

我国教育硕士专业学位论文 质量评价指标体系的构建

朱晓民,张 啾,王祎琪

(山西师范大学,山西 临汾 041004)

摘要: 构建评价指标体系旨在促进教育硕士专业学位论文质量标准形成,推动教育硕士专业学位论文朝着高质量方向发展。首先,以文献样本分析为依据获取评价指标;其次,借助德尔菲法对指标体系进行修正;最后,运用层次分析法确定了以“研究水平”“写作水平”为基本维度(包括2个一级指标、5个二级指标和30个三级指标)的教育硕士专业学位论文评价指标体系,以期为科学测量教育硕士专业学位论文的水平提供量化标准。

关键词: 教育硕士;专业学位论文;质量评价;指标体系

中图分类号: G643

文献标志码: A

文章编号: 1003-2614(2020)11-0084-06

一、引言

教育硕士专业学位论文是综合考查教育硕士在基础理论、专业知识和专业技能等方面学习和研究结果的重要依据,对教育硕士专业学位论文质量进行评价是确保教育硕士培养质量的重要手段。鉴于实际需求,1999年我国教育硕士专业学位指导委员会秘书处出台了《关于教育硕士专业学位论文标准的规定》,但内容较为宽泛,且缺少量化的评价指标,使得评价的客观性和准确性无法保证^[1]。2019年5月,全国教育专业学位研究生教育指导委员会颁布了《教育硕士专业学位论文基本要求》(以下简称《基本要求》)^[2],该文件不再提及论文标准,而是对选题、文献综述等方面提出了具体要求。评价指标体系及标准的缺失已成为制约我国教育硕士专业学位论文质量与培养质量提升的重要因素。本研究基于教育硕士专业学位论文应用性和实践性特点,在文献内容分析的基础上,运用德尔菲法及层次分析法建立教育硕士专业学位论文质量评价指标体系,以期为教育硕士专业学位论文质量评价提供依据与基本框架,促进学位论文质量标准形成。

二、教育硕士专业学位论文质量评价指标筛选与初建

(一) 教育硕士专业学位论文质量评价指标的预选

评价指标的遴选与层次归类是构建指标体系的前提和

基础^[3]。确定评价指标不仅要依据有关政策规定,突出学科特色,还要参考和借鉴已取得的成果。因此,本指标体系结合现实需要,以教育硕士专业学位特点为基础,以《基本要求》为依据进行构建。

1. 基于教育硕士专业学位特点的分析

首先,评价要遵循“实践取向”^[4]。教育硕士是国家为了培养高素质的专业化人才而设置的专业性学位,主要以培养具有职业背景的应用型中小学教师和管理者为目的,所以,实践性是其最本质的特性。《基本要求》明确说明论文选题应遵循实践性原则,紧密联系基础教育(中等职业技术教育)实际^[5]。因此,在教育硕士专业学位论文质量评价指标中凸显“实践取向”,既是对国家政策文件要求的响应与落实,也是对教育硕士专业学位本质属性的强调。其次,评价要做到“德术并重”^[6]。近年来,学术不端事件易发多发,2019年3月,《教育部办公厅关于进一步规范和加强研究生培养管理的通知》强调,“各学校要加强学术规范和学术道德教育……狠抓学位论文和学位授予管理”^[7]。因此,教育硕士专业学位论文评价指标作为衡量学术成果的标尺,应既重“术”又重“德”,做到“德术并重”。

2. 基于《基本要求》与40份文献样本的内容分析

《基本要求》和相关文献资料是指标内容来源的重要依据。40份文献资料包括25篇“教育硕士论文评价”相关文章,5份“社科类项目申报书”(国家社科基金项目申请书、全国教育科学规划课题申报书、教育人文社科项目申报书等),

收稿日期:2020-08-12

基金项目:2020年度山西省研究生教育创新计划支持课题“我国教育硕士学位论文质量评价体系建构研究”(编号:2020SY319)的研究成果。
作者简介:朱晓民,山西师范大学教育科学研究院教授,教育学博士,主要从事课程与教学、教师教育研究;张 啾,山西师范大学教师教育学院,研究方向:课程与教学评价;王祎琪,山西师范大学教师教育学院,研究方向:课程与教学评价。

10份“高校教育硕士专业学位论文评阅书”。以质性分析软件 Nvivo. 11 为主要工具,运用编码统计的方法对资料进行内容分析,以短语或句子为最小分析单元进行逐级编码、逐渐抽象,最终形成类目系统,如表1所示。

