

УДК 005.935.3 + 025.2  
ББК 65.050.2-73

**О СТАТЬЕ О.В. ФЕДОРЦА «КОЛЛЕКТИВНАЯ  
ЭКСПЕРТИЗА НАУЧНЫХ ЖУРНАЛОВ: МЕТОДИКА  
АГРЕГИРОВАНИЯ ЭКСПЕРТНЫХ ОЦЕНОК И  
ПОСТРОЕНИЯ РЕЙТИНГА»<sup>1</sup>**

**Миркин Б. Г.<sup>2</sup>**

*(Государственный университет –  
Высшая школа экономики, Москва;  
Биркбек Колледж Лондонского университета,  
Лондон, Великобритания)*

*Объясняется ненаучный характер статьи О. В. Федорца «Коллективная экспертиза научных журналов: методика агрегирования экспертных оценок и построения рейтинга»; вместе с тем само наличие методики объявляется весьма полезным. Предлагается альтернативный подход к оценке российских научных журналов, основанный на использовании таксономии предметной области и соответствующей детализации экспертных оценок.*

Ключевые слова: критические замечания, рейтинг научного журнала, коллективная экспертиза, индекс цитирования.

Статья [1] О. В. Федорца, на мой взгляд, не может рассматриваться или публиковаться в качестве научной работы, так как не содержит достаточного материала по обоснованию и вери-

---

<sup>1</sup> Данная публикация отражает работу автора в Лаборатории выбора и анализа решений ВШЭ и частично профинансирована этой Лабораторией.

<sup>2</sup> Борис Григорьевич Миркин, доктор технических наук ([bmirkin@hse.ru](mailto:bmirkin@hse.ru)).

фикации описываемого метода обработки экспертной информации. Более того, предложенный метод вызывает существенные замечания, если не возражения, с точки зрения корректности. Вместе с тем статья представляет несомненный общественный и исторический интерес с точки зрения процесса рационализации решений по такой социально значимой проблеме, как отбор корпуса журналов для реферирования.

Автор представляет метод ранжирования изданий для выбора их подмножества  $P$  для реферирования. Согласно этому методу индивиды, рассматриваемые в качестве экспертов, дают словесную оценку *некоторых* из изданий в шестибальной шкале, нуль которой означает «не включать», а остальные оценки переводятся в числовую шкалу от 0,1 (наименее предпочтительный) до 0,9 (наиболее предпочтительный) практически равномерно. При этом одни – более известные – издания получают много оценок, тогда как другие – менее известные – мало. Тем не менее, по причинам, в статье не раскрываемым, в качестве агрегированной оценки издания автор берет суммарные, а не усредненные величины экспертных оценок, и эти суммарные величины действительно-таки оказываются характеризующими не качество, а популярность изданий. По причинам, не раскрываемым в статье, автор по-прежнему избегает усредненных оценок (возможно, автор считает, что суммирование сохраняет порядковый тип шкалы, а усреднение – нет), а вместо этого формирует лексикографическое правило с тремя критериями. А именно, сначала формируются пять классов изданий по максимуму выставленных им оценок, затем внутри этих классов – издания упорядочиваются по суммарной оценке, а поскольку и это не обеспечивает достаточной грануляции, то подключается третий критерий, основанный на относительном ранге издания у эксперта, давшего максимальную оценку; вычисление относительного ранга, по-видимому, тоже не считается нарушением порядковости шкалы.

Почему взят именно этот метод? Почему равномерная или почти равномерная шкала? Почему, наконец, сумма оценок,

зависящая от их количества, а не медиана или, пусть условное, среднее? Автор не объясняет сравнительных преимуществ метода, не исследует устойчивость получаемого решения, не рассматривает качество изданий, выбранных в множество  $P$ . Я так понимаю, сам факт использования метода, а не случайных или разумных, но относящихся к индивидуальным изданиям соображений, является результатом в деятельности ВИНТИ на данном этапе.

Действительно, рассматриваемая проблема представляет существенный практический интерес. Можно предположить, что в идеальном мире существует «идеальное» множество  $P^*$  журналов для реферирования. Качество такого множества  $P^*$ , в принципе, может определяться тем, насколько эффективно выбранные издания могут быть использованы для практических и теоретических разработок в данной стране. Но это только потом, а сейчас, в качестве, так сказать, суррогата – выбора  $P$ , а не  $P^*$  – в качестве критерия могут рассматриваться мнения специалистов и метрические индексы – цитируемость и читательский спрос. Нелишним является и факт асимметрии ошибок первого и второго рода, отмеченный автором: включение неинтересного издания в  $P$  – менее вредно, чем исключение из  $P$  ценного издания.

