

ПЕДАГОГИЧЕСКИЙ ЭКСПЕРИМЕНТ: СУЩНОСТЬ, СОДЕРЖАНИЕ, ОСОБЕННОСТИ ОРГАНИЗАЦИИ

Яковлев Е.В.

***Абстракт.** Рассматриваются роль и характеристические особенности педагогического эксперимента, специфика его содержания и организации при проведении научно-педагогических исследований.*

***Abstract.** Role and characteristic features of a pedagogical experiment as well as specifics of its contents and organization in pedagogical research are discussed.*

Реализация основной функции педагогического исследования – получение новых достоверных знаний об образовательном процессе – предполагает сбор, обработку и интерпретацию научной информации, что составляет основу экспериментального подтверждения ключевых теоретических положений. Обоснованность полученных выводов в полной мере определяется организационно-содержательной состоятельностью педагогического эксперимента, позволяющего в контролируемых условиях рассмотреть исследуемое явление в «чистом виде» и зафиксировать наиболее значимые факторы, влияющие на эффективность его функционирования.

Нарушение основополагающих принципов организации и осуществления педагогического

эксперимента, методологические ошибки в его планировании могут привести к выводам, не соответствующим действительности, открыть дорогу идеям, противоречащим природе педагогического процесса. Поэтому существенное значение имеет грамотная реализация экспериментальных процедур, которая обеспечивается компетентностью исследователя в области педагогического эксперимента, предполагающей, прежде всего, понимание сути и значения экспериментального процесса, его этапов и требований к организации. Еще большую актуальность проблема осуществления экспериментальной деятельности в рамках научного исследования приобретает в связи с направленностью на изучение педагогических явлений – сложных, многоаспект-

ных, со слабо диагностируемыми и, как правило, не поддающимися однозначному оцениванию свойствами. В такой ситуации роль и значение эксперимента выходит на лидирующие позиции в любом педагогическом исследовании, а его адекватная реализация в полном соответствии с методологической природой и механизмами использования становится важнейшей задачей научного поиска.

Эксперимент как метод исследования пришел в социальные науки из естественных примерно в XVII в., став со временем основным способом проверки научных данных. Существенное влияние на становление теории педагогического эксперимента оказали работы в области теории познания (П.В. Алексеев и А.В. Панин, В.В. Ильин, К. Поппер и др.), гуманитарных наук (В.Г. Ананьев, В.Г. Афанасьев, Н. Стефанов и др.), методологии научного знания (В.И. Купцов, В.А. Штофф, Э.Г. Юдин и др.), теории эксперимента (В.С. Демин, В.И. Кефели, Т.В. Корнилова, Д. Кэмпбелл, Д. Мартин, В.В. Налимов и Т.И. Голикова, Г.В. Осипов, В.Ф. Папуловский, В.П. Пригода, Ч. Хикс и др.). Проблема педагогического эксперимента рассматривалась в работах Ю.К. Бабанского, В.И. Журавлева, В.И. Загвязинского, А.С. Казаринова, А.И. Кочетова, Н.В. Кузьминой, В.И. Михеева, Д.А. Новикова, А.И. Пискунова, Я. Скалкой, М.Н. Скаткина, В.С. Черепанова, Е.В. Яковлева и др. Однако, несмотря на широкий интерес к данной проблеме и бесспорные успехи в разработке ее теоретических положений, в настоящее время педагогическая наука и практика не имеет однозначного и общепризнанного понимания методологических основ педагогического эксперимента, путей и способов его эффективной реализации.

В общенаучном понимании эксперимент – это «исследование каких-либо явлений путем активного воздействия на них при помощи создания новых условий, соответствующих целям исследования, или же через изменение течения процесса в нужном направлении» [9, 558]. Исходя из целей педагогического исследования, более технологичным нам представляется определение, данное В. Окном, согласно которому эксперимент – это «метод научного исследования, основанный на инициировании некоторого процесса (явления) или воздействия на данный процесс, на таком регулировании этого процесса, которое

позволяет его контролировать и измерять, а также верифицировать принятые гипотезы» [7, 27].

