

Методы оценки научных результатов у нас и за рубежом (сравнительный анализ)

А. Н. Паршин

Математический институт им. В. А. Стеклова РАН

4-я конференция научных работников

Физический институт им. П. Н. Лебедева РАН

27 марта 2018

Официальные документы РФ (2013) и Великобритании (2012)

Перечень поручений по итогам заседания Совета по науке и образованию при Президенте РФ 30 апреля 2013 г. в Ленинградской области:

2г. Возможность использования показателей результативности деятельности научных организаций **экономически развитых стран** для оценки результативности деятельности российских научных организаций.

Срок – октябрь 2013 г. Ответственный: Медведев Д. А.
<http://kremlin.ru/d/18146>

Постановление Правительства РФ от 1 ноября 2013 г. № 979

4. Межведомственная комиссия ежегодно формирует минимальные значения показателей результативности для референтных групп на основании материалов Министерства образования и науки Российской Федерации, подготавливаемых исходя из: $\langle \dots \rangle$

в. показателей результативности деятельности научных организаций экономически развитых стран, осуществляющих сходные типы исследований.

К этому времени уже больше года окончательный свод правил оценки в Великобритании за 2008-2012 гг. (Research Excellence Framework, 107 страниц) был вывешен на сайте www.ref.ac.uk (Январь 2012 г.).

Система оценки деятельности научных организаций в Великобритании за 2008-2013 гг.

Название: Research Excellence Framework (REF)

Расположение: <http://www.ref.ac.uk>

Органы оценки: 35 комиссий, объединенных в 4 центральные комиссии

- ▶ Март 2010 Первое сообщение
- ▶ Июль 2011 Публикация REF 01.2011
- ▶ Конец июля 2011 Публикация наброска REF01.2012
- ▶ 5 октября 2011 конец сбора откликов
- ▶ Январь 2012 Публикация REF01.2012 (окончательный свод правил оценки)
- ▶ Октябрь 2012 Приглашение делать представления
- ▶ 29 ноября 2013 Окончание подачи представлений
- ▶ Декабрь 2014 Публикация результатов оценки

Использование импакт-фактора

53. Комиссии не будут использовать импакт-факторы журналов, рейтинги или подразумеваемую репутацию издателя при вынесении суждения об уровне представленных результатов.

Это относится ко **всем**, без исключения, научным дисциплинам.

Импакт фактор

We urge Higher Education Funding Council of England to remind RAE panels that they are obliged to assess the **quality** of the content of individual articles, not the **reputation** of the journal in which they are published.

Комитет по науке и технологиям Палаты Общин
Британского Парламента, 10-ый отчет, раздел 9, 20 июля
2004 г.

[https://publications.parliament.uk/pa/cm200304/cmselect/
cmsctech/399/39912.htm](https://publications.parliament.uk/pa/cm200304/cmselect/cmsctech/399/39912.htm)

52. Ввиду ограниченного значения цитируемости для процедуры оценки, финансирующие структуры (the funding bodies) не предлагают и не рекомендуют ВУЗам основываться на данных цитирования при отборе персоналий или результатов для включения в представления (submissions).

Центральная комиссия А

1. Клиническая медицина
2. Санитария и гигиена, Организация здравоохранения, Первая помощь
3. Вспомогательные медицинские профессии, Стоматология, Уход за больными, Фармацевтика
4. Психология, Психиатрия, Неврология
5. Биологические науки
6. Сельское хозяйство, Ветеринария, Диетология

Центральная комиссия А

52. Подкомиссии центральной комиссии А используют показатели цитируемости, при их наличии и в случае уместности их использования, как показатель научной значимости, для выработки оценки уровня представленных результатов.

54а. В случае их наличия, при необходимости, данные цитирования будут рассматриваться как позитивный показатель научной значимости результатов исследования. Они будут лишь одним из составляющих элементов для выработки рецензентами оценок уровня представленных результатов и не будут использоваться **как основной параметр для оценки.**

Центральная комиссия В

- 7. Науки о Земле и окружающей среде
- 8. Химия
- 9. Физика
- 10. Математические науки
- 11. Вычислительная техника и информатика
- 12. Авиастроение, Машиностроение, Химические и промышленные технологии
- 13. Электротехника и электроника, Metallургия, материаловедение
- 14. Строительство
- 15. Общественные дисциплины

Центральная комиссия В: данные, представляемые в комиссии

Оцениваются группы исследователей (департаменты, лаборатории, etc). Эти группы представляют следующие данные части ее сотрудников, отобранных самой группой:

1. outputs (по 4 отобранных работы за 2008-2013 гг.)
2. enviroment (приглашения на конференции, премии, почетные звания, etc)
3. impact (приложения и общественное признание вне академической среды)

Центральная комиссия В

59. Подкомиссии 7, 8, 9 и 11 отмечают, что данные цитирования широко используются в соответствующих дисциплинах и полагают, что эта практика является общепринятой. Эти подкомиссии будут использовать доступные им данные цитирования, в качестве одного из показателей научной значимости при выработке оценок уровня исследований.

