

如何考试成为教学的终极角逐： 展望今后十年高利害关系评估测试发展

Mathew Hillier, 莫纳什大学
Andrew Fluck, 塔斯马尼亚大学

这篇论文涉及对澳大利亚高等教育未来十年的高利害关系评估测试发展的展望。作者探讨为推动这项 2025 年的计算机化考试系统的一些因素。此外，论文涉及到此类机化考试的操作特征和教学能力。用途？本论文由澳大利亚政府，学习和教学国家级项目投资 50 万澳元用于未来三年的研究。

关键词: 高利害评估，预测，未来，计算机化考试

前言

未来的考试模式需要具有完全开放性。

这个设想与当今发展局限，思维方式和科技发展显得矛盾。

想象未来，学生可以灵活运用二十一世纪的“工具”来发挥自己的特长。考试方式不仅模仿，甚至有时是以工作生活和社交生活中出现的问题，以评估学生在现实世界中遇到问题应变能力同时，计算机化考试还针对现代高利害关系监管的考试，提供保障，加强其统一性。我们可以想象参加考试的学生可以轻松地获取信息库，工具以及通讯联络，犹如现实网络化世界。

真正实现这个项目很难。如此高真评估测试 (Crisp 2009, Herrington, Reeves & Oliver 2010) 会给学生提供一个非常真实，身临其境的机会，学生可以借此去展示他们二十一世纪的技能。但是如果合适的框架没有完善，涉及考试公平性的问题会接二连三地出现。

关注今日

今日的世界充满各种各样的信息源，其质量参差不齐，质量高的信息虽然存在，但大多数并非如此的。然而，面临日常生活和工作中出现的问题，人们往往有能力利用这些大量的信息和社会资源。如今，迅速给学生提供参考现有资源途径的科技已经触手可得。然而，教育机构来面临此途径给评估程序的完整性和有效性带来的不可忽视的威胁。这种现象在教育领域中的高风险评估中尤其突出。“开放型”考试体系给教育考试权威机构带来了不安，对其保险性提出了疑问。给高利害关系考试领域所带来现实社会场景的同时，也出现了棘手的问题。比如，我们坚持采用 xx，有局限性的纸质测试，或者选择题为考题模式，致使我们愈来愈远离二十一世纪更结合实际的方法。现如今，探索世界上现有的借助科技手段高效的高利害关系评估后，我们发现这些评估不过是复制纸质考试的电子版本。无非加强版本的“文档处理”，或者美化了的多项选择题，这些工具都不具备教学灵活性，对于开拓评估方法有微乎其微的作用。参考名为“论文 1.1”

更糟糕的是，这些方法局限了教学机构的视野，进入关注与教学和科技的僵局。“黑匣子”那些尤为成功的企业正是利用这些数据不断的开采如何更有效，更快捷，同时激发创新的元素。那些‘完好的整理’并随时能提取的信息资源，被锁在系统里，从而失去被利用的价值。在信息量的不断扩大的未来，比较那些信息开放型机构，那些选择‘锁住信息’的机构会面临更多的维护费用。

展望未来

反之，想象一下将来，考试可以要求学生解决真实，复杂甚至‘特别棘手’的问题。一个未来中，二十一世纪的‘谋生的工具’将提供给学生，让他们有可能使用先进的工具去更有创造性的解答问题。再想象一下，学生在解决问题过程中，在浩瀚的信息库的各个接触点都为考试老师所认知。学生在回答一个问题时所有的步骤被详细的记录下来。不仅如此，信息出处，联络方式，决策点都会以日志的方式，布局清晰，言简意赅地展现给考官、从而使考官有可能观察到学生的分析能力，以及在现实生活中解决难题的应变能力。老师可以看到所有的互动记录，这些记录会以综合整理成图表方式，并且具有溯源功能。在考试评估过程中，学生瞬间的决策，解决问题思路会被记录下来。这个过程会以图表的方式呈现给判分老师，以便于老师决定是否因为清晰的思路或决定加分。

各种不同程度的自动给分，电脑协助判分的功能也会被陆续开发出来。每一个给学生提供的问题集，单个问题，客观事实都会被记录下来，保存，并且会被标记，以配合与学院制定的期望学习成果。学习项目负责人和学生都可以参考个人和学校制定的学习目标进而评估学生学习的进展情况。这仅仅是教育分析平台一部分，这个平台还将涵盖计算机化考试变现的数据。