Определяя место эксперимента на пересечении практики и познания и выделяя в нем признаки рационального и чувственного, эмпирического и теоретического, важно рассмотреть его объективную и субъективную стороны. Объективная сторона находит отражение в сущности и содержании эксперимента, а субъективная – проявляется, прежде всего, в деятельности познающего субъекта. В частности, к субъективной стороне относятся: «1) особенности органов чувств человека, воспринимающих информацию, полученную в процессе наблюдения; 2) теоретические способности и вообще деятельность логического мышления; 3) уровень научных знаний и духовной культуры; квалификация и способности экспериментатора; 4) поставленные и сформулированные цели и задачи эксперимента; 5) сама деятельность экспериментатора» [11, 87].

Как справедливо отмечает В.А. Штофф [11, 89], основные характеристики эксперимента требуют, во-первых, изолировать исследуемое явление от влияния побочных, несущественных и затемняющих его сущность влияний и изучить его в чистом виде; во-вторых, многократно воспроизводить ход процесса в строго фиксированных, поддающихся контролю и учету условиях; в-третьих, планомерно изменять, варьировать, комбинировать различные условия в целях получения искомого результата.

Проведенный нами анализ научной литературы показал, что общий ход эксперимента в социальной сфере определяют следующие инвариантные процедуры: установление выборки, репрезентативной для данной генеральной совокупности; подбор однородных групп или пар испытуемых; выбор конкретных методик, способов и параметров измерения экспериментальных данных; проверка доступности и эффективности методик на небольшом числе испытуемых; определение признаков, по которым можно судить об изменениях в экспериментальном объекте под влиянием соответствующих воздействий; определение времени и длительности проведения эксперимента; его проведение; качественный и количественный анализ результатов эксперимента; их интерпретация; указание границ примене-

ния проверенной в ходе эксперимента системы мер.

Опираясь на исследование Ч. Хикса [10], мы выделяем три основных этапа в проведении экспериментальной работы:

формализация проблемы – постановка задачи, выбор отклика, варьируемых факторов, уровней этих факторов, подбор сочетаний уровней факторов;

планирование – определение необходимого числа наблюдений, порядка проведения эксперимента, используемого метода рандомизации, математической модели описания эксперимента;

анализ – сбор и обработка данных, вычисление статистики для проверки гипотез, интерпретация результатов.

Качество осуществления каждого этапа очень важно для эксперимента в целом. Общеизвестно, что неудовлетворительное планирование и проведение эксперимента невозможно исправить никакими методами обработки данных. В практическом отношении планирование эксперимента является мощным инструментом, позволяющим сократить число измерений, увеличить объем новой получаемой информации, уменьшить вероятность принятия ошибочных решений. При этом само по себе планирование не может улучшить качества построенной модели, оно лишь позволяет улучшить ее статистические свойства, то есть дает возможность точнее оценить параметры имеющейся модели. Следовательно, даже самое тщательное планирование не принесет желаемых результатов при неверно осуществленной формализации задачи.

Эксперимент в педагогике, имея свои особенности, в целом подчиняется общим методологическим принципам организации и проведения научного эксперимента. Однако существует и определенная специфика, обусловленная особенностями педагогического процесса и его субъектов:

высокий уровень сложности исследуемых педагогических объектов, детерминирующих многообразие связей и многофакторность явных и скрытых взаимовлияний, которые требуют учета и оценивания в процессе научно-педагогического исследования;

активное участие экспериментатора в развертывании педагогического процесса и его прямое вмешательство в развитие исследуемых явлений;

зависимость оценки результатов педагогического эксперимента от морально-нравственных установок в обществе, ценностных ориентаций педагогического сообщества и самого экспериментатора;

инновационный характер педагогического эксперимента, который является частью нововведений в целенаправленной преобразовательной деятельности педагога;

масштабность педагогического эксперимента.

