

ОРГАНИЗАЦИЯ ИССЛЕДОВАТЕЛЬСКОЙ ДЕЯТЕЛЬНОСТИ

**ДВА КОНТУРА ОЦЕНКИ РЕЗУЛЬТАТИВНОСТИ НАУЧНЫХ
ОРГАНИЗАЦИЙ В РОССИИ: ТЕКУЩЕЕ СОСТОЯНИЕ
И ПЕРСПЕКТИВЫ РАЗВИТИЯ С ТОЧКИ ЗРЕНИЯ
МЕЖДУНАРОДНОГО ОПЫТА**

© 2024 г. Д. В. Косяков^{a,*}, И. В. Селиванова^{a,**}, А. Е. Гуськов^{a,***}

^aРоссийский научно-исследовательский институт экономики, политики и права в научно-технической сфере,
Москва, Россия

*E-mail: kosyakov@sciencepulse.ru

**E-mail: i-seli@yandex.ru

***E-mail: guskov.andrey@gmail.com

Поступила в редакцию 10.10.2023 г.

После доработки 25.10.2023 г.

Принята к публикации 10.01.2024 г.

В России фактически сложились два контура оценки научной результативности организаций: экспертный, лежащий в основе категорирования научных организаций, и количественный, который по методике Комплексного расчёта публикационной результативности позволяет определять и контролировать выполнение государственного задания на фундаментальные научные исследования. В статье описана история становления этих контуров, их достоинства и недостатки, а также предпосылки для дальнейшего развития. Модельными примерами для улучшения обоих контуров рассматриваются британская система экспертизы REF (Research Excellence Framework) и норвежская модель количественной оценки. Особенности REF состоят в дифференцированной оценке результатов, импакта и среды научных подразделений, которая организуется экспертными панелями и группами. Особенности норвежской модели — экспертный подход к формированию национального белого списка научных журналов, конференций и издателей, методика расчёта публикационного индикатора и открытая база данных публикаций норвежских учёных, по которой этот индикатор определяется. Авторы полагают, что эти два примера можно рассматривать как параметрические модели оценки, адаптация которых с учётом национальных особенностей позволит актуализировать и усовершенствовать оба контура оценки научных организаций в России. Акцентируется важность регулярного ретроспективного осмысления опыта, процедур и результатов оценки организаций, что необходимо для планомерного развития этой системы на следующих этапах.

Ключевые слова: научная результативность, оценка результативности научных организаций, экспертиза, комплексный расчёт публикационной результативности (КБПР), категорирование организаций, наукометрия, норвежская модель, REF.

DOI: 10.31857/S0869587324010063, EDN: HAXNUC



КОСЯКОВ Денис Викторович — заместитель заведующего лабораторией наукометрии и научных коммуникаций РИЭПП. СЕЛИВАНОВА Ирина Вячеславовна — кандидат технических наук, научный сотрудник РИЭПП. ГУСЬКОВ Андрей Евгеньевич — кандидат технических наук, заведующий лабораторией наукометрии и научных коммуникаций РИЭПП.

Концепция оценки научных исследований имеет богатую историю, уходящую корнями в принципы нового государственного управления, идеологии управления, которая подчёркивает эффективность, подотчётность и результативность той или иной деятельности. Появление парадигмы нового государственного управления в конце 1970-х — начале 1980-х годов ознаменовало собой значительный сдвиг в административной практике. Первоначально это движение зародилось в Великобритании в администрации премьер-министра Маргарет Тэтчер, а также в различных муниципальных органах власти в США, пострадавших от экономического спада. Затем правительства Новой Зеландии и Австралии также приняли парадигму нового государственного управления, а её успешная реализация привела к включению административных реформ нового государственного управления в повестку дня многих стран ОЭСР и не только [1]. Оценка исследований и системы финансирования исследований на основе результатов (Performance-based Research Funding System, PRFS) вытекает из этой идеологии [2, 3]. Изначально эта система в значительной степени опиралась на экспертную оценку — из-за отсутствия качественных данных о научных публикациях. Однако со временем, с появлением современных технологий всё большее распространение получали библиометрические показатели и количественные оценки.

Великобритания стала пионером и в области оценки научных исследований. В 1985 г. Комитет университетских грантов (University Grants Committee, UGC) этой страны начал процесс создания национальной системы оценки университетских исследований [4]. Цель этой системы заключалась в содействии выборочному финансированию исследований — стратегии, которую UGC считал необходимой для поддержания качества университетской науки. Впоследствии эта система превратилась в сложный механизм распределения финансирования под управлением финансовых советов. В 1980-х и 1990-х годах это стало важной частью политики пересмотра отношений между государством и британской системой высшего образования.

Система оценки качества исследований в Великобритании, носящая сейчас название Research Excellence Framework (REF), послужила влиятельной моделью для аналогичных систем в различных странах мира. В Гонконге, например, Комитет университетских грантов ввёл в 1993 г. свою систему оценки исследований RAE (Research Assessment Exercise) [5]. Как и REF, гонконгский RAE оценивает качество исследований по шестибальной шкале, учитывая такие факторы, как результаты исследований, исследовательская среда и влияние исследований. Австралия в 2008 г. запустила собственный проект Excellence in Research for Australia (ERA) [3, 6]. ERA оценивает качество

исследований, проводимых в австралийских университетах, в сравнении с национальными и международными эталонами. Эта система, так же как и REF, претерпела заметные изменения. В Новой Зеландии в 2003 г. был создан Фонд исследований, основанных на результатах (Performance-based Research Fund, PBRF), с целью поощрения и вознаграждения передовых исследований в секторе высшего образования [7]. Однако его подход несколько отличается от зарубежных аналогов большим акцентом на индивидуальные результаты. В 2011 г. национальную оценку исследований, известную как Valutazione della Qualità della Ricerca (VQR), ввела Италия. VQR сочетает в себе экспертную оценку и учёт библиометрических показателей [8].

Интеграция библиометрических показателей в систему оценки исследований была впервые инициирована Фландрией (Бельгия) и Китаем. Во Фландрии библиометрические показатели стали одним из условий распределения государственного финансирования университетов в рамках BOF-key (Bijzonder Onderzoeksfonds, или Фонд университетских исследований) в 2003 г. [9, 10]. Это один из первых случаев, когда библиометрические показатели использовались для распределения финансирования на системном уровне, причём публикация в журналах, индексируемых в Web of Science, является основным фактором, определяющим распределение финансирования. Китай также выделяется ранним внедрением библиометрических показателей в оценку исследований. Правительство Китая с 1990-х годов ориентировалось на количество публикаций в международных журналах, индексируемых в основных индексах Web of Science [11]. Количественный подход к оценке результативности исследований оказал значительное влияние на исследовательскую практику в Китае, побуждая исследователей отдавать предпочтение публикациям в международных изданиях.

Норвегия выделяется разработкой собственной уникальной системы оценки исследований, основанной на белом списке журналов и издательств, известном как Норвежский реестр научных журналов, серий и издательств, разделённых на две категории качества [12]. Оценка публикаций дифференцируется по типу и категории источника, что позволяет косвенно оценить их качество. Влияние этой системы, введённой в 2005 г., заметно во многих странах, которые адаптировали свои процедуры оценки исследований на основе подобных моделей. Например, Дания использует версию системы библиометрических показателей исследований (BFI), которая была введена в 2009 г. [13]. BFI присваивает баллы научным публикациям в зависимости от их качества и престижа издания. Финляндия также черпает вдохновение в норвежской модели [14]. Эта система призвана содействовать проведению высококачественных исследований

и стимулировать конкуренцию между университетами. Внимание норвежской модели уделяется и в Швеции [15].

Использование систем оценки результатов исследований, которые выполняют такие важные функции, как распределение ресурсов, стратегическое планирование и подотчётность, может иметь целый ряд последствий, как предполагаемых, так и непреднамеренных. К положительным можно отнести повышение качества исследований, усиление подотчётности и прозрачности, возможность принятия стратегических решений как на уровне отдельных организаций, так и в национальном масштабе, содействие междисциплинарному и международному сотрудничеству. К негативным последствиям причисляют стандартизацию исследований, вызванную необходимостью выполнять требования системы оценки в определённые сроки. Это побуждает учёных нацеливаться на проведение “безопасных” исследований, результаты которых могут быть получены за относительно короткое время и будут соответствовать требованиям высокорейтинговых журналов. Такие мотивы ограничивают инновации и творчество; увеличивают административное бремя, связанное с подготовкой необходимых данных; формируют неравенство между дисциплинами, связанное с разными практиками представления результатов исследований; стимулируют чрезмерный акцент на количество в ущерб качеству и на манипуляции с целью повышения оценок. Реальные и потенциальные последствия введения тех или иных систем оценки результативности в науке свидетельствуют о важности продуманной разработки и регулярного пересмотра таких систем, чтобы максимизировать их преимущества и минимизировать недостатки.

Цель настоящей работы — сформировать и обосновать методологический базис для развития российской системы оценки результативности научных организаций. Для этого необходимо описать сложившуюся в стране национальную модель оценки науки, определить её сильные и слабые стороны, а затем наметить перспективы эволюционного развития, основанные на глубоком изучении зарубежного опыта. Мы полагаем, что в России в категоризации научных организаций, несмотря на изначально количественную методiku, возобладал экспертный подход, реализуемый посредством экспертизы, выполняемой Российской академией наук, членами ведомственных и межведомственной рабочих групп. С учётом комплексного расчёта публикационной результативности (КБПР) к настоящему времени в РФ сформировалось два контура оценки результативности организаций. По каждому из этих контуров накопился груз проблем, связанных с изменением внешней среды, адаптацией оцениваемых организаций, изначальным несовершенством методик. При их развитии необходимо учитывать богатый международный опыт,

прежде всего конкурирующих подходов — британской системы REF и норвежской модели.

В этой статье рассматриваются вопросы, связанные с оценкой результативности научных организаций и университетов. В целом национальная система оценки научной деятельности имеет более сложную структуру и включает такие элементы, как: грантовое финансирование через Российский научный фонд; аттестация научных кадров; экспертиза проектов, выполняемая Российской академией наук; программа Приоритет-2030 и некоторые другие программы поддержки научно-технологического развития. Вопросы изменения общей архитектуры российской системы оценки требуют отдельного обстоятельного исследования. По той же причине за скобками этой работы оставлен детальный анализ финансовых, административных и иных последствий проведения оценки. Следует добавить, что в российской периодике публикации по этой тематике, в которых применяются доказательные методы или глубоко анализируется существующий опыт, встречаются крайне редко.

ОЦЕНКА РЕЗУЛЬТАТИВНОСТИ НАУЧНЫХ ОРГАНИЗАЦИЙ В РОССИИ

В Российской Федерации национальная оценка результативности деятельности научных организаций была инициирована Постановлением Правительства № 312 от 8 апреля 2009 г. В соответствии с постановлением такая оценка должна проводиться федеральными органами исполнительной власти и государственными академиями наук, в ведении которых они находятся [16]. Каждая организация должна была быть отнесена к одной из трёх категорий: 1-я категория — научные организации-лидеры; 2-я категория — стабильные научные организации, демонстрирующие удовлетворительную результативность; 3-я категория — научные организации, утратившие научный профиль и перспективы развития. Результаты оценки, проводимой каждые пять лет, предполагалось учитывать при формировании планов научно-исследовательских работ, определении объёмов бюджетного финансирования и оптимизации сети научных организаций. Ведомственная оценка проводится с 2010 г., но до конца 2012 г. только часть ведомств завершила эту работу. Например, Российская академия наук провела оценку 86% подведомственных учреждений, а Российская академия сельскохозяйственных наук — только 54%. Некоторые ведомства вообще не приступили к оценке. Всего было оценено 817 из 1119 (73%) организаций [17]. Из оценённых только семь организаций получили третью категорию, то есть были признаны утратившими научный профиль и перспективы развития, а к первой категории были отнесены 546 организаций (67%).

