

作业社会化批阅工具的设计、开发与评估

王琦^{1,2} 余胜泉^{1,2}

(1. 北京师范大学 教育学部 北京 100875;

2. 北京师范大学 未来教育高精尖创新中心 北京 100875)

[摘要] 传统教学中作业批阅以全批全改和精批细改为主,造成了教师批阅压力大、作业反馈不即时、评价结果缺乏针对性、不利于开展过程性评价问题。自评、同伴互评一定程度能减轻教师批阅压力,但质量不能保证。本研究针对上述问题,提出了作业社会化批阅方法,将作业批阅权限公开化,充分利用社会化批阅专家、家长、学生的力量和社会化批阅系统中内置的批阅可信度保障模型,在保证作业批阅结果准确性基础上,降低教师的批阅压力、促进作业反馈的即时性、个性化,进而推动作业过程性评价和学生学习效率提升。在社会化批阅理念指导下本研究访谈了教师作业批阅需求,并基于学习元平台设计和开发了作业的社会化批阅系统并选取了PPT的使用课程进行实验,结果表明,作业进行访谈社会化批阅系统具有较好的实用性和易用性,其批阅结果也与教师批阅结果存在较大一致性,从而验证了作业社会化批阅方法具备实践价值。

[关键词] 社会化批阅; 作业评价; 过程性评价; 学习元

[中图分类号] G434 **[文献标识码]** A **[文章编号]** 1007-2179(2017)03-0096-09

一、前言

作业是教育教学领域的重要环节,起到了评价学生学习过程和为教师教学提供参照的作用,而作业减负和作业有效批阅、评价问题一直是教育领域关注的重点问题,但由于缺乏对学生学习过程的个性化评价方法,很多教师只能通过大量作业考查学生的知识掌握情况,给学生带来作业压力,也使教师的批阅压力陡增。教师的作业批阅方式以全批全改和精批细改为主,试想每位教师每天批阅100份作业,每份作业耗时2分钟,那么教师用于作业批阅的时间就达到3-4小时(杨建国,2008)。这种情况下教师就不能很好地对学生的作业进行过程性评价和个性化反馈,学生也不能即时地获取作业的批阅结果并进行学习巩固。久而久之,恶性循环,衍生出学生作业和教师批阅压

力大、作业反馈不即时、过程性信息得不到有效记录等问题。

基于网络学习平台的作业自动化批阅和评价能够降低教师的批阅压力,为学生提供即时的作业反馈。但这种方式主要针对选择、判断等客观题,主观题因缺少语料库和有效的内容分析算法,很难取得好的批阅效果。一些专家提出通过自评和同伴互评解决上述问题,并在Coursera和Edx等MOOC平台上进行实践,但这一方面给学生增加了批阅工作量,同时也很难保证批阅结果的权威性。

近年来,以社交网络服务(social network services,简称SNS)为代表的社会化媒体不仅构造了新的社会形态和交往方式,也创造了一种社会化生产模式(程罡等,2009)。该模式极大地提高了生产效率,越来越得到行业的认可。教育作为社会生活的重要方

[收稿日期] 2017-03-13 **[修回日期]** 2017-04-20 **[DOI编码]** 10.13966/j.cnki.kfjyyj.2017.03.011

[作者简介] 王琦,在读博士生,北京师范大学教育学部,研究方向:计算机教育应用(wangqi.20080906@163.com);余胜泉,教授,博士生导师,北京师范大学教育学部,研究方向:移动学习、泛在学习。

面 在社会化生产思路的指导下,充分利用互联网社交网络服务的优势,采用作业的社会化批阅有利于充分调动教育相关者(学科专家、教师、家长、学生和希望为教育发展做贡献的社会化批阅者)的积极性,使作业批阅成为开放的社会活动,从而减轻教师的批阅压力,促进作业的即时反馈以及学生作业过程性信息的有效记录,也能减轻学生的作业压力。

二、作业批阅解决方案评述

针对前言中提到的作业批阅中存在的诸多问题,研究者也探索并提出了相应的解决方案。

(一) 作业批阅方案

1. 网络学习平台的自动批阅

基于网络学习平台的自动批阅包括对客观题的批阅和主观题的批阅(张际平等,1997)。对客观题批阅主要依赖系统题库以及题目、答案和详解的关联关系完成并给出即时反馈。由于填空题存在同义词问题,对其批阅除利用全文匹配、关键词匹配外(史娟,2010),还需要建立同义词林,从而更加准确地批阅和反馈。

