

工程教育认证关于专家如何判定 评价机制底线要求的工作指南

一、说明

本文件不是关于中国工程教育专业认证标准的全面指南性文件，其目的仅在于对“专业是否建立面向产出的内部评价机制”的关键“形似”要求提供一个基准线，并为专家在认证实践中提出可行的应用方法。使得认证专家在按照面向产出的认证模式开展工作时，充分理解基于现有工作基础的最低要求。

为此，本文件特别强调工程教育专业认证的合格“底线”，即采用面向产出的评价方式时，专业达到合格要求的必要条件是：专业建立了面向产出的内部评价机制，并通过评价能证明自己合格。

中国工程教育专业认证仍然处在成长期，坚持底线对于确保认证工作始终沿着正确的方向发展，以及在确保认证质量的同时快速发展具有重要意义。对于认证专家，最重要的认证工作专业性水准应体现在：对于面向产出的核心要求态度明确，对于发展中的细节问题能够适当灵活处理。

二、底线概念的性质

本文提出的“底线”并非对中国工程教育认证标准的修订或扩展，甚至不是“新”解释。底线概念是基于如下考虑的操作性提法：

- 便于专家对真正的“形似”要求达成一致，为向“神似”进步提供坚实认识基础和工作基础；
- 适应当前认证工作的客观水平，在认证规模迅速扩大的背景下，帮助专家快速抓住面向产出评价方法中的关键要求；
- 突出认证合格的必要条件，有利于在认证规模扩大的同时，保持认证工作本身质量持续稳定。底线要求只能用于判定专业达不到合格要求，不能用于判定专业合格。

三、底线要求的合理性

面向产出的评价方式要求专业必须自行举证证明毕业生达到合格要求，不是由专家帮助专业“寻找”证据。证据必须能够反映学生能力，并能说明能力水平达标，因此证据只能来自于专业内部面向产出的评价机制的结果。应当承认当前认证实践中专家认可的部分证据有明显的缺陷，如：

- 仅用是否有某课程作为证据：课程并非产出；
- 主要依赖考试/考核成绩计算：成绩是否与要求的能力匹配并未经过评估；
- 用人单位反馈：样本偏少，且用人单位普遍不了解产出要求的全部内涵，反馈通常比较泛。

认证标准第四项对专业内部评价机制建设有明确要求。标准说明中用两个关键词，即评估与评价，描述了内部持续改进机制的两个核心：

- 评估：研究确定面向产出目标达成评价所必须的基础数据，并通过日常教学质量监控持续收集数据，这些数据（例如学生成绩）确实能反映能力；
- 评价：基于基础数据采用科学方法定期评价各项毕业要求达成情况。并且能够解释评价方式的合理性。

当前认证实践表明，突出对标准中内部评价机制建设有关要求是必要的。近两年认证报告中对于标准第四项提出的问题相对较少，这与此第四项 Pw（认证结论形式修改前）最多形成明显不一致。实际上目前专业建设中自我评价是最薄弱环节，这在认证报告中没有得到反映，对认证“以评促建”的初衷产生负面影响。

四、如何在认证工作中判定底线

本文侧重于如何将“底线”概念用于申请书审核和自评报告预审，现场考查可做参照使用。必须再次强调：底线概念只能用于判定某专业申请不应被受理，或者判定某专业认证应该被终止，不能用于判定专业认证合格。

申请书篇幅有限，提供的材料相对较少，因此本文主要介绍如何基于最基本的材料进行判定而不会导致误判。对于自评报告预审，因为可用材料丰富，下结论更加容易，采用的审阅方法没有本质差别，本文一并描述。

基本方法可以简称为“一体两翼”。“一体”是指对内部评价机制直接相关材料的判读，主要是该专业是否明确建立了面向产出的内部评价机制并开始正常运行，特别是机制是否真的“面向产出”。“两翼”是指两方面辅助材料：一是对产出目标进行指标点分解，二是将指标点达成任务落实到具体教学活动（课程或相关教学环节）中。通过对这两个方面材料进行判读，帮助专家判定专业建立的内部评价机制是否真实。如果指标点大多数难以衡量，或者相应支撑课程的评价材料无法判定指标点达成，则有理由认为内部评价机制实际上并未形成。