Оптимизация (то есть сокращение) сети научных организаций была основным инструментом повышения удельного финансирования на научного сотрудника без существенного изменения общих бюджетных расходов. Очевидно, что эта цель не была достигнута, в связи с чем в 2013 г. по итогам заседания Совета при Президенте РФ по науке и образованию Правительству было поручено пересмотреть принципы и методику оценки [18]. Это было сделано путём внесения изменений в упомянутое выше Постановление и разработки ряда подзаконных актов. Так был запущен второй раунд оценки, который уже в гораздо большей степени опирался на количественные индикаторы. Подготовительный этап этого раунда в основном завершился в 2014–2016 гг., а с 2017 г. сначала научным организациям, а потом и высшим учебным заведениям стали назначать категории результативности по новым правилам.

Процедура оценки была разделена на два этапа, первый из них опирался на количественную методику, разработанную специалистами Высшей школы экономики на основе показателей, которые организации стали заносить в созданную для этих целей Федеральную систему мониторинга научных организаций (ФСМНО). Методика не была апробирована на реальных данных, в связи с чем в ходе её применения обнаружился ряд проблем [19, 20], связанных как с достоверностью используемых показателей, так и с недостаточной её проработанностью — она носила излишне теоретический характер. Одной из проблем стало использование в качестве знаменателя в ряде индикаторов среднесписочного состава научных сотрудников — показателя, который имеет разное значение в академических учреждениях, отраслевых НИИ и вузах, кроме того, он подвержен сознательным манипуляциям. В части публикационной активности методика базировалась на количестве публикаций, индексируемых в базах данных Web of Science или Scopus, без внимания к типу этих публикаций и уровню издания, в которых они опубликованы.

В последующие годы методика несколько раз уточнялась, изменялись принципы отнесения организаций к референтным группам, расчёта пороговых значений показателей. Эти работы, а также верификацию значений показателей в ФСМНО в соответствии с решениями Межведомственной комиссии по оценке результативности деятельности научных организаций выполнял РИЭПП. Корректный расчёт количественной оценки вузов стал заметной проблемой в силу существенных отличий в организации исследований по сравнению с научными учреждениями, прежде всего в связи с мультидисциплинарным характером исследований и участием в них профессорско-преподавательского состава наряду с научными работниками.

Второй, экспертный этап оценки, роль которого сводилась к точечной корректировке результатов количественного этапа, стал определяющим. Благодаря решениям Межведомственной комиссии результаты формальных расчётов пересматривались, причём заметно, при этом экспертный этап в методике практически не регламентировался и был сведён к мнению членов комиссии, аргументация которых оставалась неизвестной. Обоснованность отдельных мнений можно оценить, изучив представления, поданные на заседания комиссии отсутствовавшими членами, так как они приложены к протоколам, которые доступны на сайте ФСМНО (теперь База данных результатов деятельности научных организаций, БД РД НО, <https://sciencemon.ru>).

По текущим результатам оценки, представленным на указанном сайте, категории получили 974 организации, из них 289 отнесены к первой категории, 771 — ко второй, 203 — к третьей (табл. 1). Первые организации получили категории в 2017 г., большая часть оценивалась в 2018–2019 гг. (табл. 2). Необходимо отметить, что часть из включённых в реестр БД РД НО организаций уже ликвидирована или прошла реорганизацию в виде присоединения к другому юридическому лицу.

Почти 68% категорированных организаций — научные, чуть больше 29% — учреждения высшего

Таблица 1. Годы категорирования организаций по ведомствам

Ведомство	2017	2018	2019	2020	2021	Итого
Минобрнауки России	10	441	257	47		755
Минздрав России		43	47			90
ФМБА России		19	10	1		30
Роспотребнадзор		1	22			23
Минпрос России	2		19			21
Другие	1	17	16	19	2	55
Всего	13	521	371	67	2	974

Таблица 2. Присвоенные организациям категории по ведомствам

Ведомство	1-я категория	2-я категория	3-я категория	Итого
Минобрнауки России	222	356	177	755
Минздрав России	27	55	8	90
ФМБА России	11	19		30
Роспотребнадзор	12	10	1	23
Минпрос России		8	13	21
Другие	17	34	4	55
Всего	289	482	203	974

образования. Из научных организаций третью категорию получили немногим более 17%, из организаций высшего образования — почти 31%. Формально, в соответствии с Постановлением Правительства № 312 от 8 апреля 2009 г., организации, получившие категории в 2017–2018 гг., должны были пройти переоценку в 2022–2023 гг. Однако этого не произошло. Так как для организаций отнесение ко второй или третьей категории означает лишение доступа к определённым видам финансирования, вопрос переоценки или отмены категории является остро актуальным.

КОМПЛЕКСНЫЙ БАЛЛ ПУБЛИКАЦИОННОЙ РЕЗУЛЬТАТИВНОСТИ

Изменения бюджетного законодательства, связанные с введением в 2009 г. понятия государственного задания как основного инструмента бюджетирования, привели к необходимости определения количественных и качественных показателей для контроля его исполнения. Для вида работ “фундаментальные научные исследования” в качестве количественного показателя было принято количество выполняемых проектов научных исследований, а качественного — количество научных публикаций в журналах, индексируемых в базе данных Web of Science. В 2019 г. Министерством науки и высшего образования РФ был разработан новый качественный индикатор — комплексный балл публикационной результативности (КБПР); в нём используется шкала баллов в зависимости от индексации издания, в котором опубликована научная работа, в базах данных и квартиля журнала в соответствии с импакт-фактором в Journal Citations Report (JCR) компании Clarivate Analytics.

Использование этого показателя в значительной степени снизило остроту проблемы гонки за количеством статей, приведшей к заметному снижению качества российского публикационного потока на фоне его заметного роста [21]. Введённая шкала баллов стимулирует публиковаться

в ведущих научных журналах. Однако КБПР не свободен от недостатков, в частности, это касается “штрафов на коллаборации”: статьи, подготовленные многочисленными соавторами и обычно оказывающие значительное влияние на развитие науки, оцениваются для каждой из аффилиаций крайне низко. А значит, занижаются показатели организаций, занимающихся, например, физикой высоких энергий или некоторыми направлениями медицинских исследований. Сложно однозначно оценить подход, в соответствии с которым не поощряются популярные в последние годы множественные аффилиации [22] и уравниваются работы в области гуманитарных и общественных наук, опубликованные в журналах из первого и последнего квартилей; в результате снижается эффект программ привлечения университетами ведущих зарубежных учёных. Кроме того, есть основания полагать, что использование квартилей JCR вносит элемент несправедливости в оценку научных результатов, так как в разных областях исследований различаются частота выхода и объём выпусков журналов, что приводит к неодинаковому распределению количества публикаций по результирующим квартилям [21, 23].

В настоящее время КБПР применяется для формирования и контроля исполнения государственного задания, а также в отчётах руководителей организаций, подведомственных Минобрнауки. В ряде случаев используется удельное значение этого показателя по отношению к среднесписочной численности научных сотрудников. Опыт показывает, что эта практика не лишена недостатков, так как численность подвержена искусственным манипуляциям посредством перевода части научных сотрудников на доли ставки; имеют значение и различия в роли научных сотрудников в научно-исследовательских институтах и высших учебных заведениях.

ТРАНСФОРМАЦИЯ НАЦИОНАЛЬНОЙ СИСТЕМЫ ОЦЕНКИ ОРГАНИЗАЦИЙ

По поручению вице-премьера Правительства РФ Д.Н. Чернышенко Министерство науки и высшего образования РФ в 2022 г. приступило к разработке Национальной системы оценки результативности научных исследований и разработок. Постановлением Правительства от 19 марта 2022 г. до конца 2023 г. приостановлены требования и целевые показатели, связанные с наличием публикаций, индексируемых в международных базах данных Web of Science и Scopus. Уже предложено продлить действие этого моратория на 2024 г.

Межведомственная рабочая группа, созданная при Минобрнауки, составила “Белый список” – перечень наиболее авторитетных научных журналов, в который входят около 30 тыс. российских и международных изданий. В мае 2023 г. была утверждена методика распределения этих изданий по категориям. Предполагается использовать белый список для задач мониторинга и оценки публикационной активности российских учёных, но нормативная база для этого пока не сформирована. Стоит отметить, что в белый список вошли только научные журналы, в то время как для части направлений исследований и многих российских организаций значимую роль играют публикации в материалах конференций [24], а для представителей гуманитарных наук исключительно важны монографии.

Два контура системы оценки результативности, сложившиеся к настоящему времени в России, ещё не обрели окончательные очертания и находятся в процессе активных трансформаций. Накопленный опыт, изменения, произошедшие за последние годы в области научной политики и исследовательского ландшафта в РФ, обширные зарубежные исследования по этой тематике позволяют обозначить недостатки подхода и наметить пути их преодоления. История развития российской системы оценки во многом повторяет путь, пройденный другими странами. Анализ зарубежного опыта в области оценки результативности научных организаций имеет решающее значение для выработки оптимальных методик и подходов на национальном уровне. Это позволит не только изучить лучшие практики, но и понять, какие ошибки были допущены и как их избежать.

БРИТАНСКАЯ МОДЕЛЬ ОЦЕНКИ НАУЧНОЙ РЕЗУЛЬТАТИВНОСТИ

К середине 1980-х годов в Великобритании насчитывалось 46 университетов, научные исследования в них финансировались из государственного бюджета через Комитет университетских грантов (UGC) при правительстве. Процесс распределения

финансирования научных исследований был непрозрачен, а в условиях усиления финансового давления со стороны правительства М. Тэтчер бремя сокращения финансирования распределялось по университетам неравномерно [25, 26]. Назрела необходимость введения прозрачной процедуры выбора приоритетов финансирования, чтобы, с одной стороны, не размазывать имеющиеся средства тонким слоем, а с другой – обосновать приоритеты [27]. На эту цель был ориентирован первый раунд оценки 1986 г., получивший название Research Selectivity Exercise (RSE). К этому времени оценка результативности исследований и соответствующие индикаторы уже привлекли заметное внимание исследователей и администраторов, в том числе в рамках Программы организационного управления в высшем образовании ОЭСР [28]. В рамках этого мероприятия университеты предоставили информацию о ряде таких индикаторов, в том числе: текущие показатели доходов на научные исследования и соответствующие расходы, управление ресурсами и поддержка исследований и учёных, научные планы и приоритеты. Информация предоставлялась в заявках факультетов или групп факультетов вузов, объединённых по направлениям финансирования в соответствии со статистикой в области высшего образования (37 “центров затрат”). Заявки, кроме того, содержали краткое описание основных результатов исследований, включая список из пяти лучших публикаций за последние пять лет. На этой основе была выполнена экспертная оценка заявок в тематических подкомитетах UGC в рамках непрозрачной процедуры.

Результаты оценки учитывались при распределении части ежегодного финансирования университетов и вызвали серию критических замечаний относительно недостаточной представленности отдельных дисциплин, дискриминации небольших или относительно новых факультетов, необоснованности некоторых выводов. Широкой критике общественности подвергся и каждый следующий раунд оценки. Часто подобная критика содержалась в полноценных исследовательских статьях, где в том числе предлагались альтернативные подходы. При этом координационный орган и лица, ответственные за принятие решений, гибко реагировали на критику, предложения и аргументацию [30, с. 10], что имело своим следствием достаточно серьёзные изменения в методике и процедурах (табл. 3).

