

**Всегда ли воспроизводимость важна и возможна
для научного эксперимента?**

© 2021 г. В.С. Пронских

Институт философии РАН, Москва, 109240, ул. Гончарная, д. 12, стр. 1.

E-mail: vpronskikh@gmail.com

Поступила 24.03.2021

Настоящая статья представляет собой расширенный комментарий к трем книгам Р. Лаймона и А. Франклина, посвященным методологии и эпистемологии научного эксперимента, а также их обзорной статье по вопросу воспроизводимости в эксперименте, основанной на этих книгах. Воспроизводимость научных результатов исторически считалась одним из методологических стандартов науки, что было связано с представлениями об истинности и intersubjectности результатов научного познания. Проблема воспроизводимости привлекла к себе особое внимание в последние десятилетия, когда специализированные исследования выявили, что более половины измерений в социальных науках не может быть воспроизведено; вскрылось множество случаев мошенничества в биомедицинских науках; коллективный характер субъектности в физике элементарных частиц акцентировал нестабильность получаемого большими коллаборациями знания. В статье осмысливаются дискуссии о воспроизводимости в философской литературе, что позволяет наметить различия между реплицированием эксперимента как повторением, максимально приближенным к изначальному, и собственно воспроизведением как повторным получением ранее наблюдаемого явления в существенно видоизмененной теоретико-инструментальной постановке. Также вводится понятие о реплицировании-2 как промежуточной форме между реплицированием и воспроизведением. Эти формы повторения исследования выполняют различные функции в экспериментальной практике. Показано, что разнообразные по форме способы реплицирования и воспроизведения лежат в основе набора эпистемических стратегий – методологических стандартов эксперимента, выделенных Франклином на основе многолетнего изучения научной практики. Проанализирован ряд экспериментов, в которых оказывалось достаточно однократного измерения в отсутствие эпистемических стратегий для принятия сообществом новой теории, и утверждается, что в описанных случаях теория, основанная на высокоценных принципах симметрии, оказывалась доминирующей оптикой сообщества, тогда как эксперимент выполнял лишь демонстрационную роль. Подобные примеры, на наш взгляд, указывают, что в ситуации смены научных парадигм роль эксперимента отлична от таковой в нормальной науке: требования к воспроизводимости и эпистемическим стратегиям в ней значительно ослаблены.

Ключевые слова: научный эксперимент, intersubjectность, реплицирование, воспроизводимость, эпистемические стратегии, трансцендентализм, парадигма, нормальная наука, физика элементарных частиц.

DOI: 10.21146/0042-8744-2021-8-103-115

Цитирование: *Пронских В.С.* Всегда ли воспроизводимость важна и возможна для научного эксперимента? // Вопросы философии. 2021. № 8. С. 103–115.

Is Reproducibility Always Important or Even Possible for a Scientific Experiment?

© 2021 Vitaliy S. Pronskich

*Institute of Philosophy, Russian Academy of Sciences,
12/1, Goncharnaya str., Moscow, 109240, Russian Federation.*

E-mail: vpronskikh@gmail.com

Received 24.03.2021

This article provides an extended commentary on three books by R. Laymon and A. Franklin about the methodology and epistemology of the scientific experiment, as well as their article on the issue of reproducibility of experiments. The reproducibility of scientific results has historically been considered one of the methodological standards of science, and it is associated with ideas about the truth and intersubjective nature of scientific knowledge. The problem of reproducibility has received particular attention in recent decades because specialized studies have revealed that more than half of the results from the social scientific studies cannot be reproduced; many cases of fraud in biomedical sciences have been uncovered; and the collective nature of subjectivity in elementary particle physics has accentuated the instability of the knowledge obtained by large collaborations. In reconstructing discussions about reproducibility in the philosophical literature, we distinguish between replicating an experiment by repeating it in a way that is as close as possible to the original and actually reproducing it by re-obtaining a previously observed phenomenon in a significantly modified instrumental-theoretical setting. We also introduce the concept of replication-2 as an intermediate form between replication and reproducing. These kinds of research repetitions perform different functions in experimental practice. We show that a variety of kinds of replication and reproduction are at the heart of a set of epistemic strategies: experimental methodological standards identified by Franklin based on decades of research in scientific practice. We analyze a number of experiments in which a single measurement, in the absence of epistemic strategies, was sufficient for the community to accept a new theory. In these cases, we argue, a theory based on high-value symmetry principles turned out to be the dominant lens of the community, while the experiment played a role only as a demonstration. Such examples, in our opinion, indicate that the experiment's role in a situation of shifting scientific paradigms is different from its role in normal science: the requirements for reproducibility and epistemic strategies are significantly alleviated in the former in comparison to the latter.

Keywords: scientific experiment, intersubjectivity, replication, reproducibility, epistemic strategies, transcendentalism, paradigm, normal science, particle physics.

DOI: 10.21146/0042-8744-2021-8-103-115

Citation: Pronskich, Vitaliy S. (2021) "Is Reproducibility Always Important or Even Possible for a Scientific Experiment?", *Voprosy filosofii*, Vol. 8 (2021), pp. 103-115.

Введение

В науках, в первую очередь естествознании, а также и в других областях, опирающихся на количественный анализ эмпирических данных, воспроизводимость эксперимента, то есть возможность как для того же ученого, так и для любых других индивидов

повторить некоторый опыт и при этом получить те же результаты, что и прежде, всегда молчаливо считалась методологическим стандартом. Она явно или неявно связывалась в науке с представлениями об истинности результатов естествознания, а также их intersubjectности, обеспечивающей значимость познавательного результата одного индивида для остальных. Любые отклонения от воспроизводимости результатов относились либо к невольным ошибкам экспериментатора, либо к осознанным нарушениям методики эксперимента и его научно обоснованных интерпретаций. Однако представления о том, что воспроизводимость, по крайней мере в принципе, – отличительная особенность научного метода, долгое время не подвергались сомнению. Как отмечает Б.И. Пружинин, «наши индивидуальные соображения по поводу устройства мира и его отдельных фрагментов (можно сказать, наши мнения о мире) становятся научным знанием только в том случае, если существует метод, позволяющий эти соображения (в принципе) воспроизвести» [Пружинин 2014, 136]. А.Ю. Сторожук связывает отрицательный результат попыток воспроизведения с истинностью теории, а положительный – с правдоподобием суждения (см.: [Сторожук 2008]).