Автор полностью отвергает метрические критерии, вероятно, вследствие технических проблем, связанных с их учетом. Достаточно сказать, что международные измерители цитируемости, с их явным креном в сторону англоязычных публикаций и англоязычных научных сообществ, могут иметь лишь вспомогательное значение для России; между тем, русскоязычные измерители отсутствуют. По-видимому, именно ВИНТИ с его пулом журнальных и иных изданий мог бы инициировать подобный измеритель. Мне кажется, не существует никакого более объективного критерия значимости научной работы, чем цитируемость, конечно, с учетом значимости цитирований, как это делается при подсчетах спортивных рейтингов. Думается, что наличие системы русскоязычного рейтинга научных публи-

каций было бы чрезвычайно полезно для развития российской науки в целом. Возражения, включая то, что цитирование можно организовать, мне представляются надуманными. Никто же не отменяет систему вознаграждений в зависимости от результатов из-за возможных или реальных случаев мошенничества.

Если ограничиться, как это делает автор, лишь учетом экспертных оценок, то представляется, что повысить объективность экспертного отбора можно привязкой изданий к научным направлениям. Для этого, прежде всего, необходимо иметь, т.е. разработать или же использовать, имеющийся список, а лучше того – таксономию, дисциплин, к которым относятся издания – кандидаты на включение в множество  $P$ . Это позволит повысить объективность экспертного опроса за счет его усложнения. Для каждого издания эксперт должен будет указать: (а) к каким дисциплинам и насколько оно относится, так что структура ответа может быть примерно такой: «К дисциплине А – на 40%, к дисциплине Б – на 30% и к дисциплинам В и Г – по 15%»; (б) уровень, пусть в вербальной шкале, для каждой из дисциплин из пункта (а). Наш опыт показывает, что научные работники довольно легко формулируют процентные оценки принадлежности проектов или изданий к дисциплинам. Эти оценки принадлежности нетрудно использовать при усреднениях.

Я думаю, что квантизация вербальных шкал в манере, описанной в статье – разумное дело, потому что концепция порядковой шкалы неадекватно отражает деление ее континуума человеком, тем более, экспертом. Следует, однако, позаботиться об устойчивости результатов относительно изменения интервалов между значениями шкалы, что, на мой взгляд, нетрудно, ибо достигается чисто вычислительными средствами. Более того, «неустойчивые» издания, т.е. те, факт включения которых в  $P$  зависит от незначительных изменений в количественных значениях шкалы, могут образовывать фонд кандидатов на исключение из  $P$  при сокращении финансирования.

Следующий важный вопрос – критерии использования экспертных оценок для включения издания в реферируемое множе-

ство *P*. По-видимому, предлагаемый «трехчленный» метод не адекватен из-за слишком больших возможностей манипулирования величиной максимальной оценки. Ясно, что учет особенностей индивидуальных экспертов в стиле метода, предложенного в статье, должен получить отражение в финальной оценке. Другая важная характеристика – количество высших оценок, полученных изданием, позволила бы учесть массу голосов «за». Вероятно, финальная оценка должна отражать, как минимум, эти три характеристики: среднюю оценку, массу голосов «за» и индивидуальные особенности экспертов. Формулу связи их в окончательном критерии можно было бы выбрать после эксперимента – в зависимости от устойчивости получаемых результатов.

В заключение перечислим основные тезисы данного материала:

1. Рецензируемая статья не может рассматриваться как научная публикация.
2. Необходимо развивать русскоязычный индекс цитирования.
3. Привязка опроса экспертов к таксономии научных дисциплин могла бы повысить объективность экспертных оценок.
4. Следовало бы проверять устойчивость результатов в зависимости от способа квантизации используемой вербальной шкалы.
5. В интегральном критерии ранжирования следует учитывать, как минимум, следующие три компонента: среднюю оценку, массу оценки и особенности индивидуальных шкал.

### **Литература**

1. ФЕДОРЕЦ О. В. *Коллективная экспертиза научных журналов: методика агрегирования экспертных оценок и построения рейтинга* // Управление большими системами. – 2009. – Выпуск 27. – С. 12-35.

**ON THE “COLLECTIVE EXPERT EXAMINATION OF SCIENTIFIC JOURNALS: PROCEDURE OF EXPERT JUDGEMENTS AGGREGATION AND RATING CONSTRUCTION” ARTICLE BY O. FEDORETS**

**Boris G. Mirkin**, State University – Higher School of Economics, Moscow; Birkbeck College of London University, UK, Dr. Sc. ([bmirkin@hse.ru](mailto:bmirkin@hse.ru)).

*Abstract: It is shown that the method proposed in the paper “Collective expert examination of scientific journals: procedure of expert judgments aggregation and rating construction” needs more substantiation, although the fact of using a method for the problem is a welcome step. An alternative approach to rate Russian scientific journals is proposed: a taxonomy of the subject area should be developed so that the expert evaluations could proceed along the taxonomy branches.*

**Keywords:** critical notes, scientific journal rating, collective expert examination, citation index.