警示标识 现场整洁度	合并	均属“现场规范”可观察行为	现场规范习惯
工装夹具保养 刀具磨损识别 设备点检表	合并	均指向“设备保养责任”	设备保养责任
紧急停机响应	合并	与“风险识别与处置”同义	风险识别与处置
材料节约意识	删除	赛场文件无扣分项,且非学生日常训练核心目标	—
程序备份习惯	删除	赛场存在违规风险	—
故障沟通能力 同伴沟通协作	删除	已归入世校赛“团队合作”独立维度	—
创新设计能力	删除	已归入世校赛“创新创意”独立维度	—

### 3.5. 指标体系

为实现SC-T-P模型向赛场及日常训练评分工具的转化,本研究将18项指标与世校赛官方评分维度(职业道德与行为规范、安全意识、工匠精神)进行系统映射,形成“行为-维度-评分”逻辑闭环(见表3)。该设计使裁判能够“看见行为、对应维度、直接打分”,既强化了评价的客观性,又确保了模型在高速竞赛环境下的适用性。

**Table 3.** Mapping and evaluation table of observable indicators for professional competency in equipment manufacturing and the WVCSC

**表 3.** 装备制造类职业素养观测指标与世校赛映射及评价表

序号	维度	三域分类	指标名称	世校赛映射场景 (七赛道通用或列举)	判定示例 (符合 = 1, 不符合 = 0)	数据来源	主要依据标准
1	职业道德与行为规范	S	制造现场6S与定置	现代加工技术、模具、焊接等“现场管理”评分项	操作结束后,工具按定位置位归位,废料分类投放	现场观察	GB/T 19001-2016 7.1.4

续表

2	职业道德与 行为规范	P	设备/硬件点 检与保养	智能制造设备、电梯、 无人机“赛前/赛后检 查”	按点检卡完成润滑或电 池外观检查并签字	记录审 查	企业点检规范 /TSG T5002-2017
3	职业道德与 行为规范	S	绿色制造与 环保处置	焊接烟尘、无人机电池 回收、加工冷却液更换	废切削液、废电池投入 专用收集桶	现场观 察	GB 18599-2020
4	职业道德与 行为规范	P	实时数据记 录规范	七赛道均需填写“产量 -质量-时间”原始记 录	每工序完成后即时记 录,无事后补写	记录审 查	GB/T 19001-2016 8.5.1
5	职业道德与 行为规范	S	节能与低耗 操作	无人机经济转速、焊接 经济电流、电梯待机	运行电流/功率 $\leq$ 额定 80%	现场仪 表	JT/T 325-2023 附 录 E
6	工匠精神	P	零缺陷首检 与自检	模具首件、无人机首 飞、电梯层门首件	首件完成后立即自检并 填写检验表	记录 + 测量	GB/T 2828.1-2012 4.2
7	工匠精神	S	多工位/多角 色节拍协同	智能制造产线、焊接协 作、电梯井道协同	整体节拍延迟 $\leq 60$ s 且 交接口述完整	现场计 时	企业工艺卡
8	工匠精神	P	高精度测量 与重复验证	数控、模具、无人机臂 长校核	重复 3 次读数差 $\leq 0.02$ mm 或 $\leq 1'$	工具测 量	JJG 21-2008/GB/T 3177
9	工匠精神	P	程序/参数验 证习惯	数控、机器人、无人机 飞控	新程序先空运行或仿真 $\geq 5$ s 无报警	仿真记 录	GB/T 25663-2010 6.1
10	工匠精神	C	制造图纸与 工艺识读	模具、焊接坡口、电梯 导轨安装图	坐标系或基准线设定误 差 $\leq 0.05$ mm	工具测 量	GB/T 4458.4-2003
11	工匠精神	C	基础代码/参 数输入	数控 G 代码、机器人 RAPID、无人机 Mission	输入段无语法报错且逻 辑正确	系统反 馈	企业编程规范
12	工匠精神	S	资源寿命与 库存预警	刀具、焊嘴、无人机桨 叶、电梯钢丝绳	达到预设寿命前主动报 告更换	记录审 查	企业刀具/配件管 理规范
13	工匠精神	P	工艺参数双 确认	焊接电流、电梯导轨扭 矩、无人机电池电压	启动前核对工艺卡数值 并签字	记录+ 观察	企业工艺卡/TSG T5002
14	工匠精神	P	传感器/坐标 零位校准	无人机磁罗盘、机器人 TCP、数控回零	校准后误差 $\leq$ 允许阈值 且记录	系统截 图	MZ/T 105-2020/企 业规范
15	安全意识	S	能量锁定 LOTO	数控、机器人、电梯、 焊接电源	完整挂牌上锁 + 双人 互检	现场观 察	GB/T 33579-2017 6.3
16	安全意识	S	制造异常急 停处置	七赛道“突发异响/报 警”评分点	立即按下急停并口述报 告	现场+ 记录	GB/T 33000-2016 5.3
17	安全意识	S	安全互保与 提醒	协作工位、无人机桨 区、焊接火花区	同伴高风险操作前主动 检查	互评记 录	GB/T 33579-2017
18	安全意识	P	闭锁型安全 回路验证	电梯层门、机器人护 罩、无人机机头罩	通断电阻 $\leq 1 \Omega$ 且双人 签字	仪表测 量	GB 7588-2020 7.7.3