Довольно неожиданным стал анализ, проведенный Open Science Collaboration [Franklin, Laymon 2021], продемонстрировавший, что из попыток коллаборации воспроизвести результаты ста эмпирических исследований в области психологии, опубликованных в трех ведущих психологических журналах, удачными оказались менее половины. Эти результаты вызвали бурную полемику в научном сообществе и привлекли серьезное внимание к воспроизводимости и особенностям экспериментирования в социогуманитаристике. Иную природу имел ряд скандалов, разразившихся в последние десятилетия в связи с обнаружением подделок и фабрикацией результатов в биомедицинских науках, однако он уже был явно связан с умышленными действиями, направляемыми социальными интересами исследователей, и обострившейся в последнее время конкуренцией между ними (борьбой за гранты, пожизненные позиции в университетах и высшие научные награды). Отличный от вышеперечисленных характер имели несколько ситуаций с невозможностью результатов в физике элементарных частиц. Среди них – замечательно проанализированная А. Франклином (подробнее в [Pronskikh 2015]) история открытия, «закрытия» и переоткрытия пентаварка. Они неявно указывали на изменившийся характер экспериментирования в современной мегасайенс, вызвавшей к жизни феномен коллективного познания в научных коллаборациях. Онтологические отличия коллективного экспериментатора от экспериментатора прошлого, нестабильный, аморфный и осциллирующий характер познающего субъекта в мегасайенс были акцентированы Галисоном (см.: [Галисон 2018]) и могут, на мой взгляд, служить объяснением нестабильности получаемого коллаборациями знания. Таким образом, специфика области знания (обилие неявных допущений в социогуманитарном экспериментировании), неконтролируемый рост влияния социальных интересов на научные практики и онтологические изменения в коллаборационной субъектности привели к тому, что воспроизводимость экспериментов в различных областях за последние десятилетия стала проблематичной.

Осмысливая проблему воспроизводимости, известные исследователи в области эпистемологии эксперимента А. Франклин и Р. Лаймон в 2018–2021 гг. опубликовали три книги о научном эксперименте. В первой из них подробно обсуждаются ситуационные исследования, в которых экспериментальные результаты либо воспроизводились (бозон Хиггса), либо не могли быть воспроизведены (пентаварк), а также связанные с воспроизводимостью экспериментальные проблемы [Franklin 2018]. Во второй осмысливаются эксперименты, как классические (галилеевские), так и современные, где экспериментаторы применяли различные техники и повышали точность измерений, но при этом согласованно получали отрицательные результаты (например, поиски пятой силы, эфирного ветра или суперсимметричных частиц на коллайдерах) [Franklin, Laymon 2019]. Третья книга посвящена таким экспериментам, в которых даже одиночного измерения в отсутствие какого-либо воспроизведения становилось достаточно, чтобы сообщество приняло некоторую теорию: например, открытие позитрона, экспе-

рименты Менделя с наследственностью или открытие омега-минус-гиперона [Franklin, Laymon 2021].

В публикуемых в настоящем и последующем номерах «Вопросов философии» текстах Франклин и Лаймон пытаются кратко суммировать свои идеи. Задачу настоящей статьи-комментария я вижу в том, чтобы, не пересказывая конкретные сюжеты, ввести их рассуждения в философско-методологический контекст, акцентируя внимание на истории дискуссий о воспроизводимости и основаниях эпистемических стратегий экспериментирования.

Возможны ли полностью идентичные эксперименты?

Проблемы воспроизводимости эксперимента долгое время не привлекали внимания философов. Первой заметной дискуссией на эту тему стала таковая между А. Франклином, Г. Коллинзом и Н. Картрайт. А. Франклин и его соавтор К. Хаусон выступили с обоснованием того, что воспроизведение (реплицирование) некоторого эксперимента тем же самым способом менее эпистемически значимо, чем воспроизведение его несколькими отличающимся способом [Franklin, Howson 1984]. Иными словами, если экспериментальная установка¹ основана на инструментальных теориях T , то повторение эксперимента на другой установке, основанной на теориях T' , придаст результату больше достоверности. Аргументы Франклина и Хаусона были двойки. С одной стороны, они приводили логические доводы, основывая свои умозаключения на формулах вероятности Байеса, показывая, что вероятность результата, полученного на приборах, опирающихся на T и T' , выше, чем для двух одинаковых результатов, измеренных на приборах с T . С другой стороны, авторы изучили научную практику экспериментаторов в физике высоких энергий в историческом ракурсе в той мере, в которой она изложена в ведущих научных журналах, и обнаружили, что физики на практике придают больший вес усредненным результатам, которые получены на установках различного принципа действия (T , T' , T''), например, на пузырьковых (водородной и дейтериевой) и искровой камерах. Столь же различающимися считаются результаты, измеренные на одной установке (при условии, что различные наборы данных с нее обрабатывались по разным алгоритмам, а затем усреднялись). При этом, например, если различные наборы данных с одной и той же установки анализировались единообразно, но разными группами аналитиков (как правило, незнакомых друг с другом), то такие результаты Франклин и Хаусон относили к повторениям одной и той же процедуры. Эксперименты, идентичные друг другу, авторы назвали гераклитовыми.

