

# 基于CIPP模型的四川开放大学网络化考试管理评价体系探析

吴文嘉, 何 婧

四川开放大学, 四川 成都  
Email: 704910804@qq.com

收稿日期: 2021年8月4日; 录用日期: 2021年8月18日; 发布日期: 2021年9月3日

## 摘 要

随着信息化时代的迅速发展,“互联网+教育”不断推进,各类高校均在加强信息化建设,推动网络化考试改革。作为以远程教育为主的新型高校,开放大学网络化考试从事科目梳理和学生梳理都在不断增加,本文以四川开放大学成都地区办学系统为研究对象,对开放教育网络化考试现状进行了分析,基于CIPP模型提出了网络化考试管理评价体系的模型,从学习中心、管理部门两个层面进行了问卷调研访谈,运用SPSS统计软件对调研结果进行了分析,取得了一定的成果,对于开放大学系统网络化考试质量管理和改革工作提供了参考。

## 关键词

网络化考试, 考试组织, 考试管理, 评价体系

## On the Evaluation System of Network Examination Management of Sichuan Open University Based on CIPP Model

Wenjia Wu, Jing He

The Open University of Sichuan, Chengdu Sichuan  
Email: 704910804@qq.com

Received: Aug. 4<sup>th</sup>, 2021; accepted: Aug. 18<sup>th</sup>, 2021; published: Sep. 3<sup>rd</sup>, 2021

## Abstract

With the rapid development of information age, Internet plus education is constantly advancing,

and all kinds of universities are strengthening information technology and promoting the reform of network examination. As a new type of University focusing on distance education, the subject sorting and student sorting of Open University network examination are increasing. Taking the school running system of Sichuan Open University in Chengdu as the research object, this paper analyzes the current situation of open education network examination, and puts forward the model of network examination management evaluation system based on CIPP model. Questionnaires and interviews were conducted from the two levels of learning center and management department, analyzed the survey results by using SPSS statistical software, and achieved some results, which provides a reference for the quality management and reform of networked examination in the open university system.

## Keywords

Networked Examination, Examination Organization, Examination Management and Evaluation System

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着互联网、人工智能、大数据、区块链等技术迅猛发展, 不仅改变了教与学的方式, 同样也对传统的考试形式提出了挑战。网络考试作为考试评价体系的一部分, 是考试评价迈向智能化、智慧化的重要着力点。与传统纸笔考试相比, 网络化考试突破了时间和空间的限制, 建立标准化试题库, 统考系统选题、组卷、阅卷、自动评分等, 大大减轻教师的负担, 还可使考试更趋于标准化, 更能客观、真实、全面地反映教学的实际效果, 有利于促进教学质量的提高。特别是 2019 年新型冠状病毒疫情背景下在线教育的大面积使用, 使得网络在线考试在我国的普及度和学生接受度大大提升。以国家开放教育办学体系为例, 近 3 年中, 网络考试在国家开放大学考核中的占比约为 60%。

考试作为教学质量评价的重要环节, 网络化考试的组织管理非常重要, 当前选择开放教育学习的学员大都是在职人员, 工学矛盾较为突出, 开放教育的学习中心数量众多、办学条件良莠不齐, 进而对考试组织、考务管理、学生管理等方面提出巨大挑战。如何高质量的组织学生考试并确保考试质量成为远程开放教育办学管理部门不容忽视的难点。特别是针对民办教学点众多的远程教育办学地区, 如何科学制定网络化考试的评价体系, 加强对网络化考试的监控, 提高考试管理质量成为摆在考试管理部门面前必须解决的问题。本文基于 CIPP 模型提出了网络化考试管理评价体系的模型, 通过对四川成都地区多个民办学习中心进行问卷调查研究分析, 拟建立一套科学系统的评价体系, 对于远程开放教育网络化考试改革工作决策提供参考性意见与建议, 为其他网络化考试管理改革提供参考。

