

УДК 168

DOI:

10.15372/PS20190403

В.Е. Осипов

ВОПРОСЫ ВОСПРОИЗВОДИМОСТИ

В статье дается обзор литературы по вопросу воспроизводимости результатов экспериментов постнеклассической науки. Основное внимание уделено обзору показателей воспроизводимости. Обсуждается понятие воспроизводимости в свете различных моделей величины эффекта, связь понятий «воспроизводимость величины эффекта» и «существование явления». Рассматриваются некоторые результаты проекта, осуществляемого под руководством Б. Носека. В качестве эмпирической апробации известного тезиса о невоспроизводимости результатов парапсихологических экспериментов выполняется метаанализ результатов экспериментов по дистанционному видению.

Ключевые слова: критерий воспроизводимости; показатели воспроизводимости; Брайан Носек; метаанализ; парапсихология; дистанционное видение

V.E. Osipov

ISSUES OF REPRODUCIBILITY

The article reviews publications on the problem of reproducibility of experimental results of post-nonclassical science. The main attention is paid to the review of reproducibility indicators. The notion of reproducibility is discussed in the light of different models of the effect size. The relationship of the concepts of “reproducibility of the effect size and the existence of phenomena is considered. Some results of the project led by Brian Nosek are examined. As an empirical testing of the well-known thesis of non-reproducibility of results of parapsychological experiments, we carry out meta-analysis of the results of experiments in remote viewing.

Keywords: reproducibility criterion; reproducibility indicators; Brian Nosek; meta-analysis; parapsychology; remote viewing

Общее суждение: «результаты парапсихических исследований и экспериментов невоспроизводимы повторно», – сформулированное Т.Г. Лешкевич [7, с. 133], вызывает следующий ряд вопросов.

- Все ли парапсихологические исследования не воспроизводятся или только часть из них?

- Что такое воспроизводимость и какова математическая формулировка критерия воспроизводимости?
- Как обстоят дела с воспроизводимостью в науке?

Целью настоящей статьи является попытка ответить на указанные выше и другие вопросы без претензии на полноту.

1. Обзор философской литературы

Рассмотрев более десятка доступных работ из области философии науки, в которых упоминается термин «воспроизводимость», мы не находим явного определения этого термина, но из контекста можно понять воспроизводимость как идентичность результатов оригинального и повторного (повторных) эмпирических исследований. Обсуждается специфика применения критерия воспроизводимости в классической и постнеклассической науке. Требование воспроизводимости и повторяемости, трактуемое как возможность повторения опыта в любом месте, в любое время и любым экспериментатором с одинаковым результатом, относят к разряду классической науки [16, с. 92–93].

Для социально-гуманитарных наук, относимых к разряду постнеклассической науки, характерным является ограниченный характер воспроизводимости [1; 2; 5, с. 435–436; 6; 12, с. 258], что В.В. Ильин называет деформацией критерия воспроизводимости. В частности, в работе [1] сообщается, что под руководством Б. Носека был реализован проект с участием нескольких групп исследователей, которые должны были с максимально возможной точностью повторить психологические эксперименты, описанные в статьях из очень авторитетных научных журналов за 2008 г. В итоге было обнаружено, что сходной статистической значимости удалось добиться лишь в 36 случаях из 100 и только четыре работы из 100 были воспроизведены полностью. По мнению авторов работы [1], требуется пересмотр перечня критериев научности с учетом специфики постнеклассической науки, в результате чего воспроизводимость если и будет включена в этот перечень, то не в качестве важнейшего критерия.

В связи с этим возникает два вопроса. Насколько пригоден к употреблению деформированный критерий? Если непригоден, то нельзя ли в таком случае заменить критерий воспроизводимости (в его классическом понимании) другим, близким по смыслу критерием? В работах [15, с. 52; 17, с. 73] предлагается критерий «тенденция к воспроизводимо-

сти», но, к сожалению, авторы не уточняют, каким образом следует выявлять эту тенденцию.

2. Элементы математического формализма.