Гарри Коллинз (яростный оппонент Франклина на протяжении многих лет) уточнил, что полностью идентичные эксперименты не имеют смысла для подтверждения результата, поскольку, в силу их неразличимости, не добавляют нового знания при совпадении результатов [Collins 1984]. Эксперимент может подтвердить предыдущий, только если они различаются в своей постановке². В случае же несовпадения результатов различающихся по реализации экспериментов интерпретация результатов, по мнению Коллинза, требует от экспериментаторов слишком большого объема знаний. Экспериментаторам необходимо будет убедиться, что все части их установок и теории, на которых они основаны, соответствуют друг другу, а также в том, что в обоих случаях исследуются одни и те же явления. Ссылаясь на одну из первых критических работ на эту тему [Medawar 1963], Коллинз утверждает, что физические экспериментальные работы описывают методику постановки экспериментов не настолько полно и корректно, чтобы сделать возможным их независимое повторение другими авторами. Оказывается, ни реальная логика экспериментирования, ни особенности используемых инструментальных теорий, как правило, не фиксируются в журнальных публикациях, в силу чего буквально воспроизвести эксперимент становится практически невозможным. Иначе говоря, полностью идентичные эксперименты не только не имеют смысла, но и невозможны. Недостаток же в статьях важной для реплицирования информации зачастую не только связан с утаиванием ее по коммерческим причинам

в прикладных исследованиях (см.: [Пружинин 2014, 138]), но и является принципиальным, в том числе, в исследованиях фундаментальных.

Даже если исследователь технически полно воспроизведет некоторое измерение, вне его внимания остается множество допущений, например, в части общетеоретических положений, таких как инвариантность сдвигов во времени и пространстве. Таким образом, по мнению Коллинза, не может быть полностью идентичных, гераклитовых экспериментов. При этом он не отрицает, что для опровержения результатов некоторого эксперимента (в отличие от его подтверждения) максимально точное повторение его постановки может оказаться важным. Само представление о том, что считать тем же или другим по реализации экспериментом, существенно зависит от конвенций об инструментальных теориях и теориях явления, а желание экспериментаторов варьировать эксперименты во многом определяется системой вознаграждений: полное повторение эксперимента редко удостоивается высоких научных наград. Представления ученых о сходстве и различиях экспериментов исторически изменчивы, а кроме того, важно учитывать не только установки, но и экспериментальную систему в целом, которая включает и экспериментатора. Комментарий Коллинза был не столько полемическим по отношению к идеям Франклина, сколько поясняющим, почему ученые предпочитают повторять эксперименты именно отличающимися способами, то есть он исходил как из эпистемологических (теоретические допущения), так и социологических (интересы ученых) положений.

Реплицирование и воспроизведение экспериментов

Еще большую ясность в дискуссию внесла Н. Картрайт, сделавшая акцент на необходимости различать реплицирование как реализацию всех процедур повторяемого исследования способом, максимально сходным с таковым в исходном эксперименте, и воспроизведение, означающее повторение результата с применением принципиально иной методики и (или) процедуры обработки данных (см.: [Cartwright 1991]). Отдельно Картрайт выделила рутинную проверку имплементации исходного исследования, что, очевидно, может выражаться в повторении эксперимента на исходной установке и при исходных же условиях. Последнее, как мы полагаем, имеет много сходств с так называемым набором статистики, то есть сбором данных от повторяющихся измерений, что имеет своей целью определение и увеличение точности измерения (и требует только отмеченных Коллинзом допущений о пространственно-временных симметриях). Задача реплицирования, по мнению Картрайт, состоит в проверке («корректно ли работает прибор»), хотя по ее мнению, оно не может помочь далеко продвинуться в вопросе о том, корректна ли теория прибора в целом; оно может лишь помочь удостовериться, что некоторое взаимодействие прибора с объектом исследования корректно реализовано. Полноценная же проверка инструментальной теории возможна только в ходе воспроизведения измерения, то есть измерения того же явления на установке, основанной на других инструментальных принципах (что согласуется с выводами Коллинза).

Несколько иной взгляд на реплицирование, которое в данном случае будет переходной формой между реплицированием и воспроизведением (мы назовем реплицированием-2), предлагает Питер Галисон. Он указывает на варьирование рабочих параметров уже работающей установки, как на одну из эпистемических стратегий, используемых экспериментаторами, чтобы удостовериться в надежности экспериментального результата и его устойчивости [Galison 1987]. Задачей такого варьирования он считает попытку «избавиться» от эффекта: поскольку эффект наблюдается только при некотором сочетании настроек, а не в широком рабочем диапазоне, постольку есть большая вероятность того, что эффект является аппаратным артефактом, а не физическим явлением. Более того, Галисон считает эту стратегию связанной с критерием завершения эксперимента, так как утверждает, что экспериментаторы продолжают измерение некоторого эффекта до тех пор, пока не убеждаются, что тот не может внезапно исчезнуть при некотором наборе параметров (притом что отсутствуют

инструментально-теоретические основания ожидать каких-либо изменений в наблюдаемом явлении).