需要说明的是，分解指标点并非认证规定的必须做法，对于采用其它方式建立内部评价机制的做法，只要合理就应该肯定，但专业同样要对其内部评价机制的基础数据进行评估：是否合理，是否可行。

1. 内部评价机制（一体）

主要审阅点：

- 是否有一个完整的内部评价机制的描述？
- 评价机制各相关角色是否明确，合理？
- 是否能明确该机制的数据收集过程？
- 是否已经正常运行，还是仅仅为这次认证准备的？
- 是否能明确体现面向产出，即是否对数据与产出相关性进行评估？

主要存疑点：

- 如果专业描述的内容主要是学校传统的内部质量监控机制，需要考虑机制是否已经过面向产出的改造提升；
- 如果描述的内容是近几年来流行的算分法，需要考虑分值是否能反映全体学生能力要求的达成状况；
- 如果上述两者分别描述，需要考虑是否建立了一个完整、真实的机制。

2. 对毕业要求进行指标点分解（两翼之一）

主要审阅点：

- 指标点分解能否反映正确理解了通用标准和专业补充标准相关要求？
- 指标点是否“可衡量”？
- 指标点对应课程能否支撑“可衡量”？

主要存疑点：

- 专业自己制定的毕业要求只是字面上反映标准要求（包括直接套用），但内涵理解不一致，导致实际用于评价的产出目标并不完全支持标准要求；
- 技术类要求过泛（例如设计能力），导致提供的达成证明缺乏针对性；
- 非技术类要求含义不明（例如国际视野、终身学习能力），导致无法衡量；
- 把针对“复杂工程问题”的能力案例化，导致无法衡量。

3. 毕业要求指标点的支撑课程（两翼之二）

主要审阅点：

- 对任一指标点，支撑课程是否可能实现学生能力目标？
- 支撑课程的支撑效果是否可评价？
- 相关教学活动的评分标准是否与其支撑目标关联？

主要存疑点：

- 支撑课程只是内容与特定指标点相关，甚至只是名称上有点关联，但很难实现能力目标，则不能认为课程实现支撑；
- 很难通过对课程的评价判断是否能使学生达成相应的能力要求，则不能认为课程实现支撑；具体来讲，就是相关能力要求是否在该课程的评分标准中占有比例？
- 对于以课堂授课为主的课程，其考试/考核方式是否能反映学生能力？如果仅仅体现记忆与验证能力，则不能作为支撑的依据；
- 对于实践类课程（或者实践环节比例较大的课程），其考试/考核方式是否能反映每个学生的能力？如果没有明确的评分标准则不能作为支撑；

五、基于底线判定标准项达成情况

为进一步方便专家开展认证实践，对底线相关标准项作进一步说明，并通过每项标准设置达成四级指标，提出该项标准的底线，有关注意事项如下：

本指南提出值得特别关注的“考查点”，并针对相应考查点给出关于符合程度的基本描述。所选考查点是根据其在标准中的作用以及当前认证实践中常见问题确定，并不能全面覆盖标准项，也不能替代标准要求。

针对所列考查点，给出达成的四级评价指标，并分级给出简单描述。此描述不具备标准意义，仅用于帮助专家给出对该项指标判定认证结果。四级中第1级可理解为“按标准要求没有问题”，第4级可理解为“没有达到相应标准要求，应判定为认证不能通过”。也就是说，在当前认证实践中，考虑当前认证工作阶段性特征，第3级应该理解为认证合格的“底线”。