主观题批阅的典型方法是文本相似度计算,基本过程是利用预设的分词词典对学习者的答案和标准答案进行分词,然后通过词频统计,利用向量空间模型计算两个文本的相似度(刘晓华等,2006;张均胜等,2014)。如自动文本标注(Automated Text Marker)平台引入分词词典和同义词词典,实现了对答案的语法和语义分析,在自动批阅领域取得了一定效果(曹洋等,2014)。又如托福的作文自动评分系统(e-rater)对文章进行分词和词频分析,实践检验显示该方式在一定程度上可以较好地判断学生的用词能力,但文本的语义信息还较难把握(Yigal et al., 2014; 王颖等,2012)。

2. 学生自评

自评是教师为了提高作业批阅效率经常采用的作业评价方式。自评中教师设定作业批阅规范,并给出标准答案作为参照,学生据此独立完成作业批阅(黄杰锋,2011)。该方式可以让学生参与作业批阅过程,促进对知识的反思和反馈。当前主流的MOOC平台如Coursera、Edx都引入了自评。然而,自评虽然提供了标准答案作为参照,但由于学生能力的差异以及标准答案难以覆盖所有解题方法,其

质量和公正性也难以保证。

3. 同伴互评

同伴互评是课堂教学和MOOC教学中一种重要的批阅方式。与自评相似,该方式也需要教师预设批阅标准和规范,一方面避免自评中可能出现的不公正性,另一方面让学生通过作业的批阅加强自身的反思(阮高峰等,2006)。MOOC平台的同伴互评主要通过评价校准和多个体重复评价保障评价的可靠性,即一方面让学生正式批阅前参与校准测试,校准合格后方可通过批阅准入;另一方面让多个同伴评价同一份作业来保证作业批阅结果的一致性。研究表明,同伴互评的可靠性较高,与教师批阅的结果没有显著差异(孙力等,2014),减轻了教师的批阅压力并促进了作业的即时反馈和过程性评价开展(郑志高等,2014)。

上述方式在一定程度上促进了作业批阅问题的解决,但也存在不足。比如,由于技术的限制和不同学科完整语义词典的缺乏,自动批阅功能存在不准确性,难以在教育教学中广泛应用;又如,自评一方面对学生的公正性有要求,另一方面学生不同层次的批阅能力限制了其质量;同伴互评虽然通过评价校准和多人交叉评价建立了较好的质量保证体系,但增加了学生的压力,同时也难以保证每个学生评价的公正性。因此,作业批阅在借助计算机的即时反馈和过程记录功能的基础上,仍然需要建立更完善的流程保证作业的批阅质量。

(二) 作业批阅新思路——作业社会化批阅

“作业社会化批阅”是在信息技术和学习系统理论快速发展的背景下,为解决作业批阅问题探索的新型作业批阅模式。“社会化”即以在线作业平台为支撑,以完善的批阅认证体系、批阅质量保障体系和批阅激励机制为保障,将作业批阅的权利网络化、公开化,使作业批阅、评价的主体从教师群体扩展为在线学科专家、家长、学生甚至有作业批阅兴趣的社会成员,实现作业批阅权利从教师到全社会的转变。

1. 作业社会化批阅模式

作业社会化批阅模式以网络学习平台为支撑,进行作业管理服务、过程性评价信息采集服务和数据分析服务,实现对作业社会化批阅流程的支持,包括:作业布置、完成和上传流程;基于云平台的作业

信息管理和存储流程; 作业社会化批阅流程; 作业评价结果的分析和报告反馈流程(见图1)。

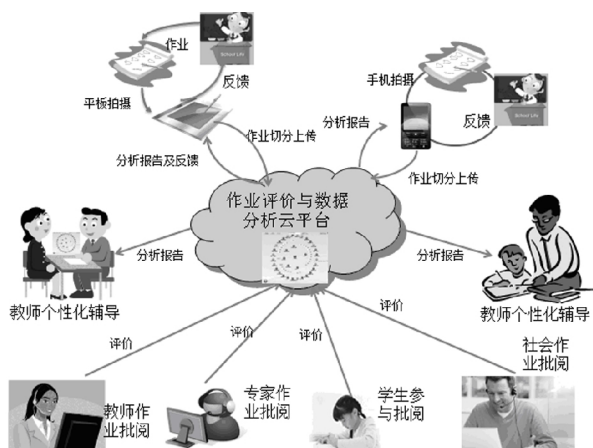


图1 作业社会化批阅模式

2. 作业社会化批阅的特点

1) 提供多样化的作业类别及多元的上传形式。作业类别包括选择、填空、问答题和探究式活动,方便教师多角度、过程性地考查学生的学习。学生可以线下把作业拍照上传,也可在线上完成作业。