Показатели воспроизводимости

Обратимся к работе Б. Носека с соавторами. Они пишут, что «не существует единого стандарта для оценки воспроизводимости» [22, р. аас4716-2], и перечисляют ряд показателей воспроизводимости, использованных в их проекте. Этот проект, в котором оценивалась воспроизводимость результатов исследований в области психологии, предполагал, в частности, следующее. Во-первых, выбрать ряд оригинальных исследований, опубликованных в трех солидных психологических журналах. Во-вторых, однократно повторить указанный ряд исследований. В-третьих, результаты оригинальных исследований сравнить с результатами повторных исследований. Из 488 статей 2008 г. командой было отобрано 111 статей для повторения исследований, но только 100 повторных исследований уложились в дедлайн для включения в сводный отчет. Таким образом, получилось 100 пар, состоящих из оригинального исследования и соответствующего ему повторного исследования.

Результаты отдельных исследований характеризуются несколькими параметрами: наблюдаемой величиной эффекта, значимостью (статистическая) величины эффекта (значима / незначима), уровнем значимости (статистической) величины эффекта (или p -значением) и др. И поскольку параметров несколько, то и показателей воспроизводимости было использовано также несколько. Оценка воспроизводимости осуществлялась с использованием пяти показателей (indicators): 1) значимости; 2) p -значений; 3) величин эффекта; 4) субъективных оценок команды экспериментаторов; 5) метаанализа величин эффекта. Значение каждого из этих показателей находилось одним или несколькими способами. Если эти «несколько способов» считать обязательными к применению, то названные пять показателей превращаются в пять групп показателей, а если эти способы считать альтернативными, то мы имеем пять показателей. Далее в настоящей работе мы будем термином «группы показателей» обозначать то, что в работе [22] обозначено термином «indicators», а термином «показатели» – различные способы (методы) оценки, объединяемые той или иной группой показателей.

Рассмотрим кратко первые три группы показателей.

Значимость. В данную группу был включен один показатель, который определялся следующим образом. Исходя из наличия двустороннего статистического критерия с уровнем значимости, равным 5%, результаты оригинальных и повторных исследований были классифицированы на значимые (при $P \leq 0,05$) и незначимые (при $P > 0,05$). Однако для четырех оригинальных исследований, в которых незначимый уровень $P < 0,06$ был интерпретирован как значимый, все соответствующие оригинальные и повторные исследования при таком же уровне значимости кодировались как значимые. Б. Носек пишет, что они проверяли гипотезу о том, что пропорции значимых результатов в оригинальных и повторных исследованиях равны, используя критерий Макнимара для номинативных парных данных. Понятно, что критерий Макнимара в данном случае есть проверка по критерию хи-квадрат отношения

$$\chi^2 = \frac{(|b - c| - 1)^2}{b + c},$$

где b – число исследований, значимых в оригинале и незначимых при повторении; c – число исследований, незначимых в оригинале и значимых при повторении.

Если проверка по критерию хи-квадрат даст незначимый результат, то принимается нулевая гипотеза H_0 , гласящая об отсутствии связи между повторением экспериментов и долей значимых результатов, и этой гипотезе будет соответствовать наличие воспроизводимости по данному показателю. В противном случае данный показатель принимает значение «отсутствие воспроизводимости».

P-значения. В эту группу было включено два показателя, при вычислении которых сравнивались центральные тенденции (central tendency) распределений значений P в оригинальных и повторных исследованиях. В первом показателе был использован T -критерий Вилкоксона (Wilcoxon signed-rank test), а во втором показателе – t -критерий (критерий Стьюдента) для связанных выборок. Понятно, что в обоих критериях для всех пар вычислялась разность значений P , составляющих пару. Использовались только те пары исследований, для которых были известны оба значения P .

Понятно также, что в случае незначимых результатов ($T_{\text{эмп}} \geq T_{\text{кр}(0,05)}$ в первом критерии и $T_{\text{эмп}} < T_{\text{кр}(0,05)}$ во втором критерии) соответствующие показатели принимают значение «наличие воспроизводимости». В случае

значимых результатов проверки по указанным критериям соответствующие показатели принимают значение «отсутствие воспроизводимости».

Величины эффекта. Б. Носек пишет о том, что с целью унификации, удобной для дальнейшего анализа, они преобразовали величины эффекта из других форм (z , F , t , χ^2) в коэффициент корреляции. При этом коэффициенты корреляции, соответствующие оригинальным исследованиям, кодировались как положительные, а у повторных исследований, где величина эффекта была противоположна величине эффекта оригинального исследования, коэффициент корреляции кодировался как отрицательный.