Есть пример, иллюстрирующий вышеприведенные различия в понимании реплицирования. Рассмотрим мысленный эксперимент по измерению температуры кипения воды. Если мы соберем установку, состоящую из газовой горелки, колбы с водой, опущенного в воду жидкостного термометра, и будем нагревать колбу на пламени горелки до тех пор, пока в колбе не появятся пузырьки, свидетельствующие о закипании воды, то показания термометра в момент закипания сообщат нам искомую величину температуры. Рутинной проверкой, в смысле Картрайт, будет повторение процесса нагревания воды до кипения в последующие моменты времени уже после ее остывания, хотя, как указывал ранее Коллинз, даже неизменность процесса во времени между последовательными закипаниями тоже уже является теоретическим допущением (и основано на независимо проверенных теориях). Такая рутинная проверка, действительно, хотя и способна установить устойчивые корреляции между зажиганием горелки и закипанием воды, еще не способна достаточно глубоко проверить инструментальные теории, на которых строится работа установки. Следующий уровень реплицирования, реплицирование-2 в смысле Галисона, будет состоять в варьировании (например, интенсивности горения газа горелки, положения термометра в объеме колбы, уровня воды в колбе). Экспериментатор варьирует параметры установки, не изменяя принципиально теоретического описания кипячения воды. Третий уровень – воспроизведение, как его понимают Франклин и Картрайт – сводится к получению и измерению того же явления (соответственно, проверке той же теории явления), но на установке, основанной на принципиально иных инструментальных теориях. В нашем примере этому могло бы соответствовать, например, кипячение воды в микроволновой печи, где тот же эффект и исследуемый процесс, а именно сообщение молекулам воды в колбе кинетической энергии, достаточной для ее закипания, достигается с помощью иной инструментальной теории, а именно дециметрового электромагнитного излучения, которое нагревает жидкость по всему объему колбы, а не за счет теплопроводности через дно и последующей конвекции, как в случае газовой горелки.

Можно задаться вопросом: замена, например, газовой горелки на электрическую плитку, жидкостного термометра на оптический или электронный, а стеклянной колбы на металлическую будут реплицированием или воспроизведением? С одной стороны, базовые инструментальные теории, такие как конвекция, теплопередача, остаются теми же, что и в исходном эксперименте, изменяются лишь табличные коэффициенты. Но, с другой стороны, изменяя принцип действия, например, нагревающего элемента, даже при сохранении способа теплопередачи, мы вводим какие-то новые теории и допущения, которые требуют новых вспомогательных теорий (например, о составе материала стенок колбы, о его поведении при нагревании, о составе газа горелки, о различиях в коэффициентах расширения жидкости в термометре и многих других). На то обстоятельство, что в эксперименте опора на какие-либо инструментальные теории требует учета неограниченного списка вспомогательных теорий и допущений практически до бесконечности, одним из первых в философской литературе обратил внимание Гарри Коллинз, назвав его «регрессом экспериментаторов». Коллинз сделал довольно широкое социологическое обобщение, утверждая, что по причине «регресса экспериментаторов» получение надежного экспериментального знания в принципе невозможно, а согласие в сообществе по поводу экспериментального результата происходит в результате компромисса, в основе которого лежат социальные, в частности карьерные, интересы ученых. Какие-либо формальные критерии проверки того, корректно ли работает установка, отсутствуют.

Однако на доводы Коллинза убедительно возразил Франклин [Franklin et al. 1989], разъяснив оппоненту, что для выхода из теоретических и инструментальных взаимозависимостей (число которых действительно крайне велико) экспериментаторы применяют калибровку приборов. Калибровка – приведение в соответствие показаний приборов, предназначенных для одной и той же функции (измерения одной и той же

величины), через согласование их шкал. Например, электронный термометр может быть прокалриброван жидкостным и давать согласованные с таковым значения. В вышеприведенном примере замены жидкостного термометра на электронный или оптический ввиду их калибровки отсутствует необходимость каждый раз учитывать все теории, на которых основано функционирование этих приборов, и их неопределенности, поскольку шкалы приборов единообразно приведены в соответствие и могут быть использованы в измерении без отсылки к структуре теорий (погрешность шкалы прибора учитывает неопределенность знания этих теорий). Также возможна калибровка измерительных приборов искусственным сигналом, например спектрометров гамма-квантов генератором импульсов с заданной при изготовлении и точно определенной частотой. Такой тестовый сигнал принимается за референтный (реперный) в обоих сравниваемых приборах, а их шкалы размечаются, исходя из соглашения об одинаковом понимании теории, ответственной за сигнал. Калибровка, как показал Франклин, снимает проблему «регресса экспериментаторов» Коллинза, поэтому мы полагаем, что замена в установке приборов на иные, включая даже основанные на различных инструментальных теориях, но приведенных за счет калибровки к одинаковым шкалам, вполне может считаться реплицированием-2. Аргумент же Коллинза о социальной природе компромисса по поводу экспериментальных результатов относится не к результатам конкретного измерения, а в более широком смысле к парадигмальным сдвигам в теории, как их понимали Лакатос и Кун. Отметим пока только, что в вышеприведенном примере кипячения воды речь шла об эксперименте «нормальной науки», не предполагающем революционного изменения теории, допустим, фазовых переходов.

Роль реплицирования в современной физике

Реплицирование-2, несмотря на его промежуточное положение между рутинным повторением и воспроизведением, играет особую роль в современном сложном эксперименте, а его игнорирование зачастую приводит к плачевным результатам. Многие в мире физики элементарных частиц помнят сделанное в 2012 г. экспериментом OPERA в ЦЕРН [Adam et al. 2011] заявление о наблюдении нейтрино, движущихся быстрее скорости света. Невозможность достижения скоростей, превышающих скорость света, постулируется специальной теорией относительности А. Эйнштейна (СТО), и наблюдение более высоких скоростей ставило бы под вопрос справедливость этой фундаментальной теории. Коллектив авторов насчитывал 174 человека. Установка эксперимента создавала поток нейтрино на ускорителе в ЦЕРН и посылая его сквозь Землю в Италию, на расстояние 730 км, в помещение, расположенное под Апеннинским горным массивом Гран Сассо. Время между выходом нейтринного потока из Женевы и его прибытием в Гран Сассо измерялось по спутниковой системе навигации с использованием оптоволоконного кабеля для быстрой передачи сигнала к компьютеру в Гран Сассо. Хотя технически эксперимент был весьма многофакторным, анализ данных был довольно прост и концептуально не выходил за рамки школьного курса физики: чтобы получить скорость нейтрино, достаточно было разделить хорошо известное расстояние от ЦЕРН до Гран Сассо на время движения нейтрино из Женевы, измеренное по спутнику. Проведя анализ, экспериментаторы обнаружили, что скорость нейтрино на 60 наносекунд опережает скорость света. Однако ряд членов коллектива обратили внимание коллег на то, что такой результат получается лишь при некотором способе подключения оптического кабеля к разьему; при других способах подключения кабеля результат получался не превышающим скорости света. Разногласия в коллективе были преодолены авторитетом одного из руководителей эксперимента, чья группа была ответственной за данное подключение, в связи с чем считалась наиболее экспертной в этом вопросе. Несмотря на то что варьирование подключения кабеля не приносило повторяющихся результатов, коллектив доложил результаты на общем семинаре ЦЕРН (где вопрос кабеля не поднимался в силу различий и узости специализаций участников), а также дал интервью газете «Нью Йорк Таймс», произведя

