

ОБЩАЯ ПЕДАГОГИКА

УДК 159.9.072

ПЕДАГОГИЧЕСКИЙ ЭКСПЕРИМЕНТ: ВОПРОСЫ НЕДОСТАТОЧНОГО ОБЪЕМА ВЫБОРКИ

Е.В. Борисова

Тверской государственной технической университет

Рассмотрены аспекты постановки и обработки результатов педагогических экспериментов с ограниченным числом объектов изучения. Показано, что при работе с узкими специалистами или при изучении специфических свойств приходится ограничиваться тем количеством объектов, которые доступны исследователю. Особенностью малых выборок является отличие от нормального закона распределения случайных ошибок, приводящее к нарушению репрезентативности результатов. Приведен пример с малым объемом выборки, обосновано объединение данных с целью построения выборки нужного объема и дальнейшего использования стандартных методов математической статистики.

Ключевые слова: *выборка, постановка эксперимента, уровни компетенций, обработка данных.*

В теории и практике научно-педагогических исследований большое значение имеет возможность оценивать не только качественные, но и количественные характеристики изучаемых явлений и процессов. Особое место в изучении вопросов и проблем образования занимает методологическая стратегия постановки и обработки результатов педагогических экспериментов. Пытаясь выявить взаимосвязи между различными признаками, подтвердить установленные закономерности, исследователь, имея, в общем-то, ограниченные ресурсы, применяет выборочный метод. И здесь с неизбежностью возникает вопрос о репрезентативности результатов, напрямую связанный с объемом выборочного массива.

Методам статистической обработки данных педагогического эксперимента посвящено достаточно публикаций различного уровня [2; 4; 9]. Однако следует иметь в виду, что, если в классической статистике исследователи оперируют большими объемами данных и при этом существует некий образец сравнения [6; 8; 12], то в педагогических экспериментах изучаемые свойства, признаки и факторы являются латентными, в лучшем случае наблюдаемыми и косвенно измеряемыми. Более того, часто экспериментальная выборка в силу объективных причин, характеризуется ограниченным числом объектов изучения [1; 9]. Причинами могут быть и специфические условия исследования, и доступность изучаемой категории выборочных единиц. Причем в такой выборке требуется выделить контрольную и экспериментальную часть, а

численность испытуемых может составлять от 5 до 10 человек, то есть представлять собой малую выборку. Закон распределения случайных ошибок в такой выборке не подчиняется нормальному распределению, и это особенность малых выборок.

Впервые с малыми выборками начал системно работать У. Госсет, специальные подходы позже предложили М. Розенблатт и Е. Парзен. В советской научной школе можно отметить Д.В. Гаскарова, В.И. Шаповалова [3]. Эти работы заложили фундамент для большинства современных методов анализа малых выборок. Вместе с тем прикладная область разработанных методов находится в сфере точных и инженерных наук. Применить разработанные алгоритмы для областей гуманитарных исследований и для педагогической науки, в частности, проблематично ввиду отсутствия однозначных рекомендаций по выбору значений используемых параметров. Отметим, что большинство статистических методов разработано еще в 30-е гг. прошлого века, когда вычислительная техника была дорога и временные ограничения на выполнение расчетов значимы. Широкое включение в арсенал современного исследователя компьютерных технологий позволяет извлекать адекватную информацию и из малого количества данных [8; 11].

