

Цитируемость, показатели цитируемости и качество исследований: обзор базовых понятий и теорий*

Дар В. АКШЕС
(Dag W. AKSNES),

Лив ЛАНГФЕЛТ
(Liv LANGFELDT)

Северный институт анализа в области инноваций, исследований и образования, г. Осло, Норвегия

Пол ВУТЕРС
(Paul WOUTERS)

Центр научных и технологических исследований, Лейденский университет, г. Лейден, Нидерланды

Цитируемость все больше используется в качестве показателя производительности в научной политике и внутри исследовательской системы. Как правило, предполагается, что цитируемость свидетельствует о влиянии исследования или его качества. Что подтверждает эти предположения и как цитируемость соотносится с качеством исследований? Эти и подобные вопросы изучаются на протяжении десятилетий наукометрических исследований. Предоставляется обзор некоторых основных актуальных вопросов, включающих теории цитируемости, трактовку и обоснованность использования цитируемости как измерения результативности. Качество исследований является многоаспектным понятием, в котором достоверность/правильность, оригинальность, научная ценность, а также общественная ценность общепринято воспринимаются ключевыми характеристиками. Изучается то, как цитируемость может затрагивать подобные разнообразные измерения качества исследований. Утверждается, что цитируемость отражает аспекты, касающиеся научного влияния и релевантности, но с определенными ограничениями. С другой стороны, нет ни одного свидетельства, подтверждающего, что цитируемость отражает другие ключевые величины качества исследований. Следовательно, рост использования показателей цитируемости в оценке исследований и финансирования может снижать внимание к этим иным величинам качества исследований, таким как надежность/достоверность, оригинальность и общественная ценность.

ВВЕДЕНИЕ

В последнее время библиометрические показатели все больше применяются в контексте оценки исследований, а также научной политики в целом. Примерами служат использование показателей цитируемости в оценке научной эффективности исследовательских групп, отделений и организаций [1]; оценке предложений в исследованиях [2]; распределении научного финансирования [3] и

приеме на работу научного персонала [4]. Измерения цитируемости являются также базовыми показателями в некоторых ранжированиях университетов, таких как Лейденское ранжирование и ранжирование академических университетов по всему миру [5].

Таким образом, показатели или метрики применяются к разнообразию целей и пронизывают многие аспекты исследовательской системы. По традиции рецензирование было «золотым стандартом» в оценке исследований. В качестве альтернативы метрики все больше внедряются сами по себе или вместе с рецензированием. Например, данные цитируемости использовались в Великобритании для информирования об их оценках рецензирования некоторыми панельными сессиями в Програм-

*Перевод Aksnes D.W., Langfeldt L., Wouters P. Citations, citation indicators, and research quality: An overview of basic concepts and theories//Sage Open. — 2019. — January-March. — P. 1-17. — <https://journals.sagepub.com/doi/pdf/10.1177/2158244019829575>

ме оценки научных достижений 2014 г. (Research Excellence Framework – REF) [6]. Возникает вопрос о достоверности и надежности цитируемости как показателя производительности. В каких контекстах и для каких целей она подходит? Эти вопросы дискутируются на протяжении последних десятилетий.

В наиболее радикальной версии утверждается, что оценка исследований, основанная на цитируемости и других библиометрических показателях, превосходит традиционный метод рецензирования. Например, Абрамо и Д'Анджело заявили: «эмпирические данные показывают, что для естественных и формальных наук библиометрическая методология намного предпочтительнее рецензирования. . . Компромиссные методы, такие как основанное на имеющейся информации рецензирование, где рецензент также может полагаться на библиометрические показатели при формировании суждения, по мнению авторов, не предлагают преимуществ, которые оправдывают дополнительные затраты: показатели не будут поддерживать составление оценок людьми, по максимуму допуская подтверждение или опровержение» [7, р. 512].

Похожие точки зрения были выдвинуты Реджибо и Рокетт* [8].

Тем не менее, применение библиометрических показателей к оценке научной эффективности всегда было противоречивым. Длительное время использование факторов журнального влияния (импакт-факторов журналов) в контекстах оценки исследований едва ли критиковалось [9-11]. Более того, внедрение показателей цитируемости также критиковалось в целом относительно их надежности как измерителей эффективности и их вероятного отрицательного воздействия на исследовательскую систему [12-14]. Например, Сеглен [15] изучал проблемы, затрагивающие анализ цитирования, и пришел к выводам, что «... темп цитируемости определяется таким множеством технических факторов, что сомнительно, оказывает ли чистое научное качество вообще какой-либо заметный эффект ...» [15, р. 226].

Говоря в широком смысле, хотя на протяжении 1970-1980 – х гг. шли интенсивные дискуссии о том, что цитируемость реально «измеряет» и как она соотносится с научным качеством (см., например, [16]), этот вопрос по -видимому получает меньше внимания в последние десятилетия. Сегодня чаще принимается как должное, что цитируемость неким образом измеряет научное влияние, одну из составляющих научного качества. Больше

внимания уделяется методологическим вопросам, таким как соответствующие методы нормализации абсолютных подсчетов цитируемости [17], помимо усовершенствования и изучения новых показателей на основе цитируемости, таких как h-индекс [18, 19]. Хотя последняя разработка внесла важный вклад в прогресс области, ограничения цитируемости, обсуждавшиеся в 1970-1980-х гг., нигде не исчезли. В научной статье библиографические ссылки служат различным целям. Авторы не просто включают библиографические ссылки из-за их чисто научного качества. Отбор ссылок определяется разными факторами, одним из которых является их релевантность изучаемой теме исследования [20]. Эти ограничения нельзя преодолеть за счет создания технически более сложных или надежных показателей.

В противовес этому базису данная статья предоставляет обзор основных вопросов, связанных с цитируемостью, показателями цитируемости и их интерпретации и надежности как измерений результативности*. Вопросу, как цитируемость может затрагивать или отражать различные аспекты понятия качества исследований, уделяется особое внимание. Эта статья написана как введение в обзор для более широкой аудитории, интересующейся этими темами. Следовательно, охват тем и литературы избирателен и не касается всех деталей. Кроме того, литература по взаимодействию между практиками цитирования и процессами оценки упоминается лишь мимоходом, и мы не обсуждаем конструктивистские и семиотические теории качества и цитируемости [22].

Данная статья структурирована следующим образом: в качестве введения описываются некоторые основные вопросы, касающиеся создания указателей ссылок. **Показатели цитируемости** отчасти сфокусированы на процессе цитирования и на том, какова роль библиографических ссылок в научной статье. Большая часть предыдущих исследований сравнивала показатели цитируемости с результатами рецензирования, также этот вопрос частично изучается в разделе **Понимание цитируемости**. Некоторые факторы, влияющие на надежность показателей цитируемости, далее описываются в **Оценке надежности**. В разделе **Цитируемость как показатель – Другие вопросы надежности** рассматривается вопрос, относящийся к цитируемости и понятию качества исследований. Качество исследований является многоаспектным понятием. Следовательно, обсуждается то, как цитируемость

* Например, они заявили: «в целом есть основания поддерживать рецензирование на основе библиометрии за рамками рассмотрения затрат. Даже простые метрики могут хорошо действовать в установлении качества по отдельным областям, в то же время обеспечивая затраты на эффективное и прозрачное рецензирование. Рецензирование по-видимому не считается гарантом качества ...» [8].

* Статья частично основана на обзорах литературы, впервые проведенных для одной из докторских диссертаций автора [21], которые были объединены и дополнены более поздними материалами. Некоторые отрывки из этой диссертации были приняты и включены в настоящую статью.

может относиться к каждому различному аспекту понятия качества. Тогда как первые четыре раздела представляют краткий обзор анализируемых вопросов, последний раздел является более аналитическим и дискурсивным. Причина этого состоит в том, что некоторые более ранние исследования изучали эту тему системно.

ПОКАЗАТЕЛИ ЦИТИРУЕМОСТИ

Развитие библиометрии как области тесно связано с созданием Указателя цитированной литературе по точным, естественным и прикладным наукам (Science Citation Index– SCI) Юджином Гарфилдом в 1961 г. [21]. Первоначально эта библиографическая база данных (БД) была преимущественно создана для целей информационного поиска, чтобы помогать исследователям в установлении релевантных статей в огромных архивах научной литературы [23]. В качестве дополнительного свойства она способна качественно анализировать научную литературу. Начиная с 1960-х гг., SCI и подобные ей БД, сейчас включенные в онлайн-продукт Web of Science, применяются в массе исследований, охватывающих разные области. Эта возможность анализа цитирования служит важным объяснением его популярности [21]. В этой БД все библиографические ссылки в индексируемых статьях регистрируются. На основе этого каждая статья может быть приписана путем подсчета цитируемости, показывающего общее число раз, когда она цитировалась более ранними статьями, зарегистрированными в БД. Подсчеты цитируемости и ее показатели затем можно вычислить для агрегированных «уровней публикации», например, представляющих научные единицы, отделения или научные области. В начале 2000-х гг. были введены БД-конкуренты, которые также включают статистику анализа цитирования, преимущественно это БД Scopus (запущенная в 2004 г.) и Google Scholar (стартовавшая в 2004 г.). Охват научной и академической литературы варьируется в зависимости от этих БД, а результаты исследований цитируемости таким образом обусловлены отдельными характеристиками этих БД и их охвата.

В последние несколько лет разрабатывается огромное число различных показателей цитируемости и идут напряженные дискуссии о соответствующих методах их подсчета, процедурах нормализации, охвате БД и качестве данных (в целях обзора см. [24, 1, 25, 19]). В число наиболее часто используемых показателей цитируемости входит показатель влияния цитируемости, нормализованный по области, число/доля высокоцитируемых статей, а также h-индекс. Первым показателем является выражение среднего числа ссылок в публикациях, нормализованное по области, году издания, а также типу документа (например, регулярная статья или обзор). То есть значение 2 говорит нам о том,

что публикации цитировались в два раза выше их среднего по области и году издания, т.е. среднего в мире [17]. Показатели, связанные с высокоцитируемыми статьями, как правило, основаны на процентиле, например, числе и доле публикаций, принадлежащих топ-1% или топ-10% наиболее часто цитируемых по их областям (отрегулированным по году издания; [26]). Другим показателем на основе ссылок является импакт-фактор журнала (JIF), который, несмотря на проблемы, недостатки, а также рекомендацию по неиспользованию его в контексте оценки исследований, остается очень популярным, если не самым популярным библиометрическим показателем [27, 9].

Существуют основные вариации в среднем темпе цитирования в различных предметных областях. Например, во многих гуманитарных дисциплинах средняя статья получает менее одной ссылки в течение 10-летнего периода, по сравнению с более 40 ссылками в некоторых биомедицинских областях (данные из Web of Science за 2005-2015 гг.). По мнению Маркса и Борнманна [28], основная причина такой разницы лежит в охвате БД. Только небольшая часть научной литературы в гуманитарных науках представлена в Web of Science, а большая часть библиографических ссылок и цитируемости не будут охвачены БД. Соответственно, средний темп цитирования внутри гуманитарных наук является более высоким при использовании других БД, охватывающих литературу лучше, таких как Google Scholar [29]. Помимо этого, среднее число и возраст библиографических ссылок, а также соотношение новых публикаций в области и общего числа публикаций играют роль, когда дело касается области различий в темпе цитирования [21].

