

· 教育管理 ·

## 基于证据推理方法的学习效果评估

吕国栋<sup>1</sup>, 曾祥娟<sup>2</sup>, 苏冬<sup>1</sup>, 修保新<sup>1</sup>

(国防科学技术大学 1. 信息系统与管理学院; 2. 人文与社会科学学院, 湖南 长沙 410073)

**[摘要]** 学习效果评估被认为是不确定环境下的多属性决策分析问题。本文首次将档案袋运用于英语学术论文阅读课程中, 并首次通过证据推理方法对学习者的档案袋信息进行评价。研究对象随机选取参与本课程的十名学习者, 通过证据推理的方法对其学习效果进行实证研究。研究表明, 课程档案袋可以有效记录学习者英语学术论文阅读课程中的表现, 并且基于证据推理的学习效果评估方法在实际教学实践中是可行的。

**[关键词]** 学习效果评估; 档案袋; 证据推理

**[中图分类号]** G642.1 **[文献标识码]** A **[文章编号]** 1672-8874(2014)01-0044-07

### Learning Effect Assessment Based on the Evidential Reasoning Approach

LV Guo-dong<sup>1</sup>, ZENG Xiang-juan<sup>2</sup>, SU Dong<sup>1</sup>, Xiu Bao-xin<sup>1</sup>

(1. College of Information System and Management; 2. College of Humanity and Social Science, National University of Defense Technology, Changsha 410073, China)

**Abstract:** Learning Effect Assessment (LEA) is a strategic issue considered as multiple criteria decision making with both quantitative and qualitative under uncertain conditions. In this paper, the analysis process and algorithm for LEA is proposed on the basis of evidential reasoning (ER) approach according to the information drawn from the course portfolio of each subject. Ten learners taking the English Academic Paper Reading course were selected as subjects of the study. The results suggest that the validity of course portfolio to record learners' English paper reading abilities and the practicability of the evidential reasoning approach used to assess the learning effect information drawn from the course portfolio.

**Key words:** learning effect assessment; portfolio; evidential reasoning approach

### 一、研究背景

学习者英语学术论文阅读能力对于其学术研究有着至关重要的作用, 在过去的英语学术论文阅读课程教学实践中, 授课教师面临着两个问题, 其一是英文学术论文的学习效果由哪几部分构成, 其二是课程结束后如何判断学习者的英语学术论文阅读能力是否有所提升, 即如何分析学习者的学习效果。

对于第一个问题, 学者认为英文学术论文能力

可以分为两个方面, 一方面是理解性阅读, 包括对长句、文章结构、学术背景等知识的理解; 另一方面是批判性阅读, 即分析论文的观点正确与否, 所提出的方法是否科学等问题。

基于对第一个问题的理解, 本文以《英语学术论文阅读》课程为背景展开了对第二个问题的研究, 研究者为本课程制定档案袋用于收集学习者学习效果信息, 并基于证据推理算法对所收集的信息进行分析和评估。在研究中, 研究者随机选取10名学习者作为研究对象, 结合访谈分析和观察

**[收稿日期]** 2013-09-16

**[作者简介]** 吕国栋(1991-), 男, 河南安阳人, 国防科学技术大学信息系统与管理学院硕士研究生。

研究,本文探索了以下三个问题:学习者参与本课程后的整体表现;证据推理方法在教学评估中应用的可行性;档案袋评估中的不足之处以及如何提升。

## 二、文献综述

### (一) 使用课程档案袋的原因分析

通常情况下,档案袋用于收集学习者多方面的学习信息并反映学习者的学习过程,在实际教学中,档案袋被认为一个非常有效的学习工具<sup>[1-3]</sup>。简而言之,档案袋主要用于发展和评估<sup>[4-7]</sup>,此外它还被用于提供各种不同的学习目的<sup>[8-9]</sup>,不同形式的档案袋所收集的信息和提供的功能也不相同<sup>[8,10]</sup>。不仅如此,学习档案袋典型的应用在于展示在一段时期内学习者所学到的信息并提供个体学习过程<sup>[11]</sup>。