В 1992 г. был принят Закон о дополнительном и высшем образовании, согласно которому в число университетов были включены политехнические и другие вузы. В связи с этим количество университетов почти удвоилось (до 84, а с 1994 г. до 89), что значительно расширило рамки оценки. Этим же законом был создан Совет по финансированию высшего образования Англии (HEFCE) и аналогичные советы в Уэльсе и Шотландии. В Северной

Таблица 3. Восемь циклов оценки научных исследований в Великобритании с 1986 по 2021 г. [30]

Год	Название	Координационный орган	Ключевые характеристики
1986	Research Selectivity Exercise (RSE)	Комитет по грантам университетов (UGC)	37 центров затрат; вопросник из четырёх частей о доходах, расходах, приоритетах планирования и результатах исследований
1989	Research Selectivity Exercise (RSE)	Совет по финансированию университетов	152 единицы оценки; 70 групп экспертной оценки; два результата на одного сотрудника
1992	Research Assessment Exercise (RAE)	Совет по финансированию высшего образования Англии (HEFCE) и аналогичные институты в Уэльсе, Шотландии и Северной Ирландии	Вузы выбирают, какие сотрудники будут представлены; 5-балльная шкала; 2800 представлений для 72 тематических панелей; введение даты переписи для исследователей
1996	Research Assessment Exercise (RAE)	HEFCE	До четырёх результатов на одного исследователя; 69 тематических панелей
2001	Research Assessment Exercise (RAE)	HEFCE	2600 работ по 69 тематическим панелям; пять зонтичных групп председателей комиссий для обеспечения согласованности
2008	Research Assessment Exercise (RAE)	HEFCE	67 подпанелей под 15 основными панелями; результаты представлены в виде профилей качества
2014	Research Excellence Framework (REF)	HEFCE	Четыре основные панели; 36 подпанелей; введение оценки воздействия (импакта) и среды
2021	Research Excellence Framework (REF)	UK Research and Innovation (UKRI)	Включены все сотрудники, несущие значительную ответственность за исследования; гибкое количество результатов

Ирландии до 2016 г. соответствующими функциями было наделено Министерство занятости и обучения, позднее — Министерство экономики. Оценки 1992–2008 гг., получившие название Research Assessment Exercise (RAE), выполнялись комиссиями, созданными и координируемыми этими четырьмя финансирующими институтами.

В 2006 г. правительство объявило о проведении нового мероприятия, которое должно было прийти на смену RAE. Оно получило название Research Excellence Framework (REF). Для упрощения, ускорения и удешевления оценки планировалось ввести новую систему, основанную на метриках. Однако позже, ввиду существенной критики, подходы были пересмотрены и рецензирование осталось основным элементом оценки [31]. Результаты REF 2014 были критически рассмотрены комиссией под руководством лорда Николаса Стерна [32], созданной министром науки и высшего образования в 2015 г. Последний по времени раунд оценки был проведён в 2021 г. Советом по исследованиям и инновациям Великобритании (UK Research and Innovation, UKRI) — вневедомственным государственным органом правительства Соединённого Королевства, который с 2018 г. распределяет

финансирование исследований и инноваций, обеспечиваемое научным бюджетом Министерства бизнеса, энергетики и промышленной стратегии.

В связи с изменениями законодательства о высшем образовании в четырёх странах, составляющих Великобританию, в оценке 2021 г. проявили себя дальнейшие терминологические трансформации. Первоначально оценка касалась только университетов, потом стал использоваться термин “высшее учебное заведение” (higher education institution, HEI), включающий отдельные колледжи и другие типы организаций, в последнее время используется ещё более широкий термин “учреждение высшего образования” (higher education provider, HEP). Предполагается, что оценке подвергаются и, соответственно, могут потом получить финансирование на научные исследования те учреждения, которые присваивают учёные степени и имеют соответствующие программы обучения. В оценке 2021 г. участвовало 157 таких организаций, подавших 1878 заявок, характеризующих результаты 76132 исследователей.

Ещё до подведения итогов последнего раунда UKRI инициировал программу реформирования системы оценки Future research assessment

programme (FRAP) [33], ориентированную на изучение возможных подходов к оценке результатов исследований в области высшего образования Великобритании. Направления работ включают в себя оценку REF 2021, анализ международной практики оценки исследований и тщательное изучение возможных моделей и подходов к оценке с тем, чтобы определить те из них, которые позволяют поощрять проведение качественных исследований и повышать оценку воздействия, а также поддерживать позитивную исследовательскую культуру, одновременно упрощая и снижая административную нагрузку на сектор высшего образования.

Практически после каждого раунда оценки поднимался вопрос о высокой стоимости, значительной административной нагрузке, длительности оценочной процедуры и рассматривалась возможность перехода на более формальную количественную модель оценки с уменьшением влияния экспертизы. Однако каждый раз эти предложения отвергались и обосновывалась необходимость именно экспертного подхода [25]. В рамках Future research assessment programme было инициировано исследование, направленное на анализ возможности использования искусственного интеллекта для оценки результатов исследований [34], которое показало, что пока это невозможно [35, 36].

ЭВОЛЮЦИЯ REF

Опыт, аккумулированный в ходе критики, обсуждений, изменений британской системы оценки за прошедшие без малого четыре десятилетия позволяет проследить основные проблемы экспертного подхода к оценке науки и способы их решения. Посмотрим через эту призму на основные составляющие REF.

К сожалению, сведения о первых раундах оценки RSE1986 и RSE1989 сохранились только в государственных архивах, а также разного рода публикациях. Результаты оценок RAE были представлены на соответствующем сайте, недоступном в настоящий момент, но сохранившемся в интернет-архиве. Богатый набор материалов, сопровождающих два раунда оценки REF, представлен на сайте <https://www.ref.ac.uk>. Неоценимым источником информации являются два отчёта, представленные в 2003 г. под руководством сэра Гарета Робертса [37] и в 2016 г. под руководством лорда Николаса Стерна [32]. Эти отчёты оказали значительное влияние на последующее развитие системы оценки.

Объект оценки. В первом раунде RSE объектом оценки стали факультеты или группы близких по тематике факультетов, представлявшие результаты в 37 тематических подкомиссий в соответствии с направлениями затрат, определёнными

в статистических отчётах. Уже в следующем раунде было введено понятие “единиц оценки” (Units of Assessment, UoA), соответствующих тематическим направлениям. В процессе консультаций были сформированы 152 единицы оценки с более детальной разбивкой по направлениям исследований, что позволило университетам подавать заявки в том числе от небольших факультетов и даже отдельных исследовательских групп и центров. В дальнейшем набор UoA постепенно уменьшался за счёт укрупнения единиц, и в последнем раунде университеты подавали заявки лишь в 34 UoA (рис. 1).

Один университет обычно подаёт не более одной заявки в одну UoA, но есть исключения. В случае, когда два и более университета подают в UoA совместную заявку в связи с наличием тесно сотрудничающей исследовательской группы, каждый из университетов может подать дополнительную заявку в ту же UoA. По решению комиссий университету может быть позволено подавать более одной заявки в один UoA, если в его составе есть несколько независимых исследовательских групп с отчётливо различающимися направлениями исследований. Также предусмотрена возможность подать несколько заявок в случае недавнего объединения вузов. Таким образом, отдельным объектом оценки обычно становится исследовательская группа, объединённая как минимум общим научным направлением. Предварительные решения по следующему раунду оценки REF 2028 оставляют структуру UoA без изменений [38].

Экспертные группы. В первом раунде оценки экспертные группы были сформированы на основе имеющихся тематических советов при UGC, в которые входили авторитетные учёные. Во втором раунде эти советы были доукомплектованы представителями индустрии, органов власти, функционерами из тематических исследовательских советов и независимых финансирующих организаций. Кроме того, приглашались учёные для заполнения пробелов в экспертизе по отдельным направлениям наук. В дальнейшем происходил постепенный переход к формированию экспертных панелей под каждый следующий раунд на основе номинаций широкого круга научных обществ, предметных ассоциаций и других заинтересованных организаций отбором на основе известности и опыта. Председатели экспертных групп отбирались финансирующими советами и организовывали процесс отбора экспертов. Требовалось, кроме того, обеспечить представительство заказчиков исследований и сообществ пользователей в сфере торговли, промышленности, правительства и государственного сектора, а также учесть необходимость наличия экспертов с коллективным знанием исследовательской деятельности в Великобритании и в учреждениях с разной историей. Для контроля за отбором экспертов стали создаваться отдельные рабочие группы.

Панель А. Медицина, здравоохранение и науки о жизни	Панель В. Физические, инженерные и математические науки	Панель С. Социальные науки	Панель Д. Искусство и гуманитарные науки
<ol style="list-style-type: none"> 1. Клиническая медицина 2. Общественное здравоохранение, службы здравоохранения и первичная помощь 3. Смежные медицинские профессии, стоматология, сестринское дело и фармацевтика 4. Психология, психиатрия и неврология 5. Биологические науки 6. Сельское хозяйство, пищевые и ветеринарные науки 	<ol style="list-style-type: none"> 7. Науки о Земле и экологические науки 8. Химия 9. Физика 10. Математические науки 11. Вычислительная техника и информатика 12. Инженерия 	<ol style="list-style-type: none"> 13. Архитектура, градостроение и планирование 14. География и экологические исследования 15. Археология 16. Экономика и эконометрика 17. Бизнес и менеджмент 18. Право 19. Политика и международные отношения 20. Социальная работа и социальная политика 21. Социология 22. Антропология и исследования развития 23. Образование 24. Спорт и физическая культура, досуг и туризм 	<ol style="list-style-type: none"> 25. Краеведение 26. Современные языки и лингвистика 27. Английский язык и литература 28. История 29. Антиковедение 30. Философия 31. Теология и религиоведение 32. Искусство и дизайн: история, практика и теория 33. Музыка, драма, танцы, исполнительские искусства, киноведение 34. Коммуникация, культурные и медиа исследования, библиотечные и информационные науки

Рис. 1. Основные панели и экспертные группы (УоА) в REF 2021

В конечном счёте за назначение председателей, их заместителей и членов экспертных групп отвечают финансирующие советы, но процесс остаётся в значительной степени открытым, все результаты публикуются. Предпринимаются определённые усилия для значительного обновления состава экспертных групп с сохранением ядра экспертов, имеющих опыт участия в предыдущих раундах. В REF 2021 в качестве экспертов участвовало 900 представителей исследовательского сообщества, 220 представителей пользователей исследований, назначенных в 34 экспертные группы, которые были скомпонованы в четыре панели. Такая двухуровневая структура впервые использовалась в RAE 2008 для общей настройки и последующей детализации принципов оценки, подходящих для как широких, так и узких направлений науки, для экспертизы междисциплинарных исследований. Это позволяет сохранять общие принципы экспертизы и при этом учитывать особенности отдельных направлений науки в рамках соответствующих экспертных групп, обеспечивая сопоставимость результатов оценки.

В среднем на экспертную группу в 2021 г. приходилось 55 заявок, 5452 результата, 187 заявлений о воздействии (impact case), группа состоит из 26 экспертов (включая председателей и их заместителей) и девяти советников — представителей от

потребителей результатов исследований (табл. 4). Но ситуация по экспертным группам может сильно различаться, в соответствии с этим различается и нагрузка на экспертов. Максимальная нагрузка в тематике “12. Инженерия” в панели В, на 35 экспертов там приходится 18207 результатов, а минимальная — в направлениях “5. Краеведение” и “29. Антиковедение”, где на эксперта приходится в среднем только 71 результат.