широкий общественный резонанс. Приблизительно через год проблема с подключением кабеля выявилась, и коллаборации пришлось выступать с опровержениями собственного результата. Этот пример показывает, что при отсутствии повторения результата экспериментом, основанным на других физических принципах (то есть воспроизведением), только реплицированием-2 (подключением оптического кабеля различными способами) также можно выявить множество проблем с инструментальными теориями. Здесь важно вспомнить аргумент Коллинза о том, что не может быть двух совершенно одинаковых экспериментов, и даже столь рутинная проверка, как другое подключение кабеля, дает совершенно другие результаты.

Этот пример, на мой взгляд, сходен с примером из большой серии экспериментов по безнейтринному двойному бета-распаду, в которой эффект наблюдался однажды и только по утверждениям небольшой группы экспериментаторов под руководством Г. Клапдора-Кляйнгротхауса, выдвигавшей из коллаборации и представившей свою интерпретацию данных [Franklin, Laumon 2019, 10–11]. Другие группы, проведя независимый анализ, не подтвердили результатов Клапдора-Кляйнгротхауса, и те перестали привлекать внимание исследователей. Поскольку алгоритмы анализа данных опираются на различный выбор настроек и допущений, подбираемых каждой группой самостоятельно, повторение анализа данных другими исследователями выступает, на мой взгляд, как вариант реплицирования-2, аналогичный подключениям кабеля OPERA, то есть как стратегия Галисона.

Множественность ролей реплицирования

Я. Хакинг считает проблему воспроизводимости «философской псевдопроблемой». Ему представляется очевидным, что «многообразие экспериментов более убедительно, чем повторение одного и того же события», а «серьезные повторения эксперимента являются попытками сделать то же самое лучше – породить более устойчивый, менее зашумленный вариант явления» [Хакинг 1998, 239], то есть не являются в буквальном смысле повторениями, а, в наших терминах, либо реплицированием-2, либо воспроизведением (другим инструментальным способом). Однако здесь важно обратить внимание на еще две функции именно буквального повторения эксперимента, реплицирования в терминах Картрайт, также отмеченные Хакингом. Одна из них академическая – «научить людей тому, как стать экспериментаторами, и отсеять тех, для кого экспериментальная работа вряд ли будет подходящей профессией» [Хакинг 1998], то есть формирование практического навыка и приобщение к культуре сообщения. Вторая – ее акцентируют как Хакинг, так и Коллинз – попытка опровержения прежнего результата. Думаю, что в отличие от первых двух функций повторения – как реплицирования-2, так и воспроизведения другим способом, где речь идет об изменении (в большей или меньшей степени) инструментальных теорий с целью повышения точности измерений – в последних двух случаях неявно подразумевается требование общезначимости результата эксперимента.

Если с позиций, более свойственных для отдельного ученого, отождествить субъект познания с индивидуальным экспериментатором, то мы могли бы сравнить такое требование общезначимости с интересубъектностью результата, которая не может быть объяснена в рамках частных наук. С другой стороны, поиск оснований единства нашего знания может быть осуществлен в рамках философии трансцендентализма. Единство воспринимаемых всеми явлений, в частности результатов измерений, может предполагать поиск трансцендентального основания. Пространственные и темпоральные характеристики результатов измерений могут быть поняты как результат действия предзаданных категорий сознания. Однако, для того чтобы договориться о сравнении результатов экспериментов, от экспериментаторов требуется также согласие об используемых теориях (как теориях явлений, так и инструментальных теориях, реализованных в приборах и установках). Для такого согласия требуется исключительно важное условие: понимание полной теоретической нагруженности эксперимента (и получаемых результатов)

«сверху», от теории явления высокого уровня, «донизу», то есть всех инструментальных элементов установки. Только такое представление дает понимание того, что и с чем возможно сравнивать и, таким образом, реплицировать или воспроизводить. Ключевой момент состоит в том, что такой подход, равно как и собственно представление о теоретической нагруженности, являются краеугольными для философии постпозитивизма. Теоретическая нагруженность хотя и указывает на методологический приоритет теоретического над эмпирическим, оказывается в научном эксперименте не только не проблематичной, но и эпистемологически плодотворной. На методологическую роль теоретической нагруженности указывают, в частности, [Tal 2011; Beauchemin 2017; Staley 2020; Пронских 2020]. Именно благодаря этой роли реплицирование и воспроизведение технически сложных экспериментов обретают методологический смысл.

Воспроизводимость в ряду эпистемических стратегий эксперимента

Франклин предложил регулярно пополняемый и обновляемый им список эпистемических стратегий, методологических стандартов, которые используют (и должны использовать) экспериментаторы, чтобы убедиться, что эксперименты дают корректные результаты [Franklin 2002]. Полагаю, что Я. Хакинг и П. Галисон также внесли существенный вклад в разработку таких стратегий. Различные виды повторения эксперимента, рассмотренные выше, занимают первые места в эпистемических стратегиях. Рассмотрим их более детально:

1. Калибровка установки, когда она воспроизводит известные явления.
2. Воспроизведение других явлений, о наличии которых заранее известно.
3. Устранение вероятных источников ошибок и попытки альтернативных объяснений наблюдаемых закономерностей.
4. Внутренняя согласованность результатов (в том числе объяснения фоновых эффектов).
5. Использование надежной и предварительно многократно проверенной теории для объяснения результатов.
6. Использование приборов, чей принцип действия основан на многократно проверенных теориях.
7. Использование статистических аргументов.
8. Манипуляция изучаемым объектом для получения предсказуемого результата в установке.
9. Независимое подтверждение.
10. Использование «слепого анализа», когда экспериментатор до конца измерений не видит результата.