档案袋评价在中国高等英语教学中已有广泛研究<sup>[12-15]</sup>,但大多集中在基础教育或者高等教育的基础层次展开,较少研究涉及到研究生这个层次。此外,实证研究多集中在写作教学,在为数不多的阅读教学研究中,也是集中在大学英语课外阅读、一般性大学英语阅读,而忽略了英语学术论文章阅读的研究<sup>[12,16-17]</sup>。

由于档案袋评价重视过程,这一点与英语学术论文章阅读循环往复的特点相契合,可以反映学习者阅读过程中的问题和收获,达到提高学习者学术论文章阅读能力的目的。同时,用档案袋来记录论文章阅读过程,展示阅读效果,有助于下一步学习者用英语撰写论文章。因此,在课程中通过档案袋收集学习者信息是科学合理的。

### (二) 档案袋信息评价方法选取

然而,基于档案袋收集到的信息,如何对其进行有效的评价是本文考虑的另外一个问题。在以往关于档案袋的研究中,很少有学者对档案袋评估方法进行探讨,并且如何评价档案袋至今仍是一个困惑<sup>[3,18]</sup>。

档案袋可以用于总结性评价和学习的实证研究<sup>[3,19]</sup>,对于档案袋的评估者而言,需要知道档案袋的评估指标和标准,然而这些评估指标经常是相互冲突的,这就给评估的效果带来一定影响。Dysthe 发现教育者并不能很自信地基于已有的实证信息对学习者进行评估<sup>[20]</sup>。此外,评价标准也是评估中的一大难处,准备的评估需要评估人员制定详细的评价标准,因此需要建立一个严格的数学评

估模型,来对档案袋所收集到的信息进行评价。

基于上述分析,本文旨在建立一种基于多种多样的学习者学习效果信息的评估模型,着重考虑两个问题:首先学习者学习效果的评估涉及方方面面,包括像成绩一样定量的信息,也包括学习者上课表现等方面的定性的信息。而一个综合性的评价需要综合所有可行的信息,这就需要将定性的信息和定量的信息转换为统一的格式并通过一定的方法进行处理;其次,学习者学习效果评估存在不完整性和不确定性。学习者的表现需要根据专家给出评价,也就是教师,但由于专家知识的局限性以及学习者个体表现的复杂性,这就很难完整而精准的评价信息。学习者学习效果评估不确定的环境下,需要处理不完整、模糊或者可能被忽略的信息。因此,需要找到合适的方法来分析不确定环境下学习者的学习效果。

而本文通过证据推理算法很好地处理了上述问题。证据推理方法是由 Yang、Singh、Xu 等学者在决策理论和 Dempster - Shafer (D - S) 证据理论基础之上设计的算法,用以处理诸如此类的问题。<sup>[21, 22]</sup>

## 三、研究设计

本研究基于该校秋季学期开设的研究生《英语学术论文章阅读与陈述》课程,学习者大多数为新入学的研究生以及博士生。本课程 70% 的时间用于学术论文章阅读教学,30% 时间用于论文章宣读教学。研究以该课程为依托,以 2012 级 10 名选修该课程的研究生为研究对象进行整体分析。

研究首先结合教学大纲和在本课程中的教学经验,通过本体知识建模的方法制定适用于本课程的档案袋,并通过 AHP 方法推算出档案袋中各部分在最后的评价中所占的权值。

在课程教学的实施过程中,教师根据学术论文章的结构 (IMRD) 展开讲授,讲完一部分内容,学习者要阅读自己的学术论文章并将阅读结果记录在自己的档案袋中。阅读部分课程讲授完后,开展为期四周的阅读陈述活动,学习者在课堂上汇报自己论文章的内容和自己的见解,由老师和其他学习者进行点评,此后,学习者根据课堂上习得的知识进行档案袋的修改,并由老师和学习者进行评价,如此修改四次,最后一次作为教师对学习者学习本课程后的学习效果的评价依据。教师根据设定的标准对 10 名学习者的最后的档案袋进行综合评价。

