

# 基于学科核心素养的学业质量评价与水平划分

陈友芳

**摘要：**在新一轮课程改革中，普通高中思想政治学业水平测试将分为合格性与等级性学业水平考试，以客观准确地度量学生思想政治学科核心素养的真实发展水平为根本目的。新型课程标准能否成为思想政治课程学业水平考试的命题依据，关键要看学业质量标准是否管用好用。而要使得思想政治学业质量标准管用好用，需要研制出一套质量水平划分科学精确、具有充分可测性和可操作性的学业质量标准。这是当前和未来思想政治学科评价研究面临的重大新问题新挑战，是思想政治课程标准研制需要突破的核心问题。

**关键词：**学科核心素养 学业质量评价 水平划分

陈友芳，华南师范大学公共管理学院教授，《经济生活》教材主编，思想政治课程标准修订组、思想政治学科核心素养测试组核心成员。

目前已徐徐拉开帷幕的新一轮普通高中课程改革正在构建新型的思想政治课程标准：以思想政治学科核心素养为灵魂，融学业质量标准与传统的课程标准于一体。这种以思想政治学科核心素养为引领的新型课程标准不仅是教材编写的指导、教师培训的大纲、课堂教学的遵循、学生学习的标准，而且是考试命题的依据。

在新一轮课程改革中，普通高中思想政治学业水平测试将分为合格性与等级性学业水平考试，以客观准确地度量学生思想政治学科核心素养的真实发展水平为根本目的。新型课程标准能否成为思想政治课程学业水平考试的命题依据，关键就看学业质量标准是否管用好用。要使得思想政治学业质量标准管用好用，就要求我们研制出一份质量水平划分科学精确、具有充分可测性和可操作性的学业质量标准。这是当前和未来思想政治学科评价研究面临的重大新问题新挑战，是思想政治课程标准研制需要突破的核心问题。笔者在参与思想政治课程标准修订和学科核心素养测试工作中对这个问题进行了一些初步思考，在此将我的不成熟想法与大家共同分享，以期抛砖引玉。

## 一、关于思想政治学科核心素养基本评价策略的思考

要构建质量水平划分科学精确、具有充分可测性和可操作性的思想政治学业质量标准，首先要搞清楚如何才能观测到思想政治学科核心素养及其发展水平。

思想政治学科核心素养，是指个体在面对复杂的、不确定的现实生活情境时，综合运用本课程的学习所孕育出来的学科知识与技能、学科思想与观念，在分析情境、应对挑战、发现问题、确认问题、思

考问题、解决问题的过程中,表现出来的参与经济、政治、社会、文化生活的关键能力和必备品格。根据普通高中思想政治课程的学科本质和功能定位,它包括政治认同、理性精神、法治意识、公共参与等基本素养。

从评价角度看,与其他核心素养一样,思想政治学科核心素养同样是不可直接观测和度量的。用信息不对称博弈理论的语言说,某个人的思想政治学科核心素养如何,是他的私人信息,外人不可直接观测。面对不可直接观测的私人信息,外人只能通过可观测变量加以推断。这是我们寻找思想政治学科核心素养水平评价方法的基本出发点。

根据信息不对称博弈理论的要求,要通过可观测变量有效地推断不可观测变量,评价者设置的可观测变量必须要至少满足“充分性条件”:充分性是指这个或这些可观测变量是不可观测变量的充分统计变量。具有充分性的可观测变量必须具备两个基本性质:一是敏感性高,二是随机性小。

敏感性高是指可观测变量与不可观测变量之间有着密切的内在关联,一旦不可观测变量发生变化,就会引起可观测变量及时的、强烈的变化。可观测变量的敏感性越高,评价者对不可观测变量的推断难度就越小。例如,要推断一个农民种地是否努力,用产量高低来衡量就比用其体重大小来衡量更加合理。虽然,一个努力种地的人也许会因为其付出更大的劳作而使得其体重下降,但是,其努力程度无疑与产量水平之间具有更强的内在关联。