Для сравнения величин эффекта служили четыре показателя. Первые два сравнивали центральные тенденции распределения величин эффекта в выборке оригинальных исследований и в выборке повторных исследований. Первый показатель использует t -критерий для связанных выборок, а второй – T -критерий Вилкоксона. В третьем показателе вычислялась доля исследуемых пар, у которых оригинальная величина эффекта была больше, чем воспроизведенная; при этом проверялась гипотеза о том, что для генеральной совокупности пар эта доля равна 0,5. В четвертом показателе вычислялось «покрытие» (coverage), или доля пар исследований, в которых точечная оценка оригинальной величины эффекта попадает в доверительный интервал повторного исследования. Известно, например, что для модели с фиксированным эффектом (модель рассмотрена ниже) при 95%-м доверительном интервале и при равном объеме выборок математическое ожидание доли попаданий составляет 83,4% [19, p. 13], что нетрудно показать. Проверка на соответствие указанной пропорции осуществлялась с помощью критерия хи-квадрат. Очевидно, что если критерий хи-квадрат не покажет различий, то это следует интерпретировать как адекватность модели с фиксированным эффектом, равенство истинных значений величины эффекта в обеих выборках (оригинальных и повторных исследований), что означает наличие воспроизводимости истинного значения величины эффекта.

Итак, в работе [22] мы находим несколько групп показателей, по которым оценивается воспроизводимость, включая такую группу, как «субъективные оценки команды экспериментаторов». При этом каждая из групп объективных показателей содержит один или несколько показателей. Таким образом, воспроизводимость предстает здесь в виде критерия, включающего ряд показателей. Среди этих показателей не выделены определяющие (участвующие в принятии решения о качественном оценивании воспроизводимости: «воспроизводимость имеется»/ «вос-

производимость отсутствует»). В работе [22] приведен ряд показателей воспроизводимости, но не дана *общая математическая формулировка критерия воспроизводимости* (например, по типу критерия пригодности в квалиметрии [8, с. 322]), которая позволяет дать интегральную оценку воспроизводимости. И если к тому же учесть, что «не существует единого стандарта для оценки воспроизводимости» [22, р. аас4716-2], то в итоге мы с точки зрения математического формализма имеем в данной области фактически произвол.

3. Метаанализ величин эффекта

Рассмотрим следующие вопросы: 1) какие характеристики или виды величин эффекта существуют; 2) как понятие «воспроизводимость величины эффекта» связано с понятиями «гетерогенность» и «гомогенность»; 3) какие существуют показатели гетерогенности; 4) как воспроизводимость величины эффекта соотносится с существованием исследуемого явления.

Каждое исследование, охваченное метаанализом, характеризуется, в частности, следующими величинами:

- i – порядковый номер исследования¹. Назначается произвольно;
- θ_i – математическое ожидание величины эффекта, которое М. Боренштейн называет истинной величиной эффекта или генеральным средним. Генеральный параметр; значение неизвестно;
- Y_i – наблюдаемая величина эффекта. Выборочный параметр; случайная величина; значение известно; является оценкой величины θ_i ;
- V_i – оценка дисперсии наблюдаемой величины эффекта. Выборочный параметр; значение известно.

В общем случае полагают [24, р. 3], что

$$Y_i = \theta_i + \varepsilon_i,$$

$$\varepsilon_i \in N(0; v_i),$$

¹ В метаанализе эти исследования называют первичными (primary studies), имея в виду, что метаанализ является вторичным исследованием, а также называют индивидуальными исследованиями (individual studies).

где ε_i – случайная ошибка величины эффекта в исследовании i (т.е. случайное отклонение Y_i от θ_i); $N(\mu, \sigma^2)$ – нормальное распределение с математическим ожиданием μ и дисперсией σ^2 ; ν_i – (истинная) дисперсия наблюдаемой величины эффекта в i -м исследовании.

В зависимости от ограничений, накладываемых на величины θ_i , различают три модели величины эффекта (или три класса моделей): модель с фиксированным эффектом, модель со случайными эффектами и модель со смешанными эффектами.

Модель с фиксированным эффектом предполагает, что все генеральные средние θ_i равны некоторому фиксированному значению θ (константа):

$$\theta_1 = \theta_2 = \dots = \theta_k = \theta,$$

где k – количество первичных исследований, рассматриваемых в данном метааналитическом исследовании. Тогда [18, р. 64–65; 21, р. 4]

$$Y_i = \theta + \varepsilon_i,$$

где θ – истинное значение величины эффекта.

