

结构化知识评价工具——概念图

赵国庆 北京师范大学知识工程研究中心

guoqingzh@163.com

摘要：本文是概念图作为结构化知识评价工具的综述性文章。主要介绍概念图作为评价工具的适用范围、评价框架、记分方法以及评价的信度和效度分析。

关键词：概念图 结构化知识 评价 信度 效度

一、绪论

信息加工心理学家把广义的知识分为两大类：陈述性知识和程序性知识。陈述性知识用于回答“是什么”的问题，程序性知识用于回答“怎么做”的问题^[1]。概念构图^[3]（Concept Mapping）是康乃尔大学的诺瓦克（J.D. Novak）博士根据奥苏贝尔(David P. Ausubel)的有意义学习理论提出的一种教学技术。概念图（Concept Map）通过节点表示概念，连线表示关系，概念-连接词-概念的三元组组成一个命题，并将最具包容力、最宽泛的概念置于顶端，最一般、最具体的概念置于底端，从而形成一个清晰的知识网络结构。将概念图作为学习者陈述性知识（尤其是结构化知识）的外化表征方式，可以在一定程度上直接反映学习者对某个知识领域的理解，找出误解。因此，概念图可以作为结构化知识评价工具，有着其特定的优势。

目前，将概念图作为评价工具的研究已经非常之多，本文将对这一领域的研究进行综述，介绍概念图作为评价工具的适用范围、评价框架、记分方法以及评价的信度和效度分析等。

二、概念图评价方法适用的范围

概念图既可以在形成性评价中使用，也可以在总结性评价中使用。在形成性评价中，学生被要求在学习过程的不同阶段制作概念图，教师使用这些图，既可以评估学习者的理解，也可以根据其修改课程方案。总结性评价可以在一个教学单元结束时实施，用来测定学生对该单元的理解，从而评定等级。

三、基于概念图的评价框架

Ruiz-Primo & Shavelson（1997）提出概念图评价需要具备三个要素：

- (a) 为了让学生提供他们在某个领域的知识结构的证据而要求学生参加的任务（a task）；
- (b) 学生的反应形式(response format)；
- (c) 准确、持续评价学生做的概念图的记分系统（scoring system）。缺少这三个要素中的任何一个，概念图都不能被称为评价工具^[4]。

构图任务在任务需求（Task demands）、任务限制（Task Constraints）、内容结构(Content structure)方面都有很大的变化。在任务需求方面，学生可能被要求去对一个不完整的概念图填空，也可能是自己抽去概念，然后构图，还有可能是去摆放写有概念的卡片、写小论文或面谈。任务限制方面，学生可能被要求去构建层次化的图形或者不是，可能允许两个节点之间使用多个连接或者不允许、要求集体构图或者个人构图等等。

学生的反应形式可能是使用纸和笔，也可能是使用计算机完成；构图者可能是学生，也可能是教师（根据学生的反应）。

记分系统可能是以图的成分结构作为评分依据，也可能是和专家图对照进行评分，还可能是这两种方法的结合，这个将在下一部分专门介绍。

John R. McClure, Brian Sonak & Hoi K. Suen（1999）也提出概念图评价由两个部分组成：概念图构图任务和概念图评估。概念图构图任务是指让学生用概念图表征他们的知识的一系列过程，概念图的评估则包括对概念图内容和结构的检查，可能是量上的计算，也可能是质上的观察^[5]。

四、概念图评分方法

1. 成分结构评分法

Novak & Gowin（1984）在《学习如何学习（Learning How to Learn）》中提出了以成分和结构为依据的概念图的评价方法。Novak & Gowin给正确的命题、阶层、分支、交叉连接和特例指定了评分的权重，每个正确的命题、阶层、分支、交叉连接和特例

的得分分别为1分、5分、10分和1分。在这里，阶层的数量揭示了概念包容的程度，分支的多少则表明了构图过程的差异，交叉连接的多少表明知识整合的程度。这种评分的技术虽然很耗时，但却能够提供构图者知识结构的大量信息^[3]。