## 2. 文献综述

推广和建立高质量的网络考试是远程教育研究领域的热点之一, 学者通常从技术层面、考试评价方式等角度对远程教育网络考试展开研究, Eurboonyanun (2020)等从系统集成和考试优化角度对人脸识别、远程监控、自动组卷等技术在网络考试系统中的实施节点、利弊进行了详细的探讨, 与此同时陈荣(2016)、陆宏(2020)从计算机自适应考试系统角度、郭涛(2020)从区块链技术角度对网上远程考试系统的技术前沿

进行了探讨。结合课程特征多元化考试形式是当前教学质量提升的重要途径。何苗(2017)从办学点在远程教育网络考试实施过程中的问题出发对网络考试顶层设计进行探讨, 宋明珠(2020)、梁雅(2019)等从考试评价机制出发, 探讨了新技术背景下的过程性考核改革路径。总之, 远程考试技术的不断成熟为考试形式改革提供了广阔的空间。但上述研究总体集中于顶层设计或制度, 忽视了对考试工作实施者的管理研究。

CIPP 模型是当前项目评估、高校课程评估中使用的模型, Stufflebeam (1967)从背景评价、输入评价、过程评价和结果评价四个方面设定 CIPP 模型, Stufflebeam (1971)将上述模型拓展至教育领域。2003 年左右很多国内学者将上述模型引入到远程教育课程评价中, 张殿尉(2016)基于 CIPP 模型探讨了高校实践教学评价指标体系, 谢娟(2017)基于 CIPP 模型构建了翻转课堂教学评价体系, 周风(2019)基于 CIPP 模型探讨了高校创新创业教育评价体系的构建。总的来说, CIPP 模型在高校教学评估中的使用较为广泛, 为将该模型拓展至高校考试管理评估奠定了坚实的理论基础。

综上, 目前国内对于远程开放教育基于网络化考试及形成性考核等网络化考试做了较系统的研究, 但是针对某个学习中心或者某个地区的网络化考试的本土研究较缺乏, 缺少区域性和典型性。也缺乏基于 CIPP 模型对考试管理评级体系方面的拓展性研究。本文从远程网络考试组织实施管理角度出发, 借鉴 CIPP 模型建立科学的考试组织管理评价体系, 并通过问卷调查方式对基层办学点和省级开放大学两个层面进行调查, 进而在全面勾勒远程网络考试组织管理的现状的同时对指标体系的构建进行有益的探索 [1]。

### 3. 开放教育网络考试管理评价体系构建及调查问卷设定

#### 3.1. 基于 CIPP 模型的网络化考试评价体系框架

结合 2020 年《远程教育服务规范》国家标准(GB/T 39050-2020)和《国家开放大学教学质量 1.0》中关于远程教育考试相关部分的规定和模块, 本文基于 CIPP 理论的四个要素和开放教育网络化考试的特点, 建立了背景分析、考前评价、考中评价、考后评价构成的开放教育网络化考试管理评价体系理论框架(如图 1 所示), 在此评价框架下又分为了 8 个模块, 如图 1 所示:

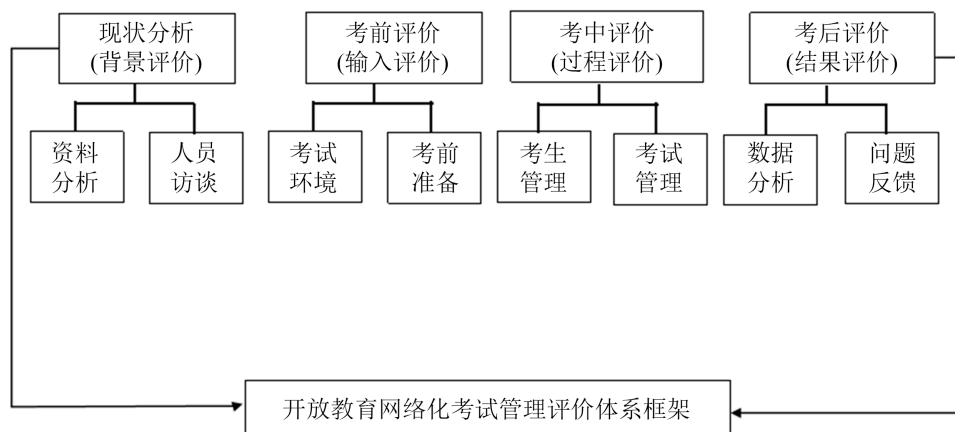


Figure 1. Framework of open education network examination management and evaluation system