## 四、档案袋设计

### (一) 档案袋指标选取

课程档案袋包括四部分内容：第一部分包括个人的基本信息和所选取论文理由；第二部分包括对论文的理解性阅读和批判性阅读；第三部分包括同

伴互评，教师反馈后的修改报告和自我反思情况；第四部分是学习者期末成绩。档案袋的完成形式为电子文档，学习者可以根据反馈反复修改。

通过基于本体建模的方法，研究给出评价学习者档案袋的各项指标及其子指标，如表1所示。

表1 档案袋的评估指标及其权重

学习者表现	一级指标 (权重)	二级指标 (权重)	三级指标 (权重)	
学习效果评估	背景介绍 (0.0357)	论文选取缘由 (0.5)		
		领域知识介绍 (0.5)		
	内容分析 (0.6625)	理解性阅读 (0.4)	语言分析 (0.2)	
			语篇分析 (0.4)	
			信息提取 (0.4)	
		批判性阅读 (0.6)	语言表达 (0.2)	
			分析深度 (0.4)	
			分析广度 (0.4)	
	同伴协作 (0.1161)	同伴互评 (0.25)		
		教师反馈 (0.25)		
		档案袋修改 (0.25)		
		自我反思报告 (0.25)		
	期末成绩 (0.1857)			

### (二) 各指标权重确定

各指标权重通过专家经验以及 AHP 分析方法给出。例如，在一级指标中，按照重要性程度专家结合自己的经验给出两两相比的重要性程度，例如内

容分析与背景介绍相比，内容分析的重要程度是背景介绍的9倍，相对应地背景介绍的重要程度就是内容分析的1/9。由此可以得到如下等级列表，如表2所示。

表2 基于 AHP 方法的指标确定

A 与 B 的重要性比值		B			
		背景介绍	内容分析	同伴协作	期末成绩
A	背景介绍	1	1/9	1/5	1/7
	内容分析	9	1	7	6
	同伴协作	5	1/7	1	1/2
	期末成绩	7	1/6	2	1

通过 AHP 分析法知，一致性比率为 0.0099，通过一致性检验，背景介绍、内容分析、小组协作、学习测试成绩的权值分别是：0.0357、0.6625、0.1161、0.1857。

通过此方法，可以得到各级子标题所占的权重，如表1所示。

## 五、基于证据推理的档案袋评估过程

学习效果评估的过程可以看作是决策分析的过程。通常而言，很难直接通过档案袋而直接给出对学习者的学习效果的整体评价，原因主要有以下四个方面：(1) 学习者的学习效果由很多种复杂的、多维度、多层级的指标组成；(2) 评价档案袋时，需要考虑多种指标，但是这些指标有时候是会

冲突的；(3) 评价是需要考虑多种多样的信息，包括定性的分析和定量的数据；(4) 信息的不确定性和模糊性，甚至部分数据难以给出，但这些信息又需要被评定。

基于证据推理算法的档案袋评价方法操作如下所示：

首先，要基于档案袋的特点，将各因素的指标分解成更小的指标，建立因素层级模型，如在第四部分提到的。

其次，要收集评价档案袋的定性和定量的信息，这些信息主要由任课老师给出，即本领域的专家。

然后，要将定性的与定量的信息通过信度结构

$$H = \{H_j, j = 1, 2, \dots, 5\} = \{\text{Worst}', \text{Poor}', \text{Average}', \text{Good}', \text{and Best}'\}.$$

表 1 中给出了所有的评价指标和评价，其中只有一项定量（期末成绩）的指标，其他的均为定性的指标。

为了分析学习者的期末成绩，研究通过专家经验给出如下等级信息：Worst'（小于等于 59），Poor'（60 和 69 之间），Average'（70 和 79 之间），Good'（80 和 89 之间），以及 Best'（90 和 100）。因此，对于各项指标，可以给出如下等级划分区间： $H_{TS} = \{h_1, h_2, h_3, h_4, h_5\} = \{59, 69, 79, 89, 100\}$ 。

如表 3 所示，学习者在考试中得到了 94.5 分，

$$A_i = \begin{matrix} & D & C & C+ & B & B+ & A & A+ \\ \begin{matrix} \text{Worst} \\ \text{Poor} \\ \text{Average} \\ \text{Good} \\ \text{Best} \end{matrix} & \begin{bmatrix} 1 & 0.6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.4 & 0.6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.4 & 1 & 0.4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.6 & 0.4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.6 & 1 \end{bmatrix} \end{matrix}$$

