

评价项目：展现理解力和技能

评价量规和评分指南

评价项目

基于项目的学习需要一种更为先进的评价方式。在这种评价方式中，学生能将学习视为一种过程并使用问题解决策略来满足或超越对项目的期望。如今，评价量规和评分指南也已被应用到课堂教学中去了，这能使学生能更好地理解什么东西正在被评价、什么标准等级正在被遵循以及哪些明显压抑的作品标准必须舍弃。当然，这个评价量规和评分指南的要旨在于监控和调整这些评价进展而不是仅仅为了评价最终结果。

评价量规和评分指南具有以下若干优势：

- 通过给学生清楚呈现对其学习的评价过程和期望方向，学生的学业成绩会得到改善。
- 学生能更好地评价自己的学习质量。
- 学生会得到更多关于自己强弱势的信息反馈。
- 学生能明白提供同伴反馈的使用标准。
- 标准是用特定的术语来确定的。
- 评价变得更为客观和一致了。
- 评价学生学业的时间缩短了。
- 教学的有效性是通过多种方法来探讨的。
- 学生的学业进展是根据基准来衡量和存档的。

作为计划的指南，评价量规和评分指南给学生提供了明确的熟练度目标。因此，如果学生手头有这份评价指南的话，他们就能在学习之前就知道不同层次的质量表现。一旦学生经常使用这种评价方式来评定其学习，那么他们就会逐渐承担更多产生最终结果的责任，而减少了质问诸如“我正在被评价吗？”这些问题。

作为一种监控当前项目进展状况的量规，评价量规和评分指南在帮助学生实现其目标方面非常便利。学生能根据这种熟练度标准衡量自己目前的学业成就，并据此提醒自己要实现的目标。

最后，作为一种总结性评价，评价量规和评分指南能被用来评估项目、学生群体或者单个学生。学生能将相同的评价量规和评分指南用于自我评价、团体评价和同伴反馈中。

为了使评价变得有效，每条量规和评分指南的语言表述必须能为广大学生所理解。使用第一人称的语言表述，有助于加强学生对评价过程的掌握。一旦学生参与了评价量规和评分指南的形成过程，那么这种作用将更为突出，因为在这一过程中，学生必须积极关注并讨论有效的绩效、作品和行为的特性，并赋予其更深的理解和见解。另外，学生也有权形成他们自己的评价标准，这样的话，其学习将变得更为聚焦和自主。

评价量规的定义

虽然评价量规呈现一个事物的多种形式和水平，但我们仍能找出它们的普遍特性，即：

- 关注测量的既定目标经常是根据**特性**或（绩效、行为或质量）维度进行界定的。
- 使用指标对不同水平的每个**特性**进行描述。

在上面的“评价项目”示例库中，所有的评价量规都呈现纵贯最高到最低的**四个质量水平**。“评价项目”的评价量规分为一般性评价量规和具体特性评价量规两种。其中，一般性评价量规概述了对象而不根据具体特性进行界定，而下面的“合作评价量规”则是根据具体特性进行界定的例子：

合作评价量规

特性	4	3	2	1
贡献	我总是积极参与讨论，为小组作贡献。我接受并完成了所有要求我的任务。我帮助本小组确定任务，带领小组实现我们的目标。	我参与小组的讨论，为小组作贡献，完成了分配给我的任务，帮助确立和实现我们的目标。	我有时需要鼓励才能完成分配给我的任务。我在确立目标和实现目标的过程中需要得到帮助。	我选择不参与。我没有完成分配给我的任务。我阻挠确定目标，并使小组远离要实现的目标。
合作	我根据主题分享很多想法，贡献很多合适的信息，鼓励其他成员共享思想观点。	得到鼓励我会分享我的想法，我愿意所有的成员都共享思想观点。	受到鼓励我时不时地分享想法，我愿意大多数的小组成员一起分享想法。	我不愿意分享我的想法，因此我对小组讨论没有什么贡献。在其他成员分享观点的时候，我常常打断他们。