Если трактовать указанную выше выборку из k исследований как выборку, содержащую одно оригинальное исследование и остальные ($k-1$), повторяющие его, то адекватность модели с фиксированным эффектом будет означать воспроизводимость истинной величины эффекта θ во всех повторных исследованиях из рассматриваемой генеральной совокупности исследований.

Модель со случайными эффектами предполагает, что каждое из рассматриваемых исследований имеет свое специфическое истинное значение величины эффекта, т.е. свое генеральное среднее θ_i . При этом полагают [18, р. 70–72], что

$$\theta_i \in N\left(\mu; \tau^2\right),$$

где μ – математическое ожидание истинной величины эффекта (общее среднее, overall mean, grand mean); τ^2 – дисперсия истинной величины эффекта. Тогда

$$Y_i = \theta_i + \varepsilon_i = \mu + \zeta_i + \varepsilon_i,$$

$$\zeta_i \in N\left(0; \tau^2\right),$$

где ζ_i – случайное отклонение θ_i от μ .

Адекватность модели со случайными эффектами будет означать, что истинные величины эффекта θ_i изменяются от исследования к исследованию. Иначе говоря, они не воспроизводятся.

Модель со смешанными эффектами (одна из ряда возможных) предполагает, что

$$\theta_i = \mu + \beta_1 x_{i1} + \dots + \beta_j x_{ij} + \zeta_i,$$

где x_{ij} – переменная-модератор j -го фактора в исследовании i ; β_j – коэффициент для j -го фактора.

Очевидно, что модель со смешанными эффектами есть модель со случайными эффектами, дополненная слагаемыми, описывающими влияние различных факторов на величину истинного эффекта. Данная модель также не предполагает воспроизводимость.

Наиболее часто используются две первые модели величины эффекта, и выбор между этими моделями осуществляют путем проверки на гетерогенность с помощью Q -теста. Для этого вычисляют

$$Q = \sum_{i=1}^k W_i Y_i^2 - \frac{\left(\sum_{i=1}^k W_i Y_i \right)^2}{\sum_{i=1}^k W_i}, \quad (1)$$

$$W_i = \frac{1}{V_i}. \quad (2)$$

Для нулевой гипотезы статистика Q приблизительно подчиняется распределению χ^2 с $df = k - 1$ степенями свободы [20, р. 181].

Если проверка по критерию хи-квадрат даст незначимый результат ($p > 0,05$ или $p > 0,10$ [18, р. 112]), то считают, что исследования имеют одно и то же истинное значение величины эффекта θ , т.е. *истинные* величины эффекта являются гомогенными, и в этом случае применима модель с фиксированным эффектом². Тогда оценкой величины θ является взвешенное среднее

² В этом месте ряд авторов делают предостережение, что незначимый результат Q -теста не является строгим доказательством нулевой дисперсии истинной величины эффекта. М. Боренштейн с соавторами объясняют, что незначимый результат может быть не толь-

$$M = \frac{\sum_{i=1}^k W_i Y_i}{\sum_{i=1}^k W_i}. \quad (3)$$

Если же результат проверки по критерию хи-квадрат окажется значимым, то считают, что истинные величины эффекта являются гетерогенными и применима модель со случайными эффектами. В этом случае величину t^2 часто оценивают по следующей формуле, предложенной Р. Дерсимонян и Н. Лэрд [20, р. 183]:

$$T^2 = \max \left\{ 0, \frac{Q - (k - 1)}{\sum_i W_i - \frac{\sum_i W_i^2}{\sum_i W_i}} \right\}, \quad (4)$$

а оценкой величины μ является следующее взвешенное среднее:

$$M^* = \frac{\sum_{i=1}^k W_i^* Y_i}{\sum_{i=1}^k W_i^*}, \quad (5)$$

ко за счет малого значения дисперсии истинной величины эффекта, но еще и за счет малой мощности Q -теста в случае малого числа исследований и/или большой дисперсии величины эффекта, наблюдаемой в первичных исследованиях [18, р. 113, 273]. М. Маркер также отмечает, что нужно с осторожностью интерпретировать результат проверки по критерию хи-квадрат ввиду его малой мощности при малом размере выборки и малом числе исследований. «Это означает, – пишет он, – что статистически значимый результат может указывать на проблемы с гетерогенностью, а незначимый результат не может приниматься как доказательство отсутствия гетерогенности» [21, р. 6]. Вместе с тем в указанных предостережениях мы не видим алгоритмической части, что, как представляется, ведет к их игнорированию во многих ситуациях на практике.

$$W_i^* = \frac{1}{V_i + T^2}.$$

Статистика Q и уровень значимости p (для Q) являются двумя показателями гетерогенности, служащими для проверки нулевой гипотезы (о том, что величины эффекта являются гомогенными).