图 1. 开放教育网络考试管理评价体系框架

本文基于 CIPP 模型的网络化考试评价体系主要包括: 3 个维度、6 个一级指标和 8 个二级指标。具体内容: 现状分析(资料分析、人员访谈)、考前管理评价(考试环境、考前准备), 考中管理评价(考生管

理、考试管理), 考后管理评价(数据分析、问题反馈)。在现状分析中, 资料分析主要是查阅对近 2 年开放教育网络化考试工作总结报告, 重点了解存在的困难和问题, 同时对教务管理人员和学习中心管理人员进行访谈调研, 听取相关意见和建议[2] [3]。

### 3.2. 考试管理评价问卷调查设定

为了更好的了解当前远程开放网络考试组织管理现状和检验本文构建的评价体系框架是否合理。本文设计了合计 15 道题目的问卷(如表 1 所示), 其中 1~3 题为定性类题目, 主要类型统计填写学点名称、岗位和在籍生规模等属性。通过对国家开放大学网络化考试管理制度文件, 四川开放大学成都地区办学系统近 2 年的考试总结报告分析进行了认真查阅, 对部分教务管理人员、学习中心管理人员开展访谈调研的基础上, 将成都地区网络化考试管理评价体系的 8 个二级指标进一步拓展为 12 道问卷题目(即 4~15 题)。本问卷题项采用李克特五级量表模式, 测评选项和得分依次为: 1 完全不需要改进、2 不需要改进、3 无所谓、4 需要改进、5 迫切需要改进[4]。

**Table 1.** List of open education network examination management questionnaire and item mean variance  
**表 1.** 开放教育网络考试管理问卷及题项均值方差一览表

维度	一级指标	二级指标	题序	均值	标准差
考前评价 (输入评价)	考试环境	考场环境的准备(电脑、软件、网络等)	4	2.39	1.114
		学生的复习准备(课程教学知识掌握)	5	3.06	1.212
	考前准备	学生考前教育(诚信教育、考试纪律宣传)	6	2.58	1.122
考中评价 (过程评价)	考生管理	考试入口的检查管理(身份核验、三证检查)	7	2.21	1.093
		考试的工作人员职责监管(考场外围、楼层)	8	2.29	1.105
	考试管理	考试的监考人员职责监管(考室内部)	9	2.32	1.085
		巡考管理(蹲考 + 飞巡)	10	2.27	1.062
考后评价 (结果评价)	数据分析	考试监考数据分析(签到数据、实考率、及格率)	11、13	2.31	1.121
		及格率分析	12	2.61	1.167
	问题分析	考试对教学反馈分析	13	2.65	1.157
		考点组织管理问题分析(内部)	14	2.46	1.136
		巡考问题分析(外部)	15	2.42	1.128

## 4. 四川开放大学成都地区网络化考试管理评价分析

### 4.1. 问卷调研基本情况

本次问卷以四川开放大学成都地区办学点为研究对象。采用问卷星进行网络匿名填写方式, 针对 40 多个学习中心发放问卷 158 份, 回收问卷 150 份, 有效问卷 147 份。

本问卷的定性类题目填写情况如表 2 所示, 从表中可以看出: 本次问卷调查较好的涵盖我校成都地区办学点和考务、监考人员, 能力较真实的反映办学点情况。表中虽然办学点属性的填写人员数量较少, 但远大于巡考人员数量, 说明开放大学单位属性的填写者较好的分散到考试组织管理相关岗位中, 避免了对巡考人员态度的集体性偏误。规模属性虽然样本分布不均, 但较为真实的反映了当前成都地区办学点的现状。