Исследовательский персонал. Отбор персонала, подлежащего оценке, был предметом постоянных споров. В первом раунде оценки не было привязки к персоналу, что вызвало критику, так как крупные факультеты или их группы получали преимущество за счёт большего числа грантов, дохода от заказных исследований, наград и призов. Во втором раунде требовалось представить детальные данные по всем исследователям (академический и научный персонал), включая фамилию и инициалы, дату рождения, текущее место работы, структуру заработной платы, учёные степени и звания, награды и т.д. Третий раунд (RAE 1992) был ориентирован на представление в заявках основных исследователей, которые отбирались вузами по их усмотрению из числа работающих на определённую комиссией дату (census date). Эта схема сохранялась до REF 2014 включительно, но в оценке 2021 г. подход изменили. В соответствии

Таблица 4. Статистика по экспертным панелям и группам (UoA) в REF 2021

Панель	Экспертная группа	Заявок	Результатов	Импактов	Экспертов	Советников
А. Медицина, здравоохранение и науки о жизни	1. Клиническая медицина	31	11991	254	33	6
	2. Общественное здравоохранение, службы здравоохранения и первичная помощь	33	4941	151	25	9
	3. Смежные медицинские профессии, стоматология, сестринское дело и фармацевтика	89	11627	393	44	17
	4. Психология, психиатрия и неврология	93	9773	326	27	21
	5. Биологические науки	42	7109	192	31	2
	6. Сельское хозяйство, пищевые и ветеринарные науки	24	3430	103	23	5
Итого А		312	48871	1419	22	5
В. Физические, инженерные и математические науки	7. Науки о Земле и экологические науки	38	4387	148	19	4
	8. Химия	40	3692	113	16	6
	9. Физика	44	5495	169	27	3
	10. Математические науки	53	6000	176	27	10
	11. Вычислительная техника и информатика	90	7294	271	26	8
	12. Инженерия	88	18207	391	35	40
Итого В		353	45075	1268	18	6
С. Социальные науки	13. Архитектура, градостроение и планирование	37	3688	127	22	10
	14. География и экологические исследования	56	4477	180	21	8
	15. Археология	23	1208	59	13	7
	16. Экономика и эконометрика	25	2232	88	19	4
	17. Бизнес и менеджмент	108	16032	504	37	14
	18. Право	68	5855	226	21	16
	19. Политика и международные отношения	56	4707	166	26	9
	20. Социальная работа и социальная политика	75	5156	222	30	7
	21. Социология	37	2669	107	26	0
	22. Антропология и исследования развития	26	1762	77	19	3
	23. Образование	83	5272	230	22	14
24. Спорт и физическая культура, досуг и туризм	61	3539	160	23	2	
Итого С		655	56597	2146	23	6

Окончание таблицы 4 на стр. 42

Таблица 4 (окончание)

Панель	Экспертная группа	Заявок	Результатов	Импактов	Экспертов	Советников
D. Искусство и гуманитарные науки	25. Краеведение	23	1429	57	20	6
	26. Современные языки и лингвистика	47	3876	154	27	10
	27. Английский язык и литература	92	6505	273	28	10
	28. История	81	5746	240	21	14
	29. Антиковедение	17	1069	48	15	2
	30. Философия	35	1708	85	12	7
	31. Теология и религиоведение	29	1246	68	15	3
	32. Искусство и дизайн: история, практика и теория	86	6375	262	32	8
	33. Музыка, драма, танцы, исполнительские искусства, киноведение	84	3680	196	25	7
	34. Коммуникация, культурные и медиа исследования, библиотечные и информационные науки	58	3176	145	20	5
Итого D		552	34810	1528	19	5
Всего		1872	185353	6361	869	309

с рекомендациями отчёта Стерна [32] в REF 2021 университеты должны были представлять всех сотрудников, ответственных за исследования, то есть тех, у кого основной работой являются исследования или преподавание и исследования и кто работает на не менее чем 0.2 ставки. Это требование привело к заметным затруднениям в подаче заявок и большому количеству разъяснений и уточнений в связи с обстоятельствами, связанными с учётом внешних совместителей, работников с почасовой оплатой, а также находящихся в длительных отпусках и т.д. Предполагается, что в REF 2028 это требование будет изменено и университеты вообще не будут подавать сведения о сотрудниках [38]. Вместо этого данные будут получены из отчётов, регулярно подаваемых в Агентство по статистике в области высшего образования (Higher Education Statistics Agency, HESA).

Результаты исследований. Как было сказано выше, в первом раунде университеты включали в каждую заявку по пять лучших публикаций. В следующем раунде RSE 1989 на каждого сотрудника, занятого на полной ставке, требовалось подать для оценки две публикации, вышедшие в период с начала 1984 г. по конец 1988 г., автором или соавтором которых был представленный сотрудник и которые явились результатом его работы в рамках

оцениваемого УоА. Это привело к тому, что одна и та же публикация с несколькими авторами могла быть упомянута несколько раз как разными университетами, так и разными УоА в одном и том же университете. Кроме того, требовалось включить в заявку полные списки всех публикаций сотрудников за 1984–1988 гг. В следующем раунде (1992) в связи с расширением списка вузов допускалось представление результатов в виде не только научных публикаций, причём предусматривалась подача не более двух результатов на каждого активного исследователя. В 1996 г. максимальное количество результатов было увеличено до четырёх, и это ограничение действовало вплоть до REF 2021.

В REF 2021 в связи с новым требованием — упоминанием всех исследователей, участвующих в заявке, были введены минимальное (не менее одного) и максимальное (не более пяти) ограничения на количество результатов на одного исследователя. При этом требовалось сообщить об общем количестве результатов, равном 2.5, умноженном на количество штатных единиц исследователей. Результаты могут быть в том числе за авторством уже не работающих в вузе на дату переписи сотрудников или быть аффилированными с другой организацией в случае, если исследователь перешёл в данный вуз позднее. Это изменение связано со

стремлением не привязывать результаты к исследователям, имея в виду, что исследователи, в том числе ведущие, могут показывать разную результативность. Предложения по поводу грядущего раунда оценки идут ещё дальше: предполагается снять требование относительно максимального и минимального количества результатов и ограничиться требованием аффилированности результатов с данным университетом на момент их опубликования или обнародования.

Несмотря на то, что ограничения на виды результатов были сняты в 1992 г., бóльшая часть представляемых результатов (84% в REF 2021) — это по-прежнему научные публикации в рецензируемых журналах (рис. 2). Тем не менее на рассмотрение были предложены различные виды результатов, большее их разнообразие отмечается в социальных науках (панель С) и особенно в искусстве и гуманитарных науках (панель D). Распределение результатов REF 2021 по видам в разрезе панелей показано на рисунке 3, а сведения о результатах, отличных от наиболее распространённых их видов, в заявках — на рисунке 4.

Импакт. В 2014 г. при проведении REF 2014 был введён новый показатель — Impact. С его помощью измеряется любое воздействие, изменение или польза для экономики, общества, культуры, государственной политики, сферы услуг, здоровья, окружающей среды или качества жизни за пределами научных кругов. В REF 2014 на Impact приходилось 20% оценки, а в REF 2021 коэффициент увеличился до 25%.



Рис. 2. Распределение результатов по типам в REF 2021

Показатель, связанный с более широким эффектом оценки, присутствовал начиная с RAЕ 2008 и назывался репутацией (esteem), но он определялся на основе ничем не подтверждённого мнения экспертов. С 2014 г. в заявки стали включать форму Impact Case Study, которая представляет собой документ, разбитый на пять пунктов и описывающий влияние проведённых исследований на разные сферы [39]. Включение показателя влияния исследований в REF 2014 изначально рассматривалось

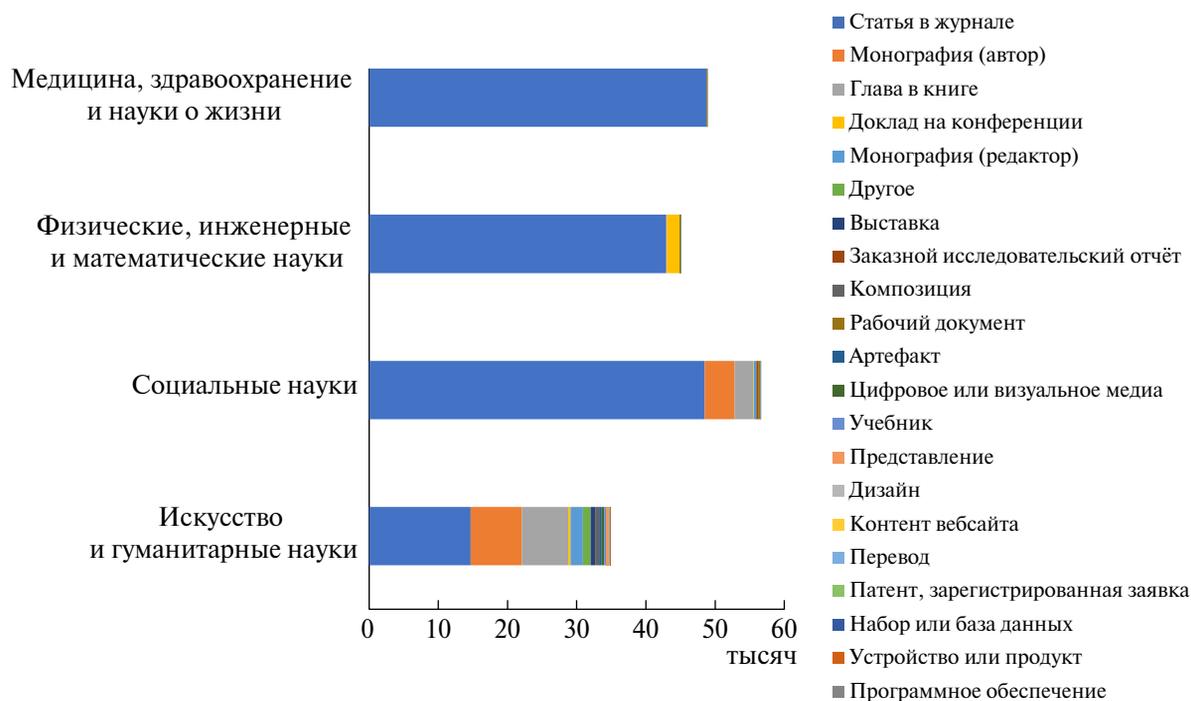


Рис. 3. Распределение результатов по типам в REF 2021, по экспертным панелям



Рис. 4. Количество результатов менее распространенных типов в REF 2021

как спорное, как ответ на требование Министерства финансов обосновывать общественную значимость финансирования университетских исследований. Методика оценки поначалу была не очень сбалансированной, поскольку воздействие должно было основываться на исследованиях, выполненных в период 1993–2014 гг., что могло поставить в невыгодное положение те подразделения, которые начали свою работу позднее [40].

Вузы должны были приводить примеры воздействия, демонстрировавшие влияние их исследований за пределами академического сообщества. Количество примеров, которые необходимо было представить каждому подразделению, зависело от количества штатных единиц сотрудников данного подразделения, при этом для самых маленьких подразделений требовалось не менее двух примеров [39]. Помимо примеров воздействия как таковых, необходимо было подготовить заявление с описанием подхода к практическому внедрению результатов исследований, стратегии, способствующей формированию воздействия. Если примеры воздействия оценивались по охвату и значимости, то описания — по тому, насколько подход и стратегия способствуют достижению цели. Такие заявления подавались по одному на общую заявку подразделения.

Предусмотрена классификация типов воздействия в соответствии с соглашением PESTLE (политический, экономический, социальный, технологический, правовой, экологический или природный), широко используемом при разработке государственной политики. К перечисленным типам были добавлены связанные с влиянием научных достижений в таких областях, как здравоохранение и культура, традиционно включаемых в социальный тип. На рисунке 5 показано

распределение примеров воздействия по их типам (областям воздействия) в REF 2021.

