

# 幼儿园园长专业知识测验工具的开发

王 萍

(西南大学 教育学部,重庆 400715)

**摘 要:**幼儿园园长专业知识是幼儿园园长专业素养的重要组成部分。考察幼儿园园长专业知识现状需要开发科学有效的测验工具。经过初始测验试题的编制、预测、检验、修订和完善等步骤,最终形成了难度适中、区分度较高、信度和效度都良好的正式测验工具。然而,在工具开发过程中,由于原创性题目较多等原因尚存在不足,需要在未来的研究中进一步探讨与反思。

**关键词:**园长专业知识;测验工具;知识测量;试题;开发

**中图分类号:**G617 **文献标识码:**A **文章编号:**2095-8129(2016)02-0118-07

幼儿园园长(以下简称“园长”)专业知识是园长专业能力与专业精神的基础,也是园长专业素养的重要组成部分。提升园长专业知识水平是促进园长专业发展的核心与基础,也是幼儿园教育质量提高的重要前提与保障。我国学者宋乃庆指出:“当前,通过对教育问题的测评、监测与督导提升教育质量,已经成为当前世界教育发展的趋势。”<sup>[1]</sup>因此,测评园长专业知识现状是提升园长专业知识水平和提高幼儿教育质量的必要前提,同时也是学前教育发展的时代诉求。然而,笔者通过梳理已有文献发现,国内外关于园长专业知识的研究较少,且多从理论层面探讨园长应具有的专业知识结构以及构建专业知识体系等问题,而对园长专业知识现实问题的考察和量化研究却少有人问津。

心理学研究已证实,专业知识属于心理特质的一种。目前,教育与心理测量理论主要有经典测量理论体系、概化理论体系和项目反应理论体系。这3种理论体系,各有长短<sup>[2]</sup>。一般而言,经典测量理论中提出的教育测验法是测量专业知识的主要方法,即按教育测验的规则,对学习者掌握某些学科的知识、技能予以数的描述的一种方法。因此,本研究依据和借鉴教育测验的方法与技术,以教育部2015年颁布的《幼儿园园长专业标准》中提出的园长专业知识体系为理论模型,立足于西南地区园长专业知识的实际情况,开发、研制了一套园长专业知识测验工具。

## 一、编制测验试题

开发园长专业知识测验工具是一项复杂的系统性工程。参考、借鉴已有相关文献<sup>[3-4]</sup>,编制初始测验试题,主要步骤有:明确测验目的、形成测验蓝本、编制测验试题。

(一)明确测验目的

园长专业知识测验目的在于测查不同民族、地区、性质的幼儿园正、副园长对《幼儿园园长专业标准》所要求的专业知识的掌握情况。就测验功能而言,该测验应具有预测、摸底的诊断性功能,即通过测验揭示园长专业知识的现实问题,为提升园长专业知识水平提供实证依据。同时,这类测验属于标准化测验的一种,要求测验工具具有较高的信度与效度。因此,编制测验试题时需从教育学、心理学、社会学、管理学等多种学科发展的视角出发,结合高校专家、优秀教研员、专家型园长、普通一线园长等不同专业层级的人员来确定、评价、修正测验题目。此外,该测验的计分标准主要以《幼儿园园长专业标准》中提出的园长专业知识目标为主要参照标准,以评估园长专业知识的实际水平。另外,辅以园长专业知识的常模(即测验平均分)为次要参照标准,以衡量不同区域、性质、职称、学历、专业等背景下园长专业知识水平的差异情况。

(二)编制测验蓝本

测验蓝本也就是通常所说的双向细目表,是描述试卷内容、层级、题量等有关构成比例的文件。园长专业知识测验蓝本在横向维度上明确了园长专业知识的内容,在纵向维度上明确了园长专业知识的目标。此外,园长专业知识的测验方法、题量等内容也被纳入园长专业知识测验蓝本当中。