1. 毕业要求项

【关键词1】可衡量

【关键词2】覆盖（通用标准中的12条毕业要求）

【关键词解释】这两个关键词在认证实践中可以结合在一起考查。先谈“覆盖”。覆盖不是拘泥于字面上的理解，而是指如果能证明专业本身的毕业要求达成，也就等于证明了通用标准的毕业要求达成。因此关键在于对要求内涵的理解（注意：如果没有对于内涵的理解，即使照抄通用标准也未必就能判断“覆盖”）以及如何衡量达成与否。再谈“可衡量”。简单地说就是专业能否提供一个要求是否达到的判定方法。这应该是一个可重复执行的方法，而不是因人而异的。（例如，“身体好”是一种要求，但不可衡量。在实践中，体检总是通过建立一个参照系，也就是分项的指标，来进行判定的）

在认证实践中，这两点都是通过对专业如何分解指标点来考查的。一个毕业要求项分解成哪些指标点反映专业对于毕业要求项的理解（专家应该考查一个毕业要求项下指标点的集合是否体现相应毕业要求的广度与深度。另一方面，每个指标点是否可衡量是相应毕业要求是否可衡量的基础。任何一个指标点如果判定为“不可衡量”，那么整个毕业要求项就不可衡量。因此专业不可能建立完整的毕业要求达成评价机制。（在认证实践中，专家对于难以判断的指标点应该要求专业提供自己采用的衡量方法，专家的责任是判断专业所说的方法是否合理并可操作。）

注意：指标点分解并不是认证标准的要求。只是从判断覆盖与可衡量两方面考虑，不分指标很难实现对毕业要求的支撑与达成判定。如果专业不分指标点也能很好地解决上述问题，就应该清晰地表述自己采用的毕业要求达成判定方法的内容以及分析其合理性。专家应该在这一表述的基础上判断是否可接受。以下考查点仅针对采用指标点分解的做法。

【考查点1】指标点分解是否能清晰的表现专业对与通用标准中12项毕业要求的理解以及本专业毕业要求对其的覆盖？

- (1) 指标点分解体现了对通用标准中12项的合理解。从深度与广度上均能反映通用标准的要求。

- (2) 指标点分解体现了对通用标准中 12 项的合理解释，广度上能覆盖通用标准的要求，但部分指标点不够具体清晰，导致达成度判定在体现通用标准要求上有不确定性。
- (3) 指标点分解体现了对通用标准中 12 项的基本理解，广度上能覆盖通用标准要求，但部分指标点含义不清晰，会影响对通用标准要求达成的判定。
- (4) 指标点分解不合理，不能反应对相应毕业要求的理解。特别是对于照抄通用标准毕业要求的专业，指标点分解往往表现为字面上的拆分。或者通用标准中某些要求明显没有被覆盖。

【考查点 2】对分解的指标点是否全部可衡量？从教学实践与认证实践来看，毕业要求指标点达成的支撑在于教学活动（为了方便以下统称课程），因此衡量也是基于课程来实现的。从这个意义上说，毕业要求中的所谓技术性要求与非技术性要求有明显的差别（这里并不想探讨技术性/非技术性的定义，只是就事论事的描述）。

技术性要求采用“纵向”分解更有利于与学校中流行的“基础/专业基础/专业”的课程分类方式匹配，适应能力的逐步提升，便于教学中贯彻指标点支撑，也便于在课程评价时考核指标点达成。这里举例说明什么是“纵向”分解。例如：“将数学、自然科学与工程基础...用于解决复杂工程问题”可以分解为基本语言与工具/建模/模型推演与分析/模型比较与评价/模型综合与创新等。又如：能针对复杂工程问题设计解决方案，并能考虑...，可以分解为：基本设计工具/功能与性能单体设计/系统设计/考虑非技术因素的工程设计/设计中的权衡与决策/设计创新等。

而非技术要求一般没有层次概念，关键是能否“说得清”，只有说清才有可能纳入教学内容，只有说清有关责任人才能评价。例如：终身学习能力，可以分解为技术理解力/总结与综述能力/提出问题的能力/在任务环境下区分知道与不知道的能力/避免习惯性干扰的能力等。