Поскольку существуют основные областные и временные различия в том, как много ссылок в среднем получает статья, на этапе становления наукометрии предполагалось, что абсолютные подсчеты цитируемости нужно нормализовать [30, 31]*. С тех пор это было стандартом регулирования по области, году издания и типу публикации при вычислении показателей цитируемости. Самым широко известным показателем является показатель влияния цитируемости, нормализованный по области, ранее известный как crown indicator (показатель – верхушка) [33], в котором учитываются все упомянутые различия. С помощью этого показателя стремятся скорректировать эффект переменных, которые считаются мешающими факторами в анализе цитирования (т. е. связаны с дисбалансом в возможностях цитирования). В последнее время много внимания уделяется

* В качестве примера, Гарфилд [32] ранее выделял, что «вместо прямого сравнения подсчета ссылок, скажем, у математика против биохимика, обоим надо ранжировать наравне с коллегами, а само сравнение проводить между сравнениями» [р. 367].

методам нормализации, вопросу о том, как усреднять научные области, используемые в нормализации, и следует ли проводить нормализации на уровне отдельной статьи или на агрегированных уровнях статей (среднее из соотношений в отличие от соотношения среднего [34, 35]). Нет общего согласия относительно того, чем является подходящий метод [36], но эмпирические исследования показали, что два различных метода нормализации, среднее из соотношений и соотношение среднего, не дают очень разных результатов, особенно на уровне стран и учреждений [37, 38].

Распределения цитируемости весьма искажены. Это искажение уже было установлено историком науки Дерекком де Солла Прайсом [39]. Наибольшая часть всех научных статей никогда не цитируется или цитируется только несколько раз в последующей литературе [21]. Наоборот, некоторые статьи имеют чрезмерно большое число ссылок, достигая сотен и даже тысяч. За последние 20 лет растет интерес к использованию верхушки, высокоцитируемых статей, в качестве показателей производительности. Ожидание заключается в том, что эти статьи представляют экстраординарно хорошие работы и поэтому могут быть использованы для установления научного превосходства, острой проблемы в научной политике [40, 41]. Имеются разные типы таких показателей; общепризнанным показателем является число или доля статей, принадлежащих топ-1% или 10% самым часто цитируемым статьям (в одной и той же области и за одно и то же время).

H-индекс был введен в 2005 г. [42] и быстро стал очень популярным библиометрическим измерением. Этот показатель учитывает как число созданных статей, так и влияние цитируемости данных статей. Согласно определению h-индекса, исследователь с h-индексом 15 имеет, по крайней мере, 15 публикаций с по меньшей мере 15 ссылками. Первоначально индекс был разработан для анализа индивидуумов, но также применялся на других уровнях, таких как исследовательские группы, отделы и учреждения. Несмотря на свою популярность, у этого показателя есть некоторые проблемы. Важнее всего то, что он не является нормализованным по области и не совершается коррекция относительно протяженности карьеры, что подразумевает – данный показатель не благоприятствует молодым ученым (см., например, [43]).

Что касается измерения частоты цитирования, то имеет значение временное измерение или временное окно. Как правило, статьи, опубликованные недавно, едва ли уже цитируются, а число ссылок со временем растет, поскольку более поздние статьи имеют больше времени для получения ссылок. В анализе цитирования используются разные временные окна в зависимости от цели и области анализа. Наиболее часто используется интервал

цитирования в 3-5 лет [44]. Это прагматический компромисс между кратковременным и долгосрочным интервалом цитирования [45]. Тем не менее, будет меняться степень, с которой краткосрочные темпы цитирования могут считаться предсказателем долгосрочных темпов [46], а использование краткосрочных окон (например, 2 или 3 года) предполагает, что вклады в фронт текущих исследований соотносятся лучше, чем долгосрочное влияние [45]. Более длинный интервал цитирования, как правило, считается более надежным, чем краткосрочный интервал. Например, Левитт и Телволл [47] утверждали, что краткие интервалы цитирования несут в себе проблему того, что статьи, опубликованные в течение года в более ранние сроки, обладают важным преимуществом (т.е. являются в среднем более высокоцитируемыми) по сравнению с публикациями, появляющимися позже в течение года. Наоборот, диспропорционально длинный временной период приводит к меньшей пользе от результатов в целях оценки. Причиной этого служит то, что в таком случае имеются данные цитирования только для статей, опубликованных много лет назад [21]. Например, применение интервала цитирования в 3 года предполагает, что статьям требуется по крайней мере 3 года, чтобы быть включенными в анализ. Таким образом, вклады из недавних лет, периода, который обычно представляет особый интерес в программе оценки исследований, не могут быть оценены.

ПОНИМАНИЕ ЦИТИРУЕМОСТИ

Вопрос о том, что ссылки «измеряют», длительное время остается важным вопросом в библиометрии. Два первопроходца в изучении ссылок, братья Коул, часто относились к ссылкам как измерению качества, хотя было дано чуть более осторожное определение во Введении их книги по социальной стратификации в науке: «Число ссылок берется, чтобы представить относительную научную важность или «качество» статей» [48, р. 21]. Даже сегодня показатели цитируемости иногда представлены как измерения научного качества (см., например, [7, 49]).

Так как цитируемость получается из библиографий в литературе, общим предположением было то, что использование ссылок как показателей научной производительности должно быть оправданным или опирающимся на поведение авторов при составлении библиографии. Уже в 1981 г. Смит жаловался: «недостаточно известно о «поведении во время цитирования» авторов – почему автор делает ссылки, почему он делает свои особые ссылки и как они отражаются или не отражаются на его реальном исследовании и использовании литературы. Когда больше станет известно о действительных нормах и вовлеченных практиках, мы будем находиться в лучшей позиции, чтобы знать,

имеет ли смысл (и каким образом) использовать анализ цитирования в различных областях применения» [50, р. 99]

Было проведено действительно много исследований по поведению при цитировании. Сошлемся на Борнманна и Даниэля [28] и Николаисена [51] для глубокого обзора данной литературы. Более недавние вклады, например, содержат работы авторов [52-54]. Грубо говоря, можно выявить две противоречивые перспективы: одна, в соответствии с которой выделяется интеллектуальная функция библиографий, а другая анализирует процесс цитирования как фундаментальный социальный процесс. Как правило, второй подход скорее сфокусирован на «внешних» и социальных факторах, чем на содержании, и в большинстве случаев ассоциируется с попытками критиковать использование ссылок как измерение производительности [21].

Роль библиографий в научной статье

Предпринятые исследования обнаружили, что роль библиографий как в цитирующем тексте, так и относительно цитируемого текста, является сложной. Например, уже в 1964 г. Гарфилд предложил 15 различных объяснений того, почему авторы цитируют другие публикации [55]. В их числе было: обеспечение дополнительного чтения, идентификация методологии, дань уважения первопроходцам, установление оригинальной публикации или иной работы, описывающей эпоним понятия, выявление оригинальной публикации, в которой идея или понятие обсуждается, выражение доверия связанным работам, критика более ранней работы, коррекция работы, поддержка утверждений, реклама предстоящей работы, обеспечение руководства для плохо распространяющейся работы, аутентификация данных и классов факта – физических констант и т.п., опровержение работ других авторов и обсуждение приоритетных утверждений.

Следовательно, текстовые функции ссылок варьируются значительно. В научной статье некоторые из библиографий будут представлять работы, важные или существенные прецеденты настоящей работы; другие могут представлять более общую дополнительную литературу [21]. Например, в обзоре литературы, опубликованном по теме в 1965-1980 гг., Смолл [56] выделил пять различий: цитируемая работа может быть (а) опровергнута, (б) только отмечена, (в) проведен ее обзор, (г) внедрена в производство или (д) поддержана цитирующей работой. Эти категории были соответственно охарактеризованы как (а) отрицательные, (б) поверхностные (формальные), (в) сравнимые, (г) используемые, (д) обоснованные. Это предполагает, что различные функции библиографий, имеющиеся в тексте, являются более сложными, чем простое обеспечение документации и поддержка отдельных утверждений.

Эти и более поздние исследования обнаружили, что библиографии имеют множество функций в научной статье. Что касается взаимосвязи между частотой цитирования и научным качеством, то образцы на соседних уровнях релевантны для рассмотрения, а не только индивидуальные статьи. Чтобы объяснить, как некоторые статьи становятся высокоцитируемыми, нужно сфокусироваться на том, как объединяются библиографии на микроуровне [21]. Как правило, научная статья структурирована как прогрессия от общего к частному [57]. Это предполагает, что Введение статьи обычно содержит библиографии на работы более общего или значимого характера в области. Аккумулятивный эффект многих статей, ссылающихся на одну и ту же общую работу, состоит в том, что подобные вклады получают очень большое число ссылок. Библиографии на высокоцитируемые публикации гораздо чаще представлены во Введении, чем в других частях публикаций [58].

Соответственно, большая часть научных публикаций содержит раздел Методы, в котором задокументированы примененные в исследовании методы. А значит, авторы цитируют, как правило, ключевые статьи, описывающие эти методы. Поэтому статьи, описывающие вообще используемые методы, могут получить очень большое число ссылок. Здесь первым примером является статья 1951 г. об измерении протена [59], которая остается самой цитируемой когда бы то ни было статьей. Эта статья сегодня является цитируемой уже свыше 305 тыс. раз в БД Web of Science [60].

Хотя важное понимание роли библиографий в научной статье достигается, аккумуляция знания одновременно затрудняется тем фактом, что различные системы классификации применяются в предыдущих исследованиях [61]. Более того, исследования часто основаны скорее на небольших массивах статей из отобранных научных областей, а результаты могут не иметь общей надежности. По мнению Борнманна и Даниэля [20], многие исследования имеют методологическую слабость и представляют результаты с меньшей достоверностью.

Поведение при цитировании

Роберта К. Мертон часто считают специалистом, предоставившим оригинальную, теоретическую основу для связи подсчетов ссылок с использованием и качеством научных вкладов [21]. С точки зрения Мертона, нормы в науке обязывают исследователей цитировать работу, на которой они строят свое исследование, и таким образом признавать или переадресовать доверие другим ученым [62]. Такие нормы поддерживаются через неформальное взаимодействие в научных сообществах и через рецензирование представленных рукописей.

Если авторы цитируют работы, которые считают полезными, часто цитируемые публикации могут, предположительно, быть более полезны чем статьи, которые едва ли кто-то процитирует. Таким образом, число ссылок может рассматриваться как мера пользы, влияния или воздействия публикации. Точно такое же объяснение может быть применено для агрегированных уровней публикаций. Чем больше ссылок привлекают публикации, например, отдела, тем больше должно быть их влияние. Существуют также присущие определенной дисциплине нормы или даже коды, отличающиеся внутри области по журналу, например, касающиеся того, как и когда цитировать, а также как много ссылок должна содержать статья [63].

Эмпирические исследования показывают, что подход Мертона к нормативной структуре науки охватывает только часть динамики [21]. Что касается процесса цитирования, это предполагает, что другие инициативы формируют модели цитирования, подобные созданию видимости чьей-то работы через самоцитирование или цитирование работы редактора журнала как попытка улучшить шансы принятия к публикации. Предыдущие исследования обнаружили множество мотиваций, функций и причин библиографий в научной коммуникации [20].

Ранние работы, рассматривающие социальные величины библиографий, были выполнены Гилбертом и позднее МакРобертсом М. и МакРобертсом Б. и др. Гилберт [64] утверждал, что цитирование («библиография») является неотъемлемым средством убеждения. Чтобы убедить научное сообщество в значении и важности своей публикации, авторы используют библиографию как риторический метод. Библиографии варьируются по силе убеждения. Следовательно, будет более убедительным цитировать авторитетную статью, и авторы стремятся отбирать библиографии, которые будут считаться авторитетными в глазах целевой аудитории.