2) 提供基于知识点的题库和对作业过程性信息的采集。学生的作业与知识点相互关联,方便通过题目作答情况了解知识点掌握程度。

3) 提供开放的批阅权限、严格的批阅准入条件和基于用户信任度的批阅质量保证机制。

4) 提供作业评价信息的在线记录和即时反馈。学生作业的错误信息、评语等在批阅中会即时提交到作业评价数据库,为学习者提供即时的评价反馈。

5) 提供学生作业情况的定期报告。如全方位的个人反馈和整体对照,可对学习者的学习情况进行个性化诊断。

3. 作业社会化批阅的教育价值

从实践角度看,社会化批阅旨在解决作业批阅存在的问题,可以促进减负政策的落地。同时,教育领域引入“社会化”的生产方式有利于让学习转变为更广泛的群建共享活动和社会交互活动,促进泛在学习的发展和社会化学习共同体的构建,同时也能减轻教师的工作压力,促进学习的过程性评价和即时反馈。

从理论的角度看,联通主义学习理论、社会化学习理论、网络化合作活动学习理论都为作业社会化

批阅实践提供了支持,而作业的社会化批阅同样发展了这些理论。联通主义强调学习是将不同的信息源和节点连接并形成交互网络的过程,学习者习得的知识不再是固化的,而是在交互网络中不断流通和主动建构得到的知识流(唐瑶等,2014)。作业社会化批阅将学科专家、教师、学生、家长连成评价网络,以学生知识评价为网络中的信息流,在网络节点与信息流的交互下促进学习者对知识的整体性理解、掌握和迁移。社会化学习理论重视交互决定论和观察学习,即学习者通过观察其他个体的行为并通过其强化结果反馈而习得新行为或对已有行为进行矫正(Bandura,1977)。网络化合作活动学习理论强调通过学习平台提供的学习活动建立师生交互、生生交互的合作关系,从而促进知识有意义建构和学习效果的有效提高(Burt,1993;陈思宇等,2013)。作业社会化批阅拓展了社会化学习,使学习者通过网络化交互批阅和反馈促进自身知识的掌握。另一方面,作业社会化批阅也拓展了网络化合作活动学习理论的应用空间和社交网络,借助网络学习平台和社会化的用户群,作业批阅从学生和教师的双向交互延伸为基于社交网络的社会化学习活动,这样更有利于学习的发展和社会化学习共同体的建立,从而提高作业批阅的权威性。

三、基于学习元平台的作业社会化批阅工具的设计和实现

为了探索更有效的作业批阅形式,本文以北京师范大学学习元平台为支撑,通过对当前作业批阅方式的分析和对已有平台中作业批阅功能存在问题的调研,设计开发了作业社会化批阅工具。

学习元是北京师范大学余胜泉教授提出的新型资源组织形式。它是学习资源表征的基本单位,支持正式学习、非正式学习以及混合式学习等学习场景,具有生成性、开放性、微型化、联通性、自组织和可进化性等特点(程昱等,2009;余胜泉等,2009;杨现民等,2010)。每个学习元相当于一节微课,其构成要素包括内容、活动和评价(王琦等,2013)。学习元平台对所有用户开放,所有学习者都可以参与感兴趣的学习元的学习,方便了社会化学习的开展;学习元是内容和活动的结合体,学生在活动中完成学习,内容和活动都以知识点为粒度,方便对学习者

进行基于知识点的过程性评价。

学习元平台为作业社会化批阅的实现提供了支撑,系统中多样化的成员为社会化批阅提供了必要的社会化关系网络。学习元中人与人、人与资源、资源与资源之间的关联属性,为过程性评价和学习者的学习提供了条件,如学习者 A 参与资源 A 的学习并在活动中取得了较差的成绩,而平台中资源 A 和知识点 A 存在关联,那么系统就可以根据该关联推理出学习者 A 对知识点 A 掌握较差,同时通过学习者 A 对资源 A 的学习行为又可以分析出相应原因。平台的消息提醒功能是社会化批阅中即时反馈的重要保障;系统中的资源关联和推荐也为学习者的个性化学习及减轻作业压力创造了条件。本研究在学习元平台基础上对作业社会化批阅工具进行需求分析和功能框架的设计。

1. 功能分析和需求调研

从功能角度看,作业社会化批阅系统为方便作业的分析应建立批阅者、学习者、作业和知识点之间的连接,同时向所有用户开放以保证社会性参与,此外要能过程性和及时评价学习者提交的多形式的作业。为使社会化批阅工具更好地适应真实的教学场景,研究者在批阅设计过程中对作业批阅过程的需求进行了调研,主要围绕以下几个方面展开:

- 1) 提供多元的作业形式、作业评价形式和涵盖多学科的题库。
- 2) 提供协同批阅者的设置功能,让教师将作业批阅任务分派给所有社会成员或指定用户,在社会化批阅者的协助下实现过程性评价和即时反馈。
- 3) 提供良好的作业批阅质量保障机制和激励机制的设置,既能让批阅结果使人信服,又能提升社会化批阅者的参与兴趣。
- 4) 提供多元化的作业提交形式。
- 5) 提供作业的可视化分析和即时反馈机制。

2. 功能架构和实现

研究者设计了作业社会化批阅的功能架构(见图 2)。

批阅工具主要包括作业布置模块、基于手机和平板的作业管理和上传模块、作业存储和管理模块、在线批阅和评价模块、评价报告生成模块。作业布置模块中,教师根据学生作业报告发布作业并设置批阅方案,创建作业、知识、学习者、批阅者之间的关

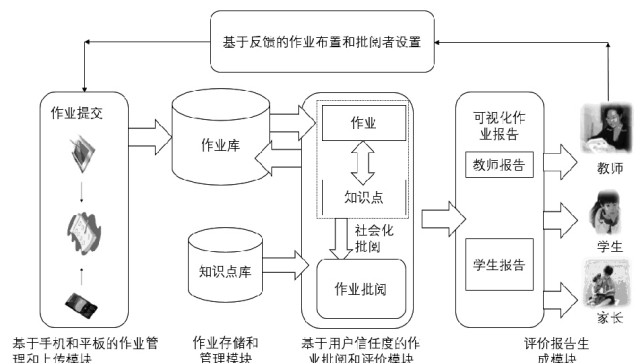


图 2 社会化批阅功能架构

联网络;基于手机和平板的作业管理和上传模块为学习者提供了泛在的作业形式;作业存储和管理模块提供了作业、题目、知识点、用户相关联的存储模式,为基于知识点的过程性评价奠定了基础。作业批阅和评价模块设置了作业批阅的基本规则、激励机制和自动化质量保障机制,同时实现了即时反馈和基于知识点的过程性评价。评价报告生成模块按照系统预设的评价模型将作业评价结果可视化并分发给学生、教师和家长。下面以实用案例的部分界面为参照进行介绍。

作业布置模块聚合了以活动题和主客观题为主的多元作业形式(见图 3)。布置的作业将提交到作业存储管理模块进行保存,同时教师需设置作业批阅方案,包括作业截止时间、社会化批阅者和批阅标准,社会化批阅者默认为所有通过批阅认证的用户,教师也可设置批阅权限,如将批阅者设置为特定专家;批阅标准包括每个题目要求的批阅者数,批阅容许误差。

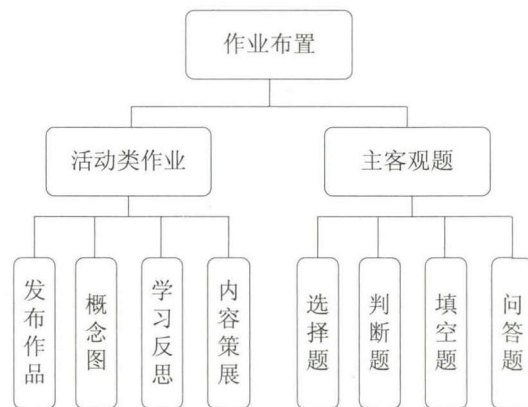


图 3 作业类型

基于手机和平板的作业管理和上传模块提供了泛在化、人性化的作业完成形式。与 MOOC 平台相比, 社会化批阅系统提供了相对灵活的作业完成和提交机制。针对操作简单的题目(选择、填空、简答) 学生可直接通过手机或 PC、手提电脑完成提交; 对于较复杂的问答题、作图题、数理类综合题, 学生可以在纸质练习册上完成后通过手机拍照上传。

作业存储和管理模块主要完成题目存储、学生作业详情存储和作业评价结果存储。题目存储即将作业布置模块提交的题目按照作业号、题号、知识点相关联的形式存入系统数据表, 一方面为基于知识点的过程性评价创造条件, 另一方面也满足教师自主创建题库的需求; 学生作业详情存储主要将学生提交的作业存入数据表并通过题号与作业题目建立关联, 方便学生作业的评价; 作业评价结果的存储实现了对批阅者评价结果的记录。