**Table 2.** List of attribute option filling  
**表 2.** 属性选项填写情况一览表

单位属性		岗位属性		规模属性	
单位性质	人数	岗位性质	人数	规模范围	人数
办学点人员	106	考点主考、副主考	61	2000 以下	106
	41	工作人员	44	2000~4000	26
开放大学管理人员		监考人员	26	4000~6000	2
		巡考人员	16	6000 以上	13
合计	147		147		147

#### 4.2. 网络考试评价体系整体改进意愿分析

首先,通过对各个选项的均值和方差来看(如表 3 所示),被调查对当前成都地区网络考试监控及质量评价体系的情况比较认同,其态度多集中于:不需要改进和无所谓状态。为利于后续分析,本文进一步将问卷个案以总分为权重的进行了个案加权处理,由此避免了由于单位属性和规模属存在样本分部不均对问卷分析结果产生的影响,提升了问卷的有效性和偏误性。此外,本问卷使用均值法将 14 个因素整合为 kq、kz、kh 变量,其中 kq 变量代表考前维度并指向 3~5 题, kz 变量代表考中维度并指向 6~9 题, kh 变量代表考后维度并指向 10~15 题。

根据问卷设计和量表设定,当被调查者 12 个题项的总分值  $tc \in [12, 24]$  时,则可认为被调查者对成都地区网络考试监控及质量评价体系非常满意,认为不需要进行改进。以此类推,当总分值  $tc \in (24, 36]$  时,可认为被调查者可以接受当前成都地区网络考试监控及质量评价体系并认为并不需要改进;当总分值  $tc \in (36, 48]$  时,可认为被调查者对成都地区网络考试监控及质量评价体系较为不满,认为需要改进;当总分值  $tc \in (48, 60]$  时,可认为被调查者对成都地区网络考试监控及质量评价体系非常不满,认为亟需改进;当总分值  $tc \in (12, 36]$  时,可以认为被调查者整体可接受成都地区网络考试监控及质量评价体系并认为不需要改进。根据上述思路,本文分属性对成都地区成都地区网络考试监控及质量评价体系的整体改进态度进行统计,结果如表 3 所示。

**Table 3.** List of respondents' overall attitudes under different attributes  
**表 3.** 不同属性下被调查者整体态度一览表

分类依据	属性	整体均值	整体态度描述	可接受不需要改进人数占比*
单位	办学点	32.08	能接受	64.6%
	开大管理人员	37.7	能接受	47.0%
岗位	主副考	32.65	能接受	65.6%
	工作人员	31.74	能接受	65.8%
	监考人员	34.22	能接受	54.2%
	巡考人员	41.13	偏向需要改进	32.6%
规模	2000 以下	31.70	能接受	67.8%
	2000~4000	33.77	能接受	54.9%
	4000~6000	47.09	需要改进	47.9%
	6000 以上	44.17	偏向需要改进	25.1%

注: \*表示总分值  $tc \in (12, 36]$  的人数占比。

从上述表格首先可以发现：巡考人员和在校生规模 6000 人以上的办学点偏向于认为当前成都地区网络考试监控及质量评价体系在可接受的状态下偏向于需要进一步改进。第二个值得注意的是：招生规模越大的办学点越倾向于改进成都地区网络考试监控及质量评价体系，随着招生规模的扩大办学点对当前成都地区网络考试监控及质量评价体系的可接受度平均下降 14%。

### 4.3. 单位属性下成都地区网络考试监控及质量评价体系的改进意愿分析

从单位属性将被调查者分类有助于从内外部视角(即：考试组织者与监控者)对成都地区网络考试监控及质量评价体系的改进意愿进行分析，也有利于今后的考试监考管理制度的制定。因此，根据办学点和开大管理人员属性对被调查者在考前、考中和考后 3 个维度的改进意愿绘制雷达图进行分析，结果如图 2 所示。