Более того, характеристики авторского поведения при составлении библиографии используются для выступления против использования ссылок в качестве показателей эффективности, например, [12, 65]. На основе эмпирических ситуационных анализов, авторы показали, что очень небольшая доля основы знания статьи (содержащей сотни или тысячи первых публикаций) действительно цитируется. Более того, цитирование противоречиво: некоторые источники цитируются достаточно часто, при каждом использовании, тогда как другие исследования никогда не цитируются, даже если используются гораздо чаще, чем высокоцитируемая работа. Соответственно, они критикуют аналитиков цитирования, которые «несмотря на ошеломляющее количество доказательств обратного... продолжают признавать традиционную точку зре-

ния на науку как привилегированное предприятие, свободное от культурных противоречий и эгоизма и соответственно продолжают обращаться с ссылками так, как если бы они были свободными от культуры измерениями» [65, p. 442].

Точки зрения ранних работ МакРобертсов привели к росту дискуссии, но их выводы, как правило, считаются слишком скоропалительными [21]. Гарфилд, например, утверждал, что будет невозможным цитировать всю раннюю литературу по отдельной теме. В соответствии с основателем SCI, тот факт, что авторы не цитируют все их влияния, не делает ненадежным использование ссылок как мер эффективности при достаточном рассмотрении литературы [66]. Хотя большая часть анализа цитирования, кажется, согласится с тем, что цитирование или библиография противоречивы, утверждается, что это противоречие не является фатальным для использования ссылок как показателей эффективности – в некоторой степени, противоречия усредняются на агрегированных уровнях. По мнению автора [67], наличие разных когнитивных значений ссылок и мотивов к цитированию совсем необязательно считает необоснованным использование ссылок как (несовершенных) измерений эффективности. Мотивы и последствия аналитически различаются.

До сих пор различные подходы нужны, чтобы не устранять друг друга. Фактически, некоторые авторы пытаются развить многоаспектный подход [68-72]. Козенс, например, делал акцент на том, что множественные объяснения ссылок предполагают, что принимаются все аспекты всех перспектив. В ходе написания статьи действия ученого могут быть ориентированы на тот или иной аспект. С одной стороны, поведение индивидуумов при цитировании подвержено влиянию внешнего давления и имеются личные мотивы, собственный интерес и т.п. в процессе цитирования; с другой стороны, существуют некоторые нормы, правила, традиции и этикет, ограничивающие масштаб и возможность применения отдельных действий. Таким образом, имеются правила поведения и взаимодействия, даже если не традиционные мертоновские нормы. Вместо стандарта («идеала») в отличие от отклонения интересным вопросом является понимание моделей и, возможно, идентификация способов связи качества с отдельными сторонами процесса цитирования.

Автор [73] ввел концептуальное различие между динамикой качества и динамикой видимости, чтобы объяснить, как на микроуровне решения о цитировании отдельных статей объединяются и приводят к высокоцитируемым публикациям. Здесь динамика качества опирается на структуру научного знания. Как правило, научный прогресс достигается через разнообразие вкладов. Некоторые представляют основные научные усовершенствования;

другие тонут в деталях. Это различие соотносится с понятиями Коула о ядре и передовом знании [74]. С точки зрения Коула, ядро знания содержит основные теории в рамках области, тогда как передовое знание является знанием, создаваемым сегодня. Большинство исследований, созданных на передовом знании, являются описательным анализом низкого уровня или представляют вклады, которые оборачиваются меньшей или краткой значимостью [75]. Поэтому большая часть того, что публикуется, как таковая, не попадает в ядро знания. Также части того, что публикуется, представляют «ступень» и не функционируют как основа для будущего развития знания. В последствии, по мнению автора работы [73], можно ожидать искаженное распределение подсчетов цитирования и разницу между областями, зависящими от взаимоотношений между эволюционирующим ядром знания и более эфемерным передовым знанием. Одновременно частота цитирования определяется иными механизмами и не является простым отражением динамики качества. Понятие динамики видимости учитывает некоторые из этих механизмов, такие как эффект победителя. Когда статья цитируется многими последующими публикациями, еще больше людей узнают об этой статье. Таким образом, ее видимость, а значит и шансы получить больше ссылок, увеличивается. Это вариант «эффекта Матфея» [76], утверждающего, что признание асимметрично в пользу авторитетных ученых. Аналогично, когда статья получает много ссылок, она приобретает статус авторитетной статьи. В свою очередь еще больше авторов будут ее цитировать, так как апелляция к существующим авторитетам может быть одной из причин цитирования статьи [64]*.

Как указывается выше, предыдущие исследования процесса цитирования не дают никакого прямого ответа на вопрос о том, что стоит за ссылками. Даже сейчас, несмотря на подробные исследования поведения при составлении библиографии, нет единой теории. Тем не менее, сохраняются некоторые общие результаты: библиографии имеют множество функций в научной статье, только небольшая доля релевантной литературы цитируется, и авторы имеют несколько мотивов для включения отдельных исследований как библиографий. В какой степени это влияет на использование ссылок как показателей эффективности, все еще является предметом обсуждения и рассматривается ниже.

* По мнению Смолла, можно предположить, что высокоцитируемые статьи представляют основные понятия, методы или эксперименты в области. Часто цитируемые статьи рассматриваются как примеры (используя терминологию Томаса Куна), посредством которых статьи цитируются, так как они представляют классическое исследование, маркер «понятия» [77], или показывают, как проводится отдельная линия исследования.

Оценка надежности

Несмотря на то, что эмпирические исследования открыли множество факторов, вовлеченных в процесс цитирования, к этому вопросу подходят и с другого угла: сравнением показателей цитирования с результатами рецензирования. В течение последних десятилетий было проведено много подобных исследований. В исследованиях данные коллегами оценки, как правило, рассматриваются как тип стандарта, в котором показатели цитирования могут подтвердиться. Основное предположение состоит в том, что должна существовать корреляция, если ссылки могут быть легитимно использованы как показатели научной эффективности. Исследования различаются по методологии и уровням изучения, выстраиваясь от отдельных статей, индивидуальных исследователей, исследовательских групп и отделов. В трех последних случаях собрание публикаций с агрегированными библиометрическими показателями, как правило, сравнивается с оценкой коллег. Таким образом, сравнительная надежность является менее прямой, фокусируясь на том, как показатели цитирования работают на агрегированных уровнях, а не на уровне отдельных статей.

Некоторые исследования анализировали рецензирование грантов с целью оценки того, цитировались ли претенденты, получившие финансирование, больше, чем не финансируемые [2, 78]. Однако в соответствии с недавним рецензированием, результаты являются неоднозначными [79]. Тогда как некоторые исследования обнаружили положительную корреляцию между финансированием и влиянием цитирования, другие задались вопросом, коррелируют ли рецензирование грантов и влияние цитирования [80].

Имеется также несколько исследований, анализирующих вопрос относительно оценок коллегами научных групп. Например, авторы работы [81] показали, что различные показатели цитирования значительно коррелировали с рейтингом коллег научных программ в физике конденсированных сред. Акнес и Такст [82] анализировали взаимосвязь между библиометрическими показателями и результатами рецензирования норвежских исследовательских групп на факультете математики и естественных наук, сообщив о позитивных, но слабых корреляциях. Другие примеры включают ван Раана [83], который анализировал корреляцию между h-индексом и несколькими стандартами библиометрических показателей с результатами оценки рецензирования для исследовательских групп в области химии в Нидерландах. Он (ван Раан) обнаружил, что и h-индекс, и нормализованный показатель влияния цитирования достаточно хорошо коррелировали с оценками коллег.

В некоторых странах национальные Программы оценки качества исследований (Research Assess-

ment Exercises – RAE) проводятся на регулярной основе. Эти оценки также стимулируют сравнительные анализы показателей цитирования и рейтинг коллег. Например, такие анализы проводились в итальянском контексте [84]. Как часть итальянской RAE, национальное агентство AVNUR анализировало соглашение между стадиями, принадлежащими журнальным статьям, путем основанного на информации рецензирования и путем библиометрических показателей. Важная степень согласованности была найдена «... поддерживающей выбор использования обеих техник с целью оценки качества итальянских научных учреждений» [84, р. 254]. Тем не менее, методологическая основа для такого вывода была опровергнута авторами [85], которые утверждают, что анализ испорчен и что основанное на информации рецензирования и библиометрия не дают одинаковых результатов. Как упоминалось во Введении, Абрамо и Д'Анджело [7] в статье, контрастирующей двумя подходами, также утверждали, что библиометрия явно является предпочтительным методом в естественных и формальных науках. Другие примеры включают работу Оппенхайма [86], который обнаружил сильные положительные корреляции между измерениями ссылок и рейтингом британских исследований в генетике, анатомии, а также археологии в Программе RAE в 1992 г. – но его выводы были подвержены критике со стороны Варнера [87]. Несколько дополнительных исследований рассматривали вопрос относительно последующей оценки качества Программой RAE и ее преемника REF [24]. Самым свежим примером является исследование, сравнивающее результаты REF 2014 г. с различными метриками [88]. Это исследование показывает, что такие метрики предоставляют кардинально разные результаты рецензирования со стороны REF. Что касается влияния ссылок, взвешенного относительно области, то коэффициент корреляции Спирмана, равный 0,28, был установлен на общем уровне, хотя со значительными вариациями в разных областях. Более того, имело место существенное снижение корреляции в самых недавних результатах. Исследование привело к выводу, что метрики не могут обеспечить эквивалентную замену для рецензирования REF. До сих пор исследование не анализировало средние оценки на уровне отделов, которые, как можно утверждать, были бы более релевантными по отношению к REF [89].

В общем можно сделать вывод, что большинство сравнительных анализов, по-видимому, находят среднее положительное соответствие, но установленные корреляции были далеки от идеала и варьировались между исследованиями. Это предполагает, что есть весьма мало эмпирической поддержки для утверждения, что метрики цитирования отражают одинаковые аспекты качества исследования или влияния как оценки рецензирования. Тем

не менее, степень, с которой корреляция кажется достаточной, зависит от контекста целей оценки.

Также имеются некоторые проблемы, связанные с фундаментом такого рода сравнительных исследований [82]. Во-первых, оценка коллегами может включать оценку факторов, выходящих за рамки научного качества, или аспекты, которые маловероятно будут отражены через подсчет цитирования. Только когда показатели ссылок используются в одном и том же контексте решения, что и рецензирования, и они изучают одинаковую величину научной эффективности, можно обоснованно их сравнивать. Эта проблема иллюстрируется в сравнительном анализе REF 2014 г., описанном выше. Здесь основой для анализа был рейтинг коллегами качества, содержащий различные элементы, такие как оригинальность, важность, точность, влияние, жизнеспособность, а также поддержка. Во-вторых, оценки коллегами могут необязательно считаться «истиной», которой должны соответствовать библиометрические измерения – коллеги могут противоречить или ошибаться в своих оценках или им может не хватать компетенции для проведения оценки [90]. Таким образом, обе методологические основы сравнения оценок, данных коллегами, и показателей цитирования, а также предположение о том, что в обеих следует ожидать корреляции, можно ставить под сомнение. Более того, панельные дискуссии все больше рассматривают измерения ссылок как часть процедуры оценки, что предполагает, что их нельзя рассматривать полностью независимыми друг от друга. Это относится к другому вопросу о том, что существует обоюдное воздействие, которое допускает, что подсчеты больших цитирований могут считаться эквивалентом научного качества. Например, по мнению автора [91], издание в журналах с высоким импакт-фактором становится независимым измерением научного качества [92]. Наконец, большое число различных измерений цитирования существует, и результат также зависит от того, какие показатели отбираются для сравнительного анализа.

ЦИТИРУЕМОСТЬ КАК ПОКАЗАТЕЛЬ – ДРУГИЕ ВОПРОСЫ НАДЕЖНОСТИ

Как очевидно из приведенного выше обзора, нет простого ответа на вопрос, что измеряют или оценивают показатели цитируемости. Ясно, что на ссылки как измерения эффективности наложены многие ограничения. Кроме фундаментальных проблем, ассоциируемых с поведением ученых при составлении мультифасетной библиографии, имеются также некоторые более специфические проблемы и ограничения показателей цитирования.