Не давая строгого определения, под педагогическим экспериментом будем понимать комплекс методов исследования, предназначенный для объективной и доказательной проверки педагогической гипотезы. Отметим, что, являясь исследовательским ядром эксперимента, гипотеза считается научно состоятельной, если отвечает требованиям простоты, точности, непротиворечивости и нетривиальности формулировки, принципиальной проверяемости (доказуемости), теоретической обоснованности. В большинстве экспериментов имеется не одна, а несколько иерархически соподчиненных, последовательно уточняющих друг друга гипотез, что в определенной степени способствует усилению достоверности полученных данных и помогает исследователю определить направление развертывания и содержание педагогического эксперимента в целом.

Сложность и многоаспектность педагогических явлений, а также разнообразие способов теоретического и эмпирического оперирования ими определяют видовую неоднородность педагогического эксперимента. В таблице представлены наиболее распространенные его классификации.

Таблица

Классификации педагогических экспериментов

Основания классификации	Виды экспериментов
Условия проведения	Естественный
	Искусственный
Цели исследования	Преобразующий
	Констатирующий
	Контролирующий
	Поисковый и др.
Количество факторов	Однофакторный
	Многофакторный
Степень контролируемости фактора	Активный
	Пассивный (регистрирующий)

Как показал анализ научной литературы, ключевыми характеристиками педагогического эксперимента являются его надежность и валидность. Надежностью называют свойство устойчивости получения данных, воспроизводимости результатов измерений, повторяемых при идентичных условиях [8, 207]. Валидность – это критерий адекватности эксперимента, характеризующий насколько точно он оценивает необходимые параметры, насколько составляющие его пробы адекватны исследуемой проблеме [8, 40].

Раскроем подробнее эти две характеристики и начнем с надежности.

Целью любого измерения является установление истинного значения измеряемой величины, т.е. существующей в действительности неискаженной величины признака, присущего данному индивиду. Однако в философии хорошо известен постулат о неизбежности погрешности измерения. Как отмечает Л. Яноши, «результат эксперимента всегда содержит ошибку, как бы тщательно ни проводились измерения» [12, 15]. Принятие этого постулата неизбежно приводит к одному из основных положений теории измерения – к тезису о невозможности знания абсолютного значения измеряемой величины. Исследователь располагает только наблюдаемым результатом измерения, искаженным под влиянием различных факторов. Можно выделить множество факторов, искажающих истинное значение измеряемой величины, например, влияние цели, задач и характера исследования, условия тестового опроса и т.д.

Итак, всякое измерение всегда содержит ошибку. Значит, задача исследователя – понизить эту ошибку до приемлемого уровня. Но точность измерения составляет только одну сторону понятия надежности. Другой важной характеристикой надежности является возможность повторить измерения с тем же результатом.

Перейдем теперь к более обширному понятию валидности. Б.П. Битинас [2] выделяет содержательную, конструктивную, конкурентную, прогностическую и перекрестную валидность.

Содержательная валидность методики измерений определяется тем, что составляющие ее признаки охватывают те области информации и деятельности, в которых проявляется измеряемое свойство. Содержательная

валидность не оценивается количественно и достигается путем составления качественной схемы сбора необходимой информации.

Конструкционная валидность указывает на то, в какой степени диагностическая методика отражает то составное свойство, для определения которого оно предназначено, и определяется, во-первых, теоретическим содержанием признаков, во-вторых, их внутренней согласованностью. Конструкционная валидность требует как можно более полного накопления информации об измеряемом свойстве.

Конкурентная валидность определяется с помощью внешнего критерия, по которому информация собирается одновременно с проверяемой методикой. Она показывает возможность замены одной диагностической методики другой.