Репрезентативность выборочных результатов означает соответствие заданных экспериментальных условий изучаемой эмпирической реальности или теоретической интерпретации причинной связи между независимой и зависимой переменными. В этом случае можно утверждать, что эффективность воздействия достигнута за счет нового учебно-воспитательного элемента. Объективные результаты получаются тогда, когда отбираются группы с одинаковыми показателями на основе констатирующего эксперимента. Количество испытуемых в экспериментальной работе имеет прямую связь с числом исследований, проводимых на каждом человеке, и зависит от его особенностей, но, работая с узкими специалистами или специфическими свойствами, приходится ограничиваться их количеством, доступным для изучения. У всех вероятностных методов построения репрезентативной выборки есть весьма существенный недостаток. Предполагается, что все элементы генеральной совокупности являются равнодоступными и случайными. Однако это не так. В реальных учебных группах, которые чаще всего являются единицами изучения в педагогических экспериментах, можно говорить лишь о квазислучайных выборках. Таким образом, все методы построения вероятностных выборок априори являются смещенными. Если экспериментатор осознает невозможность осуществления широкого эксперимента, то он конкретизирует задачу исследования и сужает ее до изучения реально возможного числа объектов и их характера. Таким образом, задачи эксперимента и число объектов, включаемых в него, тесно взаимосвязаны и могут влиять друг на друга.

Существуют ли методы для правильного определения числа

экспериментальных объектов как минимально необходимых? Если речь идет о проверке методики, то можно ограничиться одной экспериментальной и одной контрольной группами. При изучении проблем дидактики минимальное количество испытуемых в контрольной и экспериментальной группах должно быть уже не менее 60 человек. Это обусловлено тем, что только с такого объема выборки начинает рельефно проявляться закон больших чисел и, следовательно, достигается статистическая надежность исследования.

Какой выход возможен в планировании и проведении педагогического эксперимента в условиях ограниченных выборок? Сообразуясь с компетентностной парадигмой высшего образования, рассмотрим пример планирования, проведения и обработки результатов педагогического эксперимента, проведенного в рамках исследования эффективности процесса формирования профессионально-специализированных компетенций инженерно-технических сотрудников (ИТС) учебных лабораторий. Качество работы преподавательского состава (ПС) при проведении лабораторных работ или реализации модельных задач на тренажерах в немалой степени зависит от компетентности инженерно-технического состава. В свою очередь, эффективность деятельности ИТС определяется их способностями решать учебно-организационные задачи в полном объеме и с высоким качеством. Трудовые функции и действия сотрудников учебных лабораторий требуют специальных, в том числе педагогических, знаний и умений. Трудности социально-психологического, профессионального характера, с которыми сталкиваются специалисты ИТС учебных лабораторий образовательных организаций, на начальном этапе освоения новой для себя профессиональной деятельности обусловлены тем, что специальная подготовка таких специалистов не предусмотрена ни в одной образовательной программе разных уровней. Простое наличие в штатном расписании специалистов ИТС не обеспечивается компетентным содержанием их подготовки. Формирование профессионально-специализированных компетенций (ПСК) у сотрудников учебных лабораторий требует разработки и реализации программ повышения квалификации с адекватным организационно-педагогическим обеспечением данного процесса. Анализ публикаций по рассматриваемой задаче выявил лишь отдельные исследования, посвященные подготовке ИТС в компетентностном формате (Н.Г. Багаутдинова, Э.Р. Бариев, А.В. Коньшева) [5].

Программа экспериментальных воздействий в рамках указанного направления исследований разработана с учетом основных положений личностного, деятельностного и компетентностного подходов. Повышение квалификации организовано комплексно в очной и дистанционной форме, с расширенным блоком самоподготовки и самосовершенствования. Гипотеза экспериментальной работы основывалась на том, что повышению уровня ПСК у ИТС учебных лабораторий будут

способствовать проектирование содержания процесса формирования ПСК, организационное сопровождение самосовершенствования и методическое обеспечение оценки уровня ПСК.

