北京市综合实践活动学习单数据分析方案

（讨论稿）

、

北京市教育科学研究院

2018年10月

目录

[一、 基本测度框架 1](#_Toc528763871)

[1. 活动设计 1](#_Toc528763872)

[2. 活动收获 2](#_Toc528763873)

[3. 形成评价指标体系 2](#_Toc528763874)

[二、 活动基本情况 3](#_Toc528763875)

[1. 活动参加和完成情况 3](#_Toc528763876)

[2. 指导教师分析 3](#_Toc528763877)

[3. 活动评价 3](#_Toc528763878)

[三、 时间、地点、事件分析 3](#_Toc528763879)

[1. 时间分布 3](#_Toc528763880)

[2. 地点分布 3](#_Toc528763881)

[3. 内容分析 4](#_Toc528763882)

[4. 活动主题分布 4](#_Toc528763883)

综合社会实践活动学习单分析包括对活动完成情况、活动类型、活动主题、活动名称、活动简介、活动时间、活动地点、活动指导教师、活动感受和活动评分的分析。主要可以分为以下三个模块：

## 基本测度框架

### 活动设计

#### 系统度

* 将活动简介分为主题、目标、准备、实施过程、指导重点、活动评价等六个方面（暂定）
* 基于对教师填写的活动简介的文本分析，判断这六个方面是否系统和完整
* 按照完整度进行计分

#### 丰富度

* 时间丰富度
  + 活动周期内，组织的活动次数越多，则时间丰富度越高
* 地点丰富度
  + 活动地点越多，则活动丰富度越高
* 主题丰富度
  + 活动涵盖的主题范围越多，则主题丰富度越高
* 内容丰富度
  + 以学校-年级为基本单位，形成以动词为核心的活动内容表述
  + 活动内容中分词越多，则活动丰富性越强。如，主要组织同学“参观”各种博物馆，则“参观”一词在活动内容中会出现很多次，而分词后则只出现一次，则其活动丰富性就较弱。

#### 与时俱进度

* 对活动名称和活动简介分别进行TF-IDF分词后，从中提取关键词
* 邀请专家对其中关键词的新异性和时代性进行判断
* 再对各活动名称和活动简介中新异性和时代性词汇次数进行加和，形成活动与时俱进度的计分。

### 活动收获

#### 认知收获

* 对活动感受进行语义分析，邀请专家从
  + 价值体认
  + 责任担当
  + 问题解决
  + 创意物化

四个维度（暂定）对部分主观感受的表述语句贴标签

* 采用CNN深度学习网络对识别模型进行训练，得到分类准确度较高的模型
* 由机器对余下的活动感受进行分类和计分
* 得到每个学生的认知收获得分。

#### 情感收获

* 对活动感受进行分词后，根据情感词汇库，对活动感受中出现的情感词汇进行分析
* 按照词汇的极性（积极、中性、消极）进行主观感受的情感分类和强度计分
* 得到实践活动后每个同学的在七个大类上的情感收获计分。

### 形成评价指标体系

形成以活动设计和活动收获为两个主要维度，下面分别包含系统度、丰富度、与时俱进度和认知收获、情感收获等子维度及相应指标的评价指标体系。

* 邀请专家度对指标体系的合理性进行论证
* 采用层次分析法确定各指标、子维度和维度的权重
* 去量纲后，对指标按权重计分，得到每个学校的最后评价分数

## 活动基本情况

### 活动参加和完成情况

* 活动次数
* 参与人数
* 完成人数
* 10次都完成的学生占比

### 指导教师分析

* 教师平均指导学生数
* 指导老师在活动主题上的人数分布

### 活动评价

* 自评分数和教师评分的平均分
* 自评分数和教师评分的分布情况
* 教师评分和学生自评的一致性：独立样本t检验

## 时间、地点、事件分析

### 时间分布

* 学生活动频次在月份上的分布：折线图/条形图
* 团体活动中活动创建时间先于或同步于活动开始时间占比

### 地点分布

* 活动地点的分类：如学校、社区（家、家附近和社区广场）、博物馆（植物园、天文馆等）、爱国主义教育场所，自然环境等
* 区内外活动占比
* 校内外地点占比
* 自主和团体活动地点的比较
* 城市和农村活动地点的比较
* 地图可视化形式呈现全市活动所在地点

### 内容分析

* 形成以动词为核心的活动内容表述，如“参加升旗仪式”、“包饺子”等
* 自主和团体活动内容的比较
* 城市和农村活动内容的比较
* 劳动教育在活动内容中的比例和具体形式

### 活动主题分布

* 32个活动主题的频次和占比
* 自主和团体的活动主题比较
* 城市和农村的活动主题比较
* 不同年级的活动主题比较