Прогностическая валидность определяется также по внешнему критерию, но информация по этому критерию собирается некоторое время спустя после применения проверяемой методики.

Перекрестная валидность выявляется сопоставлением результатов измерения, проведенных на некоторой выборке, с результатами измерения в новой выборке испытуемых из этой же генеральной совокупности.

Приведем в заключение пример из монографии В.И. Михеева [5, 20]. «Соотношение между надежностью и валидностью лучше представить аналогией с часами: часы могут иметь точный (надежный) ход, но, будучи поставленными на неверный час, они непригодны (невалидны) для получения ответа на главный вопрос: который час».

Общие положения, на которые следует ориентироваться при организации эксперимента, в определенной степени разработаны в конкретных приложениях. Прежде всего, это выявление и обоснование критериев, а также определение уровней проявления исследуемого качества у участников эксперимента. В современных источниках под критерием понимают показатель, признак, на основе которого формируется оценка качества объекта, процесса; мерило такой оценки [3, 144]. Определение критериев и уровней их проявления определяется не только целью, предметом и содержанием педагогического исследования, но и уровнем развития диагностического аппарата, а также степенью сформированности исследовательской компетентности у са-

мого экспериментатора. При этом выбранный критерий должен быть достаточным для оценивания исследуемого качества, полным, надежным и легко проверяемым [13].

В некоторых случаях выбранный критерий может иметь комплексный характер. Поэтому для описания его содержательного наполнения исследователями строятся так называемые «критериально-целевые модели», отражающие характеристику формируемого и оцениваемого качества, критерии и показатели, по которым производятся оценивание, уровни их проявления, а также перечень необходимых для этого диагностических методик.

Второй важной проблемой эксперимента является отбор и распределение его участников по группам. При ее решении используется следующая терминология.

Генеральная совокупность – это все потенциальные испытуемые, которые могут быть привлечены к участию в педагогическом эксперименте, т.е. это вся совокупность объектов, из которой делается выборка для проведения исследования.

Выборка – это множество участников педагогического эксперимента.

Репрезентативность (от франц. *Represtatif* – показательный, характерный) – это свойство выборки, заключающееся в ее способности адекватно представлять состояние генеральной совокупности.

Поскольку данные, полученные в ходе эксперимента по выборке, будут распространяться на всю генеральную совокупность, то выборка должна, во-первых, охватывать все возможные варианты проявления исследуемого качества у испытуемых, во-вторых, иметь такое же распределение по уровням его проявления, как и во всей совокупности. Кроме того, состав участников в выборке должен соответствовать цели и содержанию исследования.

Нетривиальным является также вопрос об объеме выборки при проведении эксперимента. Например, для социологических исследований пределы выборочной совокупности колеблются от 350 до 3000 наблюдений [4, 99]. В педагогических исследованиях, как считает Ю.К. Бабанский, выделение типичных по определенным показателям контрольных и экспериментальных объектов позволяет ограничиться 250–350 наблюдениями. Отметим, что при отсутствии строгих требований к определению объема выборки, тем не менее, в

научной литературе сформулированы общие рекомендации по обеспечению необходимости и достаточности ее наполнения: 1) при сравнении двух групп данных их общая численность должна быть не менее 50 человек; 2) для снижения статистической ошибки при обработке данных количество участников эксперимента в каждой группе должно быть от 30 до 35; 3) численность испытуемых в группах должна быть примерно одинаковой; 4) чем выше изменчивость исследуемого свойства, тем больше должен быть объем выборки.

Дальнейшие действия исследователя связаны с группировкой участников эксперимента. При этом необходимо добиться, чтобы получившиеся группы испытуемых были сравнимы по оцениваемому параметру. Например, если в ходе эксперимента исследователь занимается формированием у студентов определенных знаний, то в формируемых группах распределение по уровням этих знаний на начало эксперимента должно быть статистически неразличимым. При этом степень воспитанности студентов, их ценностные ориентации или способности для данного исследования не будут иметь значения, поэтому по всем прочим показателям (кроме формируемых знаний) распределение студентов в группах может иметь даже существенные различия.