На первом этапе были сформированы экспериментальная и контрольная группы (ЭГ и КГ). В ходе констатирующего эксперимента выявлены начальные уровни компетенций ИТС. Анализ результатов показал однородность экспериментальной и контрольной групп. Формирующий эксперимент проводился в естественных условиях образовательной деятельности кафедр, имеющих в своем составе ИТС (кафедры физики, общей химии, гидравлики, вычислительной математики и др.). При этом соблюдались следующие условия: сотрудники учебных лабораторий имели равноценную базовую подготовку; начало и конец эксперимента совпадали с временными рамками естественного периода обучения. В целом в эксперименте участвовало по 15 сотрудников учебных лабораторий в каждой ЭГ и КГ, то есть практические выборки представляют малые группы. В качестве интегрального критерия выступала степень овладения ИТС интегральным комплексом ПСК, выражающаяся как способность к успешному решению задач по обеспечению образовательной деятельности кафедры. Декомпозиция интегральной компетенции на частные критерии с последующей оценкой осуществлена на основе анализа реальных задач по обеспечению образовательной деятельности кафедры, решаемых сотрудниками учебных лабораторий. Получено пять групп критериев: учебно-профессиональные, научно-исследовательские, образовательно-проектировочные, организационно-технологические и профессионально-методические компетенции с контролируруемыми частными критериями. Декомпозиция интегрального критерия на частные уровни и соответствующие им проявляемые умения и владения приведены в табл. 1. Контроль выполнения плана формирования ПСК ИТС производится методом экспертной оценки качества результатов деятельности при решении плановых задач [10]. Основанием для выставления того или иного рейтинга является степень значимости достигнутых результатов по каждому виду деятельности с применением полученных знаний, умений и навыков.

Для количественной оценки экспертам предложены условные значения рейтингов результатов деятельности ИТС (< 1 ; 1 ; > 1):

> 1 – сотрудник справился с поставленной задачей и его результат оценен вышестоящим руководством (высокий уровень);

1 – сотрудник формально справился с задачей (средний уровень);

< 1 – сотрудник не справился с поставленной задачей или справился недостаточно качественно (низкий уровень).

Оценка уровня ПСК ИТС осуществлялась по специально подготовленным индивидуальным карточкам оценивания.

Таблица 1

Содержательно-критериальные уровни компетенции ИТС

Интегральный критерий: уровень профессионально-специализированной компетенции	
Частные критерии	Качественные показатели частных критериев
Умения и владения	
Учебно-профессиональная компетенция	Согласовывать планы формирования компетенций с преподавательским составом и уточнять учебно-воспитательные цели занятий. Организовывать проведение Дня открытых дверей в демонстрационной части
Научно-исследовательская компетенция	Оказать помощь в вопросах обеспечения моделирования на базе технических и программных средств кафедры. Разрабатывать предложения по применению ТСО для совершенствования способов решения научно-исследовательских задач
Образовательно-проектировочная компетенция	Разрабатывать предложения по совершенствованию учебно-лабораторной базы. Вести учетно-отчетную документацию. Разрабатывать предложения по совершенствованию аппаратных и программных средств контроля
Организационно-технологическая компетенция	Эксплуатировать тренажеры и лабораторные установки. Взаимодействовать со службами вуза в интересах обеспечения норм труда и норм потребления, а также вести учетно-отчетную документацию
Профессионально-методическая компетенция	Применять учебно-лабораторную базу для формирования знаний, навыков, компетенций. Оценить проявления уровней компетенций на занятиях в лаборатории кафедры

По каждому критерию разработаны типовые «кассетные» задания, которые выполнялись во время самоподготовки. Замеры в естественном образовательном процессе по каждому показателю для всех участников ЭГ и КГ проводились на каждом этапе экспериментальной работы в течение всего срока обучения. Полученные результаты вносили в сводные ведомости. На основании выставленных рейтинговых оценок для каждого сотрудника определялся модальный уровень рейтинга результатов по каждому виду деятельности (с позиций детализированной компетенции). В итоге получены пять выборок данных по пяти направлениям (задачам) для КГ и ЭГ, для которых проверяется статистическая гипотеза о наличии неслучайных различий между выборками. Как правило, для проверки достоверности различий между результатами ЭГ и КГ, полученных в ранговой шкале, для каждой задачи применяют критерий Пирсона. Но тут возникает методологическая проблема, связанная с наличием ячеек, содержащих значения меньше 5, иначе проявляются малые частоты, например, результаты рейтинговых уровней по выполнению задачи № 1:

Группы	<i>n</i>	> 1	= 1	< 1
КГ	15	3	8	4
ЭГ	15	3	5	7

Этот факт является ограничением для применения критерия

Пирсона. В большинстве практических рекомендаций предлагается провести операцию по укрупнению разрядов. Но в данном исследовании эта рекомендация не может быть использована, поскольку рейтинговых уровней всего три и адекватных задаче исследования возможностей для их объединения нет. Стоит заметить, что ситуация с ограниченным числом разрядов широко распространена в современных педагогических экспериментах, что обусловлено реализацией компетентного подхода в сфере образования и выделяемыми уровнями развития компетенций (чаще всего высоким, средним, низким).

Построим и оценим план эксперимента согласно классификации Дж. Кэмпбелла. В экспериментах с участием естественных групп чаще используются квазиэкспериментальные планы, которые создаются с отступлением от схем истинных экспериментов [7]. Наиболее распространенным является план для неэквивалентных групп, представленный следующей схемой (в обозначениях Дж. Кэмпбелла):

$$\frac{O1 \quad X \quad O2}{O3 \quad O4}$$

Такой план направлен на установление причинно-следственной зависимости между переменными, однако в нем отсутствует процедура уравнивания групп (рандомизация). К проведению эксперимента в рассматриваемом случае привлекаются две реальные группы. Обе группы предварительно тестируются. Затем одна группа (ЭГ) подвергается экспериментальному воздействию, а другая (КГ) – нет. Затем обе группы повторно тестируются. Результаты первого и второго тестирования обеих групп сопоставляют. Хотя различие *O2* и *O4* могут свидетельствовать о естественном развитии и/или фоновом воздействии, такие планы достаточно популярны. Этот план аналогичен плану истинного эксперимента для двух групп с тестированием до и после воздействия. Главным источником артефактов является различие в составе групп.

Если возвратиться к конкретной экспериментальной задаче, то контроль уровней ПСК ИТС производится на основе оценки выполнения комплекса заданий, ориентированных на частные критерии. Иначе – план проведенного эксперимента можно рассматривать как вариант временной серии с повторным экспериментальным воздействием. Статистический анализ результатов такого эксперимента сходен с обработкой результатов двух объединенных групп. Используя сводную ведомость рейтинговых оценок, построим обобщенную выборку данных формирующего эксперимента (табл. 2) и оценим экспериментальную гипотезу по критерию Пирсона.

Таблица 2

Сводные данные рейтинговых оценок

Группы/ критерии	<i>n</i>	< 1	= 1	> 1
Задача 1	Учебно-профессиональная компетенция			
КГ	15	3	8	4

ЭГ	15	3	5	7
Задача 2	Научно-исследовательская компетенция			
КГ	15	5	5	5
ЭГ	15	1	6	8
Задача 3	Образовательно-проектировочная компетенция			
КГ	15	4	6	5
ЭГ	15	2	5	8
Задача 4	Организационно-технологическая компетенция			
КГ	15	4	3	8
ЭГ	15	1	2	12
Задача 5	Профессионально-методическая компетенция			
КГ	15	4	7	4
ЭГ	15	3	2	10
Интегральный критерий	Уровень профессионально-специализированной компетенции			
КГ	15	20	29	26
ЭГ	15	10	20	45

Получены ячейки (две последние строки табл. 2) со значениями, пригодными для применения типового статистического критерия. Рассчитаем эмпирическое значение критерия Пирсона по сводным данным из табл. 2:

$$\chi^2 = \frac{1}{75 \cdot 75} \sum_{i=1}^3 \frac{(20 \cdot 75 - 10 \cdot 75)^2}{30} + \frac{(29 \cdot 75 - 20 \cdot 75)^2}{57} + \frac{(26 \cdot 75 - 45 \cdot 75)^2}{71} = 10,01.$$

Табличное значение критерия Пирсона при числе степеней свободы, равном 2, составляет $\chi^2_{\text{крит}} = 5,99$, а вычисленное эмпирическое, равное 10,01, позволяет отвергнуть нулевую гипотезу на уровне статистической значимости 0,05 и принять альтернативную – о достоверности различий между КГ и ЭГ по уровням профессионально-специализированных компетенций, сформированным в результате педагогических воздействий в ходе экспериментальной работы.