基于上述理解，可以将考查点 2 分为两点，分别对应技术与分技术要求：

【考查点 2-1】对于标准中偏技术性要求，分解的指标点是否可衡量？

- (1) 指标点分解合理，便于按课程层次衡量是否达成。课程支撑也能提供可操作证明。
- (2) 指标点分解基本合理，层次比较清楚，便于说清楚衡量方法。
- (3) 指标点分解能考虑到可衡量的要求，但在若干点上衡量的可操作性不足。
- (4) 指标点分解基本上没有考虑可衡量性，只是在字面上给以分解甚至没有分解。

【考查点 2-2】对于标准中的非技术性要求，是否能够通过分解指标点使其含义清晰化。即专业基本上对于描述性名词或形容词能给出自己的解释。

- (1) 对于描述性的能力（这里指得是“团队能力”、“国际视野”、“终身

学习能力”等说法)要求能够具体列出其主要内涵,且其内涵表述是可以通过常规的考试/考核来体现的。

- (2) 对于描述性能力的要求能够具体列出其内涵,但部分内容衡量的可操作性不足。
- (3) 指标点分解能考虑到可衡量的要求,但总体上衡量的操作性偏弱。
- (4) 指标点分解基本上没有考虑可衡量行,只是字面上给以分解甚至没有分解。

2. 持续改进项

【关键词】课程质量评价

【关键词解释】面向产出的评价机制的核心是毕业要求情况达成情况评价。毕业要求达成是靠课程支撑的,所以毕业要求达成评价的基本数据来源必须是以毕业要求为参考系的课程评价。没有这样的基础数据,通过算分的方法评价毕业要求达成一定会沦为数字游戏。

【考查点 1】专业是否理解什么是标准中说的课程质量,并在内部课程评价机制建设中得以体现?

专业充分理解课程质量的衡量标准是对于相关毕业要求指标点的实际支撑情况。并能在内部持续改进机制中建立合理的课程评价标准与方法,并有足够文档表明能够稳定实施面向产出的内部课程质量评价机制。

专业对课程质量理解基本符合面向产出教学评价的要求。至少在核心课程中能够稳定实施面向产出的质量评价,评价标准与方法基本合理。

开始建立面向产出的内部课程评价机制,标准与方法还有待改进,但至少已经在主干课程上实施。

对课程质量概念的理解没有明显体现面向产出。除了应对本次认证的一些做法外,没有材料可以证明已经开始建立面向产出的内部课程评价机制。

【考查点 2】毕业要求中非技术要求专业是如何评价达成情况的?

针对相应毕业要求指标点专业有明确的定量或者定性的评价方法,并有文档证明已经能够较稳定地实施。

针对相应毕业要求指标点,专业能描述适当的评价方法并有明确的评价责任者,但相应的历史文档明显不足。

针对相应毕业要求指标点,专业能描述基本合理的评价方法,但方法的可操作性不足,能提供的实施数据偏少。

没有材料能证明,除本次认证采取的措施,专业对于非技术的毕业要求达成有实际的评价过程。

【考查点 3】专业是否建立了毕业要求达成情况评价机制?

- (1) 专业能够说清楚评价的方法以及基础数据的来源。方法合理,基础

数据可靠并对面向产出的评价有意义。

- (2) 专业有基于常规教学质量提供的基础数据的毕业要求达成情况评价机制，但数据覆盖面不够充分，评价的稳定性不明确。
- (3) 能够定时进行毕业要求达成度评价，但基础数据不足，方法的合理性说服力不足。
- (4) 除了针对本次认证采取的措施，没有材料可以证明已初步建立稳定的毕业要求达成情况评价。当前的措施也只限于少量考试分数采样基础上的简单算术计算，没有可靠的课程目标达成评价提供基础数据。