Один существенный вопрос касается охвата примененных БД, а также моделей библиографий. В социальных и гуманитарных науках публикация в книгах является более распространенной, а между-

народные журналы играют менее важную роль. Кроме того, более ранняя литература все еще остается важной, и многие научные области имеют «локальную» ориентацию [93]. Хотя охват литературы по БД цитирования улучшается (Web of Science и Scopus), охват гуманитарных и некоторых социальных дисциплин остается ограниченным [19]. Соответственно, анализу цитирования может не хватать подтверждения в этих областях; отдельные страны, например, Италия, использующая количественные показатели в своей национальной оценке исследований, не включили метрики в оценки социальных и гуманитарных наук [84].

Проблемы, связанные скорее с техническими вопросами, такими как несхожесть между исследуемой статьей и цитируемой библиографией (ошибки в написании названий журналов, именах авторов, в библиографических списках и т.п.), а также ошибки в процедурах индексирования, проведенного Clarivate Analytics (ранее Thompson Reuters) или Elsevier [45, 94], могут исказить анализ цитирования. Подобные ошибки оказывают влияние, в частности, на точность подсчетов цитирования в отдельных статьях. Большое число дополнительных специфических факторов может подрывать использование ссылок как измерений эффективности [95]. Некоторые из них относятся к процессу цитирования, например, так называемым «отрицательным» ссылкам (критика, правка и опровержение других работ), «циклом цитирования» (группы ученых, цитирующих работу другого ученого), а также значительным темпам самоцитирования. Некоторые из этих проблем носят фундаментальный характер и наследуются в любом использовании ссылок как показателей, другие могут быть решены путем строительства более совершенных показателей, тогда как иные снова могут быть менее важны на практике. Например, отрицательные ссылки стремятся стать очень редкими [96], а самоцитирования можно отрегулировать при необходимости.

Тем не менее, проблемы и ограничения анализа цитирования растут по-разному на разных уровнях агрегации [21]. Когда ссылки используются как показатели, агрегированные уровни, представляющие большее число статей и ссылок, как правило, анализируются. По мнению автора [23], это имеет важные применения: «в общем чем больший массив данных цитирования используется, тем выше уровень доверия к результатам. Анализы, полностью вовлеченные в области исследования, нации, регионы и крупные университеты, виртуально не подвержены влиянию со стороны беспокойств и предостережений относительно данных цитирования ... Уровень доверия в этих больших агрегированных уровнях достаточно высок в анализах фундаментального, базового исследования» [23, р. 206] Тем не менее, существует нехватка эмпирических

исследований, подтверждающих, что это действительно так, и возможно некоторые противоречия носят фундаментальный характер, свойственный всем измерениям ссылок, тогда как эффект от иных может стремиться к равновесию, когда рассматриваются агрегированные уровни.

Пример ограничений первого типа относится к искаженному распределению ссылок. Возникает вопрос, являются ли очень высокоцитируемые статьи порядком более влиятельной величины, чем менее высокоцитируемые статьи. В идеале можно хотеть, чтобы показатели цитирования измеряли влияние однообразно: чем выше оценки, тем «лучше» статья [36]. Тем не менее, по мнению Акнеса [73], асимметрия в распределении ссылок больше, чем может оправдать дифференциация качества в научных вкладах. Это происходит из-за того, что сюда вовлечены социологические процессы и процессы агрегации. Вначале статья может цитироваться по независимым причинам (например, используется её содержание). Позже, когда статья широко известна и получила много ссылок, социологические механизмы будут иметь особую важность (авторы, цитирующие авторитетные статьи, эффект победителя и т.п.). Некоторые статьи значительно выиграют от таких эффектов, тогда как другие проиграют.

Как описывается во Введении, существует большое число показателей ссылок, каждый с разными преимуществами и ограничениями. Из-за этого библиометриками долгое время акцентировалось, что в контекстах научной оценки необходимо использовать свыше одного показателя [97]. Например, средняя нормализованная оценка цитирования не зависит от размера и не принимает в расчет число публикаций. По мнению авторов [98], это базовая проблема такого показателя, так как он не представляет истинную продуктивность. Тот факт, что распределения ссылок чрезвычайно искажены, также наводит на вопросы, касающиеся использования среднего как показателя, а авторы работы [99] предложили использовать ранги в процентилях как непараметрическую альтернативу средствам распределения ссылок для нормализации.

ВЕЛИЧИНЫ НАУЧНОГО КАЧЕСТВА И ЦИТИРУЕМОСТЬ

Как показано выше, вопрос об отношении между ссылками и научным качеством сложен и поднимается по-разному в зависимости от анализируемой области, используемой базы данных, примененного временного интервала, показателей и т.д. Более того, качество исследования является многоаспектным понятием, а в этом разделе мы изучим этот вопрос глубже.

В качестве отправной точки можно взять три величины, различаемые Полани [100]: правдопо-

добно, оригинальность и научное значение*. С этой точки зрения хорошее исследование основано на свидетельстве и является научно правильным (правдоподобность), оно обеспечивает новое знание (оригинальность) и имеет важность для других исследований (научное значение). Самые недавние исследования добавили общественную ценность, т.е. включающую важность для общества как четвертую величину научного качества [101, 102]. Во многих программах оценки научных достижений научное качество и общественная важность / влияние рассматриваются как две независимые опоры (например, в Великобритании REF, в Нидерландах SEP и в самых недавних оценках, выполненных Научным советом Норвегии).

Отмечено, что эмпирические исследования ученых относительно научного качества связаны с множеством понятий и аспектов качества. Они идут от коррекции, точности, ясности, результативности, признания, новизны, красоты, важности, автономности, сложности и релевантности до этического/поддерживающего исследования [103-107, 102]. В целом величины могут считаться попытками создать категории на стыке такого множества критериев и аспектов.

Более того, все оценки научного качества могут зависеть от контекста, с точки зрения, например, времени оценки и перспектив экспертов относительно времени/области/разделов. Разные эксперты могут иметь неодинаковое восприятие того, что является важным и полноценным исследованием, а что, являющееся оригинальным, будет со временем меняться по определению. Также могут существовать внутренние нестыковки между величинами. Тогда как надежность и научное значение требуют некоторого соответствия предыдущим установленным нормам и предыдущему исследованию, то самое оригинальное исследование может противопоставляться этому [108, 100].

Короче говоря, тогда как правдоподобность/обоснованность, научное значение и общественная ценность и оригинальность, кажется, широко воспринимаются основными характеристиками научного качества, каждая из этих величин включает разнообразие аспектов; они могут зависеть от контекста и могут также конфликтовать друг с другом.

Ниже обсудим, как ссылки могут относиться к каждой из этих величин понятия качества. Удивительно, что эта тема редко изучается отдельно в литературе и имеется несколько исследований, анализирующих этот вопрос эмпирически. Иссле-

*Отмечено, что Полани использовал термин «научная простота», а не «качество». Качество может быть более широким термином, включающим больше аспектов, чем простота. До сих пор считается, что Полани анализировал одни и те же вопросы, релевантные нашему обсуждению научного качества и показателей цитируемости.

дования поведения при составлении библиографии предоставляют некоторые результаты непрямой релевантности. Однако из самих подсчетов цитирования нельзя обнаружить, почему отдельная статья повторно цитируется другими учеными. Общая методологическая проблема состоит в том, что множество причин библиографий нельзя дедуктивно вывести из ссылок путем «возврата». Причиной этого служит то, что способ индексирования ссылок развивался исторически и привел к потере информации о контексте цитирования в БД по цитированию [109, 110]. Многие причины ссылок в статьях поэтому становятся изъятными из данных. В результате этого ссылки нельзя сортировать в тех ссылках, которые имеют значение для воспринимаемого качества цитируемой статьи, и в тех, которые не имеют значения.

Далее проиллюстрируем это, рассмотрев различные величины, которые вместе содержат общеизвестное понятие «научного качества».

Надежность и правдоподобность

Первая величина понятия качества рассматривает правдоподобность, обоснованность и надежность исследования. Включены такие достоинства, согласно которым исследование должно быть хорошо обоснованным, построенным на научных методах, а также создавать соответствующие результаты.

Как ссылки относятся или отражают эти аспекты понятия качества – сложно оценить, так как нужно рассматривать много разных величин. Даже когда надежность и связанные академические достоинства являются аспектами, которые рассматриваются коллегами при представлении рукописи в журнал для публикации, существуют большие различия, когда речь идет о надежности и правдоподобности опубликованных исследований. Литература содержит множественные публикации, у которых надежность является низкой, результаты ненадежны или даже приводят к недостойному поведению или научному обману [111]. Последний вопрос также изучается эмпирически, показывая, что некоторые публикации, которые отозваны из-за подделки и фальсификации результатов, являются высокоцитируемыми, некоторые даже с сотнями ссылок [112]. Более того, диспропорциональная доля статей, отозванных из-за обмана, была опубликована в престижных журналах высокого импакта. Хотя такие статьи представляют очень маленькую процентную долю от общего объема научной литературы, проблема может усиливаться [112]. Рецензенты журнала честно рассмотрели эти статьи как достаточно надежные для публикации. В более общем виде, есть также признаки того, что методологическая достоверность и правдоподобность недостаточно акцентируются в обзоре рукописей для публикации [113]. Таким образом, система рецензентов не полностью гарантирует ве-

личину качества, касающуюся надежности и правдоподобности, нет признаков того, что высокие подсчеты ссылок отражают надежность.

Этот вопрос можно рассматривать с другого угла: с позиции читателя и потенциального цитирующего. Можно предположить, что в случаях, когда надежность или правдоподобность оцениваются низко, работа будет считаться не заслуживающей цитирования (т.е. будет отклонена), а в случаях, когда более одного исследования показывают одинаковые результаты, у автора есть выбор исследования для цитирования с опорой на свое восприятие того, какое из них является наиболее надежным. А значит надежность /правдоподобность – воспринимаемая при цитировании – может в определенной степени отражаться в моделях цитирования. Тем не менее, имеется мало знаний о степени, в которой понимается, что имеет место тот самый случай, и (как объясняется в разделе **Понимание цитируемости**) когда исследования поведения при цитировании идентифицировали множество факторов, которые сами по себе не ассоциируются с надежностью исследований. Следовательно, кажется невероятным, чтобы ссылки могли рассматриваться обоснованными показателями надежности публикаций.

Оригинальность и новизна

Вторая величина, оригинальность и новизна, происходят из фундаментального требования к исследованию создавать новое знание. Оригинальность может включать новую гипотезу, новые методы, новые теории и модели, а также новые результаты и может растягиваться от дополнений/улучшений выявленного знания до кардинальной новизны /сведения на нет существующего исследования.

Есть основания полагать, что исследования с высокой оригинальностью или новизной будут больше цитироваться. Например, утверждается, что открытия потенциального прорыва в науке можно идентифицировать на основе моделей цитирования [114]. Более того, Нобелевские лауреаты, которые предположительно внесли вклад в исследование с помощью экстраординарной высокой оригинальности и новизны, стремятся быть более высокоцитируемыми, чем средние ученые [115, 116], и могут публиковаться как так называемая «классика цитирования». На основе данных наблюдений Гарфилд ранее изучил вероятность использования статистики цитирования для предсказания будущих победителей [117]. Одновременно высокие подсчеты ссылок необязательно предполагают прорыв или исследование класса Нобелевских лауреатов. Чрезвычайно высокоцитируемая статья Лоури и др. [59] по измерению протеина, описанная выше, представляет собой в этом отношении интересный случай. В силу библиографических норм

данная статья, вероятно, цитируется почти всякий раз, когда используется ее метод. Но, по мнению самого Лоури, «Просто так получилось, это был пустяк, лучший или более легкий или более чувствительный, чем другие методы, и безусловно почти каждый измеряет им протеин в наши дни» [118, р. 363-364].