表1 样本内容分析结果

一级主题词	二级主题词	三级主题词
研究水平(215)	论文选题(57)	选题源于实践问题(12);选题具有理论基础(10);选题的创新性(20);选题的可行性(15)
	文献综述(35)	紧密结合研究主题(5);对相关研究分析概括(6);核心概念的界定(15);对过往研究进行评价(9)
	研究实施(123)	假设合理(8);数据真实有效(6);资料可靠(2);研究思路清晰(20);研究方法具体可行(8);研究对象具有代表性(22);研究设计科学合理(8);研究工具的适应性(8);数据分析的准确性(6);结果的可靠性(25);结论的概括性(6);讨论的科学性(4)
写作水平(111)	写作规范(62)	写作格式的规范性(40);参考文献引用规范(17);学风/学术道德规范(5)
	写作表达(49)	题目具体明确(10);英文表达水平(3);摘要的全面性(5);图表文呈现的合理性(8);概括与表述能力(15);学术语言使用(8)

注:括号内的数值代表该主题出现的总频次。

结合样本内容分析,拟将研究水平和写作水平作为一级指标。研究水平下包括三个二级指标,其中,研究实施水平是对论文研究工具与方法、研究设计与思路、研究结果和结论等多个方面的考量。写作水平下的二级指标包括写作规范水平与写作表达水平。写作规范水平主要体现为论文基本格式和学术道德的符合规范;写作表达水平是判断作者是否能准确、条理地运用学术语言进行书面表达。在确定以上指标的基础上,邀请相关领域的专家进行会谈,对二三级指标进一步细化与完善,最终确定了包括选题的实践性、参考文献格式的规范性等32个三级指标。

(二)教育硕士专业学位论文质量评价指标体系的两轮专家咨询

1. 预设评价指标体系的删选

评价指标体系的删选主要运用德尔菲法,即向与研究主题相关的领域专家发放匿名咨询函,在收回咨询函后将专家意见进行整理与分析,直至得到一致的意见为止。

参与德尔菲法的专家必须是对相关研究领域有深入、全面的认识与了解,并有主动参与的积极性与意愿。基于此,主要按照以下标准选择专家:一是国内对教育硕士论文评价及其相关问题有着清晰认识的专家;二是“研究生教育与评价”相关领域的专家;三是有5年以上指导教育硕士经验的导师(职称为副教授及以上)。参照以上标准,加之对不同批次开设教育硕士学位点的高校以及教育硕士招生学科领域的考虑,最终确定了中国教育科学研究院等研究机构或高校的业内专家,专家的研究方向涉及教育管理、学科教学(各个方向)等领域。为了保证指标设计的完善性、科学性和权威性,两轮咨询涉及的专家共有106人,咨询专家工作单位分布见表2。

表2 咨询专家工作单位分布表

设置硕士学位点批次	工作单位	人数
——	研究生教育指导委员会	1
——	中国教育科学研究院	2
第一批	北京师范大学、东北师范大学、西北师范大学、哈尔滨师范大学、陕西师范大学、华东师范大学、山东师范大学	32
第二批	首都师范大学、山西师范大学、河北师范大学、上海师范大学、云南师范大学	28
第三批	杭州师范大学、江苏师范大学、浙江大学、宁夏大学	15
第四批	重庆师范大学、山西大学	5
第五批	河北大学、湖南大学	5
第六批	淮北师范大学、闽南师范大学、青岛大学	9
第七批	中南大学、中央民族大学	4
第九批	湖北师范学院、太原师范学院、江西科技师范大学	5

注:第八批7所教育硕士培养单位为“服务国家特殊需求人才培养项目”,故未作为样本学校选取咨询专家。

咨询函主要采用问卷的形式,问卷课题组成员经过反复修改讨论,并邀请专家对问卷的表述进行审阅,最终确定了《教育硕士专业学位论文质量评价指标重要程度征询意见问卷》(以下简称《征询意见问卷》)。问卷的结构主要参考运用专家咨询法的硕博论文中所附的问卷形式,问卷共设置三部分内容:第一部分为卷首语,简要介绍本次研究的背景、目的、意义及打分标准;第二部分为指标重要度打分表:打分表仿照Likert七点量表设计,要求对各级指标分别进行打分,打分表后附有意见栏,以便专家提出修改意见,目的是为了依

据专家对每个指标重要程度的判断结果对指标进行筛选;第三部分包括职称、学科领域、工作单位、指导研究生年限等专家的基本信息和专家对指标的熟悉程度以及打分时的判断依据(包括实践经验、理论分析、直观感受、对国内外相关研究的了解等),目的是为了对专家意见的可靠性进行评判。