因此，如果学习者内容分析部分的表现  $H^{CA} = \{(H_n, \gamma_{n,i})\} = \{B-, 1\}$ ，本文可以得到  $b_i = A_i \times r_i = (0, 0.6, 0.4, 0, 0)$ ，进一步可以得到经过转化后的信息  $S^{CA} = \{(\text{Poor}', 0.6), (\text{Average}', 0.4)\}$ 。

模型转换为相同的格式<sup>[21,22]</sup>。

之后，通过证据推理的迭代算法将各个子指标信息逐步融合到其上层的指标中。<sup>[21,23]</sup>

最后，再次通过迭代算法，将所有信息融合到最上层的信息中，即对于档案袋的整体评价。

为了能够清晰地说明证据推理在本研究中的具体应用过程，本文给出对某一名学习者评价的详细过程。

### (一) 建立评估等级并转换原始信息

基于档案袋评价的学习者学习效果评价结果可以分为以下五个等级 'Worst', 'Poor', 'Average', 'Good', and 'Best', 因此五个等级可以表示为：

介于 H4 和 H5 之间，因此，通过信度结构<sup>[23]</sup>可以计算出，该信息可以表示为  $\{\beta_4 = 0.5, \beta_5 = 0.5\}$ 。因此，我们可以得到  $S^{TS} = \{(H_4, 0.5), (H_5, 0.5)\}$ 。

对于定性的信息，如学习者对于文本内容的分析情况，本文给出如下评价等级： $\{D, C, C+, B, B+, A, A+\}$ ，而其上级的评价等级为  $\{\text{Worst}', \text{Poor}', \text{Average}', \text{Good}', \text{and Best}'\}$ 。因此根据专家经验，可以给出如下转移矩阵：

### (二) 收集学习者 1 的档案袋信息

基于学习者在课程中学习表现，研究根据档案袋中所设定的指标，依此得到学习者的学习信息，如表 3 所示。

表3 学习者1的学习表现

指标一级(权重)	二级指标(权重)	三级指标(权重)	学习者学习效果信息
背景介绍 (0.0357)	论文选取缘由(0.5)		{(A, 1)}
	领域知识介绍(0.5)		{(A, 0.6), (A+, 0.4)}
内容分析 (0.6625)	理解性阅读(0.4)	语言分析(0.2)	{(B+, 0.5), (A, 0.5)}
		语篇分析(0.4)	{(B+, 0.3), (A, 0.5)} (不完全信息)
		信息提取(0.4)	{(B+, 0.4), (A, 0.6)}
	批判性阅读(0.6)	语言表达(0.2)	{(A, 1)}
		分析深度(0.4)	{(A, 0.5), (A+, 0.3)} (不完全信息)
		分析广度(0.4)	{(A, 0.5), (A+, 0.3)} (不完全信息)
同伴协作 (0.1161)	同伴互评(0.25)		{(A, 1)}
	教师反馈(0.25)		{(A, 0.5), (A+, 0.5)}
	档案袋修改(0.25)		{(A, 0.9)} (不完全信息)
	自我反思报告(0.25)		{(A, 0.5), (A+, 0.5)}
期末成绩 (0.1857)			{(94.5, 1)}

### (三) 基于证据推理算法的聚合评估

通过证据推理的方法可以得到学习者在各个部分的表现情况,如表4所示。

表4 档案袋各个部分学习者的表现情况

档案袋各部分	评价等级(信度等级)								定量数据
	定性分析								
	D	C	C+	B	B+	A	A+	未知信息	
背景介绍	0	0	0	0	0	84.62%	15.38%	0	0
内容分析	0	0	0	0	13.34%	64.90%	10.83%	10.93%	0
同伴协作	0	0	0	0	0	77.33%	20.81%	1.86%	0
期末测试	0	0	0	0	0	0	0	0	94.5