Примером статей, которые, как правило, будут считаться низкой оригинальностью и новизной, служат статьи так называемых «дублирующих исследований». Хотя такие исследования важны для надежности исследования, для проверки и демонстрации обобщенности существующих наблюдений, они скорее рассматриваются «тренировочными» упражнениями, чем основными вкладами в область [119]. Если результаты только подтверждают предыдущие исследования, у них более низкая новизна и меньше вероятность быть цитируемыми. Многие журналы, кажется, неохотно готовы к публикации дубликатов, так как они будут иметь отрицательное влияние в темпе ссылок, импакт-факторе журнала [120]. Тем не менее недавнее внимание к отсутствию дубликатов результатов в биомедицине, клинических и психологических исследованиях [121] может привести к более высокому социальному статусу дублирующих исследований.

Вышеуказанные наблюдения показывают, что не существует простой взаимосвязи между оригинальностью или новизной и ссылками. Исследования с высокой оригинальностью могут включать как основные научные разработки, так и минимальные вклады. Во втором случае статьи могут не цитироваться, так как их исследовательский вопрос является «тупиком», предполагающим, что несмотря на новый или оригинальный подход, он не несет в себе функцию положительной основы для дальнейшей работы. Это заставляет нас обсудить следующую величину научного качества – научное значение.

Научное значение

Академическая или научная важность может включать релевантность предыдущему, а также будущему исследованию – куммулятивность, также как открытие новых научных областей. Оценки важности исследования могут зависеть от обобщенности результатов, размера и общего интереса к научной области/вопросу.

Научное значение и важность являются величинами понятия качества, по отношению к которому некоторые ссылки могут иметь непосредственную связь. Это общепринято обсуждается следующим образом. Когда ученый ссылается на статью, она оказывается полезной или релевантной в некотором роде для настоящего исследования или для написания публикации. Таким образом, часто цитируемые статьи можно считать более полезными, чем публикации едва ли цитируемые или

не цитируемые вообще, и вероятно более полезными и таким образом важными в их собственном праве [21]. Это предполагает, что число ссылок можно рассматривать как меру полезности статьи, импакта или влияния на другие исследования. Одинаковое обоснование можно использовать для агрегированных уровней статей. Это типичный способ оправдания использования ссылок как показателя эффективности. Тем не менее, как обсуждалось в разделе **Понимание цитируемости**, ссылки имеют как интеллектуальные, так и социальные функции. В последнее время взаимосвязь между научным качеством и ссылками становится более сложной по мере осознания учеными необходимости увеличить свою видимость. Это становится особенно важно в условиях, когда научное финансирование сокращается, а соперничество за ресурсы обостряется. Более того, начиная с использования показателей ссылок в качестве показателей эффективности, ученые осознают, что их библиографии могут оказывать влияние на карьеру ученых, которых они цитируют. Большие числа ссылок на отдельную научную группу или отдельного ученого могут таким образом быть результатом сильной стратегии видимости или прямой или непрямои «игры в ссылки» [122]. Хотя методы стратегического цитирования не являются по определению сомнительными научными практиками (но некоторые из них определенно квалифицируются таковыми), эти процессы недооценивают надежность ссылок как показателя научного качества.

В 1983 г. Б.Р. Мартин и Ирвин [123] описали концептуальное различие между качеством и влиянием следующим образом: «качество есть свойство публикации и описанного в ней исследования. Оно описывает, насколько хорошо выполнено исследование, свободно ли оно от очевидной «ошибки» ... насколько оригинальными являются выводы и т.д.» [123, р. 70]. Влияние публикации, наоборот, определяется как «действительное влияние на окружающие исследование деятельности в определенное время». С точки зрения Б. Р. Мартина и Ирвина [123] влияние публикации теснее всего связано с понятием научного прогресса – публикация, вызывающая большое влияние, представляет основной вклад в знание во время своего опубликования. Используя эти определения, также очевидным становится то, что влияние будет более адекватной интерпретацией ссылок, чем качества. Например, даже «ошибочная» публикация может иметь большее влияние путем стимулирования дальнейшего исследования. Аналогично публикация от признанного ученого может быть более видимой и значит иметь большее влияние, владеть большим числом ссылок, даже если ее качество (с точки зрения оригинальности и устойчивости) оказывается не большим, чем у публикации менее известных ученых [124]. Влияние – это наиболее

общепотребительное понятие для того, что отражают ссылки, хотя другие понятия, такие как воздействие, важность, значимость и польза от случая к случаю тоже используются [1]. Тем не менее, использование *влияния* как наиболее подходящего понятия, как правило, оправдывается теоретическими рассуждениями, и есть несколько попыток изучить этот вопрос эмпирически или связать его с предыдущими наблюдениями относительно поведения при цитировании. Некоторые авторы стремятся решать эту проблему, применяя комбинируемое понятие *влияния ссылок*, так как это выражает методологию, используемую для измерения влияния [1]. По мнению авторов [125], влияние ссылок следует отличать от научного влияния, поскольку влиятельный ученый иногда имеет более низкую эффективность с точки зрения высокоцитируемых публикаций, чем некоторые его менее влиятельные коллеги.

Общественная ценность и релевантность

Эта величина понятия качества может включать любого рода экстранаучную релевантность, например, релевантность по отношению к здоровью, богатству или окружающей среде. Во многих установках исследование с внешним научным значением будет цениться выше, а общественная релевантность и более широкое влияние часто являются частью обзорных критериев по научным грантам финансирующих агентств [126].

Общественная релевантность во многих случаях рассматривается как нечто, гораздо труднее поддающееся измерению, чем научная релевантность или влияние [127]. Есть широко распространенное предположение, что этот вопрос нельзя адекватно оценивать через стандартные показатели ссылок, и в последнее время растущее внимание уделяется развитию методологий оценки и измерения общественной релевантности и влияния [128, 129].

Долго анализ цитирования применялся в патентных исследованиях [130]. Через анализы ссылок в патентах на научные публикации получалось знание о взаимодействии и влиянии науки на технологию. Таким образом эти исследования производят информацию об отдельном типе общественной релевантности и влияния – технологическом [131]. До сих пор основным ограничением является то, что множество инноваций не запатентовано, и патенты не подходят для оценки общественной релевантности или влияния в более широком контексте. Только очень небольшое количество публикаций, индексированных в БД Web of Science или Scopus, реально цитируется патентами [131].

Более общей причиной того, почему общественную релевантность трудно оценить через подсчеты ссылок, служит литература, индексированная в БД Web of Science и Scopus и содержащая

большинство академических и учебных публикаций*. Тогда как ссылки могут отражать внутринаучное использование, использование и применение, имеющие место наравне с другими величинами, гораздо менее вероятно будут охвачены подсчетами ссылок в таких журналах. Например, авторы [132] показали, что некоторые статьи о диабете, оцененные как обладающие большим влиянием на клиническую практику, не получили много ссылок.

Показатели цитируемости также часто рассматриваются как имеющие важные ограничения в областях применения. Например, ле Пер [133] акцентировал, «в технологии или практическом исследовании библиометрия является недостаточным средством оценки. От нее мало пользы, также часто она может приводить к ошибочным выводам» [133, р. 18]. Аналогично, исследование главным образом национального или локального интереса может часто быть плохо цитируемым в литературе, опубликованной в международных научных журналах.

Тем не менее ясно, что научные вклады с большей общественной релевантностью могут также быть высокоцитируемыми. Например, Эвард К. Прескотт и Финн Э. Кидланд получили нобелевскую премию в области экономики в 2004 г. за две статьи, которые оказали глубокое влияние на практику экономической политики вообще и денежную политику в частности [134]. Эти статьи являются высокоцитируемыми также в академической литературе. Аналогично, в 1994 г. скандинавское исследование выживаемости симвастатина (*simvastatin survival study*) обеспечило первое неоспоримое доказательство, что понижение уровня холестерина липопротеинов низкой плотности с помощью лечения статинами снижает сердечно-сосудистые заболевания и общую смертность [135]. Эта статья сегодня цитируется свыше 7700 раз в БД Web of Science. Симвастатин был разработан компанией Merck&Co. и запущен в медицинское использование в 1992 г. и оказал огромное влияние на здоровье людей [136]. До потери своей патентной защиты симвастатин был крупнейшим продаваемым Merck лекарством и вторым самым распространенным снижающим холестерин лекарством в мире. Несмотря на эти и многие другие аналогичные примеры, невозможно идентифицировать общественную релевантность из подсчетов ссылок самих по себе, но и нецитируемая или мало цитируемая публикация может внести вклад в результат большой общественной релевантности.

Как описывается выше, сегодня интерес смещается в сторону разработки альтернативных показате-

телей, которые могут охватить эти аспекты научной деятельности лучше, что может быть недооценено при использовании традиционных на основе ссылок показателей. Это включает альтметрию с использованием данных из источников социальных медиа [137] и разработку моделей для анализа влияния исследования, таких как «подход окупаемости» [138]. Новые нормы анализа цитирования также разрабатываются для анализа общественного влияния исследования. Например, влияние исследования на здравоохранение изучалось с применением данных из публикаций, цитируемых в клинических руководствах [139, 140]. Аналогично новые методы и образцы классификации ссылок вводятся в оценку того, как научные наблюдения переводятся и используются в клинической практике [141].

ЗАКЛЮЧЕНИЕ

Использование показателей цитируемости в контекстах оценки исследований в последнее время растет, как описывается ранее. Мнение, как правило, разделяемое экспертами в рамках библиометрии, кажется, состоит в том, что ссылки представляют хороший, но не идеальный показатель влияния. Однако касаясь различных ограничений, свойственных ссылкам в качестве измерений эффективности, большинство библиометриков утверждает, что библиометрический анализ не может функционировать как дополнение к рецензированию [1]. Вместе с тем, имеются также разные ограничения и недостатки в оценке коллегами [142]. Например, проведенная людьми оценка субъективна, и мнения экспертов могут оказывать влияние на недостаток знания и ограниченный когнитивный кругозор [143, 41]. Более того, рецензенты слишком дороги и медлительны*.

Основываясь на этом, часто утверждается, что библиометрический анализ может разбалансировать недостатки и ошибки в оценках коллег [21]. Таким образом, библиометрическое исследование должно рассматриваться как дополнение к оценке коллегами [44]. По мнению авторов [82], такая комбинация методов улучшит надежность оценок, осуществляемых в Норвегии. В случаях с большим рассеянием между количественной оценкой коллегами и библиометрическими измерениями эффективности, комитет по оценке должен изучать причины этих отклонений. Затем они могут установить, что их собственные оценки ошибочны или

* Тем не менее, в соответствии с сетевой страницей Scopus (<https://www.elsevier.com/solutions/scopus/content>) свыше 300 торговых журналов, ориентированных на особую индустрию, торговлю или тип бизнеса, были выбраны для охвата в Scopus.

* Например, авторы [144] утверждали, что «частное субъективное рецензирование подвержено ошибкам, противоречиво и дорого; поэтому мы должны задать вопрос, стоит ли использование рецензирования в программах, таких как программа оценки исследований и программа оценки научных достижений огромного числа затрачиваемых на него ресурсов».

что библиометрические измерения не отражают эффективность данной единицы [145].