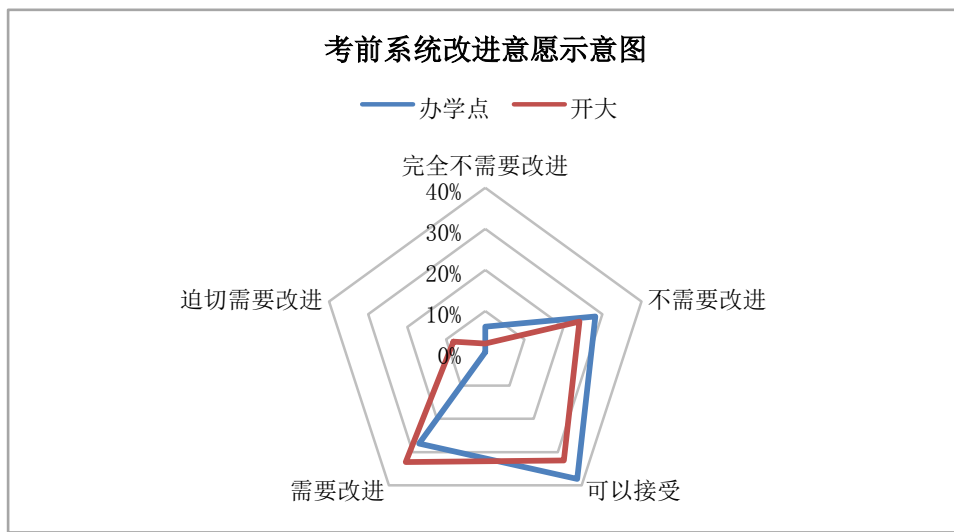


Figure 2. Schematic diagram of willingness to improve the pre examination system  
图 2. 考前系统改进意愿示意图

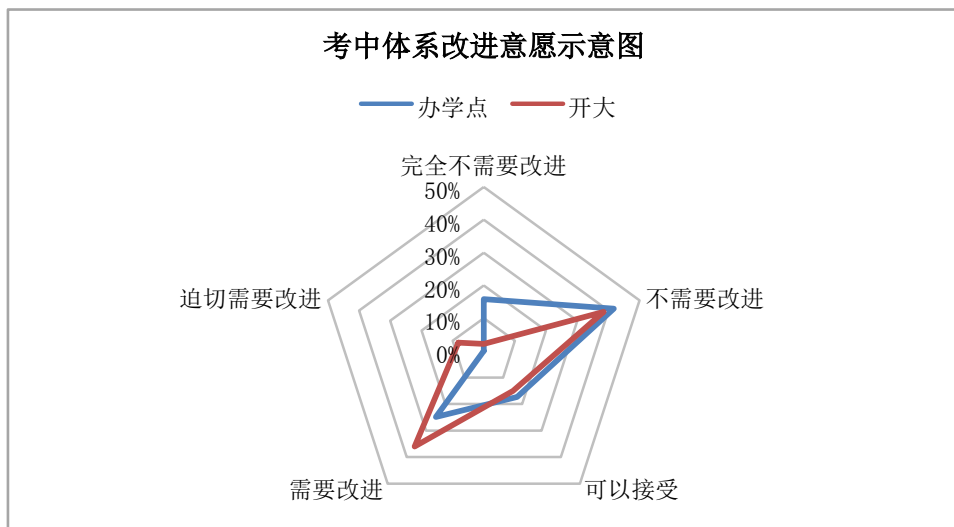


Figure 3. Schematic diagram of willingness to improve the system in the examination  
图 3. 考中体系改进意愿示意图