62b. **Отсутствие** данных цитирования не будет автоматически интерпретироваться как отсутствие научной значимости.

63. Подкомиссии 10, 12, 13, 14 и 15 считают, что сводка данных цитирования в их дисциплинах неравномерна по качеству и нередко содержит пробелы и потому не позволяет получить сбалансированную и надежную дополнительную информацию для выработки оценки уровня исследований.

Центральная комиссия С

16. Архитектура, Градостроительство, Ландшафтный дизайн
17. География, Охрана окружающей среды, Археология
18. Экономика и Эконометрика
19. Исследования по бизнесу и менеджменту
20. Право
21. Политика и международная политика
22. Социальная защита и социальная политика
23. Социология
24. Антропология, Социология развития
25. Образование
26. Физкультура и спорт, Досуг, Туризм

Центральная комиссия С

64. Подкомиссии 16, 17, 19, 20, 21, 22, 23, 24, 25 и 26 не получат в свое распоряжение и не будут использовать данные цитирования или другие типы библиометрического анализа, включая импакт-факторы журналов.

65. Подкомиссии 18 (Экономика и эконометрика) будут предоставлены доступные данные цитирования и она будет использовать их при необходимости.

67b. **Отсутствие** данных цитирования конкретного исследования не будет сказываться на его оценке.

Центральная комиссия D

- 27. Региональные исследования
- 28. Современные языки и литература
- 29. Английский язык и литература
- 30. История
- 31. Классическая филология
- 32. Философия
- 33. Теология, Религиоведение
- 34. Искусство и Дизайн: история, практика и теория
- 35. Музыка, Драматическое искусство, Танец и другие виды театрального искусства
- 36. Коммуникационные исследования, Культурология, Медиаведение, Библиотечное дело, Информационный менеджмент

Центральная комиссия D

74. Подкомиссиям центральной комиссии D не будут предоставляться и они не будут использовать данные цитирования или какую-либо библиометрическую информацию для выработки оценок.

Результаты REF (2008-2013)

- ▶ 154 университета
- ▶ 52 тыс. сотрудников (academic staff)
- ▶ Около 200 тыс. исследовательских отчетов (research outputs)
- ▶ 30% мировые лидеры (world-leading)
- ▶ 46% мирового уровня (internationally excellent)
- ▶ 20% международно признанные (recognized internationally)
- ▶ 3% признанные в стране (recognized nationally)

REF 2021 (за 2013 -2020 гг.)

Обсуждение правил новой оценки в Великобритании в 2021 г.

<http://www.ref.ac.uk/publications/2017/initialdecisionsontheresearchexcellenceframework2021.html>

Autumn 2017 Invite nominations for panel members

Spring 2018 Panels meet to develop criteria

Winter 2018-19 Publish final guidance and criteria

2020 Submission phase

2021 Assessment phase

Assessment metrics

14. Quantitative data may be used to inform the assessment of outputs, where panels consider this **appropriate** for the discipline.

Опрос директоров математических исследовательских институтов

Использование библиометрических данных
(citation index, h-index, impact-factor for journals)
в оценке научной деятельности в таких ситуациях:

- 1) Прием на работу постоянных членов
исследовательских институтов
- 2) Решение о временных позициях
- 3) Ежегодные отчеты институтов
- 4) Обследование институтов внешними
комиссиями

Ответ НЕТ на все четыре вопроса

1. Alfred Renyi Institute of Mathematics (Budapest, Hungaria)
2. Institut des Hautes Études Scientifiques (Bures-sur-Yvette, France)
3. Institut Mittag-Leffler (Stockholm, Sweden)
4. Isaac Newton Institute for Mathematical Sciences (Cambridge, UK)
5. Istituto Nazionale di Alta Matematica Francesco Severi (Roma, Italy)
6. Mathematisches Forschungsinstitut Oberwolfach (Oberwolfach, Germany)
7. Institut Henri Poincaré (Paris, France)

Ответ НЕТ на все четыре вопроса

8. Max-Planck-Institut für Mathematik (Bonn, Germany)
9. Max-Planck-Institut für Mathematik in den Naturwissenschaften (Leipzig, Germany)
10. The Abdus Salam International Centre for Theoretical Physics (Trieste, Italy)
11. Mathematical Science Research Insritute (Berkeley, USA)
12. Tata Institute of Fundamental Research (Bombay, India)
13. Weizmann Institute of Science (Rehovot, Israel)
14. Institute for Advanced Study (Princeton, USA)

Другие ответы

- 15. Research Institute of Mathematical Sciences (Kyoto, Japan) 1, 2 НЕТ 3, 4 ДА по требованию правительства
- 16. Weierstrass Institute for Applied Analysis and Stochastics (Berlin, Germany) 1, 2, 3 НЕТ 4 ДА неформально
- 17. Centre Interfacultaire Bernoulli (Lausanne, Switzerland) НЕТ на уровне центра, возможно ДА на более высоком уровне
- 18. European Institute for Statistics, Probability and Stochastic Operations Research (Eindhoven, The Netherlands) ДА как вспомогательные данные

Из письма проф. В. Бальмана, директора Института Математики Общества Макса Планка в Бонне

Our reports to external evaluation committees are not based on bibliometrical data, but discuss our research achievements as such. Of course we include lists of publications in such reports, but citation index, h-index, impact factors, and the like do not play a role.