Как правило, при формировании групп возникает необходимость выравнивания начальных параметров. Для практического решения этой задачи существует несколько методик.

Точечное выравнивание заключается в прямом сопоставлении испытуемых. Из выборки берется участник с определенными характеристиками в одну группу, для него выбирается другой участник с аналогичными свойствами в другую и т.д. Такой способ выравнивания предпочтителен для экспериментирования с малыми группами.

Частотное выравнивание сводится к отбору участников в группы в зависимости от процентного содержания проявления исследуемого качества в генеральной совокупности в целом. Например, если во всей совокупности половина состава имеет низкий уровень проявления заданного качества, то такая же пропорция должна быть сохранена в каждой группе.

Однако на практике использование данных методов выравнивания в педагогическом эксперименте не всегда возможно. Поскольку он имеет, как правило, естественный характер,

связанный с работой в реальных условиях образовательного процесса, то используется метод случайного выбора, при котором принять участие в эксперименте могут любые испытуемые или группы. Исследователь выбирает уже функционирующие академические группы (одинакового возраста, уровня развития и т.д.), и если они неэквивалентны по исследуемому параметру, то либо проводится специальная подготовительная работа по «дотягиванию» участников эксперимента до нужного уровня, либо при анализе не учитываются крайние результаты, полученные в ходе эксперимента.

Самостоятельной проблемой в организации и проведении эксперимента является его оптимизация. Ю.К. Бабанский [1], рассматривая проблемы повышения эффективности педагогических исследований, отмечает, что педагогический эксперимент следует проводить с минимально необходимым расходом времени, усилий и средств. Оптимизации, как правило, подвергаются количество экспериментальных объектов и длительность проведения эксперимента. Очевидно, что максимизация этих параметров усиливает объективность выводов, однако в то же время требует больших затрат. Данная проблема пока остается открытой в общем случае и решается в каждом конкретном исследовании индивидуально.

Оптимальная организация педагогического эксперимента предполагает его соответствие определенным критериям, которые подразделяются на статистические и динамические. К статистическим относятся требования, которым должно удовлетворять некоторое задаваемое планом эксперимента расположение экспериментальных точек в пространстве факторов (независимых переменных), подлежащих варьированию. К динамическим – выраженные математическим языком высказывания, которые представляются правомерными на уровне интуитивных представлений [6,7–8].

Перечислим наиболее распространенные критерии оптимальности эксперимента. Первая группа критериев относится к статистическим критериям, связанным с точностью оценок параметров. D-оптимальность направлена на уменьшение рассеяния оценок параметра. A-оптимальность – на уменьшение средней дисперсии оценок. E-оптимальность – на ограничение дисперсий отдельных параметров.

Вторая группа критериев представляет собой статистические критерии, связанные с ошибкой в оценке модели. G-оптимальность направлена на минимизацию максимального значения дисперсии оценки. В G-оптимальных планах нет точек, в которых точность оценки слишком низкая. Q-оптимальность направлена на минимизацию средней дисперсии оценки. Ротабельность – на упорядочение плана эксперимента, в результате которого все направления становятся равнозначными с точки зрения точности оценок. Униформность требует, чтобы дисперсия оценки была постоянной в некоторой области вокруг центра эксперимента.

К важнейшим характеристикам оптимальных планов относится следующая группа показателей, которые с некоторой натяжкой также можно отнести к критериям оптимальности. Эти характеристики в меньшей степени затрагивают специфические понятия математической статистики и более удобны в работе.

Рандомизация состоит в случайном порядке проведения эксперимента. Рандомизация производится для исключения влияния переменных, неконтролируемым образом изменяющихся во времени или пространстве. Этот критерий традиционно имеет большое значение в реальных экспериментах. Поэтому многие авторы строят теорию планирования эксперимента, отталкиваясь от идеи рандомизации.