下面这个是一般性评价量规，因此没有界定批判性思维的具体特性：

批判性思维的一般性评价量规

4	3	2	1
我能讲明自己正在从事的研究的最重要信息。	我通常能讲明什么信息是最重要的。	我有时能从一大堆混杂不重要的信息中抓住那些最重要的信息。	我通常不能讲明重要信息和不重要信息之间的差别。
我利用掌握的知识来对相应的信息做出推论、得出结论，并核实自己的这些结论是否正确。	我利用了解的知识来对相应的信息做出推论、得出结论，并通常核实自己的这些结论是否正确。	在别人的帮助下，我对相应的信息做出推论、得出结论，但有时我没有合适的理由来说明为什么是这样。	在对相应的信息做出推论方面，我存在困难。
我进行各项研究以获得新观点和新概念。	我努力学习以获得更多的新观点和新概念。	如果某个人提醒我，我就会学到更多的新观点和新概念。	我通常总是安于现状，不去烦恼着学更多的东西。
我能用说或写的方式清楚无误地解释某个主题，并阐明这样做的理由。	我能解释我的观点而且能阐明充足的理由。	我通常能解释我的观点，但并不是总有充足的理由。	我不能解释我的观点。

评分指南的定义

“评价项目”之所以将评分指南与评价量规加以，其重要原因在于评分指南包含了各个水平的分数或分级，这些有助于确定成绩。下面的这个示例看起来与评价量规非常类似，除了添上乘数以增加每个特征的权重外，：

4	3	2	1
关键的等式和数学解答 _____ x 8 = _____			
等式显示对理解力和组织的掌握。	等式显示了好的理解力和组织。	等式显示了发展中的理解力和组织。	等式是受到限制的或完全不存在的。
图表展示了一种数学表示法 _____ x 5 = _____			
图表是合乎逻辑的且有足够的细节，而且答案是显而易见的。	图表是合理的。展现出了所有部分，而且答案也是明显的。	图表是受限的，而且所呈现的答案是不完整的。	图表是不完整的或缺失的。
评价和问题洞察力 _____ x 2 = _____			
评价非常有效，所给信息促成了问题解决。	评价是相关的，所给信息有助于问题解决。	评价需进一步完善，但所给信息与问题解决相关。	评价不完整，且所给信息不能支持问题解决。

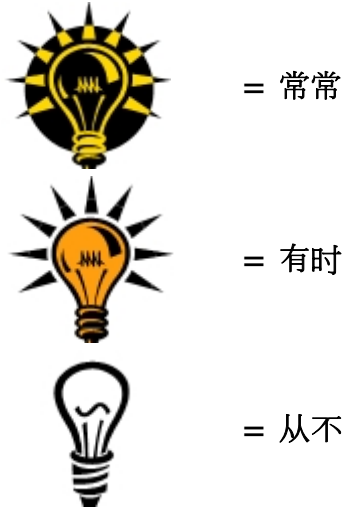
下面这个评分指南示例看似与自我评价问题列表相像，但是它包含了对预期结果进行分级的若干评价指标。


5	4	3	2	1
超出了预期结果。期望得到了实现和扩大，远远超出了项目的范围。	实现了所有的预期结果并包含了额外的促成项目的要素或细节。	实现了预期结果。	接近实现预期结果，但遗漏了一个或两个要素或细节。	没有实现预期结果且遗漏了一些核心要素。

预期结果	权重	成绩
研究论文和报告概述了农业中的基因工程。	X2	
研究论文和报告围绕基因工程农业的利益和风险，提供了一项关于社会、种族和货币问题的支持性调查。	X5	

下面的这个评分指南的例子要求小学生评估自己完成项目的优劣程度。答案提供了一些评价分级信息，但是并未如上述评价量规指标或评分指南评价指标例子那样细致描述。

评分指南图标



我工作非常努力			
我提供正确信息帮助我的小组实现正常运作。			

使用评价量规划分等级

如果要使用评价量规对绩效任务分等级，那么就需要对学生和家长进行教育，因为他们已经习惯了传统测验。与传统测验相比，利用评价量规划分的等级能展示学生的更多情况，包括知识、技能、策略和过程等，这也就要求对质量有更清楚的理解。下面列举了三种可用来对作品和绩效进行分级的评价量规。