И хотя сам Франклин относит к реплицированию в основном стратегию 9 [Franklin 2013], на мой взгляд, большее число стратегий может быть ассоциировано с повторением результата в той или иной мере. Первые две стратегии подразумевают получение в установке явления, ранее известного из измерений на других установках, прежде всего с другим принципом действия (вслед за Картрайт мы называем это воспроизведением³). Упомянутые в стратегии 7 статистические аргументы сводятся к повторению измерения не только на той же установке, но и в условиях, максимально приближенных друг к другу. Стратегия 8 может быть понята двояко: либо как манипулятивная стратегия Хакинга (манипулирование объектом исследования предсказуемым образом, то есть приготовление явления в определенных условиях), либо как реплицирование-2 (манипулирование настройками установки предсказуемым образом). Независимое подтверждение (стратегия 9) подразумевает воспроизведение явления (на установке с другим принципом действия). Остальные стратегии имеют смысл либо использования методологического потенциала теоретической нагруженности, либо исключения влияния психологических факторов (неосознанного стремления подогнать результат под ожидания), например через так называемый «слепой анализ». Полный список эпистемологических стратегий экспериментирования более логично упорядочить следующим образом:

1. Внутренняя воспроизводимость:
 - a. Реплицирование с минимальными изменениями (статистика). Цель – увеличение прецизионности измерения.
 - b. Реплицирование-2 (стратегии Хакинга и Галисона). Цель – устойчивость результата и исключения методической ошибки.
 - c. Калибровка (воспроизведение других известных эффектов). Цель – повышение прецизионности измерения и исключение методической ошибки.
2. Внешняя воспроизводимость (измерение явления на установках другого принципа действия).
3. Проверка согласованности теоретической нагруженности измерения (внутренняя согласованность результатов, надежность инструментальных теорий).
4. Устранение психологических факторов, ведущих к предвзятости экспериментаторов в отношении результатов (возникают, как правило, на этапе применения вышеперечисленных стратегий и интерпретации их результатов).

Перечисленные эпистемические стратегии собирались Франклином на протяжении нескольких десятилетий в ходе изучения практики физических лабораторий и тщательного анализа ошибок, возникавших в ходе экспериментирования. Большая часть экспериментов, изученных им, соответствовала нормальному развитию науки и решению головоломок, в ходе которых полнота опоры на методологический потенциал эпистемических стратегий приобретала особую важность. Как можно видеть, в их основе лежат различные варианты (вообще говоря, полный набор) способов как внутреннего, так и внешнего воспроизведения результатов, включающих повторение как без каких-либо аппаратных изменений, так и при частичной и полной (но контролируемой) замене инструментальных теорий. На этом фоне особый интерес представляют исторические примеры экспериментов, которые привели к принятию теории без какого-либо реплицирования (да и воспроизведения вообще).

Когда эксперименту не требуется повторение?

Первый из подобных примеров – открытие американским физиком Андерсоном позитрона в космических лучах в 1933 г. В фотографических снимках пузырьковой камеры, помещенной в магнитное поле и облучаемой космическими лучами, Андерсон обнаружил след частицы, которая по массе соответствовала электрону, но имела искривление траектории, соответствовавшее положительному электрическому заряду. Несколько ранее положительные электроны (позитроны) были теоретически предсказаны Полем Дираком. Андерсон, по его собственным воспоминаниям, хоть и не владел деталями теории Дирака, был знаком с ее основными предсказаниями. Измерение Андерсона опиралось, как и свойственно экспериментальной технике, на множество допущений и предполагало множество интерпретаций (например, можно было предположить, что на фотографии – след обычного электрона, но породившая его частица влетела с противоположной стороны камеры). Андерсон опубликовал работу, где основным результатом эксперимента стала одиночная фотография новой частицы. Вскоре Дирак был удостоен Нобелевской премии за предсказание позитрона, фактически подтвержденное единичным измерением. Это один из показательных исторических примеров того, как теория, уже принятая научным сообществом на основе ее ценных внутритеоретических качеств (симметрии уравнений теории), получила окончательное признание в форме Нобелевской премии с обоснованием в виде одиночного измерения (допускавшего в принципе различные интерпретации) без какого-либо воспроизведения. Хотя вскоре после эксперимента Андерсона были сделаны еще несколько независимых измерений того же эффекта, но они не привлекли достаточного внимания и не послужили обоснованием награды Дирака. Такая интерпретация, на мой взгляд, может служить иллюстрацией идей Куна о том, что парадигмальные сдвиги подготавливаются изменившимися взглядами научного сообщества, тогда как эксперимент, согласно Лакатосу, не является решающим в признании этих взглядов. С этой точки зрения, одиночное измерение позитрона, осуществленное Андерсоном, послужило скорее

формальным символом эмпирического подтверждения, поскольку традиции физики требовали (и требуют) эмпирического обоснования теорий. В описанном же примере роль эксперимента состоит не столько в проверке или попытке опровержения теории, сколько в демонстрации ее следствий (на такую роль экспериментов Франклин указывает во многих своих книгах, например [Franklin 2013]).

