

LC WHITE PAPER SERIES

LCP 的偏差测试

Lani Van Dusen, Ph.D.

委托方: *Leadership Circle*

LEADERSHIP
CIRCLE®

摘要

在当今文化背景中,人们对于观察和评价他人的过程中存在的潜在偏见越来越具有敏感性。从个体的评价中消除所有个人偏见是不可能的——人类天生就根据自己的经验,基于刻板印象和感知做判断。然而,通过确保测量模型和相关表现结果不会因被评估个体的群体成员身份不同而产生重大差异,可以最大限度地减少任何个人偏见对评估结果的影响。

目前已知产生测量偏差的若干原因,以及可用于确定一项评估工具是否带有这种偏见的相应测试。2012年,Reynolds 和 Suzuki 在一篇题为《心理评估中的偏见》(Bias in Psychological Assessment) (<https://onlinelibrary.wiley.com/doi/full/10.1002/9781118133880.hop21004>) 的文章中对与这些测试相关的研究文献进行了很好的回顾。

我们选择了研究界最为知名、广为接受的三种测试,以确定 LCP 在多大程度上对所有领导者同等适用,无关性别或种族归属。(请注意,在 LCP 中性别和种族是自行填写的。这些身份会用于测试 LCP 对于不同群体的相关性。)本论文的其他部分提供了这些测试的结果。

LCP 的偏差测试

委托方：

Leadership Circle

作者：

Lani Van Dusen 博士



2021年3月

简介

在当今文化背景中，人们对于观察和评价他人的过程中存在的潜在偏见越来越具有敏感性。从个体的评价中消除所有个人偏见是不可能的——人类天生就根据自己的经验，基于刻板印象和感知做判断。然而，通过确保测量模型和相关表现结果不会因被评估个体的群体成员身份不同而产生重大差异，可以最大限度地减少任何个人偏见对评估结果的影响。

目前已知产生测量偏差的若干原因，以及可用于确定一项评估工具是否带有这种偏见的相应测试。2012年，Reynolds 和 Suzuki 在一篇题为《心理评估中的偏见》（Bias in Psychological Assessment）

（<https://onlinelibrary.wiley.com/doi/full/10.1002/9781118133880.hop21004>）的文章中对与这些测试相关的研究文献进行了很好的回顾。

我们选择了研究界最为知名、广为接受的三种测试，以确定 LCP 在多大程度上对所有领导者同等适用，无关性别或种族归属。（请注意，在 LCP 中性别和种族是自行填写的。这些身份会用于测试 LCP 对于不同群体的相关性。）本论文的其他部分提供了这些测试的结果。

形成偏差

产生测量偏差的一种原因是，评估工具的内容（所含条目）侧重于知识或价值观，而这种知识或价值观仅与一组人群的经历相呼应（或对另一组人群来说具有冒犯性）。关于这一点，研究文献中有一个例子，即对学龄儿童数学技能标准化测试的分析。如果关于数学的问题用词（评估中的条目）都是基于体育方向的，这可能会导致女孩表现不佳，因为她可能在这方面的经验较少。

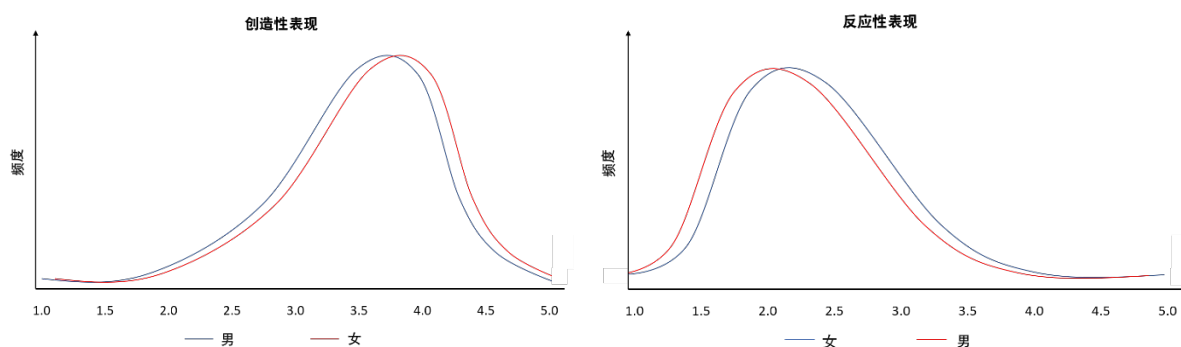
LCP 的构建基于通用的领导力模型，内容以日常业务体验和成年人互动为主。鉴于所有他评和领导者都是商业环境的一部分，并且无论他们的群体身份如何，都会有这些共同经验，因此 LCP 的内容很可能与所有领导者相关。此外，鉴于大多数组织（尤其是大型公司）的商务环境和工作小组都是由男性和女性以及来自不同种族背景的领导者组成的，工作相关的行为和态度不太可能偏向某个群体的经历而忽视另一个群体。