项目设计者，老师，经理将可以对学生技能进行评估。范围更广的课程大纲，表现模式设计，以及评估方式和课程设计的改变很有能给学习结果和学生考试带来影响。学生可以通过广泛的学习主题和二十一世纪所需技能中获取更多进步。评估可以更加具有适应性，根据学生的进步情况而逐渐加大难度。

老师可以利用多种数据分析方法，选择简单的图表方式来测量学生对某个特定问题回答结果和表现情况。疑难，有争议的问题可以被突出以便改进。有效的问题会被推到前列，其他的教育界人士可以参考。在考试进行前，考题质量会通过安全的网络，整体的考试题目开发，质量控制和核实程序被进一步审核。这个程序会确保减少考试过程中出现问题，错误或误解。

计算机化考试行政管理程序的设计考虑到操作实用性，适应那些没有技术背景管理人员和老师。从社团里的小型教育机构到大型院校，都可以通过这个经过反复思考设计出来，开放性科技的考试平台中获得先进的评估技术和科技的便利。

“开放性以标准为基准”会存储数据以适合未知未来。这样可以保证院校可以在未来直接获取细小数据。这些数据可以促进更快捷的了解学生表现情况。还有一系列未开发的数据，分析工具和介绍的工具。随着时间推移，对学生表现更深一层了解将会成为可能。通过更完整观测每一个学生对每一项问题的答复，每一个来源的参考相关的行动。

各个学院的政策可以保留现有的直接发布考试成绩的方式，同时，可以先进的科技将进一步加快对高利害关系考试结果的反馈。反馈周期也将会缩短，譬如，计算机打分将可以提供即时结果。而那些计算机不能完成的评估项目也可以很快的分配到合适的并且可以打分老师。打分老师不用在对不断增加的潦草的试卷头疼。他们只需提供从数据库里提取已存在的评语中挑取适合，并加以改善的反馈即可。通过联网现代化系统可以改进评分体系的一致性。

如何做到

在本文中所规划的未来是来源于现代高等教育领域中多方面因素（参考文章 Hillier & Fluck 2013）。此外，高等教育的“普及”也意味着本科演变成新高中。在澳大利亚，2014 年的数据显示，工作人口中至少有百分之二十八拥有学士学位（ABS 澳洲统计局 2014）。这个数字在 2004 年只有百分之二十一。同样趋势在 18-34 岁年轻成年人呈现，他们更愿意加入受高等教育的行列。根据澳洲统计局 2013 数字表明，拥有学士学位的人已经有 1976 年的百分之五增长到 2011 年的百分之二十六。这个数字的急速增加给高等教育院校带来巨大的压力。他们要保

证给大班的提供高质量的教育，同时又要面对教育基金削减的挑战（参考第七页澳大利亚大学，2015）

现代互联网让获取知识变成一件轻而易举的事情。高等教育机构已经失去了对信息垄断的角色。反之，大学应该运用他们的战略性强项和专业性知识进行批判性的分析和讨论。我们引导学生培养他们具有批判性的思维，分析问题的能力。并以适应二十一世纪现实世界的标准去评估学生的解决问题的能力。

实现这个远见卓识的目标构成了一个真正“ 系统性难题” (Rittel & Webber 1973, Ackoff 1999). 用个比喻比较计算机化考试的贯彻为“ 令人头疼” 的系统性难题恰到好处。因为计算机化考试涉及了来自不同部门的利益相关方，以及在处理问题的视角也截然不同（Linstone 1999）。这些不同的视角来源于院校内外的机构，例如，大学，政府决策人，机构经理，学生，教师，财经部门和人力资源部门，考试组织人员，课程设计者，家长，雇主，技术人员，校园设施管理，建筑和执行部门。澳大利亚的一项国家项目（名为“ 改革考试 2015）在全国聚集人员，以审视现有问题的各项细节，并在作出计划进一步在各大学发展和引领程序上和技术上的方案。这个项目同时在澳洲各州和领地开展活动，从密集型研究到着重教学实践，从大城市到地区性院校。整个项目为期 2 年，从 2016 年到 2018 年。此项目是延续塔斯马尼亚大学从 2007 年工作成果（Fluck, Pullen & Harper 2009）以及之后昆士兰大学通过 OLT 2013 到 2014 年的种子项目 (Hillier & Fluck 2014)。