1. 排列出具有代表性的园长专业知识点

借鉴已有相关研究<sup>[5]</sup>,《幼儿园园长专业标准》遵循“专业职责-专业知识”的建构思路,明确指出园长应具备的专业职责和专业知识。本测验试题的编制维度是依据《幼儿园园长专业标准》提出的园长6项专业职责,即规划幼儿园发展、营造育人文化环境、领导保育教育、引领教师成长、优化内部管理、调试外部环境。《幼儿园园长专业标准》涉及的专业知识点繁多,为使测验内容精简到可操作的范围,本研究将6项专业职责下的每条专业知识细化为2~6个具体内容,进而编制成对园长专业知识点相对重要性排序的调查问卷。依据三角互证原理,旨从宏观专业引领以及专业实践角度出发,向一线园长(25名)、高校专家(5名)、区县教研员(5名),发放问卷并回收有效问卷,利用excel软件进行数据管理及统计分析,最后统计出18个代表性知识点,作为园长专业知识测验蓝本的内容。

表 1 园长专业知识中的代表性知识点

| 园长专业职责   | 园长专业知识中的代表性知识点                                             |
|----------|------------------------------------------------------------|
| 规划幼儿园发展  | 熟悉《幼儿园工作规程》《幼儿园教育指导纲要(试行)》;了解国内外学前教育发展的基本趋势;掌握制定幼儿园发展规划的理论 |
| 营造育人文化环境 | 具有良好的品德与艺术修养;了解幼儿园文化建设的基本理论;掌握幼儿身心发展特点                     |
| 领导保育教育   | 掌握国家关于不同年龄阶段幼儿发展的目标;熟悉幼儿园一日活动的组织与实施;了解国内外幼儿园保教发展动态         |
| 引领教师成长   | 把握专任教师职业素养要求;熟悉幼儿园教师专业发展各阶段的规律和特点;掌握园本教研、合作学习等学习型组织建设方法    |
| 优化内部管理   | 熟悉幼儿园管理的基本知识;掌握幼儿园园舍规划方法与实务;掌握幼儿园卫生保健方法与实务                 |
| 调试外部环境   | 掌握幼儿园与家长有效沟通的策略和方法;熟悉社区教育资源的功能;掌握家园共育的知识和方法                |

2. 确定各专业知识点的目标

借鉴已有研究成果,本研究将“把知识的掌握水平分为识记、理解、运用三个层次”<sup>[2]</sup>。通过访

谈具有学前教育以及教育管理专业背景的高校专家、专家型园长、优秀教研员等,确立了从学科发展角度和教育领导与管理实践的角度建立园长专业知识目标层级。访谈中的专家、学者普遍认为:园长作为幼儿园发展的核心人物,应更多达成“理解”类目标,以确保幼儿园发展方向的正确性以及观念的引领性;对于园所文化、游戏活动、环境创设等实践性、应用性较强的领域应达成“运用”目标;对于安全、卫生、环境设置等具体要求则需达到“识记”目标。

### 3. 确定测验方法、题目数量、测验时间和评分方法

本研究主要采用笔试测验法。由于学术研究的伦理性、人文性、无利害性等特点,要求测验对象匿名填写,测验题目数量适中,完成时间不宜过久。遵循覆盖全部知识点、控制答题时间等原则,题目数量稍多于代表性知识点的数量。通过高校专家、专家型园长等人对试题难度进行评估,确定了测验时间为15~20分钟。此外,为方便计分与统计分析,本研究采用0/1分值法,即客观题答对得1分,答错得0分;主观题答对一题计1分,答错为0分,最高分为所要求的答题数目。

