



针对国际认证泌乳顾问® (IBCLC®) 认证考试的通过分数研究

针对
国际泌乳顾问考试委员会® (IBLCE®) 而进行

2016 年 6 月

编制者:

Lawrence J. Fabrey 博士
高级副总裁, 心理测量学

Lily Chuang 女士
副心理测量师

Copyright © 2016 国际泌乳顾问考试委员会® (IBLCE®)。所有权。版权所有。未经 IBLCE 书面许可, 不得以任何形式或通过任何手段 (电子或机械, 包括影印或录制) 或者任何信息和检索系统对该出版物的任何部分进行复制或传播。

www.goAMP.com

介绍

该研究的目的是为国际泌乳顾问考试委员会® (IBLCE®) 国际认证泌乳顾问® (IBCLC®) 认证考试确定一个通过分数（分数线）。在确定通过分数的过程中，制定了通过考试的最低标准，考生必须达到该标准才能获得认证状态。

与任何专业的执业标准设置过程一样，需要某种类型的评价。然而，至关重要的是，在确定通过分数时所涉及的评价是由合格的专家做出的，这些专家对考试的预期使用情况有充分的了解，并且在内容领域拥有必要的知识和经验，了解应该合理地期望什么级别的能力。此外，评价应该以一种有意义的方式呈现，考虑测试的格式和目的。该报告描述了通过分数程序的方法和结果——归因于 Angoff（1971 年）——并记录了针对考试的预先确定的通过/不合格截断分数的适当性。

方法论

Angoff 技术是由 AMP，即 PSI 公司 (PSI/AMP) 推荐的，并由 IBLCE 选择作为估算通过/不合格截断分数的程序。Angoff 技术的潜在哲学观点是，标准集应该与项目难度有关，特别是接近分数线的考生的预期困难。因此，该技术需要评委以针对每个测试题目呈现预期的表现评定。完成 Angoff 技术需要几个步骤；用于完成这些步骤的程序如下。

1. **评委的选择。**在分数研究小组任职的评委包括 11 位主题内容专家。他们被挑选出来，以在反映考生人数的潜在相关特征上考虑适当的平衡，如特殊专长领域和地理分布等。
2. **评定程序培训。**在 2014 年 11 月举行的一次会议上，讨论了通过分数研究的目的和程序。培训过程包括对几个题目的小组讨论。对于每一个这些题目，评委都提供了独立的评定，并公布了他们的评定。评委们讨论了在确定评定时所考虑的因素，特别是那些与相对较高或较低的评定相关的因素。
3. **定义能力标准。**评委们参与了关于最低能力程度的执业医师 (MCP) 定义的讨论，即能够展示足够的知识以通过 IBCLC 考试的个人。一般来说，MCP 有足够的知识来安全和胜任地执业。评委们讨论了对于 MCP 将会特别容易或者难于掌握的内容领域，并且记录了讨论的结果。
4. **呈现评定。**评委们被要求为每个题目提供答案和评定，然后根据其回答的适当性，如有需要，审查答案要点并修改他们的评定。例如，特别建议评委要考虑的是，关于提供了初始评定时其回答不正确的题目，他们的评定可能过高的可能性。所有评价都是独立做出的；然而，邀请了评委以确定哪些题目适合进行讨论。完成初始评定后，讨论了一些题目，这导致了对评委评定的一些细微修改。

分析和结果

评委的单个通过分数估算很相似，范围从低到高是 80 分到 85 分。评委估算的平均分数为 82.25 分（应用于 175 个题目的测试），代表原始分数的通过分数 144 分。

计算了鉴别指数值，以分析单个评委所做的评定与其他评委评定合计之间的关系，这类类似于题目-总分相关性。鉴别值（范围从 0.71 到 0.85）代表评定具有很好的可靠性，提供了一个证据来源，证明对评定者的培训是有效的。

结论

IBLCE 的指定代表讨论了该研究的结果，以确定要用于 2016 年春季 IBCLC 考试表的通过分数。给出了研究结果，并考虑了调整是否是适当的，例如，通过应用一个置信区间估算或者考虑与形式相关的统计参数（例如，平均 p 值）。计算了测试后等值结果以帮助进行比较——相对于之前的基础形式，这些新形式的难度。该“现实检查” (Livingston and Zieky, 1982) 的目的是验证 IBLCE 对考试的目的和预期用途的期望，以及考试结果的一致性。讨论之后，IBLCE 官员一致认可将 128 实施为 2016 年春季 IBCLC 考试表原始分数的通过分数。

进行该报告中记录的通过分数研究是因为，使用了基于 2015 年完成并在其他地方记录的国际执业分析研究所建立的已更新的考试规范（即详细内容大纲）。为 2016 年春季的 IBCLC 考试表建立的原始通过分数，将作为新的能力标准，直到完成下一个执业分析。这并不意味着 128 分的原始分数将被用作每个未来考试表的通过分数，更确切地说，统计等值程序将被用于确定在那些未来考试表上将代表相同知识量的通过分数。因为每个考试表上都引入了新题目，并且那些题目的难度直到执行表格和审查题目分析之后才知道，所以通过所需的原始分数可能会更高或更低，以反映那些新考试表的难度。IBLCE 将审查和使用等值程序的结果，以批准每个考试表的通过分数。

参考文献

Angoff, W. H. (1971). Scales, norms, and equivalent scores. In R.L. Thorndike (Ed.), *Educational Measurement* (pp. 508-600). Washington, DC: American Council on Education.

Livingston, S.A. & Zieky, M.J. (1982). *Passing scores - A Manual for Setting Standards of Performance on Educational and Occupational Tests*. Princeton: Educational Testing Service.