第一轮共发放41份问卷,回收40份有效问卷。指标的遴选综合考虑专家打分、修改意见和因子分析结果等三部分内容。首先,对回收问卷进行数据分析,计算出均值、满分率、标准差、变异系数,并结合专家意见的可靠性确定出筛选标准:剔除专家权威系数 ≥ 0.8 、重要性打分均值 < 5.25 (75%的删选率)的指标。经课题组成员讨论,删除“选题的理论性”(5.20)、“文献综述的创新性”(5.10)这两项指标。“附录的必要性”均值为5.18(< 5.25)变异系数为0.23,经讨论后认为,教育硕士论文多为应用型,附录呈现的内容较为重要,考虑到是由于指标表述不清导致专家评分较低,故将指标改为“附录的合理性”。其次,结合专家意见栏内容对指标进行修改,将“问题的提出”改为“选题的新颖性”,将反映学术道德的主要指标确定为“参考文献引用的规范性”和“重复率符合规定”。最后,结合因子分析结果对指标的结构进行修改,对“写作规范水平”与“写作表达水平”下三级指标的结构进行调整,使得“写作规范水平”体现基本格式的规范和学术道德的规范,其下包括题目格式的规范性等7个三级指标;“写作表达水平”体现论文各部分及整体的组织、概括与表达能力等,其下包括题目的明确性等8个三级指标。

2. 正式评价指标体系的确定

基于第一轮问卷的数据分析,在对预设指标进行删选与调整后形成第二轮《征询意见问卷》,第二轮问卷主要有三个方面的变化:一是指标数由第一轮的32个调整为30个;二是专家数量和范围扩大,在第一轮的基础上增加到106人,涵盖的学科和单位更广;三是标示个别项的含义并对指标设置的意图作简要说明。第二轮咨询共发放问卷106份,其中,有效问卷104份,咨询问卷数据处理结果显示:各项指标的均值最小为5.92,单项赋值满分率最高值约达72.12%,标准差最大值为1.11,变异系数均小于0.2(取值范围为0.097-0.187),通过分析问卷统计数据以及专家意见反馈信息,可以判定专家意见趋向一致、协调程度高,说明专家对指标体系较为认可,故不再进行第三轮咨询。结合专家“指标表述尽量保持一致”的意见,第二轮问卷结果分析后仅对“文献研究水平”下三级指标的表述进行修改,将“清晰条理陈述文献”改为“文献引述的明晰性”,将“分析成就与不足”改为“文献评论的客观性”,最后邀请5名专家对形成的教育硕士学位论文质量评价指标体系进行审阅,最终形成了包含2个一级指标、5个二级指标、30个三级指标的教育硕士学位论文质量评价指标体系,见表8。

三、教育硕士学位论文质量评价指标体系的检验

(一) 专家意见可靠性分析

两轮专家咨询成员主要为评价领域的专家和教育硕士导师,且涉及多所高校,在一定程度上说明了专家意见的可靠性。此外,还可以通过以下四个方面来分析:

1. 专家积极性

第一轮发放41份问卷(回收率和有效率均为97.6%),第二轮发放106份问卷(回收率99.1%,有效率98.1%),说明专家对研究的配合度较高。

2. 专家意见的集中程度

第一轮M值为5.69,满分频率K值为25.92%;第二轮M值为6.40,K值为55.8%。第二轮M值接近满分,K值超过50%,由此可见,第二轮专家意见的集中性较高。

3. 专家意见的变异系数

第一轮中变异系数 C_v 介于0.143-0.287之间,因此对“附录的必要性”“研究综述的创新性”等指标作出修改。第二轮中 C_v 介于0.097-0.187之间,小于0.2,表明第二轮咨询专家之间的一致性更高。

4. 专家的权威程度

经计算,两次咨询的专家权威系数 C_r 分别为0.819、0.862,均大于0.8,表明专家对指标的判断和取舍主要建立在理论与自身实践的基础上^[8]。专家的可靠性较高。

(二) 评价指标体系信度分析

内部一致性及折半信度是检验量表可靠性的常用指标。如表3所示,总量表与分量表的克隆巴赫Alpha均大于0.8,说明本研究所编制的研究工具《征询意见问卷》的内在一致性较高。借助SPSS24.0软件将量表随机分成两部分(不等长),两部分之间的相关性较强,相关性为0.885。进一步计算各部分内在一致性,同样较好,说明我国教育硕士学位论文质量评价指标体系的信度较高。