Bibliometric data are a dangerous weapon (all the more so in the hands of non-experts) and are completely unreliable measures of the quality of research. This has been made clear in a number of investigations, for example by the IMU, which seem conclusive to us.

Дополнительные материалы I

В 2008 г. был опубликован доклад Р. Адлер, Д. Юинг, П. Тейлор, Статистики цитирования, подготовленный по просьбе Международного математического союза (IMU), объединяющего национальные математические сообщества 77 стран и организующего каждые 4 года международные математические конгрессы, на которых присуждается наиболее известная математическая премия - медаль Филдса. Этот доклад посвящен использованию и злоупотреблениям данными цитирований при оценке научных исследований. См. русский перевод в сб. Игра в цифры или как теперь оценивают труд ученого, М., МЦНМО, 2011, 6-38
(<https://www.mccme.ru/free-books/bibliometric.pdf>).

Дополнительные материалы II

Математические подразделения CNRS (Centre National de la Recherche Scientifique, объединяющий во Франции более тысячи исследовательских лабораторий, около 50-ти из них математические) и DFG (Deutsche Forschungsgemeinschaft, грантовый фонд, один из крупнейших в Германии) не используют в своей работе библиометрические данные (сообщения автору во время личных контактов в 2016 г.).

Закон Гудхарта

Чарльз Гудхарт (р. 1936) , известный экономист, профессор Лондонской школы экономики и политических наук, работал в Банке Англии:

«Когда какой-нибудь экономический или социальный показатель становится целью для проведения политики в экономике или социальных областях, он перестает быть показателем, который заслуживает доверие или адекватен тому, что мы хотим получить» (1975).

Закон Кемпбелла

Дональд Кемпбелл (1916-1996), американский психолог и социолог: «Когда результаты тестов становятся целью образовательного процесса, они с одной стороны теряют свою ценность в качестве индикаторов уровня этого процесса, а с другой стороны, искажают образовательный процесс в нежелательных направлениях» (1976)

<http://www.wmich.edu/evalctr/pubs/ops/ops08.pdf>

"Чем больше какой-либо количественный социальный индикатор используется для принятия социальных решений, тем больше он будет подвержен давлению коррупции и тем с большей вероятностью он будет деформировать социальные процессы, для мониторинга которых он был предназначен, и увеличивать их коррупционную составляющую". Earl R. Babbie, The Basics of Social Research, Wadsworth, 2013. p. 388

Примеры:

- ▶ Примером, когда эти закономерности проявились в полной мере, является многолетняя история внедрения в России Единого государственного экзамена (ЕГЭ)
- ▶ Более свежий пример - принуждение к публикационной активности независимо от потребностей развития самой науки. Имеется публикация в «Science» (China's Publication Bazaar, Science 29 Nov 2013: Vol. 342, Issue 6162, pp. 1035-1039), где подробно обсуждается публикационный бум в Китае. Нам часто говорят, как мы отстаем от Китая. Здесь можно понять, что нам не так страшно отставать от Китая в этом вопросе, пускай он убегает вперед если хочет.

НАШИ ДНИ: Формула-1 (январь 2018)

Формула ФАНО для оценки ОПУБЛИКОВАННОЙ в 2017 г. статьи научной организации (оценка качества!!, как написано в информационной системе ФАНО):

$$(\text{оценка статьи}) = (\text{вес базы данных}) \cdot \frac{1}{\text{число авторов}} \cdot$$

$$\left(\sum_{\text{авторы}} \frac{1}{\text{число аффилиаций автора}} \right)$$

Таблица весов баз данных

WoS CC: Q1	Q2	Q3	Q4	noQ
27	9	3	1	1
Scopus not WoS	WoS RCI	журналы БАК	Other	
1	0,75	0,5	0	

НАШИ ДНИ: Формула-2 (ФАНО, март 2018)

При расчете **дополнительного** числа ПЛАНИРУЕМЫХ в 2018 г. публикаций используются следующие правила:

- ▶ 1 статья в журналах 1 или 2 квартиля (Q1 или Q2) базы Web of Science =
- ▶ 2 статьям в журналах 3 или 4 квартиля (Q3 или Q4) базы Web of Science
- ▶ 3.3 статьям в иных журналах
- ▶ 1 монография = 15 статьям в иных журналах

За пять лет ЧИНОВНИКИ так и не потрудились выяснить как происходит оценка научных результатов в экономически развитых странах!