通常,可观测变量与不可观测变量之间不会是一一对应的关系。如果二者的关系是一一对应的,从评价角度看,等于不可观测变量本质上是可观测的。例如,人们经常说“一份耕耘一份收获”。如果事情真的如此,外人就可以从观测到的收获水平完全准确地推断出耕耘水平:一份收获意味着一份耕耘,两份收获意味着两份耕耘,三份收获意味着三份耕耘……。然而,对于评价来说,麻烦就在于一份耕耘并不一定带来一份收获。如果运气不好,也许一份耕耘带来的收获会小于一份收获;如果运气好,也许一份耕耘带来的收获会大于一份收获,因为收获的大小除了要受到耕耘程度的影响外,还要受到许多随机因素的影响,例如,农民的收获就可能受到天气、自然灾害、种子质量、化肥质量等因素的影响。因此,要充分地推断出农民种地的努力程度,就必须尽可能地排除随机因素的干扰,这意味着:1. 评价者必须选择相关性强的可观测变量。例如,多年的平均产量水平无疑就比单个年份的产量水平更能准确反映农民的努力程度。2. 评价者必须引入多个可观测变量对不可观测变量加以推断。例如,如果不仅只看该农民实现的产量,而且还看气候、当地其他农民的产量、土地肥沃程度等,就显然比单一只看该农民实现的亩产量高低更能准确推断出该农民的努力程度。因为,一个在气候不佳年份能够种出不亚于平常年份的亩产量的农民,通常要付出很高的努力程度,或者一个在相对不肥沃的土地上能够种出不亚于耕作肥沃土地的亩产量的人,通常也要付出更高的努力程度。因此,综合这些可观测变量能够更好地排除随机因素的干扰,更准确地推断不可观测变量的水平。

当然,无论怎样努力排除,评价者无法做到完全排除随机性,这既是因为受制于评价技术,也是因为受制于评价成本。因此,任何评价或推断都是不完美的。

综合敏感性高、随机性小这两个基本性质,要充分地推断出思想政治学科核心素养的真实发展水平,评价者必须认识到:1. 评价核心素养发展水平的基础性变量应当来自于具有可观测性的行为及其结果。一个人的核心素养水平怎么样,会通过其行为及其结果表现出来,一个人的核心素养与其行为表现之间的内在关联性无疑要高于与其他可观测变量之间的关系。作为评价者,他的核心任务之一就是要从诸多具有可观测性的行为表现中筛选出具有高度敏感性的行为表现特征。

2. 用于评价某一个核心素养发展水平的可观测变量应该是一群或一组,即通过多个可观测变量及同一个可观测变量的出现频率来加以更准确地推断。这意味着至少两点:第一,用于推断某个核心素养发展水平的行为表现变量应该是一群或一组。例如,我们要充分推断一个人的某方面人品,可能既需要观测他在单位的表现,也需要观测他在社交中的表现,还需要观测他在家庭生活中的表现。又例如,如果

一个人能够屡屡发表高质量的论文，显然比偶尔只发表一两篇高质量论文要更能反映其学术研究能力。第二，用于评价某个核心素养发展水平的可观测变量除了相关的行为表现群以外，还必须参考其他变量。正如评价农民种地的努力程度还要考虑气候、土地肥沃程度等因素，评价学生某个学科核心素养的发展水平也同样要考虑其他变量：如情境复杂程度与新颖程度、任务的难度。例如，在更复杂或更新的情境中或在更难的任务中，能够展现出同样行为表现，无疑意味着更高的素养水平。

3. 由于评价者不能完全排除随机性的干扰，因此，任何学科核心素养发展水平的评价都不可能是完美的。再好的评价也只是对真实素养发展水平的一个近似，我们只能追求尽可能地近似于真相。

因此，基于信息不对称博弈理论的启示，思想政治学科核心素养的学业评价实际上就是要求评价者要按照敏感性强和随机性小等原则，通过精心、巧妙、科学地设置复杂程度和新颖程度不等、难度不等的任务情境以诱导出能够充分指向某个核心素养及其发展水平的行为表现群。因此，思想政治学业质量标准的核心要素包括：情境、学科内容、学科任务、行为表现。具体而言，就是通过复杂程度和新颖程度不同的具体情境中，让学生运用学科知识和方法完成某一项或几项任务难度不等的学科任务，依据所展现出来的行为表现群来推断学生某一素养的发展水平。

## 二、关于思想政治学业质量水平具体划分标准的思考

根据前一部分的思考，要根据思想政治学业质量推断学科核心素养发展水平，应该考虑三个维度：情境的复杂程度与新颖程度、行为表现群或学科任务的完成质量、学科任务的难度。与此相关，思想政治学业质量标准研制需要解决的另一个关键问题则是：如何根据基础性可观测变量—行为表现群和参考性可观测变量—情境复杂程度与任务难度，具体区分学业质量，进而判断思想政治学科核心素养水平的高低呢？这是一个蕴含着大量评价学理论研究 with 实证研究难题的问题。笔者接下来将就学业质量标准中如何具体刻画思想政治学业质量水平谈谈自己的思考。

为了不使得问题变得过于复杂，笔者暂时忽略情境复杂程度与新颖程度、任务难度这两个维度，谈谈如何对任务完成质量进行划等。

假设思想政治学业质量水平划分为4级，1级最低，4级最高，每一级均代表着整体学科核心素养的一个层级。因此，为了便于刻画思想政治学业质量的不同水平层级，学科任务完成质量也分为四级。