使用单个指标的具体特性评价量规

通过给不同的绩效水平赋予分值，我们可以把“具体特性的评价量规”转化为“用于分级的评分指南”。一个特性的可能分值应该考虑到不同特性的相对分值。下面的这个例子，对一个特性的不同水平进行了评价，但并没有把不同特性的相对比重考虑进去。

学习日志评分指南

	10	9	8	7
数据	我的数据是详尽、深刻且完整的。	我的数据是完整的。	我的数据比较完整，但是可能遗漏了一些方面。	我的数据是不完整的，遗漏了一些重要的方面。

	5	4.5	4	3.5
组织	我的数据组织得相当好，可以迅速便捷地找到自己所需要的信息。如果其他人需要这些信息，他们也能够找到这些信息。	我对数据进行了组织，这使我能够找到自己所需的信息。	我计划对数据进行组织，有时很难找到自己所需的信息。	我的数据没有得到精心组织。我很难找到自己所需的信息。
外观	我的日志整洁且具吸引力，而且我的书写也很容易阅读。	我的日志是整洁的，而且我的书写容易阅读。	我的部分日志比较杂乱，有时我的书写难以阅读。	我的日志很杂乱，而且我的书写经常难以阅读。
				总分 19/20

使用多个指标的具体特性评价量规

使用多个指标的评价量规对各个绩效水平赋予分值，这就要求对特性进行更为全面的审视。同一特性在不同水平上的指标未必都是同等重要的。一个特定作品可能在某个特征水平上满足一些指标，但在其他特征水平上则满足另外一些指标。有时，不同水平的指标数量也并不保持一致。一些组成成分，例如创造性地诠释，尽管在高水平的评价量规上表现突出，但在较低的水平上并不出现。

当使用详细的评价量规对整个绩效任务分级的时候，你应该围绕或者强调适用于具体学生学习的指标。然后通过给每个特性分配点数或者观看整体的学习质量，并使用专业判断进行分级。

下面的这个例子表明：使用“每个特性的每个水平都有多个指标”的评价量规对绩效任务进行分级。可见，这种分级带有部分主观性，因为教师必须给出一个看起来合适的分数，而没有将具体分数叠加。如果这个分数是内在一致且公平的，那么学生就会适应这种分级，并喜欢给这种评分类型提供的更多反馈。

海报

	4	3	2	1
内容 (可能60分) 你的得分 57/60	<p>我的海报表明了自己对相关概念的深入理解。</p> <p>我的海报有一个目的并表达了一个主题，提到了一些与话题相关的重要且有趣的内容。</p> <p>我海报的主题和目的是令人惊奇的、原创的和有意义的。</p>	<p>我的海报表明了自己对主要概念的理解。</p> <p>我的海报有一个目的并表达了一个主题，提到了一些与话题相关的重要内容。</p> <p>我海报的主题和目的是有意义的。</p>	<p>我的海报表明了自己对概念的理解存在问题。</p> <p>我的海报试图反映一个目的和主题，提到了一些与话题相关的重要内容，但是主题比较肤浅，目的也有些模糊和含混不清。</p> <p>我海报的主题和目的是可预料到的。</p>	<p>自己对概念的理解存在明显的问题。</p> <p>我的海报没有反映与话题相关的一个主题或者目的。</p>
设计	我有效地且富有创造性地运用了线条、形状、规模	我有效地使用了线条、形状、规模、结构	因为我以众所周知的方式运用了线条、形状、规模	因为我对线条、形状、规模、结构和色彩的运用比