从图1中,可以得到对学习者的整体学习效果的评价信息,其结果可以表示为{(Worst, 0), (Poor,

0,) , (‘Average’, 0.0372), (‘Good’, 0.3568), (‘Best’, 0.5291)}, 并且包括 0.0769 的因为信息缺失而到来的不确定信息。

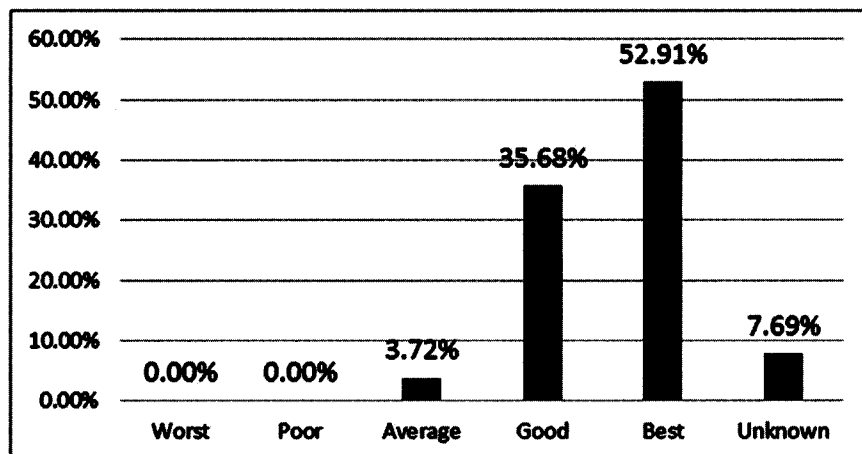


图1 基于证据推理算法的学习者1的学习效果评价信息

## 六、学习者整体情况分析

根据上述的学习效果评估过程，本文对研究所选取的 10 名研究对象的学习效果逐一进行分析，得到最后的评估结果，如表 5 所示。进一步，本文基于效用函数和排序规则<sup>[23]</sup>，对这 10 名研究对象的学习效果进行排序，得到如下结果： $S_4 > S_6 > S_9 > S_2 > S_1 > S_{10} > S_5 > S_3 > S_8 > S_7$ 。

表5 各学习者的评估结果

研究对象	期望效用			排序
	最小值	平均值	最大值	
S1	0.7393	0.0064	0.0064	5
S2	0.8153	0.0385	0.0385	4
S3	0.6593	0.0422	0.0422	8
S4	0.9406	0.001	0.001	1
S5	0.6624	0	0	7
S6	0.8841	0.0249	0.0249	2
S7	0.4226	0.0143	0.0143	10
S8	0.6112	0.0042	0.0042	9
S9	0.8366	0.0257	0.0257	3
S10	0.6821	0.0154	0.0154	6
Group	0.7232	0.0207	0.0207	

## 七、结论与讨论

学习效果评估在课堂教学中可以认为是一个战略性的决策分析过程，对于教学设计和学生发展都有着极其重要的影响。考虑到教师难以准确地把握学习者整个学习过程的所有信息，因此本文将学习效果评估过程看成是不确定环境下多属性决策过程。为了收集学习者的学习效果信息，在教学中，研究采用档案袋的教学辅助资源，并提出通过证据推理算法的策略来评价从档案袋中所收集到的定性、定量甚至不完全的信息。

本文基于中国某著名高校研究生《英文学术论阅读》课程，随机选取 10 名学习者作为研究对象，此外通过观察研究，访谈的研究成果来辅助证明通过证据推理算法所分析的学生学习效果信息的科学性和可行性。研究认为，在研究生《英文学术论阅读》课程中使用档案袋，学生整体表现较好，并且通过反复的修改档案袋，可以有效促进学生对于文章的理解，提升学生的阅读水平。具体表现在：

(1) 档案袋能够有效记录学习者阅读课程中的表现，以提供综合性评价的依据；

(2) 档案袋能够帮助改变完全接受式的英语论阅读习惯，提高了对论文进行批判性解读的能力；

(3) 档案袋部分地促进了同学之间与教师之间的交流。

此外，本文为今后的教学工作和学生评估提供了科学的方法，通过证据推理理论，对档案袋信息

评价的手段给予了技术上的支持,也为社会科学类研究分析提供了新的参考思路。但是,研究中尚存在一定的问题,例如档案袋指标难以确定,尽管通过AHP法进行分析,但是仍存在一定的不足,但是通过本次研究,得到了初始的数据,可以通过神经网络学习的方法,进一步确定合理的权值,研究将在下学期的课程中进行修正。