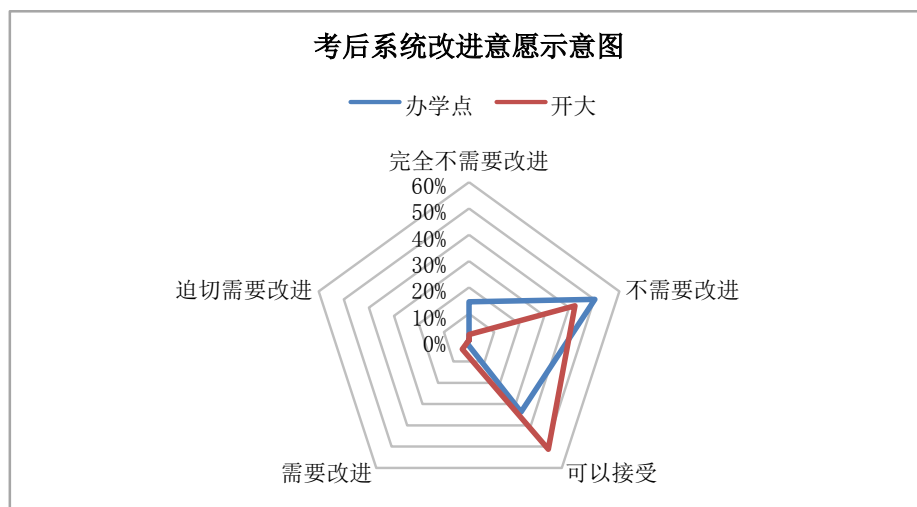


Figure 4. Schematic diagram of willingness to improve the system after examination  
图 4. 考后系统改进意愿示意图

从图 3 所示, 办学点被调查者认为当前成都地区网络考试监控及质量评价体系中考前部分可以接受的约为 40%、不需要改进的约为 30%, 与此同时开放大学被调查管理人员中超过 30%认为该部分需要改进、30%认为当前状态可以接受。从图 4 可看出, 办学点和开大管理人员存在较大分歧并且对当前的现状接受度均较低。双方人员中约 40%的人均认为当前考中系统不需要改进、20%的人均认为当前考中系统可以接受, 仅 15%左右的办学点人员认为需要改进但开大管理人员认为需要改进的比例接近 40%, 特别是办学点中有超过 15%的认为该部分完全不需要改进。从图 4 可以发现, 办学点和开大管理人员对于考后系统的满意较为一致, 超过 50%的办学点认为该部分不需要改进, 超过 50%的开大管理人员认为该部分可以接受。

#### 4.4. 岗位属性下成都地区网络考试监控及质量评价体系的改进意愿分析

本问卷将岗位分为: 主副主考、考点工作人员、考点监考人员、巡考人员 4 个类型, 上述划分有利于从考点实际工作出发, 对成都地区网络考试监控及质量评价体系进行微观分析, 更有利于出台针对性措施提升考试工作质量。

在实际工作中, 岗位人员的工作效率和工作情况通常存在“从众”现象, 因此在该部分分析时主要依从“频率”角度, 从众数和峰值角度展开分析。根据上述思路, 岗位属性下对当前成都地区网络考试监控及质量评价体系的代表性态度如表 4 所示。

Table 4. List of representative attitudes under position attributes

表 4. 岗位属性下代表性态度一览表

维度		主副考	工作人员	监考人员	巡考人员
考前	第一峰值	无所谓(46%)	不需要改进(34%)	<i>不需要改进(37%)</i>	需要改进(51%)
	第二峰值	×	可以接受(31%)	<i>需要改进(33%)</i>	可以接受(35%)
考中	第一峰值	不需要改进(40%)	<i>不需要改进(46%)</i>	<i>不需要改进(41%)</i>	<i>需要改进(45%)</i>
	第二峰值	可以接受(20%)	<i>需要改进(31%)</i>	<i>需要改进(30%)</i>	<i>不需要改进(28%)</i>
考后	第一峰值	不需要改进(47%)	不需要改进(52%)	不需要改进(55%)	可以接受(67%)
	第二峰值	可以接受(37%)	可以接受(31%)	可以接受(35%)	×

注: 表格中斜体字表示态度存在分歧。

以岗位维度来看表 4 可以发现, 就主副考而言其对整个考试评价体系均处于可以接受或不要改进状态, 而工作人员和监考人员的第一代表性态度倾向均与主副主考基本一致, 但考前、考中环节中部分工作人员和监考人员的第二代表性态度则出现分歧。巡考人员整体认为考前、考中环节需要改进, 与前述人员产生分歧。