Novak & Gowin提出的这种评分方法得到了学者们的进一步研究。Mintzes和他的同伴(Pearsall, Skipper & Mintzes, 1997)采用同样的成分结构对概念图进行评分，只不过评分的权重有所调整。还有一些学者(Luckie, 2001, NSF proposal)在努力寻求概念图成分结构评分法的自动化实现^[6]。

2. 相似度评价法

Ruiz-Primo & Shavelson (1996)描述了一种将学生概念图与专家概念图进行比较的方法。专家图可以由教师或是由领域专家创建，还可以是由一组教师或领域专家共同创建。比较的过程需要定义好，范围可以从命题比较到整体结构的比较。计算机处理技术可以用来简化这一比较过程，CRESST和其他一些地方的研究者们已经证实了这种可行性(Chung, Herl, Klein, O'Neil & Schachter, 1997; Herl, O'Neil, Chung, Dennis & Lee, 1997; O'Neil & Klein, 1997)。这些自动记分系统都是典型的以命题匹配为基础的，仅仅限于概念集和连接词集。整体结构比较难以自动化，因此往往需要借助人工评判^[6]。

3. 综合评定法

一些研究者们尝试综合运用基于成分的和基于专家图匹配的两种方法。一个这方面的例子是在使用传统的基于成分的记分的时候，通过设定命题的更高的权重，整合与标准图的比较，而这些命题是专家们认为最为关键的。Rye & Rubba (2002) 报告了一个基于成分的概念图记分系统，但使用专家图来评定学生图中的命题^[6]。

概念图的记分方法方面有两大问题。一是如Novak & Gowin的方法的传统方法不管是在评定概念图成分的正确性和重要性方面，还是在构建标准图方面，都非常耗时，并且需要专家的输入。另一个问题是概念图作为评价工具的心理测量特性(在第五部分详述)^[6]。

第一个问题目前提到的至少有两种解决办法：发展计算机化的评分方法和简化的评分技术。(e.g., Chung, Herl, Klein, O'Neil & Schachter, 1997; Herl et al., 1997; O'Neil et al., 1997; Luckie, 2001)。Shaka & Bitner (1996)提供了一个简化的概念图记分模型，该模型以Novak & Gowin的记分模式作为起点，提供重要图形特征的简化分析方法。在他们的研究中，包括命题、分支、阶层、举例、交叉连接等图形特征被指定一个从0到4的等级，而不是用来记分或特征化。记分的这种简化可能是概念图评价应用中的典型。

Kinchin, Hay & Adams (2000)提出了另一种方法，用来分析概念图的整体结构，而不对概念、连接和命题去做详细分析^[6]。

五、概念图评价方法的信度和效度

概念图作为结构化知识的评价工具，与其他的评价方式相比表现出强大的优势。但要从心理测量学的角度去看待概念图作为评价工具的能力，就不可避免去分析这种评价方式的可靠性（信度）以及有效性（效度）。

1. 信度

信度是指测量的可靠性、一致性和稳定性程度^[2]。由于实际上知识在不断发生变化，一般不考虑概念图评分在时间上的一致性（测试-再测信度），而更多关注多人评分的一致性（相互信度，inter-rater reliability）和内在一致性信度^[6]。

John R. McClure, Brian Sonak & Hoi K. Suen (1999)认为概念图的评价中误差产生的来源包括以下几个方面：(a)学生在概念构图技能方面的差异；(b)评分员对领域知识理解上的差异；(c)评估的一致性。第三种误差来源很大程度上取决于所选的具体评价方式，前两个因素可以通过构图技能的训练和挑选对评估主题熟悉的评分员来使误差降低到最小。他们让两个评分员在六种不同的评分方式（整体评分，带有专家图的整体评分，关系评分，带有专家图的关系评分，结构评分，带有专家图的结构评分）下对学生提交的概念图进行评分，比较这六种评分方式下评分的一致性^[5]。