作业批阅和评价模块是社会化批阅工具的核心, 该模块实现了基于用户信任度的评价和质量保证。MOOC 系统中同伴互评和自评存在质量保障流程复杂、批阅者压力大等问题, 作业的社会化批阅工具针对该问题设置了作业批阅规则(见图 4) 和基于用户信任度的批阅质量保障体系(见图 5)。学习者提交的作业可以通过两个批阅流程实现批阅质量保证, 即教师和专家批阅的结果直接通过并存储到数据库, 社会化批阅者的批阅结果则需要通过基于用户信任度的质量保障体系计算批阅分数和误差。其计算公式如下所示:

$$uScore = \sum (Si * Trusti) / (\sum Trusti)$$

其中, $Trusti$ 为用户 i 的信任度, Si 为第 i 个批阅者的批阅分数。

$$ErrorA = \sum (Si - uScore) / n$$

其中, Si 为用户 i 批阅的分数, n 为批阅人数, $uScore$ 为作品实际得分。

如果误差小于教师设置的批阅误差, 则批阅结果入库存储, 否则请求教师和专家复议。其中用户的批阅结果基于用户的批阅信任度, 用户信任度越高, 批阅权限越高, 在结果中所占的权重越大。当用户批阅信任度超过系统设置的专家信任度下限(默认为 0.9) 时, 普通批阅者可升级为批阅专家。本系统中用户 i 的批阅信任度 $Trusti$ 通过用户 i 积分 ($score$)、用户 i 作品评价的概况 ($evaluate$) 以及用户

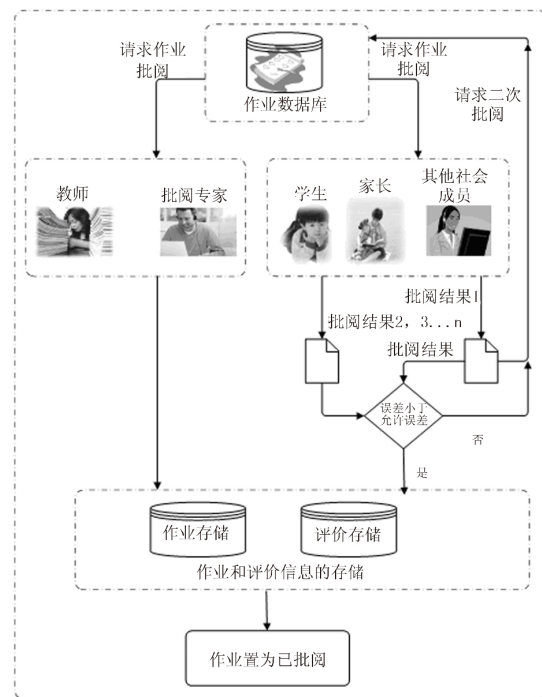


图 4 批阅规则

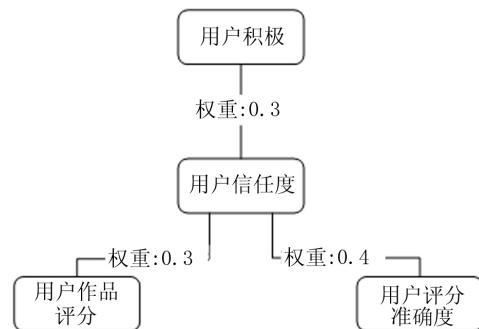


图 5 基于用户信任度的批阅质量保障体系

i 对他人作品评价的准确度 ($review$) 来衡量, 默认计算方案为:

$$Trusti = 0.3 * score + 0.3 * evaluate + 0.4 * review$$

其中, 用户 i 积分 ($score$) 为用户 i 系统积分与专家批阅者系统积分下限的比率, 如专家批阅者积分下限为 6000 分, 该用户系统积分为 $userScore$, 那么 $score$ 的值则为 $userScore / 6000$ (超过 1 则为 1)。 $evaluate$ 通过用户 i 作业的质量衡量, 如 $mean$ 为用户 i 完成作品的平均分, 而 100 分为专家作品的分数, 那么该用户 i 的 $evaluate$ 值为 $mean / 100$ 。 $review$ 值主要通过用户 i 日常评价行为的准确度来衡量, 如用户 i 共批阅过 n 份作品, 其中 m 份作品的评价

误差小于教师设定的临界误差,那么该用户的 review 值为 m/n 。社会化批阅通过上述批阅规则和基于用户批阅信任度的质量保障机制,实现了对学生作业批阅相对智能的质量保障,这也是对传统作业批阅和 MOOC 平台中互评、自评机制的改进。

评价报告的生成和输出模块为学生、家长和教师提供了即时的作业反馈报告。报告按照系统预设的评价模型为用户输出可视化作业评价结果(见表一)。

图 6-8 给出了本研究中班级作业完成详情、班级作业评价结果反馈报告和某学生作业反馈报告界面图。其中图 6 以列表方式展现了学生的作业完成情况,图 7 展示了班级每个学生作业题目的完成情况、批阅情况、平均分以及方差,图 6 和图 7 下方通过可视化方式对用户数据进行了展现,图 8 针对某学生的作业进行了可视化展现。查看详情为教师提供了作业更详细的信息。