3. 课程体系项

【关键词】支持

【关键词解释】课程能不能支持相应毕业要求的达成是衡量课程体系是否能达到认证标准要求最核心的一点。而课程体系支持毕业要求必须从两个层面考查：一是整个体系是否覆盖全部毕业要求，二是每门课程是否实现其在体系中的作用。

谈到课程的支持，要避免陷入纠缠文字的误区。课程体系对毕业要求的覆盖是操作性概念，不是理论性概念，更不是要探讨汉语特点字词含义。从教学评价操作性角度看课程对于毕业要求指标点的支持含义如下：如果说指标点 A 由课程 1, 2, 3 支持,那么在认证实践中只要考查结果认为这三门课从面向产出的意义上合格,就可以认为指标点 A 达成可以认定;不必去纠缠是否其它课程也与 A “有关”。(指标点对应的课程过多其实在评价时是不可操作的,这其实也可以作为一个专业缺乏合理的内部评价机制的间接证据。)

特别需要说明,标准中关于“复杂工程问题”的要求,主要的支持应该是所有课程都应该将培养学生“解决复杂工程能力”作为背景目标,少数高年级课程,特别是实践性较强的课程要给学生展示解决复杂问题能力的机会。从认证实践来看,对课程体系的评价是培养“解决复杂工程问题的能力”是否落实的重点。但不是要每门课中包含复杂工程问题。二是基础课必须加强分析推理(解题)能力培养;专业核心课必须加强分析设计能力培养,少数实践课要突出综合解决问题的能力(这要以合理的评分标准作为保障)

【考查点 1】整个课程体系是否能覆盖全部毕业要求的实现

- (1) 支持矩阵完全覆盖所有指标点,课程大纲能够解释该课程在矩阵中的位置。
- (2) 支持矩阵完全覆盖所有指标点,主要课程大纲能够解释该课程在矩阵中的位置。
- (3) 支持矩阵完全覆盖所有指标点,部分课程大纲明显难以解释该课程在矩阵中的位置,这些课程的数量和分布对整体支持的影响有限。
- (4) 支持矩阵不能完全覆盖所有指标点,或者明显有若干课程的大纲并不能解释其在矩阵中的位置。

【考查点 2】课程是否能满足在课程体系中所承担的任务？（附注：专家在分别考查毕业要求中某一项的达成情况时，可参照本考查点对该项的支持课程回答是否能支持的问题，并以做出具体判断。）

- （1） 承担毕业要求指标点支持的课程大纲明确反映相关要求，课程内容与授课方式确实能够实现目标要求，作业有与课程目标相适应的量与深度；课程成绩（如果平时成绩占有一定比例，包括平时成绩）能反映毕业要求的达成与否。
- （2） 核心课程大纲明确反映相关毕业要求，课程内容与授课方式能够实现目标要求，作业有与课程目标相适应的量与深度；课程成绩（如果平时成绩占有一定比例，包括平时成绩）基本反映毕业要求的达成与否，特别是及格要求应与毕业要求基本一致。
- （3） 至少能够证明专业对于课程如何支持毕业要求有正确的理解，并能开始在内部评价中贯彻，特别是注意到考试要求与毕业要求的对应。
- （4） 教师对课程支持毕业要求缺乏理解与认同，专业没有具体措施推动课程教学向毕业要求支持度改进。

【考查点 3】实践类课程或者环节是否体现对毕业要求的支持？

- （1） 有学分的实践环节均有明确、公开并区分学生个体的评分标准；及格标准与毕业要求中与本环节相关要求基本一致；有材料证明在教学实践中稳定地采用标准评分。
- （2） 有学分的实践环节均有明确、公开的评分标准；及格标准与毕业要求中与本环节相关要求基本一致；在教学实践中标准尚未能稳定有效的对所有同学执行。
- （3） 实践环节有明确公开的标准，但及格标准与毕业要求中与本环节相关要求的关联不够清晰，实施不够稳定。
- （4） 实践环节的评分标准与相关毕业要求脱节。