表3 总量表可靠性统计

	克隆巴赫 Alpha	基于标准化项的克隆巴赫 Alpha	项数
总量表	0.971	0.972	30
分量表	0.929	0.932	15
	0.960	0.963	15

(三) 评价指标体系效度分析

效度(Validity)指测量的有效性,主要包括内容效度和结构效度^[9]。内容效度指数(content validity index,CVI)是能够反映内容代表性的量化指标,包括条目水平的I-CVI和量表水平的S-CVI^[10]。经统计计算,两轮专家咨询I-CVI分别处于0.885-1.000、0.914-1.000之间,全体一致S-CVI分别为0.895、0.925,平均S-CVI分别为0.875、0.931。

时,通过一致性检验^[17]。以“研究水平”下二级指标的C值为例(见表7),一致性比例为0.0059,小于0.1。采用同样步骤检验其他各级指标,对不满足条件的数据进行修正,使得

一致性比例均小于0.1,最终由15位专家的判断矩阵加权算术平均得出最终的权重结果,见表8。

表7 集结后的判断矩阵——研究水平

论文选题水平	文献研究水平	研究实施水平	Wi	一致性比例	对“教育硕士学位论文质量评价体系”的权重
1	2.1667	0.8333	0.3833	0.0059	0.7500
0.4615	1	0.4861	0.1913		
1.2	2.0571	1	0.4254		

表8 教育硕士学位论文质量评价指标权重

一级指标	权重	二级指标	权重	三级指标	权重
研究水平 A	0.75	论文选题水平 A1	0.3833	选题的新颖性 A11	0.1240
				选题的可行性 A12	0.1103
				选题的实践性 A13	0.0532
		文献研究水平 A2	0.1913	核心概念的界定 A21	0.0641
				文献引述的明晰性 A22	0.0379
				文献评论的客观性 A23	0.0415
		研究实施水平 A3	0.4254	问题提出的明确性 A31	0.0220
				研究思路(计划)的严密性 A32	0.0253
				研究样本的代表性 A33	0.0202
				研究方法的契合性 A34	0.0383
				研究工具的科学性 A35	0.0422
				研究范式的合理性 A36	0.0308
				研究结果的可靠性 A37	0.0524
				讨论与反思的深度 A38	0.0463
对策与建议的实践性 A39	0.0415				
写作水平 B	0.25	写作规范水平 B1	0.4744	题目格式的规范性 B11	0.0215
				摘要格式的规范性 B12	0.0178
				关键词格式的规范性 B13	0.0174
				正文格式的规范性 B14	0.0192
				参考文献引用的规范性 B15	0.0126
				附录的合理性 B16	0.0096
				重复率符合规定 B17	0.0205
		写作表达水平 B2	0.5256	题目的明确性 B21	0.0202
				摘要的概括性 B22	0.0156
				关键词的代表性 B23	0.0162
				论文主体部分内容呈现的合理性 B24	0.0215
				参考文献来源的权威性 B25	0.0204
				基础语言表达的准确性 B26	0.0127
				学术语言与表达的准确性 B27	0.0151
图表文使用的合理性 B28	0.0097				

(二) 权重结果的分析

评价指标体系包括研究水平(0.75)和写作水平(0.25)两个一级指标,因为研究水平是论文的核心,是高质量论文的基础和根本,所以如权重测算结果显示,研究水平的权重远大于写作水平,但同时需注意运用合适的写作技巧。只有使用符合学术规范的格式体例、文从字顺表意准确的语句,才能将研究水平呈现出来;只有二者相得益彰,协同发挥各自的价值,才能造就出高质量的学位论文。

研究水平包括论文选题水平(0.2875)、文献研究水平(0.1435)及研究实施水平(0.3191)。其中,论文选题水平和研究实施水平较高,究其原因发现:论文选题水平的权重较高是由于选题是论文研究与写作的前提和基础,教育硕士学位论文不仅要明确提出基础教育领域中一个真实的、具有研究价值的问题,还要保证选题符合自身专业特点和专业能力^[18]。

研究实施水平权重最高是因为研究实施是开展研究的核心环节,是教育硕士论文质量最根本的保障。清晰的研究思路、科学的研究方法、可靠的数据来源、严谨的分析逻辑是产生有价值研究成果的根本保障。