В Программе оценки научных достижений (REF) 2014 г. анализ цитирования проводился для 11 из 36 разграниченных областей подпанелей, большая часть которых принадлежала научным областям о жизни и по физике [6]. В отчете о роли метрик в оценке и управлении исследованием рекомендуется, чтобы «количественные данные – в частности относительно опубликованных результатов – продолжали иметь место в уведомлении относительно оценок рецензирования научного качества. Этот подход успешно применялся в REF 2014 г., и мы рекомендуем, чтобы он продолжал использоваться и улучшаться в дальнейших программах [6]. Однако в то же время он предупреждает: «библиометрики, как правило, рассматривают темп ссылок как представителя измерения научного влияния или влияния на релевантные академические сообщества. Но это только одна из величин научного качества. Качество нуждается в рассмотрении как многоаспектное понятие, которое нельзя охватить каким-либо одним показателем, и то, какая величина качества должна быть приоритетной, может варьироваться в зависимости от области и цели».

Как видно из обсуждения данной статьи, это важный момент, так как ссылки не способны охватить все аспекты понятия качества. Следовательно, рост использования показателей ссылок в оценке исследований и финансирования может подразумевать меньшее внимание на такие величины научного качества как надежность/правдоподобность, оригинальность и общественное значение.

С введением несколько десятилетий назад показателей на основе ссылок в контекст научной оценки часто появляются контрверсии, окружающие применения [109]. Использование библиометрических показателей в целях оценки иногда встречается оппозицией со стороны научных сообществ. Например, ученые обеспокоены возможным отсутствием честности, особенно в случае, если оценки имеют последствия для научного финансирования. Оценки, которые важны или негативны, часто порождают протесты, хотя это применяется ко всем оценкам независимо от использованных методов [146]. В то же самое время другие специалисты приветствуют использование показателей ссылок. Недавний отчет о применении метрик в REF также показывает, что имеется огромное разнообразие мнений внутри академического и научного сообществ [6].

Нет признаков того, что использование ссылок как показателей эффективности в будущем ослабнет. На этом фоне важно разумное использование показателей. Показатели ссылок можно легко использовать неправильно или применять в контекстах, где у них отсутствует подтверждение или

надежность. Растет беспокойство относительно этой проблемы, также как и о потенциальном негативном влиянии научных метрик на научное сообщество. Это демонстрируется примером в публикации Лейденского манифеста, содержащего 10 принципов измерения научной эффективности [10] и Декларации Сан-Франциско об оценке исследований, которые намерены предотвратить практику использования импакт-фактора журнала ... как суррогатной меры качества отдельных научных статей, чтобы оценивать отдельные вклады ученого или при приеме на работу, продвижении или решениях о финансировании» [9, p. 869].

Приходим к выводу, что ссылки отражают – с важными ограничениями – аспекты, относящиеся к научному влиянию и релевантности, но нет свидетельства, что ссылки отражают другие ключевые величины научного качества. Отсутствует четкий путь к лучшему регулированию напряжения между административными потребностями в простых измерениях и более легкими методами оценки и запросом ученых о честных и всесторонних оценках научного качества. Показатели на основе ссылок не могут предоставить достаточно тонкие или надежные измерения качества при использовании в изоляции. Одновременно также имеются проблемы с системой рецензирования. Тем не менее, описанная во Введении точка зрения относительно того, что библиометрическая оценка превосходна по сравнению с традиционным методом рецензирования, не подтверждается в нашем мнении. Рецензирование применяется во многих разных контекстах, из которых оценки коллегами рукописей, представленных в журнал и издательства, вероятно, являются самыми фундаментальными. Что касается такой оценки, то показатели цитируемости едва ли имеют какую-либо релевантность. В более общем смысле показатели цитируемости, кажется, приносят мало пользы в оценке прочности / правдоподобности, оригинальности, общественной ценности исследования.

Финансирование. Авторы приводят сведения о получении следующей финансовой поддержки относительно исследования, авторства и/или публикации этой статьи: данное исследование было профинансировано Научным советом Норвегии, грант № 256223 (theQ-QUESTcentre).

ЛИТЕРАТУРА

1. Moed H. F. Citation analysis in research evaluation. — Dordrecht, The Netherlands: Springer, 2005.
2. Cabezas-Clavijo A., Robinson-Garcia N., Escabias M., Jimenez-Contreras E. Reviewers' ratings and bibliometric indicators: Hand in hand when assessing over research proposals? // PLoS ONE. — 2013. — Vol.8. — No. 6, e68258. — doi:10.1371/journal.pone.0068258
3. Carlsson H. Allocation of research funds using bibliometric indicators—Asset and challenge to Swe-

dish higher education sector // *InfoTrend*. — 2009. — Vol. 64, No. 4. — P. 82-88.

4. *Holden G., Rosenberg G., Barker K.* Bibliometrics: A potential decision-making aid in hiring, reappointment, tenure and promotion decisions / G. Holden, G. Rosenberg, K. Barker (Eds.), *Bibliometrics in social work* (pp. 67-92). — New York, NY: Routledge, 2005.

5. *Piro F. N., Sivertsen G.* How can differences in international university rankings be explained? // *Scientometrics*. — 2016. Vol. 109. — P. 2263-2278.

6. *Wilsdon J., Allen L., Belfiore E., Campbell P., Curry S., Hill S., . . . Johnson B.* The metric tide: Report of the independent review of the role of metrics in research assessment and management. — 2015. — <http://www.hefce.ac.uk/pubs/rereports/year/2015/metrictide/>

7. *Abramo G., D'Angelo C. A.* Evaluating research: From informed peer review to bibliometrics // *Scientometrics*. — 2011. — Vol. 87. — P. 499-514. — doi:10.1007/s11192-011-0352-7

8. *Regibeau P., Rockett K. E.* Research assessment design and the role of bibliometrics. — 2016. — <http://voxeu.org/article/using-bibliometrics-gauge-research-quality>

9. *Cagan R.* The San Francisco Declaration on research assessment // *Disease Models & Mechanisms*. — 2013. — Vol. 6. — P. 869-870. — doi: 10.1242/dmm.012955

10. *Hicks D., Wouters P., Waltman L., de Rijcke S., Rafols I.* Bibliometrics: The Leiden Manifesto for research metrics // *Nature*. — 2015. — Vol. 520. — P. 429-431.

11. *Seglen P. O.* From bad to worse: Evaluation by journal impact // *Trends in Biochemical Sciences*. — 1989. — Vol.14. — P. 326-327.

12. *MacRoberts M. H., MacRoberts B. R.* Problems of citation analysis: A critical review // *Journal of the American Society for Information Science*. — 1989. — Vol. 40. — P. 342-349.

13. *Osterlob M., Frey B. S.* Ranking games // *Evaluation Review*. — 2015. — Vol. 39. — P. 102-129. — doi: 10.1177/0193841x14524957

14. *Weingart P.* Impact of bibliometrics upon the science system: Inadvertent consequences? // *Scientometrics*. — 2004. — Vol.62. — P. 117-131.

15. *Seglen P. O.* Citation rates and journal impact factors are not suitable for evaluation of research // *Acta Orthopaedica Scandinavica*. — 1998. — Vol.69. — P. 224-229. — doi:10.3109/17453679809000920

16. *Cronin B.* The citation process: The role and significance of citations in scientific communication. — London, England: Taylor Graham, 1984.

17. *Waltman L., van Eck, N. J., van Leeuwen T. N., Visser M. S., van Raan A. F. J.* Towards a new crown indicator: Some theoretical considerations // *Journal*

of Informetrics. — 2011. — Vol. 5. — P. 37-47. — doi: 10.1016/j.joi.2010.08.001

18. *Bornmann L., Daniel H. D.* What do we know about the h index? // *Journal of the American Society for Information Science and Technology*. — 2007. — Vol. 58. — P. 1381-1385. — doi:10.1002/asi.20609

19. *Waltman L.* A review of the literature on citation impact indicators // *Journal of Informetrics*. — 2016. — Vol. 10. — P. 365-391. — doi: 10.1016/j.joi.2016.02.007

20. *Bornmann L., Daniel H. D.* What do citation counts measure? A review of studies on citing behavior // *Journal of Documentation*. — 2008. — Vol. 64. — P. 45-80. — doi:10.1108/00220410810844150

21. *Aksnes D. W.* Citations and their use as indicators in science policy: Studies of validity and applicability issues with a particular focus on highly cited papers (Doctoral thesis). — University of Twente, Enschede, The Netherlands, 2005.

22. *Wouters P.* Quality and impact in research evaluation / S. Kuhlmann D. Simon, J. Stamm (Eds.). *Handbook of science policy studies*. — Springer. (в печати).

23. *Welljams-Dorof A.* Quantitative citation data as indicators in science evaluations: A primer on their appropriate use / M. S. Frankel, J. Cave (Eds.), *Evaluating science and scientists: An East-West dialogue on research evaluation in post-communist Europe* (pp. 202-211). — Budapest, Hungary: Central European University Press, 1997.

24. *de Rijcke S., Wouters P. F., Rushforth A. D., Franssen T. P., Hammarfelt B.* Evaluation practices and effects of indicator use—A literature review // *Research Evaluation*. — 2016. — Vol. 25. — P. 161-169. — doi: 10.1093/reseval/rvv038

25. *Vinkler P.* The evaluation of research by scientometric indicators. — Oxford, UK: Chandos Publishing, 2010.

26. *Waltman L., Schreiber M.* On the calculation of percentile-based bibliometric indicators // *Journal of the American Society for Information Science and Technology*. — 2013. — Vol. 64. — P. 372-379. — doi: 10.1002/asi.22775

27. *Bornmann L., Marx W., Gasparyan A. Y., Kitas G.* Diversity, value and limitations of the journal impact factor and alternative metrics // *Rheumatology International*. — 2012. — Vol. 32. — P. 1861-1867. — doi:10.1007/s00296-011-2276-1

28. *Marx W., Bornmann L.* On the causes of subject-specific citation rates in Web of Science // *Scientometrics*. — 2015. — Vol. 102. — P. 1823-1827. — doi:10.1007/s11192-014-1499-9

29. *Harzing A. W., Alakangas S.* Google Scholar, Scopus and the Web of Science: A longitudinal and cross-disciplinary comparison // *Scientometrics*. — 2016. — Vol. 106. — P. 787-804. — doi: 10.1007/s11192-015-1798-9