### (三)编制测验试题

测验由试题构成。编制高质量的试题是开发有效测验工具的核心。

#### 1. 编制试题及答案

已有研究建议,预测试卷中的“初选题目的数量应是测验计划数量的2~3倍”<sup>[6]</sup>,以便在预测后依据预测结果进行删减与修改。因此,研究者针对每项测试内容,编制出2~3道试题,共编制40道预测试题,其中包括单选题32道、是非题7道、简答题1道。单选题测验内容广泛,填写简单,是测验的主要题型;是非题主要测验园长的“理解”能力;主观题主要测验园长对专业知识的“运用”能力。此外,编制试题主要采用了两种方式:一是编制原创性题目,占题目总数的77%,原创性题目的来源,有的是对一线园长以及学前教育专家访谈材料的整理及分析,有的是学前教育已有的研究成果,还有的是政策文本的观点和要求等;二是采用或改编园长任职资格考试试题,约占23%。因现有考察园长专业知识的试题及相关研究较少,故本研究中的试题大多为原创性题目。

#### 2. 效度系数检验

本研究主要采用内容效度的逻辑判断法对其进行效度检验。具体方法主要借鉴国际上通用的由劳希提出的内容效度系数(Content Validity Ratio,CVR)的概念及其计算方法<sup>[7]</sup>。首先,选取30位学前教育专家(包括20名高校专家与10名专家型园长)对每项代表性知识点试题的重要性进行评价。评价分3个等级,A为“非常重要”,B为“有用但不重要”,C为“没有用”。然后,回收数据,运用公式 $CVR = (n_c - N/2) / (N/2)$ ,计算出每道题的内容效度系数(其中, $n_c$ 表示选择A的专家人数,N表示专家总数),并将所有题目的内容效度系数进行均值计算。最后,经过计算发现,专家效度达到0.34,大于劳希提出的30名专家的最小内容效度系数为0.33的数值。也就是说,平均每道题中选择A(“非常重要”)的专家数达到了20人,这说明试题内容对所调查的园长专业知识的代表性程度较高。

#### 3. 组合预测试题

依据劳希效度系数检验结果,对试题进行初步修改,充分评估题目的难易程度,依据双向细目表各维度指标由易到难、循序渐进的出题原则,合成预测试题。此外,为确保预测实施的相对标准化,编制的测试说明书对测验填写要求、时间控制等都作出了详细的规定与说明。

二、预测与修订测验试题

(一)预测试题的发放与回收

为使预测样本性质与正式测验的样本性质一致,研究者主要向西南地区的重庆市、四川省和云南省发放有效的纸质和网络测验试卷 50 份,回收有效问卷 48 份,回收率为 95.56%。其中,市区幼儿园园长占总数的 53.49%,区县幼儿园园长占 39.53%,乡镇幼儿园园长占 6.98%;公办幼儿园园长占调查总数的 53.49%,民办幼儿园园长占 46.51%;正园长占调查人数的 37.21%,副园长占 62.79%。具体情况见图 1。

(二)预测试题的质量指标检验

依据教育统计与测量的相关理论与方法,研究者先后对预测试题的难度、区分度、信度、效度等指标进行检验。选择、保留、修改和删除相应试题,最终形成较高质量的测验试题。

1. 检验难度 P 值

测验试题难度系数是评价测试题质量的指标之一,也是筛选测试题的主要依据。“由于其计分方法的不同,所以其难度的估计方法也不同”<sup>[8]</sup>。结合本研究 0/1 计分方式,用答对该题的人数与所有被试人数的比来表示 P 值。依据已有教育测量学的观点,测试题难度值范围适宜在 0.2~0.8,且平均难度系数在 0.5 左右。通过难度系数计算,共 23 道题的难度系数 P 值的范围在 0.2~0.8,保留了难度系数检验合格的题目,剔除或修改了难度系数检验不合格的题目。