又假设学业质量标准是服务于纸笔测试。那么，我们首先必须筛选出适合或能够纸笔测试的思想政治学科任务。根据思想政治课程的本质特性与功能定位，思想政治课程的基本学科任务可以归纳为以下四项：分类与描述、解释与论证、预测与选择、辨析与评价。

分类与描述是指对真实社会生活情境中的事物、现象与问题的特征、表现、范围按照某个维度进行刻画、比较和分类，以识别事物、现象问题的性质与种类归属。

解释与论证是指对真实社会生活情境中的事物、现象与问题，运用学科方法与技能分析原因，探究不同变量之间的关系，以寻求对世界的理解；或运用理论逻辑和实证材料对探究结论与理论命题进行合乎逻辑和科学要求的检验、理解。

预测与选择是指结合具体的社会生活情境，运用科学的方法和原理对行为、事物、现象、问题的结果或影响进行分析和预测；根据约束条件和目标设计出合理可行的方案，比较不同方案的优劣利弊并作出理性选择。

辨析与评价是指结合具体的社会生活背景，根据某个维度对理论、事物、现象的价值与功能进行分析和评价，辨识要素、事物、概念、价值之间的关系；合理运用相关理论和方法，对不同观点与立场、不同利益诉求进行价值判断，进行辩护和批判。

应该说，上述四项学科任务基本上已经涵盖了思想政治课程几乎全部的学习任务。那么如何衡量这些学科任务的完成质量呢？笔者认为，评价者可以考虑从如下几个方面具体区分某一学科任务的完成

质量。

1. 描述、预测与评价维度的运用力。根据学科核心素养的发展水平，从用一个维度(视角)进行分析到能从多个维度(视角)进行分析，从只能用常见的、熟悉的维度进行分析到能够从不常见不熟悉的维度进行分析，从只能用给定的(常见)维度进行分析到自己能够选择合适的维度进行分析，无疑体现着学科核心素养水平的不同。

因此，从描述、预测与评价的维度看，我们依次可以从如下四个层级观测某项学科任务的完成质量与思想政治某个学科核心素养的发展水平：从只能用常见的一个维度进行分析，到能够从常见的多个维度进行分析，再到能够用新的维度进行分析，再到能够根据任务要求恰当选择合适的维度。

2. 学科知识的理解力。无论是解释世界还是实践应用，首先必须要求对学科知识具有良好的理解和掌握。理解力是评价学科任务完成质量的重要指标。从评价角度看，我们依次可以从如下四个层级观测对学科知识的理解力水平：从只能够准确表述学科知识原理，到能用恰当事例加以阐释，再到能够知道理论原理成立的条件且对学科知识的理解是结构化的，再到能够提出批判性意见。

3. 学科知识的应用力。思想政治课程学习的基本目的之一是通过运用学科知识解决实践问题，以更好地参与社会实践。因此，学科知识的应用力也是评价学科任务完成质量的基本指标。从评价角度看，我们依次可以从如下四个层级观测学科知识的应用力水平：从只能够正确运用一个知识，到能够正确运用多个知识，再到能够综合运用知识，再到能够最优地或创造性地运用知识。

4. 命题或方案的论证力。无论是认识、解释世界还是实践应用，均要涉及对原理、探究结论或方案的论证，论证其是否科学合理或是否合理可行。因此，论证力是度量学科任务完成质量的又一基本指标。我们可以依次从以下四个方面区分论证力的水平高低：从只能用典型事例加以论证，到能够用学科原理加以逻辑论证，再到能够选择合适的(事例或理论)证据加以论证，再到不仅能够选择合适的证据，而且具有结论与证据的检验意识和能力。

5. 方案的设计与选择力。当我们应用思想政治学科知识与方法解决各种公共与非公共的实践问题时，需要我们不仅要有效应用学科知识，而且还要能够根据决策目标与约束条件设计并选择出合理可行的解决方案。因此，根据实际情境中的决策目标与约束条件，设计与选择方案的能力是评价学科任务完成质量的又一项基本指标。我们依次可以从如下四个方面观测相关任务或实际问题解决能力的高低：从只能够设计出一项合理可行的方案，到能够设计出多项合理可行的方案，再到不仅能够设计多项合理可行的方案，而且还能够清楚各个方案的优缺点，再到能够根据实际情形从多个方案中选择出最佳方案。

虽然上述五个方面的具体划等是否科学合理还有待于实证研究，但笔者要表达的一个核心思想是：维度的运用力、学科知识的理解力与应用力、命题或方案的论证力、方案的设计与选择力不仅是评价思想政治学科任务完成质量或学业质量的基本指标，而且它们是能够清晰地量化评价的。尽管这种量化评价也许不是基数意义上的，而是序数意义上的。

(责任编辑：王葳)