## 4. 小样本实证检验

### 4.1. 信效度检验

#### 4.1.1. 内容效度(Content Validity)

内容效度主要检验量表条目是否全面、准确地覆盖了所要测量概念的全部内容范围。一般邀请 5~7

名前期未参与研究的专家对条目逐条打分(4级: 1 = 不相关, 4 = 高度相关), 计算 I-CVI (条目水平内容效度指数) = 3~4 分的专家人数/专家总数; S-CVI/Ave (量表水平) = 所有条目 I-CVI 的平均值, 可用来证明所测条目无遗漏、无多余, 解决条目是否齐全的问题。

本研究采用内容效度(CVI)进行检验, 邀请 6 名职教专家(未参与德尔菲者)对 18 项指标进行 4 级相关性评定(1 = 不相关, 4 = 高度相关)。计算 I-CVI (条目水平效度指数), 保留 I-CVI  $\geq 0.83$  的指标, 最终 18 项指标均达标, S-CVI/Ave = 0.92, 表明内容效度良好[17]。

#### 4.1.2. 评分者效度(Inter-Rater Reliability)

评分者信度主要检验不同评分者对同一对象使用同一工具时, 打分结果的一致程度。通常找 2~3 名裁判, 用所开发的工具独立观察同一批学生, 用 Krippendorff's  $\alpha$  或 ICC 计算一致性, 可用来证明工具不因人而异, 解决测得稳不稳的问题。

本研究选取湖南省职业院校技能大赛(装备制造赛道) 3 个赛项(n = 18 名学生), 由 3 位裁判独立使用本工具打分, 计算 Krippendorff's  $\alpha = 0.81$ , 表明评分者一致性良好[18]。

#### 4.1.3. 结构效度(Construct Validity)

结构效度是检验量表是否真正测到理论上要测得维度, 通过收集 50 份以上的实测数据, 并做探索性因子分析, 用来证明测得准, 解决测得是不是那回事的问题。

本研究采用“探索性因子分析(EFA)”对 18 项指标进行降维(n = 52), KMO = 0.78, Bartlett 球形检验显著( $p < 0.001$ ), 提取 3 个公因子, 累计方差解释率为 71.4%, 与理论维度(职业道德与行为规范、工匠精神、安全意识)基本吻合, 表明结构效度可接受。

## 4.2. 指标权重确定

为提升工具在现场评分中的实用性, 研究引入“层次分析法”对 18 项指标进行权重赋值。

#### 4.2.1. 专家构成与判断矩阵构建

邀请 9 位具备职业院校技能大赛执裁经验的装备制造类专家, 采用 1~9 标度法对 3 个维度及 18 项指标进行两两比较, 构建判断矩阵(表 4)。使用 Yaahp 12.0 软件进行一致性检验, 所有矩阵的一致性比率 CR 均小于 0.1, 满足一致性要求。