Таким образом, недостаточная численность единиц, составляющих частный объект, предполагает более тщательный подбор и применение известных методов, использование приемов предварительного и вторичного анализа. Экспериментальная выборка не должна в обязательном порядке составлять некий процент генеральной совокупности. В реальной практике ее объем определяется множеством различных обстоятельств, среди которых и организационные, и экономические. Строгие математические правила постановки и проведения педагогического эксперимента могут быть не столь жесткие, как в случае работы с техническими и естественнонаучными объектами. Прием обоснованного объединения данных позволяет формировать выборки нужного объема и далее пользоваться арсеналом математической статистики для анализа и оценки исследовательских гипотез. Однако «смягчение» общих принципов должно быть обусловлено спецификой исследования с указанием границ применимости как методов, так и выводов.

Список литературы

1. Алексеев Н.Г. Методические проблемы проведения массовых исследований уровня знаний учащихся: статистика и динамика // Проблемы педагогической квалиметрии. Вып. 1. М., 1973. 1. С. 28–39.
2. Архангельский С.И. и др. Вопросы измерения, анализа и оценки результатов в педагогических исследованиях. М.: Знание, 1995. 43 с.
3. Гаскаров Д.В., Шаповалов В.И. Малая выборка. М.: Статистика, 1978, 248 с.
4. Гласс Дж., Стэнли Дж. Статистические методы в педагогике и психологии. М.: Прогресс, 1978, 496 с.
5. Гузев М.С., А.А. Фасоля, Г.Ф. Утробин, Е.В. Смирнов, Роль и значение учебных лабораторий в учебно-воспитательном процессе военной образовательной организации // Мир образования – образование в мире. 2016. № 4(64). С. 151–159.
6. Лапко А.В., Шарков Н.А. Непараметрические методы обнаружения закономерностей в условиях малых выборок. // Приборостроение. 2008. № 8. С. 62–67.
7. Кэмпбелл Дж. Модели экспериментов в социальной психологии и прикладных исследованиях. М.: Прогресс, 1980. 390 с.
8. Самойленко А.П., Горбунова Е.Б. Метод восстановления плотности вероятностей прогнозируемой случайной величины по укороченной выборке данных // Нелинейный мир. 2015. № 6. С. 10–17.
9. Скалкова Я. Методология и методы педагогического исследования. М.: Педагогика, 1989. 224 с.
10. Черепанов В.С. Экспертные оценки в педагогических исследованиях. М.: Педагогика, 1989. 224 с.
11. Яковлев Е.В. Квалиметрический подход в педагогическом исследовании // Педагогика. 1999. № 3. С. 49–54.
12. Sayer A. Method in Social Science: A Realist Approach. L.: Hutchinson, 1984. 271 с.

PEDAGOGICAL EXPERIMENT: THE PROBLEM OF INSUFFICIENT VOLUME OF SAMPLE

E.V. Borisova

Tver State Technical University

The aspects of the formulation and processing of the results of pedagogical experiments with a limited number of objects of study are considered. It is shown that when working with narrow specialists or when studying specific properties, one has to be limited to the number of objects that are available to the researcher. A feature of small samples is the difference from the normal law to the distribution of random errors, which leads to a violation of the representativeness of the results. Accepting a reasonable combination of data allows you to form samples of the required volume and then use the arsenal of mathematical statistics to analyze and evaluate research hypotheses.

Key words: *sampling, experiment setting, levels of competence, data processing.*

Об авторе:

БОРИСОВА Елена Владимировна – доктор педагогических наук, профессор, профессор кафедры высшей математики ФГБОУ ВО «Тверской государственный технический университет» (170026, г. Тверь, наб. Аф. Никитина, 22), e-mail: elenborisov@mail.ru