以考试维度分析表 4 可以发现, 就考前部分而言: 巡考人员对考前系统中需要改进的代表性意愿高于其他人均约 10%以上; 监考人员的态度出现内部分歧且势均力敌, 37%的监考人员认为不需要改进而 33%的监考人员认为需要改进。就考中部分而言, 更考点工作人员、监考人和巡考人员内部均出现了分歧且二者的人数占比悬殊不大, 超过 40%的工作人员和监考人员均认为该部分不需要改进, 同时也有超过 30%的认为该部分需要改进; 巡考人员的第一典型态度是需要改进(占比 45%), 但也有近 30%的认为该部分不需要改进。就考后部分而言, 所有岗位的人员的态度大致趋同, 认为可以接受或不需要进行改进。

#### 4.5. 规模属性下成都地区网络考试监控及质量评价体系的改进意愿分析

“小点多、大点少”是本次问卷调查样本成都地区办学点的真实写照, 因此在多数制原则下进行的问卷分析容易出现“投票悖论”, 即重要信息由于投票人数少而未被体现。在规模属性下, 本问卷样本中在校生规模 2000 人一下的被调查者占比高达 72%, 因此, 本文将在规模属性视角下对问卷进行分析。

在规模属性下, 主考对当前成都地区网络考试监控及质量评价体系考前、考中、考后 3 个方面的态度与多数制原则下的态度出现分化。如表 5 所示, 在考前和考后两个部分, 以在校生 4000 人为临界值主副考人员的态度存在分歧, 规模越小的办学点主副考越倾向于可以接受或不需要改进, 而规模越大的办学点越态度鲜明的认为需要改进。

**Table 5.** List of attitudes of principal and associate examiners under scale attribute  
**表 5.** 规模属性下主副考态度一览表

	考前		考中		考后	
	整体均值	整体态度描述	整体均值	整体态度描述	整体均值	整体态度描述
2000 以下	2.75	偏向可接受	2.4	偏向不需要改进	2.26	偏向不需要改进
2000~4000	2.90	偏向可接受	2.55	偏向可接受	2.23	偏向不需要改进
4000~6000	4	需要改进	4	需要改进	3	可以接受
6000 以上	3.94	偏向需要改进	3.8	偏向需要改进	3.7	偏向需要改进
不考虑规模	2.91	偏向可接受	2.24	偏向不需要改进	2.22	偏向不需要改进

同理, 在岗位属性分析中出现的工作人员、巡考人员态度差异在规模属性下也呈现一定程度的分歧, 其中表现最为显著的是考中环节, 如图 5 所示, 规模从小到大排列后(注: 4000~6000 人规模工作人员样本缺失), 工作人员对考中系统的代表性态度中不需要改进的人数占比呈现“中高低”走势、需要改进的人数占比呈现“低低高”走势。并且与整体均值相比存在小规模的态度一致、6000 以上规模与整体均值表现相悖的情况。

### 5. 开放教育网络化考试管理评价体系的构建建议

#### 5.1. 优化完善网络化考试管理制度

考试是检验教学质量的最终环节, 制度化、规范化非常重要, 四川开放大学考试管理部门应该原有考试管理制度的基础上, 充分利用基于 CIPP 模型的网络化考试管理评价体框架, 对网络化考试管理制度



(考试组织管理流程、巡考流程、考试评价)进行优化, 在充分了解学习中心负责人、监考人员意见的基础上, 更多应该考虑成都地区办学系统教学、教务、考务管理部门、巡考人员的态度和建议。