Table 4. Consistency test results of expert judgment matrix

表 4. 职专家判断矩阵一致性检验结果

矩阵名称	阶数 n	$\lambda_{max}$	$CI = (\lambda_{max} - n)/(n - 1)$	RI	$CR = CI/RI$	是否通过
Goal 维度	3	3.032	0.016	0.58	0.028	✓
安全意识指标	4	4.074	0.025	0.90	0.028	✓
工匠精神指标	9	9.271	0.034	1.12	0.030	✓
职业道德与行为规范指标	5	5.063	0.016	1.12	0.014	✓
最大 CR	—	—	0.034	—	0.038	✓

#### 4.2.2. 权重确定

最终权重结果如表 5 所示。安全意识维度权重最高, 工匠精神次之, 职业道德维度权重为 0.23。在指标层面, 能量锁定 LOTO (0.21)、“零缺陷首检意识” (0.18)与“异常急停处置” (0.15)位列前三, 反映出装备制造赛道对“安全规范”与“精度控制”的高度关注。

**Table 5.** Indicator weight table  
**表 5.** 职指标权重表

一级维度	权重	二级指标	子权重	组合权重
安全意识	0.42	能量锁定 LOTO	0.50	0.21
		异常急停处置	0.36	0.15
		安全互保与提醒	0.10	0.04
		闭锁型安全回路验证	0.04	0.02
工匠精神	0.35	零缺陷首检与自检	0.51	0.18
		多工位/多角色节拍协同	0.15	0.05
		高精度测量与重复验证	0.12	0.04
		程序/参数验证习惯	0.09	0.03
		制造图纸与工艺识读	0.05	0.02
		基础代码/参数输入	0.03	0.01
		资源寿命与库存预警	0.03	0.01
		工艺参数双确认	0.03	0.01
		传感器/坐标零位校准	0.03	0.01
		职业道德与行为规范	0.23	制造现场 6S 与定置
设备/硬件点检与保养	0.23			0.05
绿色制造与环保处置	0.20			0.05
实时数据记录规范	0.17			0.04
节能与低耗操作	0.13			0.03

#### 4.2.3. 加权评分模型

最终形成如下加权评分公式，用于赛场快速计算职业素养得分：

$$\text{职业素养总分} = \sum(\text{指标得分} \times \text{子权重}) \times \text{维度权重}$$

该模型在 2025 年楚怡杯职业院校技能大赛选拔赛试点应用，裁判反馈“评分更快、维度更清晰、争议更少”。

#### 4.3. 小样本实证检验

为检验 SC-T-P 职业素养评价工具的区分效度与实用性，本研究于 2025 年 9 月在湘潭市某职业院校开展小样本实证。在机械工程系举行职业院校技能大赛装备制造数控加工技术校赛选拔赛中，由 3 名具备执裁资格的裁判对 34 名二年级选手(男 29, 女 5, 平均 18.6 岁)独立打分。高分组与低分组(各 27%)在“能量锁定 LOTO”“零缺陷首检”两项得分差异显著( $p < 0.05$ ,  $t = 2.31$ )；3 位裁判 Krippendorff's  $\alpha = 0.81$ , AHP 权重 CRmax = 0.04 未修正。小样本显示工具区分效度可接受、稳定性良好，当前样本量有限，未来研究需显著扩大实证检验的样本规模，探索性因子分析的样本量应达到指标数的 10 倍以上，后续将扩大样本继续验证。

### 5. 结束语

本研究构建了 SC-T-P 四重嵌套模型，为解决复杂技能领域的职业素养评价提供了新的理论范式。该模型“情境 - 能力 - 特质 - 绩效”的嵌套逻辑，不仅为装备制造赛道开发了可操作的评价工具，其思想

内核也为其他相关领域的评价体系开发提供了可借鉴的框架。通过反向德尔菲法,初步构建了包含 18 项观测指标的装备制造类职业素养评价体系,为职业院校技能竞赛的素养评价提供了理论框架与实践依据。然而,本研究专家样本仅覆盖广东、浙江、四川、河南四省的 16 位专家,在区域代表性与样本规模上存在局限,此为本研究需要明确指出的边界。