30. *Schubert A., Braun T.* Relative indicators and relational charts for comparative assessment of publication output and citation impact // *Scientometrics*. — 1986. — Vol. 9. — P. 281-291.
31. *Schubert A., Glänzel W., Braun T.* World flash on basic research: Subject field characteristic citation scores and scales for assessing research performance // *Scientometrics*. — 1987. — Vol. 12. — P. 267-292.
32. *Garfield E.* Citation indexing—Its theory and application in science, technology and humanities. — New York, NY: John Wiley, 1979.
33. *van Raan A. F. J.* Measuring science: Capita selecta of current main issues / H. F. Moed, W. Glänzel, U. Schmoch (Eds.), *Handbook of quantitative science and technology research* (pp. 19-50). — Dordrecht: Springer, 2004.
34. *Opthof T., Leydesdorff L.* Caveats for the journal and field normalizations in the CWTS (“Leiden”) evaluations of research performance // *Journal of Informetrics*. — 2010. — Vol. 4. — P. 423-430. — doi: 10.1016/j.joi.2010.02.003
35. *Waltman L., van Eck N. J.* Source normalized indicators of citation impact: An overview of different approaches and an empirical comparison // *Scientometrics*. — 2013. — Vol. 96. — P. 699-716. — doi: 10.1007/s11192-012-0913-4
36. *Ioannidis J. P. A., Boyack K., Wouters P. F.* Citation metrics: A primer on how (not) to normalize // *PLoS Biology*. — 2016. — Vol. 14, No. 9, e1002542. — doi: 10.1371/journal.pbio.1002542
37. *Larivière V., Gingras Y.* Averages of ratios vs. ratios of averages: An empirical analysis of four levels of aggregation // *Journal of Informetrics*. — 2011. — Vol. 5. — P. 392-399. — doi: 10.1016/j.joi.2011.02.001
38. *Waltman L., van Eck N. J., van Leeuwen T. N., Visser M. S., van Raan A. F. J.* Towards a new crown indicator: An empirical analysis // *Scientometrics*. — 2011. — Vol. 87. — P. 467-481. — doi: 10.1007/s11192-011-0354-5
39. *Price D. J. d. S.* Networks of scientific papers // *Science*. — 1965. — Vol. 149. — P. 510-515.
40. *Langfeldt L., Benner M., Sivertsen G., Kristiansen E. H., Aksnes D. W., Borlang S. B., . . . Pelkonen A.* Excellence and growth dynamics: A comparative study of the Matthew effect // *Science and Public Policy*. — 2015. — Vol. 42. — P. 661-675. — doi: 10.1093/scipol/scu083
41. *van Raan A. F. J.* The Pandora’s box of citation analysis: Measuring scientific excellence — The last evil? // B. Cronin, H. B. Atkins (Eds.), *The web of knowledge: A Festschrift in honor of Eugene Garfield* (pp. 301-319). — Medford, NJ: American Society for Information Science, 2000.
42. *Hirsch J. E.* An index to quantify an individual’s scientific research output // *Proceedings of the National Academy of Sciences of the United States of America*. — 2005. — Vol. 102. — P. 16569-16572. — doi: 10.1073/pnas.0507655102
43. *Alonso S., Cabrerizo F. J., Herrera-Viedma E., Herrera F.* h-Index: A review focused in its variants, computation and standardization for different scientific fields // *Journal of Informetrics*. — 2009. — Vol. 3. — P. 273-289. — doi: 10.1016/j.joi.2009.04.001
44. *Council of Canadian Academies.* Informing research choices: Indicators and judgment: The expert panel on science performance and research funding. — 2012. — <https://www.scienceadvice.ca/reports/informing-research-choices-indicators-and-judgment/>
45. *Leydesdorff L., Wouters P., Bornmann L.* Professional and citizen bibliometrics: Complementarities and ambivalences in the development and use of indicators—A state-of-the-art report // *Scientometrics*. — 2016. — Vol. 109. — P. 2129-2150.
46. *Baumgartner S. E., Leydesdorff L.* Group-Based Trajectory Modeling (GBTM) of citations in scholarly literature: Dynamic qualities of “transient” and “sticky knowledge claims.” // *Journal of the Association for Information Science and Technology*. — 2014. — Vol. 65. — P. 797-811. — doi: 10.1002/asi.23009
47. *Levitt J. M., Thebwall M.* A combined bibliometric indicator to predict article impact // *Information Processing & Management*. — 2011. — Vol. 47. — P. 300-308. — doi: 10.1016/j.ipm.2010.09.005
48. *Cole J. R., Cole S.* Social stratification in science. — Chicago, IL: The University of Chicago Press, 1973.
49. *Durieux V., Gevenois P. A.* Bibliometric indicators: Quality measurements of scientific publication // *Radiology*. — 2010. — Vol. 255. — P. 342-351. — doi: 10.1148/radiol.09090626
50. *Smith L. C.* Citation analysis // *Library Trends*. — 1981. — Vol. 30. — P. 83-106.
51. *Nicolaisen J.* Citation analysis // *Annual Review of Information Science and Technology*. — 2007. — Vol. 41. — P. 609-641. — doi: 10.1002/aris.2007.1440410120
52. *Camacho-Minano M. D. M., Nunez-Nickel M.* The multilayered nature of reference selection // *Journal of the American Society for Information Science and Technology*. — 2009. — Vol. 60. — P. 754-777. — doi: 10.1002/asi.21018
53. *Thornley C., Watkinson A., Nicholas D., Valentine R., Jamali H. R., Herman E., . . . Tenopir C.* The role of trust and authority in the citation behaviour of researchers // *Information Research: An International Electronic Journal*. — 2015. — Vol. 20, No. 3. — P. 1-17.
54. *Willett P.* Readers’ perceptions of authors’ citation behavior // *Journal of Documentation*. — 2013. — Vol. 69. — P. 145-156. — doi: 10.1108/00220411311295360
55. *Garfield E.* Can citation indexing be automated? / *Essay of an information scientist* (Vol. 1, pp. 84-90). — Philadelphia, PA: ISI Press, 1977.
56. *Small H. G.* Citation context analysis / B. Dervin, M. Voigt (Eds.), *Progress in communication sciences* (Vol. 3, pp. 287-310). — Norwood, NJ: Ablex, 1982.

57. *Law J.* The heterogeneity of texts / M. Callon, J. Law, & A. Rip (Eds.), Mapping the dynamics of science and technology (pp. 67-83). — London, England: Macmillan, 1986.
58. *Voos H., Dagaev K. S.* Are all citations equal? Or, did we op. cit. your idem? // The Journal of Academic Librarianship. — 1976. — Vol. 1, No. 6. — P. 19-21.
59. *Lowry O. H., Rosebrough N. J., Farr A. L., Randall R. J.* Protein measurement with the Folin phenol reagent // Journal of Biological Chemistry. — 1951. — Vol. 193. — P. 265-275.
60. *Van Noorden R., Maber B., Nuzzo R.* The top 100 papers // Nature. — 2014. — Vol. 514. — P. 550-553.
61. *Liu M. X.* Progress in documentation the complexities of citation practice: A review of citation studies // Journal of Documentation. — 1993. — Vol. 49. — P. 370-408. — doi:10.1108/eb026920
62. *Merton R. K.* Foreword. / E. Garfield (Ed.), Citation indexing — Its theory and application in science, technology, and humanities (pp. v-ix). — New York, NY: John Wiley, 1979.
63. *Amsterdamska O., Leydesdorff L.* Citations: Indicators of significance? // Scientometrics. — 1979. — Vol.15. — P. 449-471.
63. *Hellqvist B.* Referencing in the humanities and its implications for citation analysis // Journal of the American Society for Information Science and Technology. — 2010. — Vol. 61. — P. 310-318. — doi:10.1002/asi.21256
64. *Gilbert N. G.* Referencing as persuasion // Social Studies of Science. — 1977. — Vol. 7. — P. 113-122.
65. *MacRoberts M. H., MacRoberts B. R.* Problems of citation analysis // Scientometrics. — 1996. — Vol. 36. — P. 435-444.
66. *Garfield E.* Validation of citation analysis // Journal of the American Society for Information Science. — 1997. — Vol. 48. — P. 962-963.
67. *Luukkonen T.* Citations in the rhetorical, reward, and communication systems of science (Doctoral thesis). — University of Tampere, Finland, 1990.
69. *Cozzens S. E.* What do citations count? The rhetoric-first model // Scientometrics. — 1989. — Vol.15. — P. 437-447.
70. *Glaser J., Laudel G.* Integrating scientometric indicators into sociological studies: Methodical and methodological problems // Scientometrics. — 2001. — Vol. 52. — P. 411-434. — doi:10.1023/a:1014243832084
71. *Leydesdorff L.* Words and co-words as indicators of intellectual organization // Research Policy. — 1989. — Vol.18. — P. 209-223.
72. *Luukkonen T.* Why has Latour's theory of citations been ignored by the bibliometric community? Discussion of sociological interpretations of citation analysis // Scientometrics. — 1997. — Vol. 38. — P. 27-37.
73. *Aksnes D. W.* Characteristics of highly cited papers // Research Evaluation. — 2003. — Vol.12. — P. 159-170.
74. *Cole S.* Making science: Between nature and society. — London, England: Harvard University Press, 1992.
75. *Cole S.* The role of journals in the growth of scientific knowledge / B. Cronin, H. B. Atkins (Eds.), The web of knowledge: A Festschrift in honor of Eugene Garfield (pp. 109-142). — Medford, NJ: American Society for Information Science, 2000.
76. *Merton R. K.* The Matthew effect in science // Science. — 1968. — Vol. 159. — P.56-63.
77. *Small H. G.* Cited documents as concept symbols // Social Studies of Science. — 1978. — Vol. 8. — P. 327-340.
78. *Hornbostel S., Bohmer S., Klingsporn B., Neufeld J., von Ins M.* Funding of young scientist and scientific excellence // Scientometrics. — 2009. — Vol. 79. — P. 171-190. — doi:10.1007/s11192-009-0411-5
79. *Wouters P., Thelwall M., Kousha K., Waltman L., de Rijcke S., Rushforth A., Franssen T.* The metric tide: Literature review: Supplementary report I to the independent review of the role of metrics in research assessment and management. — 2015. — <http://www.hcfce.ac.uk/pubs/rereports/year/2015/metrictide>
80. *Bornmann L.* Scientific peer review // Annual Review of Information Science and Technology. — 2011. — Vol.45. — P. 199-245.
81. *Rinia E. J., van Leeuwen T. N., van Vuren H. G., van Raan A. F. J.* Comparative analysis of a set of bibliometric indicators and central peer review criteria: Evaluation of condensed matter physics in the Netherlands // Research Policy. — 1998. — Vol. 27. — P. 95-107.
82. *Aksnes D. W., Taxt R. E.* Peer reviews and bibliometric indicators: A comparative study at a Norwegian university // Research Evaluation. — 2004. — Vol. 13. — P. 33-41. — doi: 10.3152/147154404781776563
83. *van Raan A. F. J.* Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups // Scientometrics. — 2006. — Vol.67. — P. 491-502. — doi: 10.1556/Scient.67.2006.3.10
84. *Ancaiani A., Anfossi A. F., Barbara A., Benedetto S., Blasi B., Carletti V., . . . Sileoni S.* Evaluating scientific research in Italy: The 2004-10 research evaluation exercise // Research Evaluation. — 2015. — Vol. 24 — P. 242-255. — doi:10.1093/reseval/rvv008
85. *Baccini A., De Nicolao G.* Do they agree? Bibliometric evaluation versus informed peer review in the Italian research assessment exercise // Scientometrics. — 2016. — Vol. 108. — P. 1651-1671. — doi: 10.1007/s11192-016-1929-y
86. *Oppenheim C.* The correlation between citation counts and the 1992 research assessment exercise ratings for British research in genetics, anatomy and ar-