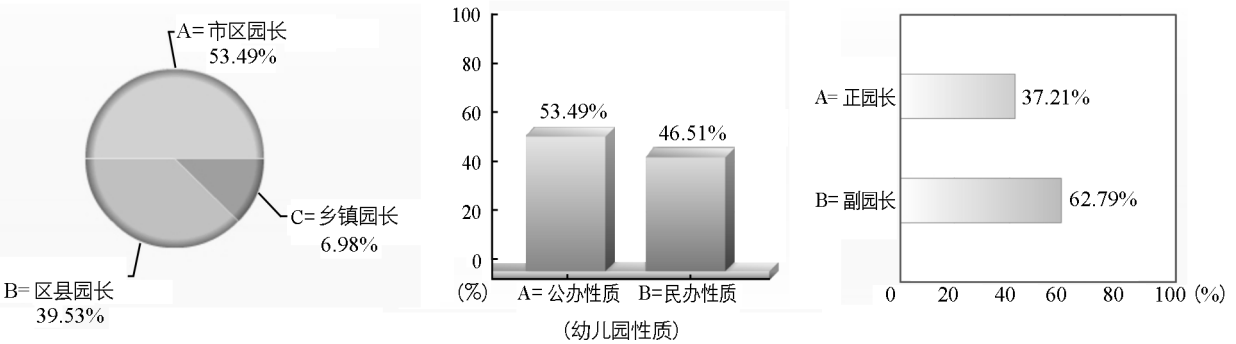


图 1 园长专业知识预测试题回收情况

2. 检验区分度 D 值

区分度是指测验对被试实际水平的区分程度,同样也是筛选测试题的重要指标。本研究主要采用分组法估计区分度。具体做法是:将测验总分由高到低排序,抽选出前 12 名的高分组和后 12 名的低分组,进而求出高分组与低分组答对该题的人数比率之差,用公式表示为: $D = PH - PL$ 。有专家(R. L. Ebel)认为:应删去区分度低于 0.20 以下的测试题;修改并提高区分度值在 0.20~0.29 的测试题;适当修改并保留区分度值在 0.30~0.39 的测试题;保留区分度值大于 0.40 的测试题。在预测试题的区分度检验中,共有 21 道题的 D 值(区分度)大于 0.20,大于 0.40 的题目共 18 道。研究者保留了 18 道较高区分度( $D \geq 0.40$ )的测试题,删去 D 值小于 0.20 的题项,并依据试题分析结果,调整和修改区分度在 0.20~0.39 的题项。

3. 检验信度系数 rtt

信度估计方法有多种,研究者主要采用分半信度的卢农(Rulon)公式估计法进行信度检验。具

体做法为:首先剔除不符合质量检验难度系数(  $P$  值)和区分度(  $D$  值)的试题;然后在 40 道预测试题中,筛选出 16 道难度系数在 0.2~0.8 之间且区分度大于或等于 0.40 的试题;再将 16 道试题均匀分为奇数题项与偶数题项(详见表 2);最后经卢农(Rulon)公式计算,得出预测试题的分半信度系数为 0.73。综合多位学者的观点——“整份量表最低的内部一致性信度系数要在 0.70 以上,最好能高于 0.80”<sup>[9]</sup>,该测试题的分半信度系数表明,该调查工具的信度值较高,整个量表可以接受,其测试结果稳定可靠,但仍需进一步修改。

表 2 园长专业知识测验试题的信度分析

| 测验试题维度<br>(专业知识类型) | 题目数量 | 题目序号           | Alpha 信度系数 |
|--------------------|------|----------------|------------|
| 规划幼儿园发展            | 4    | 1、2、3、4        | 0.641      |
| 营造育人文化             | 5    | 5、6、7、8、9      | 0.753      |
| 领导保育教育             | 5    | 10、11、12、13、14 | 0.688      |
| 引领教师成长             | 4    | 15、16、17、18    | 0.734      |
| 优化内部管理             | 3    | 19、20、21       | 0.705      |
| 调试外部环境             | 3    | 22、23、24       | 0.762      |
| 所有项总计统计量           | 24   |                | 0.802      |

(三)修订并形成正式测验试题

首先,确定正式测验试题题型。对预测试题进行难度系数、区分度、信度和效度的质量检验,结果显示,判断题只能检测最低层次的认知目标,且受随机猜测的影响较大,故舍去该类题型。依据测验内容需要,将部分判断题改编为填空题。因此,正式使用的测试题题型包括单选题、填空题和简答题 3 种题型(如图 2)。