所有学生的作业详情

学习者姓名	完成情况	最后更新时间	操作
丁国柱	已完成	2015-03-29 22:19:18.0	查看详情
xiaomao	未完成	暂未提交	查看详情
tmnsbf	已完成	2015-03-28 20:16:40.0	查看详情
smqh	已完成	2015-03-30 17:10:19.0	查看详情
langyao	未完成	暂未提交	查看详情
胡智杰	已完成	2015-03-31 22:25:07.0	查看详情
周志2012	已完成	2015-03-29 22:23:46.0	查看详情
汪丹	已完成	2015-03-29 16:18:31.0	查看详情
pansheng	已完成	2015-03-29 16:49:20.0	查看详情

总共: 9 当前: 1-10
 ● 已完成
 ● 未完成

图 6 班级作业完成统计

班级作业详情列表:

作业标题	完成情况	批阅情况	平均分	方差	操作
ppt操作	7/9	0/9	92.1428/100	38.12185	查看详情
提交你最满意的ppt	7/9	0/9	89.57143/100	25.102121	查看详情

总共: 2 当前: 1-10

图 7 班级作业评价统计

tmnsbf的作业详情:

题号	作业标题	完成状态	提交时间	得分	操作
1	ppt操作	已完成	2015-03-28 20:16:40.0	92.0	查看详情
2	提交你最满意的ppt	已完成	2015-03-28 19:28:47.0	88.0	查看详情

总共: 2 当前: 1-10

图 8 学生作业评价结果

3. 系统关键模块的技术路线

本批阅工具后台功能的实现主要通过 java + mysql 形式,前端展示采用 html + css 进行布局,同时为增强用户使用系统的体验,研究者采用了 javascript + ajax 技术,实现了即提交即更新即反馈。此外,在基于平板和手机的作业上传模块,批阅工具还通过 android 技术开发手机客户端实现拍照上传。作业批阅和评价两个模块涉及作业批阅结果,批阅工具通过百度 echarts 插件实现对评价数据的可视化。

四、系统应用分析

(一) 科技接受度分析

批阅工具开发完成并经过试用、改进后,研究者建立了一个批阅活动“PPT 的使用”,并邀请北京师范大学现代教育技术研究所成员进行了二次实验,通过改进的科技接受度问卷(5 点量表)进行该批阅工具的科技接受度分析,其中 1-6 题测试的是认知有用性,7-13 题测试的是认知易用性。笔者在试用中通过 SPSS20.0 对问卷的信效度进行了分析,认知有用性维度的 Cronbach's Alpha = 0.883,认知易用性的 Cronbach's Alpha = 0.825,说明问卷的信效度较好。

表一 评价反馈模型内容

对象	学生	教师	家长
即时评价报告	知识认知能力(总分、各个知识点得分、各个题目得分、得分率最低的题目和知识点)、个体差异比较(班级平均分、各分数段人数、排名、每个知识点平均分、每个题目平均分)	个体评价详情(与学生评价报告相同)、班级评价详情(作业均分和方差、每个知识点的均分和方差、每个题目的均分和方差、得分率较低的题目和知识点)、特殊群体详情	学生总分、排名、每个知识点的得分和排名
阶段性评价报告	成绩变化趋势(每次作业分数、排名)、知识认知情况(每个知识点得分、排名)、可寻求帮助的对象(针对得分较差的知识点提供)	个体成绩变化趋势(每次作业分数、排名)、班级成绩变化趋势、特殊群体成绩变化趋势	学生一个阶段的作业得分变化和排名变化

表二 科技接受度问卷调查结果

	样本数	平均值	方差
1. 这样的评价活动让学习评价形式更丰富	10	4.30	0.675
2. 这样的评价活动对评价学习更有帮助	10	4.20	0.422
3. 这样的评价活动提供的评价形式对我知识的巩固(学习反馈更即时和有针对性)更有帮助	10	4.10	0.568
4. 这样的评价形式有利于我(通过知识点的评价分析和自评、互评)更好地发现学习中知识点的不足	10	4.20	0.632
5. 这种方式能很大程度上减轻教师的批阅压力	10	4.40	0.699
6. 这样的评价活动比传统评价形式更有效(结果准确、权威)	10	4.10	0.568
7. 使用这种方式进行作业评价时系统操作并不困难	10	4.00	0.471
8. 我使用这样的评价方式并不需要花费过多的时间和精力	10	4.00	0.816
9. 本评价活动对我来说是清楚容易理解的	10	4.10	0.316
10. 我很快便了解了本评价活动的操作方式	10	4.00	0.943
11. 评价活动的操作对我来说没什么困难	10	4.20	0.422
12. 活动的操作界面(系统)很容易使用	10	3.80*	0.919
13. 总体而言,本系统比较容易使用	10	4.00	0.667