Следующим примером эксперимента, не требовавшего воспроизведения для признания теории, может служить открытие частицы омега-минус-гиперона в физике высоких энергий. Эта частица, состоящая из трех кварков, была предсказана М. Гелл-Манном и Ю. Неemanом в рамках кварковой модели, основанной на соображениях симметрии и предлагавшей систематику элементарных частиц, составленных из разных комбинаций кварков. Омега-минус-гиперон являлся недостающим элементом систематики, и его экспериментальное обнаружение могло легитимизировать всю систематику как способ схематизации элементарных частиц и заодно лежащую в ее основе кварковую модель. Согласно описанию Лаймона и Франклина, физики проводили измерения и обработку данных только до тех пор, пока не обнаружили указания на омега-минус-гиперон, причем одна из ключевых характеристик частицы, определяющая ее положение в теоретической схеме, – спин – была ими выражена лишь приблизительно (только то, что он не меньше, чем $\frac{3}{2}$). Несмотря на отсутствие внутреннего и внешнего воспроизведения, схема незамедлительно получила признание в научном сообществе. Как и в случае открытия позитрона Андерсоном, теория, обладающая значительными внутренне-теоретическими достоинствами и унифицирующим потенциалом (кварковая модель), была предварительно интериоризирована научным сообществом, как предпочтительная концептуальная схема, поэтому даже одиночное и теоретически нагруженное измерение обеспечило ее полное принятие. Это стало возможным, как и в случае позитрона, поскольку роль эксперимента здесь заключалась в артикуляции теории, в демонстрации предсказанного теорией феномена, а не в попытке ее опровержения.

Обсудим также третий пример из книги Франклина и Лаймона [Franklin, Laymon 2019], уже из области биологии. В генетических экспериментах Менделя, как показывают авторы, первые, ставшие убедительными для сообщества эмпирические результаты были (как выяснилось при ретроспективном анализе) статистически необоснованными и не основывались на количестве образцов, достаточном для обеспечения статистической достоверности. Несмотря на это, аргументы Менделя были основательно и убедительно сформулированы, а также широко представлены научному сообществу, в связи с чем вопросов по поводу статистической значимости данных не возникало. В результате произошло принятие открытых им закономерностей наследования признаков, которое предшествовало воспроизведению экспериментов другими учеными. Уже в XX в. были предприняты попытки воспроизвести эксперименты Менделя, и ряд исследователей утверждал, что подобные эксперименты не могли дать статистически обеспеченных результатов, а его выводы были едва ли не сфальсифицированы за счет преднамеренного отбора данных, укладывающихся в новую концепцию. Отметим, что в одной из своих более ранних книг Франклин приводит пример отбора и исключения данных и в физике (см.: [Franklin 2013]). В частности, в экспериментах Милликена по определению заряда электрона ученый использовал только около 15% измеренных данных, отбросив остальные как не вписывающиеся в его схему (а в современных экспериментах в физике ядра и элементарных частиц отбор и исключение данных как при их наборе, так и при обработке – основной метод анализа эмпирических данных). Таким образом, стратегии отбора и исключения не являются специфичными для биологии, а распространены в различных разделах естествознания.

Пример закономерностей наследования, открытых Менделем, на мой взгляд, отличен от примеров с позитроном и омега-минус-гипероном в том, что он указывает на ведущую роль коммуникаций и стратегий презентации результатов в науке: будучи широко и аргументированно представленными в сообществе, закономерности, открытые ученым, сформировали мнение научного сообщества, результатом чего стало возникновение парадигмы, основанной на теории наследования. Так же как и два предшествующих примера из области физики, пример опытов Менделя хорошо вписывается в модель

смены парадигм Куна, в которой интересам и мнению научного сообщества отводится лидирующая роль. Во всех трех случаях эксперименты использовались для демонстрации и артикуляции теории, в связи с чем каких-либо подтверждений в форме воспроизведения теорий не требовалось (подобно этому, в демонстрационных экспериментах в школьной и вузовской аудитории изучаемый эффект нередко показывается однократно).

Заключение

Рассмотренные здесь точки зрения на проблему воспроизведения в науке позволяют сделать вывод о том, что полное повторение (реплицирование) эксперимента вряд ли возможно, поскольку опирается на неустранимые допущения и может иметь смысл только либо как попытка опровержения прежнего результата, либо как особый случай набора статистики с целью снижения погрешности результата. Особое же методологическое внимание привлекает повторение эксперимента в измененной постановке (частично или полностью), что связывается многими исследователями с независимым подтверждением теории явления, которое, как правило, и называется воспроизведением. Многолетние изыскания Франклина в области методологии экспериментов в современной физике позволили ему выделить целый ряд методологических стандартов эксперимента, названных им эпистемическими стратегиями, которые в значительной мере опираются на различные формы воспроизведения и реплицирования и позволяют повысить надежность эмпирических результатов. При этом множество исторических примеров показывает, что если роль эксперимента состоит не в попытке опровержения, а в подтверждении и артикуляции высокоценной для общества новой парадигмальной теории, то в контексте научной революции значимость воспроизведения в принятии результата отходит на второй план по сравнению со стратегиями демонстрации и коммуникации.

Примечания

¹ Комплекс приборов, объединенных одной научной задачей или теорией явления, для проверки которой он используется.

² Еще Альберту Эйнштейну приписывали фразу: «Нельзя повторять одно и то же снова и снова и ожидать другого результата». Обычно эта фраза понимается как комментарий к вероятностной природе квантовой механики, однако мы полагаем, что она является хорошей иллюстрацией воспроизводимости результатов вообще.

³ Разнообразие и неопределенность терминологии (а именно, что понимать под реплицированием, репродуцированием, воспроизведением и т.д.), которой была отмечена эта область с 1980-х гг., может быть охарактеризована как «ситуация предстандарта» [Крушанов 2006], которая указывает на становление предметной области. Несмотря на бурный рост числа публикаций в мире на тему воспроизводимости (в первую очередь, в биомедицине), в силу различия стандартов в различных областях науки, становление терминологии по воспроизводимости, очевидно, еще не полностью завершилось. Например, в методологии биомедицинских наук одна из таксономий воспроизводимостей предложена в: [Goodman, Fanelli, Ioannidis 2016].