Liu & Hinchey (1996)发现Novak & Gowin(1984)的记分系统在不同的成分之间关联度相对较低，然而命题、层级、交叉连接和举例的记分可能正是结构化知识量度的不同方面。人们得出了与信度测量的各种各样的结论，包括认为它们是“合理的” (Shavelson & Ruiz-Primo, 2000; West et al., 2000)，如Shavelson & Ruiz-Primo, 2000; Herl et al.(1999)认为基于命题的记分优于基于结构

的记分，West et al(2000)认为有经验的评分者评价的信度较高^[6]。

2. 效度

效度是指测验的有效性或准确性^[2]。对概念图作为评价工具的效度的关注主要集中在同期效度和结构效度两个方面。

同期效度是指测验结果与同时或短期内组织的针对相同内容的其他测验结果的相关程度。

大量的实验表明：概念图作为评价工具时，同期效度会因为课程不同而有所变化，也会因为使用群体的知识水平的差异而不同^[6]。

结构效度是指测验结果与理论构想或特质的一致程度^[2]，在这里也是指概念图与其他有意义学习测量手段的一致程度。Novak,

Gowin & Johansen (1983) 认为概念图可以用来测量与传统评价手段不同的知识方面，所以与那些评价方法一致性不高不足为奇^[6]。

总体来说，测量的指标越细致，测量结果的信度和效度就越高。研究者们认为(e.g., West et al; 2000; Shavelson & Ruiz-Primo, 2000), 从心理测量学的角度来看，概念图作为评价工具，信度和效度已经达到可以接受的范围^[6]。

六、 概念图作为评价工具的争议

虽然建立专家图，并将其作为学生构建的概念图的评分依据是基于概念图的评价的重要方式，受到广泛的关注，但专家图是否存在以及是否能够作为评分依据，却也处于激烈的争论之中。

Robert Abrams在《An Overview of Concept Mapping》一文中指出：根据个人经验，对诺瓦克式概念构图的研究有两个学派，一个学派认为概念图对于测试学生对特定主题的理解很有用处，该学派认为构建“专家”图是可行的，通过比较学生的概念图与专家图在多大程度上相近，可以对学生的概念图进行评分；另一学派则认为概念图主要只对图的创建者是有用的，构图的过程为学生在特定主题上进行反思提供了结构化空间，这样一来，学生能够清楚的表达他对于主题的想法。Robert Abrams本人就倾向于后一种学派^[7]。

七、 结论

概念图作为结构化知识评价工具，与传统评价方法相比有着特殊的优势，有着巨大的潜力。构图者的前期训练、评分员对评价领域的掌握程度、评价框架的设计以及评分方法的选择都会影响评价的效果。专家图作为评分依据也依然存在着一些争议，是否采用需要根据具体的情况做出选择。

参考文献

- [1] 邵瑞珍, 皮连生, 吴庆麟.教育心理学[M].上海: 上海教育出版社, 1999, (6): 58-63
- [2] 金焱, 王刚.教育评价与测量[M].北京: 教育科学出版社, 2002, (12): 125-129
- [3] Novak, J. D. & Gowin, D. B.Learning how to learn[M]. New York: Cambridge University Press.1984
- [4] Ruiz-Primo, M.A., Schultz, S.E., & Shavelson, R.J.. Concept map-based assessment in science: Two exploratory studies. CSE Technical Report 436, National Center for Research on Evaluation, Standards, and Student Testing, Center for the Study of Evaluation, Graduate school of Education & Information Studies, UCLA, Los Angeles, CA. 90024-6511.1997
- [5] John R. McClure, Brian Sonak, Hoi K. Suen. Concept Map Assessment of Classroom Learning: Reliability, Validity, and Logistical Practicality[J].Journal Of Research In Science Teaching,1999, Vol. 36, No. 4: 475-492
- [6] J. W. Coffey, M. J. Carnot, P. J. Feltovich, J. Feltovich, R. R. Hoffman, A. J. Cañas, J. D. Novak. A Summary of Literature Pertaining to the Use of Concept Mapping Techniques and Technologies for Education and Performance Support[Z]. <http://www.ihmc.us/users/acanas/Publications/ConceptMapLitReview/IHMC%20Literature%20Review%20on%20Concept%20Mapping.pdf> . 2003.