Большинство примеров воздействия в той или иной степени касаются более чем одной области. Примеры охватывают очень широкий спектр влияния, включая воздействие внутри организаций, на национальную политику, на международные соглашения и на общественность. Не было обнаружено никаких свидетельств того, что какой-либо конкретный тип воздействия оценивался особым образом или получал большее количество баллов, чем какой-либо другой. Примеры воздействия, получившие высокие баллы, как правило, содержали четкие доказательства масштаба и значимости влияния научных результатов. Исследования, лежащие в их основе, соответствовали пороговому требованию международного превосходства, а также содержали убедительное описание того, как исследование привело к воздействию на жизнь социума. Высокие оценки воздействия были получены как в небольших, так и в крупных заявках, а количество предлагаемых случаев не давало преимуществ конкретным организациям.

Исследовательская среда. В рамках REF 2021 понятие “исследовательская среда” (environment) означает условия для поддержки исследований и обеспечения их результативности в каждом представленном подразделении. Среда составляет 15% от общего результата, присуждаемого каждой заявке, и оценивается по двум критериям: жизнеспособность и устойчивость.

Вузы должны представить описательные данные об условиях поддержки исследований и их результативности в каждом подразделении, а также данные о доходах от исследований в натуральной форме и защищенных диссертациях. Максимальный

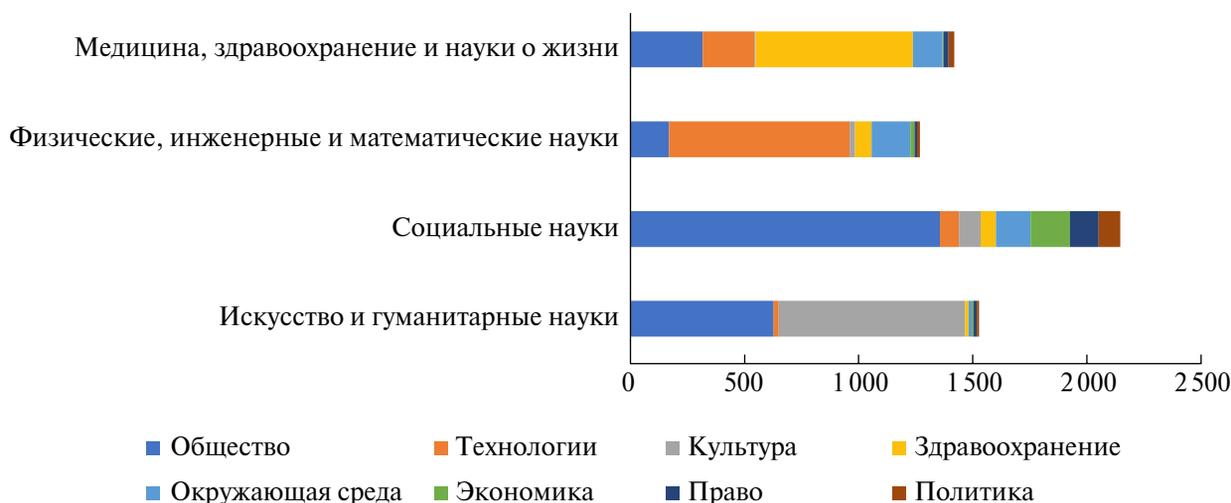


Рис. 5. Распределение количества примеров воздействия по их типам (областям) и экспертным панелям в REF 2021

объём описательной части определялся количеством полных ставок сотрудников подразделения.

В рамках пилотного проекта по оценке институциональной среды вузы также должны были подготовить отчёт на институциональном уровне. Описания институционального уровня использовались подгруппами для обоснования и контекстуализации их оценки, но не участвовали в оценке непосредственно.

Оценка среды, как указано в руководстве REF, во многих случаях оказалась в меньшей степени основанной на фактических данных, чем результаты или воздействие. В отличие от тематических исследований воздействия (где ожидалось наличие подтверждающих доказательств и проводился аудит), в заявках на оценку исследовательской среды требовалось мало подтверждающих доказательств. Заявления об исследовательской среде содержали пять частей, из которых оценивались четыре последних: обзор; стратегия исследований; персонал, включая штатное расписание, развитие персонала и студентов-исследователей; доходы, инфраструктура и материально-техническая база; сотрудничество и вклад в развитие научного направления.

Оценка сводилась к попытке ответить на вопрос “Похоже ли это на отличное место для работы, в котором должны хорошо себя ощущать как опытные, так и начинающие исследователи и которое будет способствовать развитию научного направления”. В RAE 2008 присутствовала, кроме того, оценка исследовательской среды, но она была обобщённой. Начиная с REF 2014 отдельно оцениваются четыре части заявления, которые затем суммируются для получения соответствующего профиля качества. Не все части заявления подтверждаются цифрами и фактами, ряд экспертов отмечает, что на итоговую оценку может оказывать сильное

влияние стиль написания заявления, его соответствие предложенным критериям, которое не всегда удаётся верифицировать. Тем не менее по результатам оценки REF 2021 эксперты отметили заметный прогресс в организации исследовательской среды, поддержке междисциплинарных исследований, развитии кадрового потенциала и вовлечении в научную работу начинающих исследователей и студентов.

Результат оценки. Вплоть до RAE 2001 включительно экспертная группа выставяла оценку заявке в целом, что вызывало затруднения экспертов при определении оценки по шкале из пяти баллов (звёзд). В связи с этим в 1996 г. были добавлены дополнительные “профили качества” — 3a, 3b, 5*. Это не решило проблему, кроме того, возникли опасения, связанные с наличием выигрышных стратегий по исключению из заявки слабых исследователей и другими искусственными способами повышения качества заявки.

В соответствии с рекомендациями отчёта Роберта [37] в RAE 2008 было принято решение оценивать каждый результат по отдельности и суммировать их для получения “профиля качества”, была принята шкала оценок 0–4 звезды. В случае, если у активного исследователя не набиралось четырёх результатов, нехватка оценивалась как очень слабый результат с 0 звёздами. Были определены условия, при которых этот порог мог быть снижен для определённых категорий исследователей, например, находившихся в длительных отпусках. В REF 2014 этот подход распространили на оценку примеров воздействия и отдельных элементов заявления об исследовательской среде.

Итоговые оценки результатов, импакта и среды имеют вид графиков, показывающих распределение количества соответствующих элементов

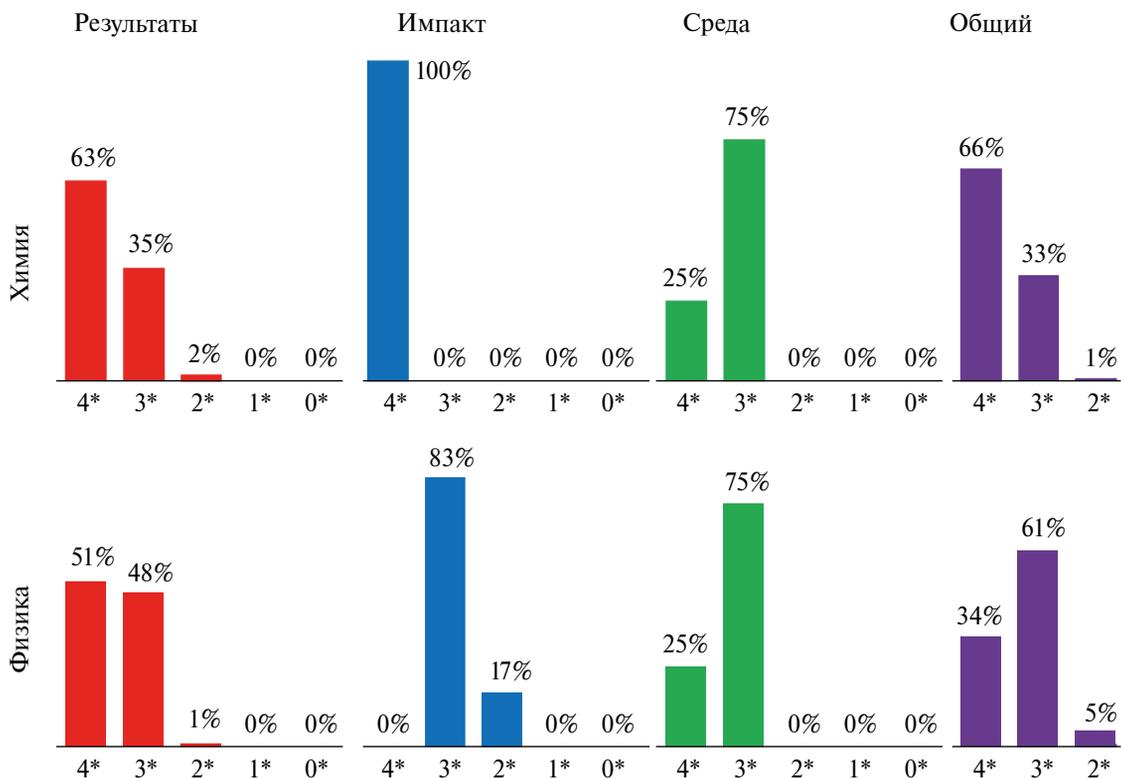


Рис. 6. Пример результатов оценки одного из университетов по двум направлениям (UoA) в REF 2021

оценки по количеству присуждённых звёзд. Профили качества комбинируются в общий, с учётом весовых коэффициентов: в REF 2021 60% приходилось на результаты, 25% на импакт и 15% на исследовательскую среду (рис. 6). При этом общий профиль качества и/или единая оценка для университета не формируется, хотя может быть выведена из профилей качества по отдельным заявкам (UoA).

Ключевые особенности REF. Можно суммировать ключевые особенности британской системы, сформированные на основе опыта восьми раундов оценки следующим образом:

- объект оценки — не организации, а подразделения в рамках тематических групп;
- иерархическая организация экспертных панелей и групп позволяет выполнять тонкую настройку принципов оценки в зависимости от специфики области исследования;
- оцениваются научные результаты; широкое воздействие на общество, политику; кадровый потенциал; исследовательская культура и среда;
- в качестве научных результатов могут быть представлены не только публикации; оценке подлежат не все, а лишь лучшие результаты; определены минимальное и максимальное количество результатов на одного исследователя;
- используются детальные данные об исследовательском персонале, принадлежность к которому не связывается с наименованием должности;
- разработана общая шкала оценки, которая уточняется в рамках тематических экспертных панелей и групп;
- эксперты оценивают каждый компонент отдельно: результат, пример воздействия, каждую составляющую исследовательской среды. Полученные оценки комбинируются в профиль качества, представляющий собой частотное распределение оценок разного уровня, что позволяет избежать когнитивных ошибок, связанных с усреднением;
- результаты оценки, представленные в открытом виде, обеспечивают возможность общественного информирования, анализа и сопоставления, формирования репутации, контроля за эффективностью расходования средств на научные исследования;
- процедура и результаты оценки являются предметом общественного обсуждения, анализируются в рамках заказанных финансирующими организациями исследований, в соответствии с ними корректируются принципы и методология следующего этапа;
- обсуждение, аналитика, опыт подготовки материалов, участие в экспертных советах стали хорошими механизмами трансфера лучших практик,

которые повышают уровень организации и качество научных исследований в национальном секторе исследований и разработок.