- chaeology // *Journal of Documentation*. — 1997. — Vol. 53. — P. 477-487.
87. *Warner J.* A critical review of the application of citation studies to the research assessment exercises // *Journal of Information Science*. — 2000. — Vol. 26. — P. 453-460.
88. *Higher Education Funding Council for England.* The metric tide: Correlation analysis of REF 2014 scores and metrics: Supplementary report II to the independent review of the role of metrics in research assessment and management. — 2015. — https://www.dcsience.net/2015_metrictideS2.pdf
89. *Traag V., Waltman L.* Systematic analysis of agreement between metrics and peer review in the UK REF. — 2018. — [arXiv preprint arXiv:1808.03491].
90. *Rip A.* Qualitative conditions of scientometrics: The new challenges // *Scientometrics*. — 1997. — Vol. 38. — P. 7-26.
91. *Wouters P.* Beyond the Holy Grail: From citation theory to indicator theories // *Scientometrics*. — 1999. — Vol. 44. — P. 561-580. — doi: 10.1007/bf02458496
92. *Rushforth A., deRijcke S.* Accounting for impact? The journal impact factor and the making of biomedical research in the Netherlands // *Minerva*. — 2015. — Vol. 53. — P. 117-139. — doi:10.1007/s11024-015-9274-5
93. *Ossenblok T. L. B., Engels T. C. E., Sivertsen G.* The representation of the social sciences and humanities in the Web of Science — A comparison of publication patterns and incentive structures in Flanders and Norway (2005-9) // *Research Evaluation*. — 2012. — Vol. 21. — P. 280-290. — doi: 10.1093/reseval/rvs019
94. *Moed H. F.* The impact-factors debate: The ISI's uses and limits // *Nature*. — 2002. — Vol. 415. — P. 731-732.
95. *Seglen P. O.* Citations and journal impact factors: Questionable indicators of research quality // *Allergy*. — 1997. — Vol. 52. — P. 1050-1056.
96. *Catalini C., Lacetera N., Oettl A.* The incidence and role of negative citations in science // *Proceedings of the National Academy of Sciences of the United States of America*. — 2015. — Vol. 112. — P. 13823-13826. — doi:10.1073/pnas.1502280112
97. *van Raan A. F. J.* Advanced bibliometric methods to assess research performance and scientific development: Basic principles and recent practical applications // *Research Evaluation*. — 1993. — Vol. 3. — P. 151-166.
98. *Abramo G., D'Angelo C. A.* A farewell to the MNCS and like size-independent indicators // *Journal of Informetrics*. — 2016. — Vol. 10. — P. 646-651. — doi: 10.1016/j.joi.2016.04.006
99. *Bornmann L., Mutz R.* Further steps towards an ideal method of measuring citation performance: The avoidance of citation (ratio) averages in field-normalization // *Journal of Informetrics*. — 2011. — Vol. 5. — P. 228-230. — doi: 10.1016/j.joi.2010.10.009
100. *Polanyi M.* The republic of science: Its political and economic theory // *Minerva*. — 1962. — Vol. 1. — P. 54-73.
101. *Gulbrandsen J. M.* Research quality and organizational factors: An investigation of the relationship (Doctoral thesis). — Norwegian University of Science and Technology, Trondheim, 2000.
102. *Lamont M.* How professors think: Inside the curious world of academic judgment. — Cambridge, MA: Harvard University Press, 2009.
103. *Aksnes D. W., Rip A.* Researchers' perceptions of citations // *Research Policy*. — 2009. — Vol. 38. — P. 895-905. — doi: 10.1016/j.respol.2009.02.001
104. *Bazeley P.* Conceptualising research performance // *Studies in Higher Education*. — 2010. — Vol. 35. — P. 889-903. — doi:10.1080/03075070903348404
105. *Hemlin S.* Quality in science: Researchers' conceptions and judgments (Doctoral thesis). — University of Göteborg, Sweden, 1991.
106. *Hug S. E., Ochsner M., Daniel H. D.* Criteria for assessing research quality in the humanities: A Delphi study among scholars of English literature, German literature and art history // *Research Evaluation*. — 2013. — Vol. 22. — P. 369-383. — doi:10.1093/reseval/rvt008
107. *Martensson P., Fors U., Wallin S. B., Zander U., Nilsson G. H.* Evaluating research: A multidisciplinary approach to assessing research practice and quality // *Research Policy*. — 2016. — Vol. 45. — P. 593-603. — doi: 10.1016/j.respol.2015.11.009
108. *Luukkainen T.* Conservatism and risk-taking in peer review: Emerging ERC practices // *Research Evaluation*. — 2012. — Vol. 21. — P. 48-60. — doi: 10.1093/reseval/rvs001
109. *Wouters P.* The citation culture (Doctoral thesis). — University of Amsterdam, The Netherlands, 1999. — <http://garfield.library.upenn.edu/wouters/wouters.pdf>
110. *Wouters P.* The citation: From culture to infrastructure / B. Cronin, C. R. Sugimoto (Eds.), *Beyond bibliometrics: Harnessing multidimensional indicators of scholarly performance* (pp. 47-66). — Cambridge: MIT Press, 2014.
111. *Fanelli D.* How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data // *PLoS ONE*. — 2009. — Vol. 4, No. 5, e5738. — doi: 10.1371/journal.pone.0005738
112. *Fang F. C., Steen R. G., Casadevall A.* Misconduct accounts for the majority of retracted scientific publications // *Proceedings of the National Academy of Sciences of the United States of America*. — 2012. — Vol. 109. — P. 17028-17033. — doi: 10.1073/pnas.1212247109
113. *Lee C. J.* Commensuration bias in peer review // *Philosophy of Science*. — 2015. — Vol. 82. — P. 1272-1283. — doi:10.1086/683652

114. *Winnink J. J., Tijssen R. J. W., van Raan A. F. J.* Theory-changing breakthroughs in science: The impact of research team work on scientific discoveries // *Journal of the Association for Information Science and Technology*. — 2016. — Vol. 67. — P. 1210-1223. — doi: 10.1002/asi.23505
115. *Gingras Y., Wallace M. L.* Why it has become more difficult to predict Nobel Prize winners: A bibliometric analysis of nominees and winners of the chemistry and physics prizes (1901-2007) // *Scientometrics*. — 2010. — Vol. 82. — P. 401-412. — doi: 10.1007/s11192-009-0035-9
116. *Wagner C. S., Horlings E., Whetsell T. A., Mattsson P., Nordqvist K.* Do Nobel laureates create prize-winning networks? An analysis of collaborative research in physiology or medicine // *PLoS ONE*. — 2015. — Vol. 10, No.7, e0134164. — doi: 10.1371/journal.pone.0134164
117. *Garfield E., Welljams-Dorof A.* Of Nobel class: A citation perspective on high impact research authors // *Theoretical Medicine*. — 1992. — Vol. 13. — P. 117-135. — doi: 10.1007/bf02163625
118. *Garfield E.* Is citation analysis a legitimate evaluation tool? // *Scientometrics*. — 1979. — Vol. 1. — P. 359-375.
119. *Everett J. A. C., Earp B. D.* A tragedy of the (academic) commons: Interpreting the replication crisis in psychology as a social dilemma for early-career researchers // *Frontiers in Psychology*. — 2015. — Vol. 6. P. 1152. — doi: 10.3389/fpsyg.2015.01152
120. *Martin G. N., Clarke R. M.* Are psychology journals anti-replication? A snapshot of editorial practices // *Frontiers in Psychology*. — 2017. — Vol. 8. — P. 523.
121. *Ioannidis J. P. A.* Why most published research findings are false // *PLoS Medicine*. — 2005. — Vol. 2. — P. 696-701. — doi: 10.1371/journal.pmed.0020124
122. *Biagioli M., Lippman A.* Gaming Metrics. Beyond Publish or Perish: Metrics and the new Ecologies of Academic Misconduct. — MIT Press. (в печати).
123. *Martin B. R., Irvine J.* Assessing basic research: Some partial indicators of scientific progress in radio astronomy // *Research Policy*. — 1983. — Vol. 12. — P. 61-90.
124. *Martin B. R.* The use of multiple indicators in the assessment of basic research // *Scientometrics*. — 1996. — Vol. 36. — P. 343-362.
125. *Waltman L., van Eck N. J., Wouters P.* Counting publications and citations: Is more always better? // *Journal of Informetrics*. — 2013. — Vol. 7. — P. 635-641. — doi: 10.1016/j.joi.2013.04.001
126. *Langfeldt L., Scordato L.* Assessing the broader impacts of research: A review of methods and practices (NIFU Working Paper 8/2015). — Oslo, 2015 (April).
127. *Martin B. R.* The research excellence framework and the “impact agenda”: Are we creating a Frankenstein monster? // *Research Evaluation*. — 2011. — Vol. 20. — P. 247-254. — doi: 10.3152/095820211x13118583635693
128. *Bornmann L.* Measuring the societal impact of research // *EMBO Reports*. — 2012. — Vol. 13. — P. 673-676. — doi:10.1038/embor.2012.99
129. *Bornmann L.* What is societal impact of research and how can it be assessed? A literature survey // *Journal of the American Society for Information Science and Technology*. — 2013. — Vol. 64. — P. 217-233. — doi: 10.1002/asi.22803
130. *Meyer M.* Does science push technology? Patents citing scientific literature // *Research Policy*. — 2000. — Vol.29. — P. 409-434. — doi:10.1016/s0048-7333(99)00040-2
131. *van Raan A. F. J.* Patent citations analysis and its value in research evaluation: A review and a new approach to map technology-relevant research // *Journal of Data and Information Science*. — 2017. — Vol. 2. — P. 13-50.
132. *Hanney S. R., Home P. D., Frame I., Grant J., Green P., Buxton M. J.* Identifying the impact of diabetes research // *Diabetic Medicine*. — 2006. — Vol. 23. — P. 176-184. — doi: 10.1111/j.1464-5491.2005.01753.x
133. *le Pair C.* Formal evaluation methods: Their utility and limitations // *International Forum on Information and Documentation*. — 1995. Vol. 20, No.4. — P. 16-24.
134. *Dymond, L. H.* A recent history of recognized economic thought — Contributions of the Nobel laureates to economic science. — Lulu Publishing Services, 2015.
135. *Pedersen T. R., Kjekshus J., Berg K., Haghfelt T., Faergeman O., Thorgeirsson G., . . . Grundstrom I.* Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: The Scandinavian Simvastatin Survival Study(4S) // *The Lancet*. — 1994. — Vol. 344. — P. 1383-1389.
136. *Li J. J.* Triumph of the heart: The story of statins. — New York, NY: Oxford University Press, 2009.
137. *Weller K.* Social media and altmetrics: An overview of current alternative approaches to measuring scholarly impact / I. M. Welp, J. Wollersheim, S. Ringelhan, M. Osterloh (Eds.), *Incentives and performance: Governance of research organizations* (pp. 261-279). — Cambridge, UK: Springer, 2015.
138. *Donovan C., Hanney S.* The “payback framework” explained // *Research Evaluation*. — 2011. — Vol.20. — P. 181-183. — doi: 10.3152/095820211x13118583635756
139. *Grant J., Cottrell R., Cluzeau F., Fawcett G.* Evaluating “payback” on biomedical research from papers cited in clinical guidelines: Applied bibliometric study // *British Medical Journal*. — 2000. — Vol. 320, Article 1107. — doi:10.1136/bmj.320.7242.1107
140. *Lewison G., Sullivan R.* The impact of cancer research: How publications influence UK cancer clinical

guidelines // *British Journal of Cancer*. — 2008. — Vol. 98. — P. 1944-1950. — doi: 10.1038/sj.bjc.6604405

141. *Jones H., Hanney S.* Tracing the indirect societal impacts of biomedical research: Development and piloting of a technique based on citations // *Scientometrics*. — 2016. — Vol. 107. — P. 975-1003. — doi: 10.1007/s11192-016-1895-4

142. *Chubin D. E., Hackett E. J.* Peerless science: Peer review and U.S. science policy. — Albany: State University of New York Press, 1990.

143. *Lee C. J., Sugimoto C. R., Zhang G., Cronin B.* Bias in peer review // *Journal of the American Society for Information Science and Technology*. — 2013. — Vol. 64. — P. 2-17. — doi:10.1002/asi.22784

144. *Eyre-Walker A., Stoletzki N.* The assessment of science: The relative merits of post-publication review, the impact factor, and the number of citations // *PLoS Biology*. — 2013. — Vol. 11, No. 10, e1001675. — doi: 10.1371/journal.pbio.1001675

145. *van Raan A. F. J.* Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises // *Scientometrics*. — 1996. — Vol. 36. — P. 397-420.

146. *Luukkainen T.* Quantitative techniques in evaluation in Western Europe / M. S. Finkel, J. Cave (Eds.), *Evaluating science and scientists: An East-West dialogue on research evaluation in post-communist Europe* (pp. 115-131). — Budapest, Hungary: Central European University Press, 1997.