促使项目贯彻而发展的 10 年战略计划不是改革，更像是一个进化的过程。因为在面对如此大规模的程序上，政策上和技术上的转变面临着不可低估的困难。必须要得到的参与者同一目标，共同参与。从纸笔考试到取代纸笔监督考试时代，再到最终运用开放性方式到高利害关系评估，一个过渡战略在计划当中以确保循序渐进的引入，并为这个发展过程存在反复性做好准备。这个过程针对于与文化背景相关的不同反应态度。在提供职业发展，科技建设发展的同时，要确保其符合教学的需求，带有稳定性，行之有效，具有合法性。最重要的是确保这些发展的说服力，打赢“ 心理战” 。

以下图一路线图总结这个未来项目，以

图一路线图高利害评估的未来

| | 当今 | 2015 年-2020 年 | 2020 年-2025 年 | 2015 年以后 |
|-----------|-----------------|--|--|---|
| 高利害关系考试模式 | 纸质 | 取代纸质 - 学生可以选择不用手写，而用打字输入（用 USB 闪存盘保存自带电脑的信息） 一些取代纸质考试的形式开始出现。 | 后纸质考试普及所有的问题和材料一律数字化， 考试学生需要人手一台计算机参加评估的各个项目。 | 完全计算机化，因特网是考生使用一系列的软件按到输入设施。 |
| 网络联络 | 无 | 严格受限制的网络，后纸质考试中采取限定网络内问题回答 | 在有选择的资源里许可脱机和在线混合方式考试 你 登入链接 | 开放性因特网链接，所有的登入活动被记录下来，包括所有信息沟通，时间，资源利用。 |
| 评估的真实性 | 考题书面陈述式，有单色的示意图 | 带有彩色的示意图，配有视频，更接近现实场景的陈述 | 高真实性，数据引导仿真题目 | 实时链接，获取全球数据库信息 |

| | 当今 | 2015 年-2020 年 | 2020 年-2025 年 | 2015 年以后 |
|-------------|--------------------------|---------------------------------|--|------------------------------|
| 考生身份确认 | 考官进行考生本人与身份证件照片对比 | 继续练习，通过手动设备并于区域数据库链接 | 继续练习，电脑自配相机间断地拍摄记录键盘使用者 | 继续练习，采用双重身份确定，指纹同脸部确认功能。 |
| 提供，允许带入考场材料 | 学生可以带进教室限定范围的书籍，电子计算器，文具 | 数字化工具开始代替现有纸质材料比如 PDFs 便携式文档格式。 | 电子书籍，高清照片，影像，虚拟？，包括所有软件工具（开放性资源）。 | 增加多样性，有科目针对性的软件工具。 |
| 评估工作流程 | | 延续现有方式，加入数字化回答可以复制，存档和发送邮件 | 数字化回答，扩展到使用科目软件创建数据文档计算机化工作流程，存储，标注所有答案。 | 除了数字化回答文档，每个学生表现衡量数据，沟通交流记录。 |
| 成效衡量 | 评价解决问题方案的质量，书写式程序 | 方法相同，对选择答复项目进行分析 | 方法相同，但分析更加具体例如，测量回答每个问题所需时间，可以赢得分数 | 更进一步分析敲击键盘、平面触摸频率 - 尤其 |
| 连续不断的评估改进程序 | 与去年同期曲线形比较控制总体考试的难度 | 一些数据还可以提供对个别试题的总体简易难度做出评判。 | 个别试题的辨别力和可靠性也会被评分。 | 试题评分会考虑到学生在评估中采取的所有的互动内容。 |

在此要申明一点的是，计算机化和网上考试已经在市场上应用，但是这种考试模式在高等教育中的应用还是微乎其微。原因是主要是它们受限于教学的能力。我们在此建议的是一个电子考试的“平台”，不仅仅是应用软件或网络平台。我们主张将“计算机环境”下所需的二十一世纪整套的工具提供给每个学生，最开始采用无网络连接，或有限链接，但是我们展望一个可以全面登录，可以控制的，却完全开放的，因特网链接的考试。

这个设想并不是触不可及。塔斯梅尼亚学历管理机构已经对 11 到 12 年级（相当于中国高二到高三），考入高等教育的测试中采取大行动，替代纸质考试，采用全面公开，链接因特网的计算机化考试。（TQA，2013）评分也是用平板电脑 iPad 操作的。国际上，芬兰正在着手实行意向名为“Digabi”的项目，目标在 2019 年大学入学考试电子化。

结论

澳大利亚处在决定高利害评估的关键时刻。院校不能为仅仅每半年举办一次的考试而投资大量的计算机设备。所以，自带仪器便成为这个文化转变的关键点，同时这个提议也带来对有效性和可靠性的一系列挑战。但是，考虑到成果，努力将会得到回报。学术性的奖励更可信，而且采用更强有力的软件工具对课程改革带来更多的机会。更多信息请参考国家“考试变革”（2015）项目。