未来的研究计划主要围绕三个方向展开:在验证研究方面,将前往江苏、湖北等装备制造产业聚集省份开展实证研究,覆盖不同地区、不同办学层次的职业院校,并纳入更多地域的专家参与德尔菲法和权重确定,以检验模型的信度与效度,并优化指标体系,增强研究结论的普适性和可靠性;在效标构建方面,将积极对接企业,追踪毕业生职业发展路径,旨在建立“模型评分-企业绩效”的关联模型,从而验证其预测效度,构建从赛场到企业的闭环效标;在技术赋能方面,将探索开发基于计算机视觉的自动评分系统,利用摄像头与 AI 识别技术实现 18 项行为的自动判定,以期降低人工裁判成本,提升评价效率,推动职业素养评价向智能化、数字化方向发展,最终服务于职业院校的精准育人及学生的全面发展。

## 参考文献

- [1] 教育部办公厅关于印发《全国职业院校技能大赛执行规划(2023-2027 年)》的通知[Z/OL]. 中华人民共和国教育部政府门户网站. [http://www.moe.gov.cn/srcsite/A07/zcs\\_yxds/s3069/202303/t20230331\\_1053799.html](http://www.moe.gov.cn/srcsite/A07/zcs_yxds/s3069/202303/t20230331_1053799.html), 2025-07-16.
- [2] 王寿斌. 高职生职业素养量表的开发与验证[J]. 职业技术教育, 2021, 42(25): 68-73.
- [3] 徐国庆. 行为观察法在职业素养评价中的应用局限与改进[J]. 中国职业技术教育, 2022(18): 45-49.
- [4] European Training Foundation (2024) Behaviourally Anchored Rating Scales for Skill Competitions. ETF Publications.
- [5] 装备制造行业发展如何? 工信部谋划新一轮装备制造业稳增长政策[Z/OL]. 中研普华\_中研网. <https://www.chinairm.com/hyxx/20241204/110010601.shtml>, 2025-02-20.
- [6] 中国政府网. 新中国 75 年经济社会发展成就系列报告之三[R]. [https://www.gov.cn/lianbo/bumen/202409/content\\_6973433.htm](https://www.gov.cn/lianbo/bumen/202409/content_6973433.htm), 2024-09-10.
- [7] 国家制造强国建设战略咨询委员会. 《中国制造 2025》重点领域技术路线图[R/OL]. <https://www.cae.cn/cae/html/files/2015-10/29/20151029105822561730637.pdf>, 2026-02-24.
- [8] 教育部. 关于印发《职业教育专业教学标准(2025 年版)》的通知[Z]. 2025.
- [9] 世界职业院校技能大赛组委会. 2025 年世界职业院校技能大赛总决赛评分要素[EB/OL]. <https://sicsve.cdp.edu.cn/webapp/accessory/ueditor/sicsve.cdp.edu.cn/upload/file/2025-06-30/1751247001068040052.pdf>, 2025-06-11.
- [10] Lave, J. and Wenger, E. (1991) Situated Learning: Legitimate Peripheral Participation. Cambridge University Press.
- [11] 赵志群. 职业教育工学结合一体化课程开发指南[M]. 北京: 清华大学出版社, 2009.
- [12] McCrae, R.R. and Costa, P.T. (2008) The Five-Factor Theory of Personality. In: John, O.P., Robins, R.W. and Pervin, L.A., Eds., *Handbook of Personality: Theory and Research*, Guilford Press, 159-181.
- [13] Kirkpatrick, D.L. and Kirkpatrick, J.D. (2006) Evaluating Training Programs: The Four Levels. 3rd Edition, Berrett-Koehler Publishers.
- [14] Raven, J. (2020) The COMET Competency Model: A Framework for Vocational Excellence. *Vocations and Learning*, 13, 211-231.
- [15] 刘育锋, 王屹, 赵文平. 反向德尔菲法在职业教育标准制定中的应用与实证——以“1+X”证书标准为例. 中国职业技术教育, 2022(14): 58-64.
- [16] 赵志群, 白滨, 乔鹤. 职业能力测评方法手册: COMET 方案[J]. 北京大学教育评论, 2019, 17(1): 165-176.
- [17] Davis, L.L. (1922) Instrument Review: Getting the Most from a Panel of Experts. *Applied Nursing Research*, 5, 131-135.
- [18] Hayes, A.F. and Krippendorff, K. (2007) Answering the Call for a Standard Reliability Measure for Coding Data. *Communication Methods and Measures*, 1, 77-89. <https://doi.org/10.1080/19312450709336664>