(单选) 根据《幼儿园教育指导纲要(试行)》,请选出说法正确的一项( )

A.《纲要(试行)》第一次提出对早期阅读的提倡和重视

B.《纲要(试行)》强化了对具体知识与技能的要求,提法变得比较具体和明确

C.《纲要(试行)》的内容框架包括:说明、教育内容与要求、组织实施、教育评价

D.《纲要(试行)》指出了五大领域包括:健康、语言、社会、科学、美术

(填空) 贵园建设特色幼儿园文化是 \_\_\_\_\_

(简述) 贵园的特色文化体现在哪些方面?请您列举 3 项。

(1) \_\_\_\_\_

(2) \_\_\_\_\_

(3) \_\_\_\_\_

图 2 园长专业知识正式测验题型

其次,筛选和调整各项测试题。通过对预测试题的质量检验,筛选并保留 15 道难度适中、区分度较高、处于不同维度(园长专业职责)的单选题。遵循难度系数、区分度、信度和效度之间的相互关系原则,并基于访谈结果以及对预测结果的分析,调整原有测试题题干的表述,改编、新增 14 道单选题、2 道填空题、3 道简答题,旨在使试题的平均难度系数调整为 0.50。这样既可达到区分度最大值(  $P = 1.00$  ),还能提高测试信度。正如 R.L.艾伯在 1962 年指出的:“要想达到理想的测验信度,提高试题的区分度是一个好方法。”<sup>[10]</sup>

最后,经二次预测后形成正式测试题。在第二次预测中,发放有效试卷 120 份,回收试卷样本量 112 份,回收率为 93.33%。其中,市区园长占 52.68%,区县园长占 35.71%,乡镇园长占 11.61%;公办幼儿园园长占 31.25%,民办幼儿园园长占 68.75%。对回收试卷进行统一的赋值评分,并进行难度、区分度、信度和效度检验。分析统计结果显示:该套试题平均难度系数为 0.45,且各题难度系数值(  $P$  )均在 0.2~0.8 范围当中,说明该测试题难度适中;区分度(  $D$  )的平均值大于 0.3,且 80%的试题区分度大于 0.4。借助统计软件 SPSS17.0 将数据进行登录、统计、管理,得出正式测试题整份问卷的 Cronbach’s Alpha 信度系数为 0.802(见表 2),高于各子维度的 Alpha 系数。

由表 2 可知,各分量表 Alpha 系数在 0.641~0.762 之间,其中,营造育人文化环境、引领教师成长、优化内部管理、调试外部环境 4 个子维度的 Alpha 信度高于 0.70,根据信度检验的相关理论与知识,可判定该测试题的信度可接受。二次预测各项质量指标,结果显示,该测试题效度较高,可以真实、有效地测量出园长专业知识的实际水平。

### 三、研究不足与建议

园长专业知识测验工具的开发是对园长专业知识进行量化研究的有益尝试,同时也是一项复杂、难度较大的系统性工程。尽管研究者已经尽量做到客观、真实、科学地编制测验工具,但在工具编制过程中也难免存在一些不足,需要进一步探讨与反思。

#### (一)原创性题目较多,可能影响测验效度

该研究中 77%的测试题为原创性题目,尽管研究者在编制测试题时,始终遵循三角互证原则,但由于参与研究者的专业知识背景不同、价值观念存在差异等因素,致使该测试题的主观差异性较大。因此,研究者无法断定,该套测验试题能否被广大学前教育专业的科研工作者以及实践工作者普遍认同。此外,该测验试题的效度检验主要采用劳希效度系数检验法,其实质为内容效度检验的逻辑分析方法以及测试题目与测试蓝本相比较的方法。虽然本研究测验的效度系数符合试卷的效度要求,但由于参与工具编制的专家专业知识背景、价值观念不同等因素,致使劳希效度系数检验方法也带有一定的主观性,可能会影响测验效度。为解决此问题,研究者除需进行自上而下的内容效度检验以外,还需增加自下而上的效度检验,以提高测验试题的效度。