参考文献

- ABS (2013) Young Adults: Now and Then, 4102.0 - Australian Social Trends, April 2013, Australian Bureau of Statistics
<http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/4102.0Main+Features40April+2013>
- ABS (2014) Non-school qualification at Bachelor Degree level or above, persons aged 20-64 years, 6227.0 - Education and Work, Australia, May 2014, Australian Bureau of Statistics
<http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/6227.0May%20201>
- Oxford University Press, New York
- Binkley, M., Erstad, O., Hermna, J., Raizen, S., Ripley, M., Miller-Ricci, M., & Rumble, M. (2012). Defining Twenty-First Century Skills. In Griffin, P., Care, E., & McGaw, B. Assessment and Teaching of 21st Century Skills, Dordrecht, Springer. <http://www.atc21s.org/>
- Crisp, G. (2009). Towards Authentic e-Assessment Tasks. In G. Siemens & C. Fulford (Eds.), Proceedings of EdMedia: World Conference on Educational Media and Technology 2009 (pp. 1585-1590). Association for the Advancement of Computing in Education (AACE). <http://www.editlib.org/p/31689/>
- Fluck, A., Pullen, D., & Harper, C. (2009). Case Study of a Computer Based Examination System. Australasian Journal of Educational Technology, 25(4), 509–523. <http://ajet.org.au/index.php/AJET/article/view/1126>
- Herrington, J., Reeves, T. C. & Oliver, R. (2010). A guide to authentic e-learning. New York: Routledge
- Hillier, M & Fluck, A. (2014) Transforming Exams: Processes and Platform for e-Exams in Supervised BYOD Environments, Australian Government Office for Learning and Teaching. <http://www.olt.gov.au/resource-transforming-exams-byod-environments>
- Hillier, M., & Fluck, A. (2013). Arguing again for e-exams in high stakes examinations. In H. Carter, M. Gosper, & J. Hedberg (Eds.), Electric Dreams (pp. 385–396). Macquarie University. Retrieved from <http://www.ascilite.org.au/conferences/sydney13/program/papers/Hillier.pdf>
- Linstone, H.A. (1999) Decision making for technology executives, Artech House, Boston
- Rittel, H. & Webber, M. (1973) Dilemmas in a General Theory of Planning, Policy Sciences, 4, pp. 155-169.
- Tasmanian Qualifications Authority (2013) Board Meetings. <http://www.tqa.tas.gov.au/10270> & <http://www.tqa.tas.gov.au/23727>
- Universities Australia (2015) Universities Australia Pre-Budget Submission 2015–16. Canberra, ACT.
- Von Zansen, A. (2015) Digabi Newsletter II/2015. http://translate.google.com.au/translate?hl=en&sl=sv&u=https://digabi.fi/%3Fflang%3Dsv_se&prev=search
- Transforming Exams (2015). 'e-Exam System' project, <http://transformingexams.com>

Formal Citation in published proceedings:

Hillier, M. & Fluck, A. (2015). A pedagogical end game for exams: a look 10 years into the future of high stakes assessment. In T. Reiners, B.R. von Konsky, D. Gibson, V. Chang, L. Irving, & K. Clarke (Eds.), Globally connected, digitally enabled. Proceedings Australasian Society for Computers in Learning in Tertiary Education (ascilite), Perth, Australia, 29 Nov – 2 Dec (pp. 463-470). <http://www.2015conference.ascilite.org/wp-content/uploads/2015/11/ascilite-2015-proceedings.pdf>

注释：所有出版论文都经查阅并经过双盲重审程序。

Project information 项目信息: <http://transformingexams.com>

Contact 联系电子邮件(用英文): [mathew.hillier\[at\]gmail.com](mailto:mathew.hillier[at]gmail.com) <http://mathewhillier.com>

Original English version 原始英文版:

http://transformingexams.com/files/hillier_fluck_2015_exam_futures.pdf



CC BY-NC-ND / 署名-非商业性使用-禁止演绎

署名 — 您必须给出适当的署名，提供指向本许可协议的链接，同时标明是否（对原始作品）作了修改。您可以用任何合理的方式来署名，但是不得以任何方式暗示许可人为您或您的使用背书。

非商业性使用 — 您不得将本作品用于商业目的。

禁止演绎 — 如果您再混合、转换、或者基于该作品创作，您不可以分发修改作品。