从表二可以看出,认知有用性六道题目旨在验证系统在解决前文所述作业批阅问题的效果,即能否有效减轻教师的批阅压力,促进作业的即时反馈和过程性评价,促进基于知识点的作业评价和学生知识点的有效巩固。结果显示,六道题目平均得分在4分以上,说明社会化批阅工具可以较好地解决上述问题,丰富了作业评价的形式,实现了社会化批阅的预期目标。在认知有用性的访谈中,用户普遍反映相对于传统的作业评价、MOOC的自评互评等,作业的社会化批阅具有以下优势:1)在减轻教师压力的同时,对作业的评价更即时,更具针对性,也更能揭示学习者知识点方面的不足;2)批阅工具设置了较严格的批阅质量保障机制,评价结果更权威;3)可视化评价结果有利于教师和学习者更直观地发现学习中存在的问题;4)以作业批阅和评价为基础,建立了以教师、学科专家、家长和学生为主要关系节点的关系网络和评价共同体,有利于充分利用社会力量促进学习者的评价和发展。认知易用性的七道题目旨在验证系统是否容易理解和使用,结果显示只

有第12题分数低于4分,说明系统总体运用便捷,不会增加用户使用的认知负荷。用户访谈显示,用户对第12题的回答是,界面链接层级跳转较多,容易产生困惑,需要对易用性进行完善。

(二) 系统应用效率度检验

作业的社会化批阅实验结束后,研究者通过采集实验过程中社会化批阅者评价的数据和教师评价的结果,进行了批阅工具的一致性效率度检验。本检验选取“PPT的使用”这一内容,要求学习者综合运用Powerpoint的功能完成PPT的制作并进行上传和批阅,系统设置的作业批阅允许误差为3分。实验过程共10名学生参与,9名学生有效完成学习任务并提交作业,研究者邀请教师、实验室研究生、校外人员以及参与作业过程的学习者对已完成的作业进行社会化批阅,并通过后台数据统计功能统计了社会化批阅的结果和教师批阅的成绩(见表三)。

研究者通过两种方式进行系统的一致性验证:一种通过SPSS20.0的Spearman检验进行分析,另一种基于批阅允许误差临界值对批阅差异值(教师

表三 教师批阅分数和社会化批阅均分比较

作业编号	1	2	3	4	5	6	7	8	9
社会化批阅均分	75	83	80	82	94	76	77	72	83
教师批阅分数	75	81	80	88	95	76	85	76	88

表四 Spearman 相关检验结果

		教师批阅	批阅系统
Spearman's rho	教师批阅	相关一致性	1.000
		显著性(双尾)	0.857**
		N	9
	批阅系统	相关一致性	0.857**
		显著性(双尾)	0.003
		N	9

批阅分数和社会化批阅均分的差) 进行一致性比较(如果差异值超过临界误差,那么该作业批阅结果不一致)。Spearman 相关检验结果显示,相关一致性值为 0.857,显著性为 0.003(见表四)。这说明,教师批阅结果和社会化批阅均分存在一致性相关。从差异值来看,9 份作业中,5 份作业差值小于临界误差 3 分,一致程度为 56%,其他 4 份作业的差异值在 5 分以内,说明社会化批阅工具在一定程度上能保证批阅的一致性。从上述两种方法的分析结果看,作业社会化批阅工具的批阅一致性效度较好,可以在减轻教师压力的基础上保障批阅的权威性,为促进学习者的作业评价和系统的大规模应用提供了契机。

五、总结与展望

本文利用基于学习元平台的作业社会化批阅工具对作业的社会化批阅形式进行了探索,结果表明,相对于传统作业批阅形式和 MOOC 中的互评、自评,作业的社会化批阅通过引入社会化批阅群体降低了教师的批阅压力、促进了作业的即时反馈和过程性评价;同时社会化批阅工具通过严格的批阅质量保证体系,使作业批阅结果保持了较高的可信度;最后系统通过与知识点关联的评价体系为学习者、教师、家长提供了基于学习过程的可视化作业反馈,让学习者能够清晰地认识自己的学习情况,促进学习者的自我反馈和自我提高。

本研究也存在不足:一方面社会化批阅工具的易用性仍需提高;另一方面,尽管通过相对社会化的使用群体进行了工具试用,并通过系统的质量保障机制保证批阅结果的准确性,但参与人数较少,且一些批阅者由于缺少批阅数据导致批阅信任度较低,造成批阅信任度较高的教师在批阅结果中所占比重相对较大,因此在真实场景中进行更大规模的验证