#### (二)研究的非利害性,可能影响测验信度

由于遵循学术研究的伦理性、非利害性等原则,研究对象匿名填写测试题,数据结果也仅用于研究分析,不参与对研究对象(园长)的考核与奖惩等。因此,研究对象在填写问卷或解答测试题的过程中,难免存在一定的随意性。这可能会导致同一研究对象在不同时间、地点进行测验时所填写的结果不一致,影响测试题信度。此外,园长专业知识测验工具的开发并非一蹴而就,而是要经过多次预测问卷的发放、回收与分析,但是由于每次测验对象不一致、测验环境有变化等因素,测验试题信度可能会受到影响。为解决此问题,研究者可借助当地教委、教研员等行政力量,尽可能创造统一、规范的测验环境。

#### (三)标准化测验工具本身不可避免的局限性

借助经典测量理论开发的园长专业知识测验工具,是对园长专业知识进行量化的第一步,也是测量园长专业知识必不可少的重要探索。然而,园长专业知识标准化测验工具因其答案的唯一性,以及测验知识覆盖面有限等不足,难以全面揭示错综复杂的园长专业知识现状。此后,研究者可遵循多主体评价、过程性评价的基本原则,将园长专业知识的标准化评价、自我评价以及他人评价等多种评价方式结合起来,可以更为真实、全面地揭示园长专业知识现状。

#### 参考文献:

[1] 宋乃庆,杨欣,王定华,等. 学生课业负担测评模型的构建研究——以义务教育阶段学生为例[J]. 西南大学学报(社会科学版), 2015,41(3):75-81.

[2] 马云鹏,赵冬臣,韩继伟. 教师专业知识的测查与分析[J]. 教育研究,2010(12):70-76.

[3] DOWNING S. Twelve steps for effective test development [M]//DOWNING S,HALYDYNA T. Handbook of test development. New Jersey: Lawrence erlbaum associates,2004;3-26.

[4] LINN R. The standards for educational and psychological testing: guidance in test development[M]// DOWNING S,HALYDYNA T. Handbook of test development. New Jersey: Lawrence Erlbaum Associates, 2006;27-38.

[5] 易凌云. 幼儿园园长专业标准的构建原则与基本内容[J]. 学前教育研究,2014(5):30-36.

[6] 金瑜. 心理测量[M]. 上海:华东师范大学出版社,2001;323.

[7] LAWSHE C H. A quantitative approach to content validity [J].Personnel Psychology,1975,28(4):563-575.

[8] 王孝玲. 教育测量;修订版 [M]. 上海:华东师范大学出版社,2011;105.

[9] 吴明隆. 问卷统计分析实务-SPSS操作与应用[M]. 重庆:重庆大学出版社,2010;244.

[10] 王汉澜. 教育测量学[M]. 开封:河南大学出版社,1987;109.

Development of Tools for Surveying the Professional Knowledge of the Principal of Kindergarten  
——A Case Study of the Southwest Region of China

Wang Ping

(Department of Education, Southwest University, Chongqing 400715,China)

**Abstract:** The professional knowledge of the principal of kindergarten is one of the main elements of kindergartens' professional quality, and to understand the status of kindergartens' professional knowledge needs to develop effective testing tools. However, developing testing tools is complex and systematic. After compiling questions, forecasting, testing the test questions quality, and constantly improving the test questions, eventually, we got the formal test's difficulty value (P) between 0.2 to 0.8, with a mean value of 0.45. The discrimination degree(D) of the 80% test questions was greater than 0.4, with the mean value greater than 0.3. The test's Content Validity Ratio(CVR) was 0.34, and the split half reliability coefficient was 0.82 for the test. So the test tool is of high quality, and can measure the professional knowledge of the principal of the kindergarten effectively. But in the process of the developing the testing tool, there are still some problems that need to be discussed in the future.

**Key words:** professional knowledge of the principal of the kindergarten; test tool; knowledge measurement; exam questions; develop

责任编辑 邱香华