НОРВЕЖСКАЯ МОДЕЛЬ

Финансирование научных исследований в вузах на основе метрик результативности было впервые внедрено в Норвегии в 2002 г. в рамках общего тренда в европейских странах, поддержанного Европейской комиссией [41]. Оно было основано на таких индикаторах, как количество присуждённых докторских степеней, объём внешнего финансирования, количество исследовательского персонала на постоянных ставках. Однако и Министерство образования и исследований, и вузы хотели бы использовать индикаторы, в большей степени отражающие публикационную результативность. В связи с этим комитет, созданный Норвежской ассоциацией организаций высшего образования, по поручению министерства разработал в 2005 г. новый индикатор [42], который был впервые использован для определения финансирования в 2006 г. Этот индикатор рассчитывается для государственных университетов и университетских колледжей Норвегии за исключением нескольких вузов в области искусства, архитектуры и дизайна и части частных вузов, а также исследовательских центров и лабораторий.

Основными компонентами норвежской модели являются [41]:

- национальный регистр научных журналов, серий и издателей,
- национальная база данных с библиографическими записями о рецензируемых научных публикациях,
- публикационный индикатор.

Национальный регистр научных изданий и издателей [43] составлен и обновляется на основе решений авторитетного совета. Изначально он создавался на базе трёх источников: списка источников Web of Science, списка научных журналов национальной библиотеки Норвегии Norart (Norwegian and Nordic index to periodical articles) и базы данных издателей академической литературы BIBSYS [42].

Регистр научных изданий и издателей и его разбиение на уровни качества обновляются каждый год с учётом предложений зарегистрированных в системе пользователей. В процессе пересмотра Национальным советом по научным публикациям (Det nasjonale publiseringutvalget, National Board of Scholarly Publishing, NPU) лучшие из них попадают на уровень 2 (наиболее престижные в мире журналы и издатели [44]), удовлетворяющие базовым критериям — на уровень 1, а те, которые признаны не соответствующими требованиям, — на уровень 0. Цель уровня 2 состоит в том, чтобы

стимулировать университеты поощрять своих исследователей к поиску наиболее компетентных рецензий и широкой читательской аудитории [45]. Решения о включении в уровень 1 принимаются на заседаниях NPU три раза в год и имеют ретроспективный эффект, то есть баллы за публикации в этих изданиях назначаются уже в текущем году. Номинации на включение в уровень 2 выполняются отдельными тематическими панелями NPU [34] с учётом предложений исследователей один раз в год и применяются с лагом в два года. Уровень также может быть пересмотрен в сторону снижения, издания уровня 1 могут быть перемещены на уровень X для повторного рассмотрения. Первая версия этого разбиения формировалась по сложной многоуровневой схеме с тем, чтобы учесть издания на норвежском языке, преимущественно в области гуманитарных наук.

Уровень 2 должен быть ограничен издательскими каналами (научными журналами, сериями, ежегодниками, сайтами и книжными издательствами), которые воспринимаются как ведущие в широком предметном контексте. В них публикуются наиболее значимые работы исследователей из разных стран, в совокупности они издают примерно пятую часть (20%) научных публикаций по данному направлению исследований. Пятая часть рассчитывается как доля от всех мировых публикаций в тех предметных областях, где международная статистика журналов доступна и достаточно полна. По остальным областям оценка проводится с использованием статистических данных норвежского сектора высшего образования, которые сравниваются с международной статистикой с учётом книгоиздания. Тематические панели ежегодно получают статистику публикаций по своему направлению для определения этой границы.

Национальная база данных с библиографическими записями о рецензируемых научных публикациях. База данных формируется на основе национальной информационной системы Cristin (Current research information system in Norway) [46], поддерживаемой Sikt — Норвежским агентством по общим сервисам в образовании и исследованиях. В ближайшее время планируется объединить эту систему с создаваемым Национальным научным архивом.

Каждая вносимая в базу данных публикация должна соответствовать следующим четырём критериям [47]: представлять новые идеи в научном формате, позволяющем верифицировать результаты исследования и/или использовать их в новой исследовательской деятельности; на языке и с распространением, которые делают публикацию доступной для соответствующей научной аудитории; в издательском канале (журнал, серия, книжное издательство), который представляет авторов из нескольких институтов и организует независимое рецензирование рукописей перед публикацией.

Таблица 5. Веса публикаций в зависимости от вида издания и уровня в регистре

Вид издания	Уровень 1	Уровень 2
Монография	5	8
Статья в периодическом издании или серии (с ISSN)	1	3
Статья в сборнике, глава в книге (ISBN)	0.7	1

Информационная система *Cristin* включает профили исследователей и информацию о научных проектах. Реализован прикладной программный интерфейс для авторизованного доступа к данным системы, что позволяет разрабатывать внешние публичные и институциональные системы и использовать накопленные данные для расчётов, создания информационных и аналитических материалов.

Публикационный индикатор. Публикационный индикатор организации — это сумма баллов, присваиваемых отдельным публикациям, с аффилиациями этой организации с использованием полного фракционного счёта — по авторам и их аффилиациям [42]. Аффилиационная доля вычисляется делением единицы на количество авторов, делением полученной авторской доли на количество его аффилиаций и последующим суммированием аффилиационных долей авторов по организациям. Аффилиационная доля умножается на вес публикации в зависимости от вида издания и его уровня в национальном регистре (табл. 5), что обеспечивает сравнимость показателя с учётом публикационных практик в различных областях исследований на уровне организаций [45]. Фракционный счёт “штрафует” крупные коллаборации, и для того чтобы снизить влияние этого эффекта, в первоначальной версии методики использовалось ограничение — фракционная доля аффилиации не могла быть менее 1/10. В 2015 г. на основании оценки методики, выполненной в Датском центре анализа исследований при Орхусском университете, в методику были внесены небольшие изменения, в частности ограничение в 1/10 для аффилиационной доли было заменено на вычисление квадратного корня из неё и установлен дополнительный коэффициент 1.3 для публикаций в международных коллаборациях.

Норвежский публикационный индикатор был положительно воспринят научным сообществом [48], так как позволил учитывать научные публикации во всех областях исследований по сравнению с индексируемыми в ведущих базах данных и обеспечил учёт разных публикационных практик в различных дисциплинах на основе более

широкого круга изданий и введения второго уровня, формируемого тематическими панелями NPU. Указывается на высокий уровень прозрачности и предсказуемости норвежской модели [49], в то же время высказываются опасения, что норвежские учёные могут сделать особый упор на публикациях в изданиях уровня 1 с тем, чтобы набрать больше очков. Однако исследования показывают, что введение публикационного индикатора привело к общему росту публикационного потока, росту доли публикаций в ведущих изданиях, расширению коллабораций [41, 50]. Отмечается, что за счёт хорошего дизайна индикатора и его сравнительно незначительного влияния на объёмы финансирования его использование не привело к существенным негативным изменениям публикационного поведения норвежских учёных или принятию направленных на искусственное раздувание индикатора управленческих решений [51]. В некоторой степени это может объясняться высоким уровнем нетерпимости к недобросовестному профессиональному поведению в норвежском обществе [52]. Хорошим знаком служит тот факт, что введение индикатора и сопутствующее увеличение публикационного потока не привело к снижению средних показателей цитируемости публикаций, в отличие от эффекта, наблюдавшегося после запуска системы финансирования на основании оценки результативности в Австралии [12, 53].

Ключевые особенности Норвежской модели. Успех Норвежской модели основан на её сравнительной простоте.

- Экспертный подход к формированию и категорированию Национального регистра научных журналов, серий и издателей с учётом мнений исследовательского сообщества, высокая степень публичности процедур и результатов отбора и категорирования.

- Выравнивание шансов на публикацию в лучших изданиях для представителей разных дисциплин благодаря включению национальных изданий и закрепления порога по доле публикаций в издательских каналах верхнего уровня в общем объёме по каждой тематике.

- Публично доступная база данных публикаций норвежских учёных, по которой рассчитывается публикационный показатель, возможность верификации расчётов, проведения аналитических исследований.

- Простой и понятный публикационный показатель, снижающий штрафы за коллаборации, накладываемые фракционным счётом, поощряющий международные коллаборации, учитывающий высокую значимость качественных монографий в ряде областей исследований.

ПУТИ СОВЕРШЕНСТВОВАНИЯ РОССИЙСКОЙ СИСТЕМЫ ОЦЕНКИ РЕЗУЛЬТАТИВНОСТИ

Два контура системы оценки результативности, сложившиеся к настоящему времени в России, пока не обрели окончательные очертания. Накопленный опыт, изменения, произошедшие за последние годы в области научной политики и исследовательского ландшафта в РФ, богатый зарубежный опыт позволяют отметить некоторые недостатки и наметить пути их исправления.

Количественная оценка. В развитие белого списка научных журналов, подготовленного межведомственной рабочей группой при участии РЦНИ и соответствующей информационной системы, на первом этапе его необходимо дополнить сериальными изданиями сборников трудов конференций. Далее нужно предусмотреть переход к формированию этого списка на экспертной основе, для чего следует создать соответствующий экспертный комитет, включая тематические группы из авторитетных представителей различных научных областей. Необходимо нормативно предусмотреть принципы ротации этого совета. Работа совета должна поддерживаться информационной системой, обеспечивающей публичный доступ к самому списку, соответствующей аналитике, возможность подачи и отслеживания предложений зарегистрированными исследователями. Основы такой системы уже заложены в РЦНИ. Рейтингование изданий, разделение их на уровни качества должно максимально обеспечивать равные возможности публикации в высокорейтинговых изданиях представителям разных научных дисциплин.

Методику расчёта комплексного балла публикационной результативности следует адаптировать к белому списку, нужно также компенсировать последствия фракционного счёта для крупных коллабораций, применяя вычисление вклада организации в публикацию как квадратный корень из фракционной доли, страновой фракционный счёт [54] или иные подходы.

Необходимо усовершенствовать практику ежегодных публикационных отчётов организаций через создание национальной информационной системы, объединяющей данные об исследователях, организациях, публикациях и научных проектах, — системы класса CRIS (Current Research Information System). Открытый доступ к данным этой системы позволит осуществлять общественный контроль и информирование о научных результатах отечественных учёных представителей бизнеса и органы государственной власти, обеспечит широкие возможности формирования производных информационных систем и интеллектуальных сервисов по поиску экспертов в определённых тематиках исследований.

Таким образом, основными компонентами количественной оценки должны стать:

- экспертный комитет для отбора и категорирования научных изданий, контроля за соблюдением научной этики, объединяющий тематические экспертные группы и состоящий из авторитетных представителей различных областей знания;
- национальный белый список изданий, разделённый на несколько уровней, формируемый на экспертной основе с учётом предложений исследовательского сообщества;
- публикационный показатель, адаптированный к белому списку и поощряющий коллаборации — новый комплексный балл публикационной результативности;
- национальная CRIS — информационная система, аккумулирующая данные о российских научных организациях, исследователях, научных публикациях и проектах.

Для снижения рисков и негативных последствий недобросовестного поведения отдельных учёных и научных коллективов необходимо информировать научное сообщество о непродуктивных публикационных практиках, контролировать соблюдение научной этики. С этой целью могут создаваться соответствующие советы или комиссии, выполняющие в том числе и выборочные контрольные функции. Эту задачу можно возложить на экспертный комитет по отбору и рейтингованию изданий. Конечно, необходимо, чтобы нарушения научной этики имели последствия как для отдельных учёных, так и для их организаций-работодателей. Такие последствия, например, могут состоять в штрафах, применяемых к публикационному показателю организации.