测量错误

另一个主要的偏差产生原因是评估不能公平地测量所有群体。这并不意味着评估不会发现不同群体的平均表现存在差异，理解这一点尤为重要。事实上，群体之间可能存在真正的差异，而评估应该足够灵敏，得以找出这些差异。例如，Okoyi、Smith、Harvey 和 Sherman（2009 年）进行的一项题为《少数民族领袖的领导风格》（Leadership Styles of Ethnic Minority Leaders）、发表在美国心理学会的公报（<https://www.apa.org/pi/oema/resources/communique/2009/08/ethnic-leaders>）上的研究，说明了不同种族的领导者在领导技能和方法上存在差异。

评估是否存在偏差取决于是否存在“估计值的系统性错误”。也就是说，LCP 必须系统地高估或低估特定群体在一个或多个维度的得分。任何一个维度上的平均表现可能不同，但相平均值周围的分布应该是相似的。

当我们使用 LCP 数据对此进行测试时，我们发现所有群体的分布曲线都非常相似。以男性和女性领导者的创造性和反应性表现分布曲线为例。有一个平均差异——即女性领导者创造性曲线的峰值稍微偏右，反应性曲线峰值略偏左，表明女性领导者的平均表现更高（与我们的其他研究结果一致）。然而，分布的整体形状是相似的，在创造性和反应性之间观察到的差异在性别之间是一致的。（注意有很多女性领导者的得分低于男性领导者的平均水平，也有很多男性领导者的得分高于女性领导者的平均水平，这表明该工具在评估中并没有一直将女性表现定义为更好。）



差异预测效力

任何标准化测量方法的目标都是为了能够准确预测结果。如果存在测量偏差，则对不同群体的预测准确性（或力度）会有所不同。为了对 LCP 进行测试，我们将各个维度的发展表现与男性和女性领导者以及具有不同种族身份的领导者的整体领导效能（LE 对比度分数）相关联。

基于 LCP 模型，我们可以预测，随着领导者培养更多的创造性技能，他们的效能应该会提高（正相关），而随着领导者消除/减少反应性倾向，他们的效率应该会提高（负相关）。从附表中可以看出，这些预测的有效性很强，并且所有群体的结果都非常相似。

这一发现表明，无论领导者身份背景如何，LCP 所测量的发展表现在他们的总体效能中都起到了重要的作用。换句话说，LCP 可用于预测领导者的效能，而不会偏向任何群体。

总的来说，我们的测试表明 LCP 评估没有显示出对于性别或种族的系统性偏见。LCP 是个有效的工具，能够为不同群体领导者的领导力发展提供洞察。

群体身份：	创造性 与 LE 的相关度	反应性 与 LE 的相关度
男性	$r = .93, (p < .001)$	$r = -.71, (p < .001)$
女性	$r = .94, (p < .001)$	$r = -.76, (p < .001)$
美国印第安人	$r = .91, (p < .001)$	$r = -.65, (p < .001)$
亚洲人/太平洋岛人	$r = .94, (p < .001)$	$r = -.79, (p < .001)$
非洲裔美国人	$r = .94, (p < .001)$	$r = -.73, (p < .001)$
拉丁裔美国人	$r = .93, (p < .001)$	$r = -.70, (p < .001)$
白种人	$r = .94, (p < .001)$	$r = -.74, (p < .001)$

WIRE 简介

WIRE 是一家成立于 1978 年的独立承包机构，为其他组织提供咨询服务，服务对象为需要提高或评估员工绩效、需要评估针对员工绩效的项目或产品的组织。也就是说，我们提供技术支持，以及从二次分析调查到国内和国际综合性研究的全方位服务。如需了解更多关于 WIRE 的信息，可以到我们的官网查询：www.worldwideinstitute.biz。