Ссылки – References in Russian

- Галисон 2018 – Галисон. П. Коллективный автор // Вопросы философии. 2018. № 5. С. 93–113.
- Крушанов 2006 – Крушанов А.А. Ситуации предстандарта в динамике научного познания // Вопросы философии. 2006. № 8. С. 54–66.
- Пронских 2020 – Пронских В.С. Проблемы измерений: современные дискуссии и модели // Успехи физических наук. 2020. № 190. С. 211–219.
- Пружинин 2014 – Пружинин Б.И. Знание как ценность (Этюд по культурно-исторической эпистемологии) // Гуманитарные исследования в Восточной Сибири и на Дальнем Востоке. 2014. № 1. С. 135–141.
- Сторожук 2008 – Сторожук А.Ю. Воспроизводимость экспериментов с точки зрения философии активности // Философия науки. 2008. № 3 (38). С. 59–76.
- Хакинг 1998 – Хакинг Я. Представление и вмешательство. Введение в философию естественных наук. М.: Логос, 1998.

References

- Adam, Thomas et al. (2011) Measurement of the neutrino velocity with the OPERA detector in the CNGS beam, *arXiv: 1109.4897v1* [hep-ex].
- Beauchemin, Pierre-Hugues (2017) "Autopsy of measurements with the ATLAS detector at the LHC". *Synthese*, Vol. 194, pp. 275–312.
- Cartwright, Nancy (1991) "Replicability, Reproducibility, and Robustness: Comments on Harry Collins", *History of Political Economy*, Vol. 23, No. 1, pp. 143–155.
- Collins, Harry M. (1984) "When Do Scientists Prefer to Vary Their Experiments?", *Studies in History and Philosophy of Science*. Part A, Vol. 15, pp. 169–174.
- Franklin, Allan (2002) "The Role That Experiment Plays in Physics", *Theoria-Revista de Teoria Historia y Fundamentos de la Ciencia*, Vol. 17, No. 2, pp. 221–242.
- Franklin, Allan (2013) *Shifting Standards: Experiments in Particle Physics in the Twentieth Century*, University of Pittsburgh Press, Pittsburgh.
- Franklin, Allan (2018) *Is It the 'Same' Result: Replication in Physics*, Morgan & Claypool Publishers, San Rafael.
- Franklin, Allan, Howson, Colin (1984) "Why Do Scientists Prefer to Vary Their Experiments?", *Studies in the History and Philosophy of Science*, Vol. 15, pp. 51–62.
- Franklin, Allan et al. (1989) "Can a Theory-Laden Observation Test the Theory?", *British Journal for the Philosophy of Science*, Vol. 40, No. 2, pp. 229–251.
- Franklin, Allan, Laymon, Ronald (2019) *Measuring Nothing, Repeatedly: Null Experiments in Physics*, Morgan & Claypool Publ., San Rafael.
- Franklin, Allan, Laymon, Ronald (2021) *Once Can Be Enough. Decisive Experiments, No Replication Required*, Springer, Cham.
- Galison, Peter (1987) *How Experiments End*, University of Chicago Press, Chicago.
- Galison, Peter (2003) "The Collective Author", Galison, P., Biagioli, M., eds., *Scientific Authorship: Credit and Intellectual Property in Science*, Routledge, New York, Oxford, pp. 325–353 (Russian Translation).
- Goodman, Steven N., Fanelli, Daniele, Ioannidis, John P.A. (2016) "What Does Research Reproducibility Mean?", *Science Translational Medicine*, Vol. 8, No. 341, pp. 12.
- Hacking, Ian (1983) *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*, Cambridge Univ. Press, Cambridge (Russian Translation).
- Krushanov Alexander A. (2006) "Pre-standard situations in the dynamics of scientific knowledge", *Voprosy filosofii*, Vol. 8, pp. 54–66 (in Russian).
- Laymon Ronald, Franklin Allan (2021) "Replication. Part I", *Voprosy filosofii*, Vol. 8, pp. 116–129.
- Medawar Peter B. (1963) "Is the scientific paper a fraud?", *Listener*, Vol. 70, pp. 377–378.
- Pronskikh, Vitaly S. (2015) "Review of Shifting Standards: Experiments in Particle Physics in the Twentieth Century by Allan Franklin", *Philosophy of Science*, Vol. 82, No. 4, pp. 727–730.
- Pronskikh, Vitaly S. (2020) "Measurement problems: contemporary discussions and models", *Physics-Uspekhi*, Vol. 63, No. 2, pp. 192–200 (in Russian).
- Pruzhinin, Boris I. (2014) "Knowledge as a value (Etude on cultural-historical epistemology)", *Humanities Research in the Russian Far East*, Vol. 1, pp. 135–141 (in Russian).
- Staley, Kent W. (2020) "Securing the empirical value of measurement results", *The British Journal for the Philosophy of Science*, Vol. 71, No. 1, pp. 87–113.
- Storozhuk Anna Yu. (2008) "Reproducibility of Experiments in Terms of Activity Philosophy", *Philosophy of Sciences*, Vol. 3, No. 38, pp. 59–76 (in Russian).
- Tal, Eran (2011) "How Accurate Is the Standard Second?", *Philosophy of Science*, Vol. 78, No. 5, pp. 1082–1096.

Сведения об авторе

ПРОНСКИХ Виталий Станиславович – кандидат физико-математических наук, кандидат философских наук, ассоциированный научный сотрудник сектора Философии естественных наук Института философии РАН.

Author's Information

PRONSKIKH Vitaly – PhD (JINR, Physics and Mathematics), PhD (Philosophy), Associate Research Fellow in the Department of Philosophy of Natural Sciences, Institute of Philosophy RAS.