<p>(可能 15分)</p> <p>你的得分 13 /15</p>	<p>、结构和色彩等使得我的海报变得有趣、富有吸引力且具有意义。</p> <p>我所使用的图像与内容相关且对海报的整体意义做了补充。</p> <p>我的海报结构均衡，所有组成部分相辅相成，传递一个核心的信息。</p>	<p>和色彩等使得我的海报变得有趣、整洁和富有吸引力。</p> <p>我所使用的图像与海报的主题相关。</p> <p>我的海报结构比较均衡，绝大多数组成部分相辅相成。</p>	<p>、结构和色彩等，所以我的海报未能吸引观众的注意。</p> <p>一些图像与海报的内容相关。</p> <p>我的海报结构有点不够均衡，有些成分与整体信息无关。</p>	<p>较随意、没有费心，所以我的海报不具有吸引力和感染力。</p> <p>我的海报结构不够均衡，其组成部分未能相辅相成呈现统一信息。</p> <p>我的图像很少与整个海报相关或者我的海报没有图像。</p>
<p>创造性 (可能 15分)</p> <p>你的得分 12 /15</p>	<p>我以非同寻常的、令人惊奇的和合适的方式使用文本、图表和排版，使之在海报中发挥交流的作用。</p>	<p>我以非同寻常的和合适的方式使用文本、图表和排版，使我的海报有趣且具吸引力。</p>	<p>我尝试以非同寻常的方式来使用文本、图表和排版，但是并非一直有效。或者：我海报的样子和主题通常都能够预料到的。</p>	<p>在我海报中所纳入的非同寻常的成分是不合适或无效的。或者：我根本没有努力使自己的海报包含非同寻常的成分。</p>
<p>规范 (可能 10分)</p> <p>你的得分 10 /10</p>	<p>我的海报在写作规范方面没有问题。</p> <p>在适当的时候，我的海报有效地运用了规范以使我的海报更加有趣和富有意义。</p>	<p>我的海报在写作规范方面有少许问题，但是不影响读者理解其意义。</p>	<p>我在写作规范上有点问题，影响了海报的意义。</p>	<p>写作规范上的多个错误影响了我我的海报的意义。</p>
<p>等级 Grade : 92/100</p>				

一般性评价量规

一般性评价量规也可以用来分级，但是并不是对个体具体特征进行阐述。它几乎不能给学生提供有关其表现的具体信息。

下面是一个关于有说服力演讲的评价量规，它概述了被评为不同等级演讲的状况。它并没有给出特定的反馈，例如哪种公众演说技能可以得到改善，但是却为学生呈现了一种总体观念，即他们的演讲与标准相比差距在哪。

说服力演讲

A	B	C	D
学生使用了有效的公开演说技能，例如眼神交流、体态语、声音等。	学生在大多数时间内使用了有效的公开演说技能，例如眼神交流、体态语、声音等。	学生有时未能使用有效的公开演说技能。	学生很少使用有效的公开演说技能。
学生表现出对听众的关注。	学生表现出对听众的一些关注。	学生所表现出的对听众的关注比较有限。	学生所表现出的对公众的关注几乎没有。
学生使用诸多可信的事实和资料以说服听众。	学生使用一些可信的事实和资料。	学生所使用的有些事实是不可信的。	学生所提的内容不包括事实或者事实是不可信的。
演讲以吸引听众的开场白开始，并以较好的结束语收尾。	演讲以开场白开始、以结束语收尾。	演讲没有开场白或者没有结束语。	演讲遗漏了开场白和结束语。

帮助学生和家长理解“使用评价量规划分等级”

那些习惯于根据测验划分等级的学生和家长都需要进行学习，以了解根据评价量规来划分等级的方法比传统测验更能反映出多方面的知识、技能和进程。根据评价量规来划分等级从来不能像统计多项选择测试的答案一样那么精确。我们之所以这样期望，是因为要求具备较深理解力和更高状态思维的工作，并不能像描述和评价事实或基本技能那样简单。探讨如何使用评价量规来分级是十分重要的，特别是在学生过去从未使用过这种方法的情况下。通过展示评价量规的例子以及等级是如何计算，教师就不会对此产生误解，而学生和其家长也会了解根据评价量规进行分级的推理过程。

如果我们对学生的学习的某个单元已进行了多元评价，也同样可避免一些关于等级的焦虑。这种多元评价通过经常地让学生了解自己的强弱项，使得学生为项目中的最终等级做好准备。如果学生在特定的技能、策略或主题上存在困难，那么最终的评价结果将不会是他们第一次听到的。

使用评价量规和评分指南的主要原因是要提高工作质量。它们对期望做出清晰的界定，促使学生对其所做的工作负责。通过使用评价量规，学生们能够清楚地说出自己所学到和所了解的知识，并能明确了解自己如果想要取得成功还需要做什么。另外，评价量规和评分指南这种评价方式能够为学生、教师和家长提供畅通的交流。这种交流使得所有参与人员都能理解这种期望并能保证学生的学习和成功。