Экспертная оценка. Экспертная оценка организаций является комплексным трудоёмким процессом и должна производиться не чаще, чем раз в пять лет. Проведение раунда экспертной оценки следует сопровождать формированием экспертного совета. Имея в виду различные практики в разных тематических дисциплинах, необходимо, чтобы в совет входили экспертные группы, составленные из представителей соответствующих дисциплин. Для учёта дисциплинарных различий важно также, чтобы общие принципы оценки, закреплённые на верхнем уровне, последовательно детализировались и уточнялись по отраслям и направлениям исследований в экспертных советах соответствующего уровня. Международный опыт позволяет предложить *трёхуровневую структуру*, включая на первом уровне непосредственно экспертный совет, на втором уровне от трёх до пяти отраслевых групп (медицина, здравоохранение и науки о жизни; другие естественные, инженерные и математические науки; общественные науки; гуманитарные науки и искусство), а на третьем уровне 30–40 подгрупп по отдельным

направлениям исследований. Такой подход позволит в должной степени учитывать дисциплинарную специфику. Экспертная оценка может проводиться и в монотематических научных организациях в рамках одной подгруппы, и в федеральных исследовательских центрах, работающих по нескольким научным направлениям, и в крупных университетах, для которых характерен широкий научно-дисциплинарный спектр.

Наиболее традиционный объект экспертной оценки — результаты, в качестве которых обычно рассматривают научные публикации. Однако, на наш взгляд, ограничивать формы и виды результатов только публикациями не следует; необходимо обеспечить возможность представлять также аналитические отчёты, изобретения, программы и базы данных, информационные системы и веб-сайты, а для вузов культуры актуальными результатами могут быть театральные постановки. Адекватно оценить все результаты организации за длительный период вряд ли можно, поэтому оценке необходимо подвергать только лучшие достижения, отобранные самими исследовательскими коллективами. Количество таких результатов можно ограничить как сверху, чтобы обеспечить выполнимость оценки, так и снизу, чтобы был представлен вклад всех членов научного коллектива. Подобные ограничения могут быть привязаны к списочному составу исследователей, чтобы учитывать размеры коллективов.

Важным элементом оценки должно стать понимание воздействия, которое научные достижения оказывают на общество, экономику или политику. В последнее время много говорится о необходимости внедрения результатов исследований в экономику, но не все они могут использоваться таким образом. Общая оценка влияния, в том числе культурного, социального, политического, позволит расширить представление о прикладном характере исследований практически на все научные направления. Едва ли возможно формализовать метрики и индикаторы такого импакта, поэтому оптимальным будет формирование нарративных изложений воздействия результатов работы научных коллективов в адекватных для экспертизы форме и объёме. Максимальное и минимальное количество отдельных примеров воздействия может быть увязано с численностью оцениваемого коллектива. Для адекватной оценки импакта в состав экспертных рабочих групп следует включать представителей промышленности, органов государственной власти, других потенциальных заказчиков.

Наука и высшее образование связаны с формированием человеческого капитала, поэтому важно оценивать условия для развития кадрового потенциала, профессионального роста, проведения исследований, созданные в научном коллективе. Предметом оценки может быть нарративное

описание этих условий, а свидетельством высокого уровня могут выступать и формальные показатели, связанные с подготовкой кадров высшей квалификации, успехами в привлечении конкурсного финансирования, объёмами заказных прикладных исследований.

Очевидно, что тематические группы экспертов должны оценивать заявки не в целом от организации, а от её составных частей соответствующей тематической направленности. В случае междисциплинарных исследований ведущая по тематике группа экспертов должна привлекать к оценке представителей других экспертных групп. Таким образом, оцениваемые организации будут подавать несколько заявок в разные тематические группы от соответствующих подразделений. В случае выполнения комплексных проектов возможна подача совместных заявок несколькими организациями.

Экспертная оценка отдельных элементов — результатов, импакта, среды — должна выполняться не менее чем двумя экспертами по балльной шкале и впоследствии комбинироваться в общую оценку по каждой тематической группе. Оптимальным способом комбинации на основе международного опыта является формирование профилей (или спектров) качества, отражающих распределение элементов оценки по баллам или их диапазонам. С учётом трёх направлений оценки такой результат будет представлять собой многовекторную комплексную меру, позволяющую в том числе не потерять первоклассные результаты на общем среднем фоне.

Таким образом, основными компонентами экспертной оценки должны стать:

- методика формирования заявок на оценку, с помощью которой определяется необходимое количество результатов и заявлений о воздействии в зависимости от численности исследователей в коллективе, общие принципы экспертизы, соответствующая информационная система;
- принципы и механизм формирования экспертного совета, его отдельных направлений и групп в соответствии с тематиками исследований с участием представителей исследовательского сообщества, бизнеса, органов государственной власти;
- методика формирования профилей качества из отдельных экспертных оценок результатов, импакта и среды, информационная система для публичного доступа к результатам оценки.

Важно, чтобы процедура и результаты оценки стали предметом общественного обсуждения, анализировались в рамках заказных исследований, в соответствии с ними корректировались принципы и методология следующего этапа. В этих целях результаты оценки должны быть представлены в открытом виде, чтобы обеспечить возможность общественного информирования, анализа и сопоставления, формирования репутации,

общественного контроля за эффективностью расходования средств на научные исследования. Необходимо отметить, что обсуждение, анализ, опыт подготовки материалов, участие в экспертных советах являются действенными механизмами трансфера лучших практик, могут способствовать повышению общего уровня организации и качества научных исследований.

* * *

Несмотря на сформировавшийся в российском научном сообществе и во власти запрос на новую систему оценки научной деятельности, пока здесь не удалось существенно продвинуться. На конец 2023 г. подготовлен лишь белый список журналов, механизмы применения и обновления которого пока не до конца ясны. Предпринято несколько безуспешных попыток формально квантифицировать критерии оценки уровня готовности технологий. Даже если вопреки мировой практике эту задачу и удастся решить, то только с привлечением трудоёмкой экспертизы и лишь для части научных дисциплин.

Причины отсутствия конструктивных идей следует искать в большом дефиците аналитических работ по уже проведённым в России процедурам оценки. За очень редкими исключениями отсутствуют исследования и аналитические отчёты как о результатах категорирования научных организаций, так и о методике комплексного расчёта публикационной результативности. За прошедшее время этот опыт так и не был должным образом осмыслен, не сформирован интеллектуальный задел для его дальнейшего развития.

Эта ситуация накладывается на естественный фактор сменяемости управленческих команд в Министерстве науки и высшего образования РФ. Каждый новый цикл оценки фактически приходится на его новый состав, который вынужден решать эту задачу как бы с чистого листа. Для этого формируются рабочие группы, дискуссии в которых нередко сводятся к обмену личным опытом и борьбе мнений при слабой доказательной базе. Как правило, это позволяет избежать откровенно неудачных решений, но вряд ли достаточно для разработки сколько-нибудь сложной системы, эффективно отвечающей на вызовы времени. В этих условиях целесообразно обратить внимание на богатый зарубежный опыт и адаптировать его с учётом национальной специфики. Рассмотренные примеры из Великобритании и Норвегии представляют собой модельные системы, характеризующиеся рядом параметров. Меняя эти параметры, при необходимости добавляя новые элементы, можно осуществить их тонкую настройку под нужды российской науки.

Стоит ли сохранять и развивать сложившуюся в России двухконтурную систему? Да, поскольку оба контура необходимы. Количественная оценка

без значительных трудозатрат позволяет определить базовый уровень научной результативности и задаёт планку, нахождение ниже которой является признаком слабости (хотя для этого могут быть и объективные причины). Она используется для формирования и контроля выполнения государственного задания на фундаментальные научные исследования. Однако практика показывает, что ограничиваются только количественными методами нецелесообразно. В частности, это подтверждает норвежский опыт, в котором количественная оценка занимает ограниченное место и сочетается с менее формализованными подходами.

Экспертиза позволяет углубиться в плохо квантифицируемые аспекты исследований, содержательно изучить и оценить вклад организации в науку и социально-экономическое развитие страны. Она необходима для создания адекватной системы целеполагания в управлении исследованиями в университетах и научных организациях, сместив фокус с достижения показателей на увеличение качества научных результатов и их влияния. Пока что достойной альтернативы экспертному подходу не найдено.

Наконец, два контура компенсируют недостатки друг друга. Экспертиза сдерживает от чрезмерного манипулирования показателями, тогда как количественная оценка задаёт рамку для эксперта и ограничивает возможности для излишне субъективных выводов.

В заключение ещё раз отметим, что, поскольку оценка результативности носит в значительной мере формирующий характер, она должна транслировать в исследовательское сообщество адекватные ценности и предотвращать или хотя бы не поощрять неэтичное поведение, недобросовестное отношение к проведению исследований. Так как национальные системы оценки способствуют горизонтальному распространению практик по исследовательским коллективам и организациям, то важно, чтобы подобные практики были продуктивными и соответствовали международным стандартам научной деятельности. Тогда национальная система оценки будет способствовать развитию культуры научной деятельности, повышению качества российских исследований и усилению их вклада в социально-экономическое развитие страны.

ЛИТЕРАТУРА

1. *Gruening G.* Origin and theoretical basis of new public management // *International Public Management Journal*. 2001. V. 4. № 1. P. 1–25. [https://doi.org/10.1016/S1096-7494\(01\)00041-1](https://doi.org/10.1016/S1096-7494(01)00041-1)
2. *Hicks D.* Performance-based university research funding systems // *Research Policy*. 2012. V. 41.