写作水平包括“写作规范水平(0.1186)”和“写作表达水平(0.1314)”。写作规范水平和写作表达水平主要把题目、摘要、关键词、论文主体、参考文献、附录这几个论文必备的环节作为参考点,要求在符合基本格式规范的基础上体现学位论文写作的简洁性和严密性。写作规范性水平反映了对教育硕士论文写作的基本要求,主要从写作格式规范和学术道德规范两个方面进行考查,写作格式规范方面“题目格式的规范性(0.0215)”最高,学术道德规范方面“重复率符合规定(0.0205)”权重较高,二者都是反映作者学术态度的重要指标。写作表达水平反映了对教育硕士论文写作更高一

级的要求,是对教育硕士书面表达的逻辑性、学术性作出考量,其中,“论文主体部分内容呈现合理(0.0215)”权重较大,这就要求在评价论文时,要重点关注论文正文中观点的明确性、概括性和可信性以及观点之间的逻辑关系。

五、小结

本研究以教育硕士学位论文为研究对象,结合其实践性特点,确定了以“研究水平”“写作水平”为基本维度(包括2个一级指标、5个二级指标和30个三级指标)的教育硕士学位论文评价指标体系。研究先后采用了德尔菲法和层次分析法建立指标体系,并对指标体系进行信效度检验,经检验,指标体系具有一定的适用性、科学性及可操作性,接下来将进一步对指标体系进行实证检验,并借助计算机技术开发线上评阅系统。教育硕士学位论文评价指标体系既能为教育硕士培养单位制定论文质量标准提供依据,也能为其他研究者制定具有教育硕士专业学科特色的评价指标体系提供参考。近年来,针对专业型硕士毕业论文,国家提出“可将研究报告、规划设计、产品开发等作为主要内容,以论文形式表现”^[19]。而本研究是根据教育硕士学位论文常见形式构建的质量评价指标体系,因此,建立不同形式的教育硕士学位论文评价体系是今后研究需要解决的问题,需要更多专家和学者去研究与探索。

参考文献:

- [1]唐 姝.教育硕士学位论文评价标准研究概述[J].
亚太教育,2016(3):276.
[2][5]全国教育专业学位研究生教育指导委员会关于公布

《教育硕士学位论文基本要求》的通知[Z].教指委(2019)9号.

- [3]胡恩华,陈沛然,顾桂芳.全日制专业硕士教育评价指标体系的构建[J].研究生教育研究,2015(1):77.
[4][6]戴 云,曹如军.目标、理念、路径:新时代高校教育硕士培养方略[J].黑龙江高教研究,2019(9):28-28.
[7]教育部办公厅关于进一步规范和加强研究生培养管理的通知[Z].教研厅(2019)1号.
[8][14]王 辉,赵霞霞,司晓悦.高校中层领导干部考核指标体系研究——基于德尔菲法和层次分析法的应用[J].东北大学学报:社会科学版,2019(2):198-199.
[9]王庆辉,杨荣昌,孙莱祥.高校本科专业评估指标体系的信度与效度研究[J].黑龙江高教研究,2017(7):60.
[10]刘 可.如何进行内容效度的检验[J].护士进修杂志,2010(1):38.
[11][12][13]王志祥,柏传志.基于SPSS的基础教育的测量与评价[M].苏州:苏州大学出版社,2016:138,167,159.
[15]Mitchell,V.W.The Delphi Technique:an Exposition and Application[J].Technology Analysis & Strategic Management,1991(3):339.
[16][17]许树柏.层次分析法原理 实用决策方法[M].天津:天津大学出版社,1988:20-22,58.
[18]苏春景.教育硕士学位论文质量保障初探[J].中国高等教育,2012(6):55.
[19]教育部,人力资源社会保障部.关于深入推进专业学位研究生培养模式改革的意见[Z].教研(2013)3号.

Quality Evaluation Index System Construction of Education Master Degree Dissertation in China

ZHU Xiao - ming , ZHANG Jiu , WANG Yi - qi
(Shanxi Normal University , Linfen 041004 , China)

Abstract: The purpose of constructing the evaluation index system is to promote the formation of quality standards for the master's degree thesis in education, and to promote the development of the master's degree thesis in education toward high quality. Firstly, the evaluation index is obtained based on the analysis of literature samples; secondly, the index system is modified by the Delphi method; finally, the analytic hierarchy process is used to determine the evaluation index system of professional master's degree thesis for education taking "research level" and "writing level" as basic dimensions (including 2 first-level indexes, 5 second-level indexes and 30 three-level indicators), with a view to providing quantitative standards for the scientific measurement of the professional degree thesis for Education Master.

Key words: Master of Education; professional degree thesis; quality evaluation; index system