实验是下一阶段研究的重点。此外,在真实教育情境中,如何通过激励机制促使社会成员积极参与社会化批阅过程仍值得探究。最后,如何充分利用社会化批阅结果进行相应的学习推荐,从而为教师和学生减负也是未来探索的重要方面。

[参考文献]

- [1] Bandura, A. (1977). Social learning theory [J]. *Journal of Organizational Behavior Management*, 3(4): 55-63.
- [2] Burt, R. S. (1993). *Structural holes: The social structure of competition* [M]. Harvard University Press.
- [3] 曹洋, 成颖, 裴雷(2014). 基于机器学习的自动文摘研究综述[J]. *图书情报工作*, 58(18): 122-130.
- [4] 陈思宇, 黄甫全, 曾文婕(2013). 新兴网络化合作活动学习的三大类型及其启示[J]. *中国电化教育*, (7): 51-56.
- [5] 程罡, 余胜泉, 杨现民(2009). “学习元”运行环境的设计与实现[J]. *开放教育研究*, 15(2): 27-36.
- [6] 黄杰锋(2011). 从“行动”到“状态”的跨越——高三学生“自批自评”作业模式的有效性研究[J]. *现代中小学教育*, (3): 68-69.
- [7] 刘晓华, 刘建舟(2006). 智能化网络作业系统的设计与实现[J]. *现代远程教育*, (3): 54-56.
- [8] 阮高峰, 林叶郁(2006). 同侪辅助学习模式及网络互助学习群体构建实践[J]. *中国电化教育*, (11): 34-37.
- [9] 史娟(2010). 考试系统中填空题自动评分的设计与实现[J]. *计算机应用与软件*, 27(9): 197-199.
- [10] 孙力, 钟斯陶(2014). MOOC 评价系统中同伴互评概率模型研究[J]. *开放教育研究*, (5): 83-90.
- [11] 唐瑶, 余胜泉, 杨现民(2014). 基于学习元平台的教育内容策展工具设计与实现[J]. *中国电化教育*, (4): 87-93.
- [12] 王琦, 余胜泉(2013). 从学习元看微课评价的设计[J]. *中国教育网络*, (10): 22-25.
- [13] 王颖, 李振阳(2012). 国外二语写作中电子反馈模式的研究评述[J]. *外语电化教学*, (4): 11-16.
- [14] 杨建国(2008). 数学作业批改模式初探[J]. *课程教材教学研究: 中教研究*, (4): 51-52.
- [15] 杨现民, 余胜泉(2010). 泛在学习环境下的学习资源信息模型构建[J]. *中国电化教育*, (9): 72-78.

[16] Yigal, A., & Jill, B. (2006). Automated essay scoring with e-rater ®; v. 2.0[J]. The Journal of Technology, Learning and Assessment, 4(2): 18-21.

[17] 余胜泉 杨现民 程罡(2009). 泛在学习环境中的学习资源设计与共享——“学习元”的理念与结构[J]. 开放教育研究, 15(1): 47-53.

[18] 张际平 张琴珠(1997). 计算机与教育[M]. 电子工业出版社.

[19] 张均胜 石崇德 徐红姣 高影繁 何彦青(2014). 一种基于短文本相似度计算的主观题自动阅卷方法[J]. 图书情报工作, (19): 31-38.

[20] 郑志高 张立国 张春荣(2014). Xmooc 的学习评价方法调查研究[J]. 中国电化教育, (11): 44-49.

(编辑: 李学书)

Design, Development, and Evaluation of Homework Social Review Tool

WANG Qi^{1 2} & YU Shengquan^{1 2}

(1. Faculty of Education, Beijing Normal University, Beijing 100875, China)

2. Beijing Advanced Innovation Center For Future Education, Beijing 100875, China)

Abstract: Traditionally, it is always the teacher who conducts homework review carefully one after another. In some aspects, this condition leads to teachers' high pressure. Moreover, due to the high quantity of homework for teachers' to review, reasonable feedbacks get delayed while the quality of them becomes lower. As a result, the evaluation does little to benefit learning and the development of instruction. Self-evaluation and peer review can help improve the condition, but they not guarantee the quality of the evaluation while increasing students' stress. Social review makes homework review an open, social activity for all social members including teachers, experts, parents and even students with the help of the auto quality guarantee mechanism created by the social review system. All social members can participate in the process of homework review if they are certificated in the system. And with the help of social reviewers, teachers become more relaxed and the feedback becomes more timely, individual and process oriented. This will finally benefit homework evaluation and students' learning. In this study, the researchers investigated teachers' needs and developed a social review system based on Lcell platform and conducted an experiment to verify the result of the system. The result demonstrates that the social review system provides practical value and is authoritative. It can be helpful in educational practice.

Key words: social review; homework evaluation; process-oriented evaluation; Learning Cell platform