- № 2. P. 251–261. <https://doi.org/10.1016/j.respol.2011.09.007>
3. *Dougherty K.J., Natow R.S.* Performance-based funding for higher education: how well does neoliberal theory capture neoliberal practice? // *Higher Education*. 2020. V. 80. № 3. P. 457–478. <https://doi.org/10.1007/s10734-019-00491-4>
 4. *Henkel M.* The modernisation of research evaluation: The case of the UK // *Higher Education*. 1999. V. 38. № 1. P. 105–122. <https://doi.org/10.1023/A:1003799013939>
 5. *Mok K.H.* Enhancing quality of higher education for world-class status: Approaches, strategies, and challenges for Hong Kong // *Chinese Education and Society*. 2014. V. 47. № 1. P. 44–64. <https://doi.org/10.2753/CED1061-1932470103>
 6. *Crowe S.F., Watt S.* Excellence in Research in Australia 2010, 2012, and 2015: The Rising of the Curate's Soufflé? // *Australian Psychologist*. 2017. V. 52. № 6. P. 503–513. <https://doi.org/10.1111/ap.12248>
 7. *Chatterjee B. et al.* The spectacle of research assessment systems: insights from New Zealand and the United Kingdom // *Accounting, Auditing and Accountability Journal*. 2020. V. 33. № 6. P. 1219–1246. <https://doi.org/10.1108/AAAJ-01-2019-3865>
 8. *Abramo G., D'Angelo C.A.* The VQR, Italy's second national research assessment: Methodological failures and ranking distortions // *Journal of the Association for Information Science and Technology*. 2015. V. 66. № 11. P. 2202–2214. <https://doi.org/10.1002/asi.23323>
 9. *Luwel M.* Performance-based Institutional Research Funding in Flanders, Belgium // *Scholarly Assessment Reports*. 2021. V. 3. № 1. P. 3. <https://doi.org/10.29024/sar.29>
 10. *Engels T.C.E., Guns R.* The Flemish Performance-based Research Funding System: A Unique Variant of the Norwegian Model // *Journal of Data and Information Science*. 2018. V. 3. № 4. P. 45–60. <https://doi.org/10.2478/jdis-2018-0020>
 11. *Shu F., Liu S., Larivière V.* China's Research Evaluation Reform: What are the Consequences for Global Science? // *Minerva*. 2022. V. 60. № 3. P. 329–347. <https://doi.org/10.1007/s11024-022-09468-7>
 12. *Aagaard K., Bloch C., Schneider J.W.* Impacts of performance-based research funding systems: The case of the Norwegian Publication Indicator // *Research Evaluation*. 2015. V. 24. № 2. P. 106–117. <https://doi.org/10.1093/reseval/rvv003>
 13. *Deutz D.B. et al.* Quantitative quality: a study on how performance-based measures may change the publication patterns of Danish researchers // *Scientometrics*. 2021. V. 126. № 4. P. 3303–3320. <https://doi.org/10.1007/s11192-021-03881-7>
 14. *Mathies C., Kivistö J., Birnbaum M.* Following the money? Performance-based funding and the changing publication patterns of Finnish academics // *High Educ.* 2020. V. 79. № 1. P. 21–37. <https://doi.org/10.1007/s10734-019-00394-4>
 15. *Hammarfelt B.* Taking Comfort in Points: The Appeal of the Norwegian Model in Sweden // *Journal of Data and Information Science*. 2018. V. 3. № 4. P. 84–94. <https://doi.org/10.2478/jdis-2018-0023>
 16. Постановление Правительства РФ от 8 апреля 2009 г. № 312 “Об оценке и о мониторинге результативности деятельности научных организаций, выполняющих научно-исследовательские, опытно-конструкторские и технологические работы гражданского назначения”. 2009. <https://sciencemon.ru/documents/6> (дата обращения 05.10.2023).
 17. Информационная справка к Совету при Президенте Российской Федерации по науке и образованию “Об оценке результативности деятельности научных организаций, подведомственных федеральным органам исполнительной власти и государственным академиям наук за 2010–2012 годы”. 2013. <https://sciencemon.ru/documents/2> (дата обращения 05.10.2023).
 18. Поручение Президента Российской Федерации от 30 апреля 2013 г. № Пр-1144 по итогам заседания Совета при Президенте Российской Федерации по науке и образованию. 2013. <https://sciencemon.ru/documents/3> (дата обращения 05.10.2023).
 19. *Гуськов А.Е., Косьяков Д.В., Селиванова И.В.* Методика оценки результативности научных организаций // *Вестник Российской академии наук*. 2018. № 5. <https://doi.org/10.7868/S0869587318050092>
 20. *Kosyakov D., Guskov A.* Research assessment and evaluation in Russian fundamental science // *Procedia Computer Science*. 2019. V. 146. P. 11–19. <https://doi.org/10.1016/j.procs.2019.01.072>
 21. *Kosyakov D., Guskov A.* Reasons and consequences of changes in Russian research assessment policies // *Scientometrics*. 2022. V. 127. № 8. P. 4609–4630. <https://doi.org/10.1007/s11192-022-04469-5>
 22. *Guskov A.E., Kosyakov D.V., Selivanova I.V.* Boosting research productivity in top Russian universities: the circumstances of breakthrough // *Scientometrics*. 2018. V. 117. № 2. P. 1053–1080. <https://doi.org/10.1007/s11192-018-2890-8>
 23. *Liu W., Hu G., Gu M.* The probability of publishing in first-quartile journals // *Scientometrics*. 2016. V. 106. № 3. P. 1273–1276. <https://doi.org/10.1007/s11192-015-1821-1>
 24. *Kosyakov D.* Analysis of the abnormal growth in the number of Russian publications in conference proceedings in Scopus // *Scientific and Technical Information*. 2023. № 4. P. 13–24. <https://doi.org/10.36535/0548-0019-2023-04-3>
 25. *Phillimore A.J.* University research performance indicators in practice: The University Grants Committee's evaluation of British universities,

- 1985–86 // *Research Policy*. 1989. V. 18. № 5. P. 255–271. [https://doi.org/10.1016/0048-7333\(89\)90053-X](https://doi.org/10.1016/0048-7333(89)90053-X)
26. *Sizer J.* The impacts of financial reductions on British universities: 1981–84 // *Higher Education*. 1987. V. 16. № 5. P. 557–580. <https://doi.org/10.1007/BF00128422>
27. *Lee F.S., Pham X., Gu G.* The UK Research Assessment Exercise and the narrowing of UK economics // *Cambridge Journal of Economics*. 2013. V. 37. № 4. P. 693–717. <https://doi.org/10.1093/cje/bet031>
28. *Sizer J.* In Search of Excellence —Performance Assessment in the United Kingdom // *Higher Education Quarterly*. 1988. V. 42. № 2. P. 152–161. <https://doi.org/10.1111/j.1468-2273.1988.tb01811.x>
29. *Bekhradnia B. et al.* Research Evaluation: Past, present and future. Higher Education Policy Institute (HEPI), 2022.
30. *Curry S., Gadd E., Wilsdon J.* Harnessing the Metric Tide: indicators, infrastructures and priorities for UK responsible research assessment. Research on Research Institute, 2022.
31. *Smith S., Ward V., House A.* ‘Impact’ in the proposals for the UK’s Research Excellence Framework: Shifting the boundaries of academic autonomy // *Research Policy*. 2011. V. 40. № 10. P. 1369–1379. <https://doi.org/10.1016/j.respol.2011.05.026>
32. Building on Success and Learning from Experience. An Independent Review of the Research Excellence Framework. UK Government. 2016. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/541338/ind-16-9-ref-stern-review.pdf (дата обращения 30.01.2023).
33. Future Research Assessment Programme. <https://www.ukri.org/about-us/research-england/research-excellence/future-research-assessment-programme-frap/> (дата обращения 30.01.2023).
34. *Thelwall M. et al.* Can REF output quality scores be assigned by AI? Experimental evidence. Wolverhampton, UK: Statistical Cybermetrics and Research Evaluation Group, University of Wolverhampton, 2022. P. 141.
35. *Thelwall M. et al.* Predicting article quality scores with machine learning: The U.K. Research Excellence Framework // *Quantitative Science Studies*. 2023. V. 4. № 2. P. 547–573. https://doi.org/10.1162/qss_a_00258
36. *Singh Chawla D.* AI system not yet ready to help peer reviewers assess research quality // *Nature Index*. 2022. <https://doi.org/10.1038/d41586-022-04493-8>
37. Review of research assessment. Report by Sir Gareth Roberts to the UK funding bodies. UK Government. 2003. P. 44. <https://web.archive.org/web/20070720232304/http://www.rareview.ac.uk/reports/roberts.asp> (дата обращения 30.01.2023).
38. Research Excellence Framework 2028: Initial decisions and issues for further consultation. Research England, 2023. <https://www.ukri.org/publications/ref2028-initial-decisions-and-issues-for-further-consultation/> (дата обращения 30.01.2023).
39. FAQs on the REF 2021 Impact case study database. <https://ref.ac.uk/guidance-on-results/impact-case-study-database-faqs/> (дата обращения 30.01.2023).
40. *Pidd M., Broadbent J.* Business and Management Studies in the 2014 Research Excellence Framework: Business and Management Studies in 2014 REF // *Brit J Manage*. 2015. V. 26. № 4. P. 569–581. <https://doi.org/10.1111/1467-8551.12122>
41. *Sivertsen G.* The Norwegian Model in Norway // *Journal of Data and Information Science*. 2018. V. 3. № 4. P. 3–19. <https://doi.org/10.2478/jdis-2018-0017>
42. A Bibliometric Model for Performance-based Budgeting of Research Institutions. Norwegian Association of Higher Education Institutions, 2004.
43. Norwegian Register for Scientific Journals, Series and Publishers. <https://kanalregister.hkdir.no/publiseringskanaler/Forside> (дата обращения 05.07.2023).
44. Costas Boletsis. Level 1 & Level 2 publications — Norwegian Scientific Index. 2015. <https://boletsis.net/level-1-level-2-publications-norwegian-scientific-index/> (дата обращения 05.07.2023).
45. *Ahlgren P., Colliander C., Persson O.* Field normalized citation rates, field normalized journal impact and Norwegian weights for allocation of university research funds // *Scientometrics*. 2012. V. 92. № 3. P. 767–780. <https://doi.org/10.1007/s11192-012-0632-x>
46. Cristin. <https://www.cristin.no> (дата обращения 05.07.2023).
47. *Sivertsen G.* The Norwegian Model in Norway // *Journal of Data and Information Science*. 2018. V. 3. № 4. P. 3–19. <https://doi.org/10.2478/jdis-2018-0017>
48. *Schneider J.W.* An Outline of the Bibliometric Indicator Used for Performance-Based Funding of Research Institutions in Norway // *Eur Polit Sci*. 2009. V. 8. № 3. P. 364–378. <https://doi.org/10.1057/eps.2009.19>
49. *Ahlgren P., Colliander C., Persson O.* Field normalized citation rates, field normalized journal impact and Norwegian weights for allocation of university research funds // *Scientometrics*. 2012. V. 92. № 3. P. 767–780. <https://doi.org/10.1007/s11192-012-0632-x>
50. *Bloch C., Schneider J.W.* Performance-based funding models and researcher behavior: An analysis of the influence of the Norwegian Publication Indicator at the individual level // *Research Evaluation*. 2016. P. rvv047. <https://doi.org/10.1093/reseval/rvv047>
51. *Aagaard K.* How incentives trickle down: Local use of a national bibliometric indicator system // *Science and Public Policy*. 2015. V. 42. № 5. P. 725–737. <https://doi.org/10.1093/scipol/scu087>

52. *Skivenes M., Trygstad S.C.* When whistle-blowing works: The Norwegian case // *Human Relations*. 2010. V. 63. № 7. P. 1071–1097. <https://doi.org/10.1177/0018726709353954>
53. *Schneider J.W., Aagaard K., Bloch C.W.* What happens when national research funding is linked to differentiated publication counts? A comparison of the Australian and Norwegian publication-based funding models // *Research Evaluation*. 2016. V. 25. № 3. P. 244–256. <https://doi.org/10.1093/reseval/rvv036>
54. *Гуськов А.Е., Косяков Д.В.* Национальный фракционный счёт и оценка научной результативности организаций // *Научные и технические библиотеки*. 2020. № 9. P. 15–42. <https://doi.org/10.33186/1027-3689-2020-9-15-42>

TWO CIRCUIT ASSESSMENTS OF THE PERFORMANCE OF SCIENTIFIC ORGANIZATIONS IN RUSSIA: CURRENT STATE AND DEVELOPMENT PROSPECTS FROM THE POINT OF VIEW OF INTERNATIONAL EXPERIENCE

D. V. Kosyakov^{1,#}, I. V. Selivanova^{1,##}, A. E. Guskov^{1,###}

¹*Russian Research Institute of Economics, Politics and Law in the Scientific and Technical Sphere, Moscow, Russia*

[#]*E-mail: kosyakov@sciencepulse.ru*

^{##}*E-mail: i-seli@yandex.ru*

^{###}*E-mail: guskov.andrey@gmail.com*

In Russia, in fact, two contours for assessing the scientific performance of organizations have developed: the expert one, which underlies the categorization of scientific organizations, and the quantitative one, which, according to the methodology of the comprehensive calculation of publication productivity, makes it possible to determine and monitor the implementation of the state assignment for fundamental scientific research. The article describes the history of the formation of these circuits, their advantages and disadvantages, as well as the prerequisites for further development. The British REF (Research Excellence Framework) examination system and the Norwegian quantitative assessment model are considered as model examples for improving both circuits. REF features include a differentiated assessment of the results, impact and environment of scientific units, which is organized by expert panels and groups. The features of the Norwegian model are an expert approach to the formation of a national white list of scientific journals, conferences and publishers, a methodology for calculating the publication indicator and an open database of publications of Norwegian scientists, from which this indicator is determined. The authors believe that these two examples can be considered as parametric assessment models, the adaptation of which taking into account national characteristics will make it possible to update and improve both contours of the assessment of scientific organizations in Russia. The importance of regular retrospective reflection on the experience, procedures and results of assessing organizations is emphasized, which is necessary for the systematic development of this system at the next stages.

Keywords: scientific productivity, performance assessment of scientific organizations, examination, integrated calculation of publication performance (CPPR), categorization of organizations, scientometrics, Norwegian model, REF.