Насыщенность плана характеризуется количеством проведенных измерений. Если число измерений равно числу неизвестных параметров, которые нужно оценить, то план называется насыщенным. Очевидно, планы с меньшим числом измерений не позволяют найти единственные оценки всех параметров.

Композиционность отражает структуру проведения эксперимента. Композиционный план позволяет разделить эксперимент на несколько этапов и постепенно переходить от простых моделей к более сложным.

Простота обработки относится к косвенным характеристикам и отражает свойства математической модели. Чем она проще, тем меньше вычислений и меньше их сложность. При отсутствии специальных средств обработки числовой информации этот показатель часто становится одним из самых важных.

Мы не будем дальше продолжать список критериев оптимальности. Развитие теории планирования эксперимента происходит быстрыми темпами, и, как следствие, появляются

новые критерии, предназначенные для решения более частных проблем. При этом зачастую невозможно совместить различные критерии. План оптимальный с точки зрения одного критерия может оказаться неудовлетворительным с точки зрения другого. Построить планы, удовлетворяющие одновременно многим критериям оптимальности, удастся только для отдельных моделей. Обычно приходится идти на компромисс, выбирая план, близкий к оптимальному, по нескольким критериям. Проблемы построения таких планов в общем случае довольно сложны и пока недостаточно разработаны. Однако педагога в первую очередь интересует не сама по себе теоретическая модель, а выводы, которые позволяет сделать ее изучение. Отсюда больший интерес не к абстрактным характеристикам планов, а к перспективам их применения. В этой связи на первое место выходят те планы, которые позволяют уменьшить вероятность ошибки при проверке гипотез.

Таким образом, педагогический эксперимент, являясь обязательным компонентом научно-педагогического исследования, выступает важнейшим средством подтверждения полученных теоретических выводов и требует тщательного планирования.

СПИСОК ЛИТЕРАТУРЫ

1 Бабанский, Ю. К. Проблемы повышения эффективности педагогических исследований:

Дидактический аспект. – М.: Педагогика, 1982. – 182 с.

2 Битинас, Б. П. Многомерный анализ в педагогике и педагогической психологии. – Вильнюс: Науч.-исследоват. ин-т школ Мин-ва просвещения Лит ССР, 1971. – 374 с.

3 Вишнякова, С. М. Профессиональное образование: Словарь. Ключевые понятия, термины, актуальная лексика. – М.: НМЦ, 1999. – 538 с.

4 Здравомыслов, А. Г. Методология и процедура социологических исследований. – М.: Мысль, 1969. – 205 с.

5 Михеев, В. И. Моделирование и методы измерений в педагогике: научн.-метод. пособие. – М.: Высшая школа, 1987. – 200 с.

6 Налимов, В. В., Голикова, Т. И. Логические основания планирования эксперимента. – М.: Металлургия, 1976. – 128 с.

7 Оконь, В. Введение в общую дидактику / пер. с польск. – М.: Высшая шк., 1990. – 382 с.

8 Психологический словарь / под ред. В. В. Давыдова, А. В. Запорожца, Б. Ф. Ломова и др. – М.: Педагогика, 1983. – 448 с.

9 Философский словарь / под ред. И. Т. Фролова. – М.: Политиздат, 1987. – 590 с.

10 Хикс, Ч. Основные принципы планирования эксперимента. – М.: Мир, 1967. – 406 с.

11 Штофф, В. А. Проблемы методологии научного познания. – М.: Высшая школа, 1978. – 271 с.

12 Яноши, Л. Теория и практика обработки результатов измерений. – М.: Мир, 1965. – 462 с.

13 Goodenough J.B., Gerhart S.L. Toward a Theory of Test Data Selection. IEEE Transactions on Software Engineering, 1975, SE-1, № 2, p. 156-193.