### [参考文献]

- [1] Wade, R. C. and D. B. Yarbrough (1996). "Portfolios: A tool for reflective thinking in teacher education?" *Teaching and teacher education* 12(1): 63 - 79.
- [2] Shulman, L. (1998). "Teacher portfolios: A theoretical activity."
- [3] Smith, K. and H. Tillema (2003). "Clarifying different types of portfolio use." *Assessment & Evaluation in Higher Education* 28(6): 625 - 648.
- [4] Centra, J. A. (1993). *Reflective Faculty Evaluation: Enhancing Teaching and Determining Faculty Effectiveness*. The Jossey - Bass Higher and Adult Education Series, ERIC.
- [5] Doolittle, P. (1994). *Teacher portfolio assessment*, ERIC Clearinghouse on Assessment and Evaluation.
- [6] Tillema, H. H. and K. Smith (2000). "Learning from Portfolios: Differential Use of Feedback in Portfolio Construction." *Studies in educational evaluation* 26(3): 193 - 210.
- [7] Snadden, D. (1999). "Portfolios - attempting to measure the unmeasurable?" *Medical Education* 33(7): 478 - 479.
- [8] Wolf, K. and M. Dietz (1998). "Teaching portfolios: Purposes and possibilities." *Teacher Education Quarterly* 25(1): 9 - 22.
- [9] Zeichner, K. and S. Wray (2001). "The teaching portfolio in US teacher education programs: What we know and what we need to know." *Teaching and teacher education* 17(5): 613 - 621.
- [10] Forster, M. and G. Masters (1996). "Portfolios. Assessment Resource Kit (ARK)."
- [11] Tanner, R., et al. (2000). "Piloting portfolios: Using portfolios in pre - service teacher education." *ELT Journal* 54(1): 20 - 30.
- [12] 黄剑茹. 以阅读档案促进课外阅读的探究[J]. *英语教师*, 2008(04).
- [13] 张琳. 档案袋评价在大学英语教学中的运用[J]. *山西财经大学学报*, 2008(S2).
- [14] 刘长伟. 电子档案袋评价在英语交际教学中的应用[J]. *宿州学院学报*, 2009(01).
- [15] 邓兆红,何世潮. 档案袋评价与大学英语教学[J]. *阜阳师范学院学报(社会科学版)*, 2007(03).
- [16] 徐锦芬,李红,李斑斑. 大学生英语阅读能力自我评价的实证研究[J]. *解放军外国语学院学报*, 2010(05).
- [17] 洪民,詹先君,赵景梅. 复合型档案袋评价对外语阅读能力的影响[J]. *外语界*, 2011(02).
- [18] Burroughs, R. (2001). "Composing standards and composing teachers the problem of national board certification." *Journal of teacher education* 52(3): 223 - 232.
- [19] Snyder, J., et al. (1998). "The inherent tensions in the multiple uses of portfolios in teacher education." *Teacher Education Quarterly* 25: 45 - 60.
- [20] Dysthe, O. (2002). Professors as mediators of academic text cultures: An interview study with advisors and masters degree students in three disciplines in a Norwegian university. *Written Communication*, 19(2), 485 - 536.
- [21] Yang, J. - B. and M. G. Singh (1994). "An evidential reasoning approach for multiple - attribute decision making with uncertainty." *Systems, Man and Cybernetics, IEEE Transactions on* 24(1): 1 - 18.
- [22] Yang, J. - B. and D. - L. Xu (2002). "On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty." *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on* 32(3): 289 - 304.
- [23] Yang, J. - B. (2001). "Rule and utility based evidential reasoning approach for multiattribute decision analysis under uncertainties." *European Journal of Operational Research* 131(1): 31 - 61.

(责任编辑:陈勇)