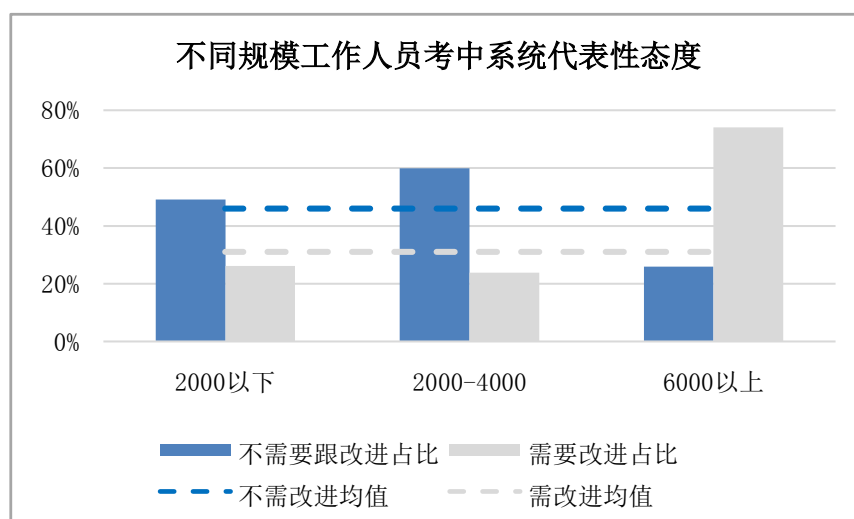


Figure 5. Representative attitude of staff of different sizes in the examination system  
图 5. 不同规模工作人员考中系统代表性态度

## 5.2. 加强对于考中环节监控

考中环节是整个网络化考试组织实施的核心环节, 要想确保考试工作质量, 保证考试公平公正, 减少学习中心有组织的替考、学员个人违纪舞弊行为发生, 四川开放大学考试管理部门应不断推进考试工作制度化、规范化、科学化进程, 加强对考试的监控, 确保考试平安顺利进行, 可以制定《网络化考试质量评价表》, 从考前准备、考中监督、考后评价三个当面对各考点网络化考试开展质量评价工作。选派熟悉教学教务业务、责任心强的教师担任巡考人员, 考前对巡考人员进行系统的培训, 确保网络化考试巡考工作质量。

## 5.3. 对评价指标设权重, 分类施策管理

从问卷分析可知, 从内外部视角看: 考试组织者(办学点属性)和考试监督者(开放大学属性)在多数环节均达成共识、评价一致但在考中环节分歧较大, 从规模属性看: 规模不同使得考试组织者之间的观点容易产生分歧。因此, 今后在网络化考试管理评价体系构建时应适当增加考中维度里一级、二级指标的权重, 降低其他部分的权重, 从而更好达到测评目的。其次, 在网络化考试管理评价体系构建时应充分考虑规模属性的影响。建议在实际评价时根据在校生规模酌情实施分类评价模式, 或根据在校生规模对参评个案适当赋予权重, 进而客观、科学的考试组织者进行评估, 更好的以评促改、以评促建[5]。

## 5.4. 用好考后评价的数据分析

四川开放大学考试管理部分、教学部门应该对考试及格率、实考率等数据进行认真分析, 对各考点的考试管理工作进行量化考评排名, 对网络化考试工作进行定期总结研判, 对在考前准备、考试组织、系统管理、人证核验、监考履职、考风考纪等方面存在的问题进行梳理汇总, 及时反馈各学习中心并督促其整改, 对于考试中反馈的试题、知识点等教学问题, 应及时通过教学部门反馈给课程责任教师, 完善教学资源建设并加强对学生的教学支持服务[6]。

## 基金项目

四川广播电视大学 2020~2021 年度科研项目年重点课题《四川电大直属学院网络考试监控及质量评价体系研究》(KTKYC2020002Z)。

## 参考文献

- [1] 周风, 曾增, 李龙. 基于 CIPP 模型和层次分析法的高校创新创业教育评价体系的构建[J]. 宿州学院学报, 2019, 34(9): 90-92.
- [2] 谢娟, 张婷, 程凤农. 基于 CIPP 的翻转课堂教学评价体系构建[J]. 现代远程教育研究, 2017(5): 17-19.
- [3] 杨瀚涛. 高校信息化发展中的考试信息化探索[J]. 产业与科技论坛, 2014(17): 103-104.
- [4] 宋明珠, 胡嫣然. 浅谈现代远程教育考试的改革措施[J]. 教育教学论坛, 2020(40): 345-346.
- [5] 梁雅. 远程教育考试评价的现状和改革初探[J]. 当代教育实践与教学研究, 2019(4): 72-74.
- [6] 张殿尉, 刘佳杰. 基于 CIPP 模式的高校实践教学评价指标体系研究[J]. 中国成人教育, 2016(9): 110-113.