Другие два показателя – T^2 и $T = \sqrt{T^2}$ оценивают соответственно такие параметры гетерогенности, как дисперсия истинной величины эффекта τ^2 и стандартное отклонение истинной величины эффекта $\tau = \sqrt{\tau^2}$.

Пятым и последним показателем гетерогенности является следующая дескриптивная статистика:

$$I^1 = \left(\frac{Q - df}{Q} \right) \times 100\%, \quad (6)$$

которая показывает, какой процент от общей дисперсии составляет дисперсия истинной величины эффекта. Ориентиры для интерпретации таковы: 0% – гетерогенность отсутствует, 25% – низкая гетерогенность, 50% – средняя гетерогенность, 75% – высокая гетерогенность [3, с. 157]. Если $Q < df$, то по формуле (6) мы получим отрицательную долю, что лишено смысла. В этом случае принимают $I^2 = 0$, иными словами, расчет ведут по следующей формуле:

$$I^2 = \max \left\{ 0, \left(\frac{Q - df}{Q} \right) \times 100\% \right\}. \quad (7)$$

Учитывая рассмотренные выше модели величины эффекта, делаем следующие выводы относительно *воспроизводимости величины эффекта*.

1. Об идентичности *наблюдаемых* значений величины эффекта говорить не приходится в силу того, что в них присутствуют случайные составляющие.

2. Идентичность *истинных* значений величины эффекта вполне возможна (в модели с фиксированным эффектом) и проверяется с помощью Q -теста с возможным привлечением других показателей.

3. В ситуации, описываемой моделью со случайными эффектами, не воспроизводятся ни наблюдаемая, ни истинная величины эффекта.

4. Если истинная величина эффекта не воспроизводится (не фиксирована), то это не значит, что она равна нулю, что исследуемое явление не существует. Эффект может существовать, но быть изменчивым и описываться такими параметрами, как математическое ожидание и дисперсия истинной величины эффекта. Как известно, отрицание случайных явлений характерно для механистической картины мира [13, с. 774]. Здесь имеют место отрицание случайных эффектов и неприятие вытекающего из этого отсутствия воспроизводимости.

5. Несмотря на возможно малую мощность Q -теста, альтернативы ему не видно, поскольку статистика Q по определению является индуктивной, а статистики F^2 и T^2 – дескриптивными, а потому вспомогательными для цели выявления различий между распределениями.

4. Некоторые результаты исследований по проекту под руководством Б. Носека

Из-за ограниченного объема статьи проанализируем результаты проекта, представленного в работе [22], лишь по трем показателям воспроизводимости, рассмотренным во втором разделе настоящей работы.

«*Покрытие*». В кратком изложении текста статьи сказано, что 47% оригинальных значений величин эффекта попали в 95%-й доверительный интервал величин эффекта повторных исследований [22, р. 943]. В полном тексте статьи поясняется [22, р. аас4716-4], что множество пар исследований, в которых могла быть вычислена стандартная ошибка, в зависимости от типа статистик было разбито на два подмножества. В первом подмножестве оказалось 73 пары исследований, в тридцати из которых (41,1%) было попадание в доверительный интервал. Во втором подмножестве оказалось пары 22 исследований, в пятнадцати из которых (68,2%) было попадание. Итоговый процент «покрытия» был вычислен как среднее взвешенное:

$$\frac{73 \cdot 41,1\% + 22 \cdot 68,2\%}{73 + 22} = 47,4\%.$$

Теперь, исходя из того, что общее число участвующих пар исследований равно $73 + 22 = 95$, и $95 \cdot 0,474 = 45,03$ есть эмпирическая частота попадания в доверительный интервал, а $95 \times 0,834 = 79,23$ есть теоретическая частота попадания в интервал, составляем таблицу сопряженности (табл. 1).

Таблица 1

Частоты попадания в 95%-й доверительный интервал
и вне его по показателю «покрытие»

Ряд	В интервале	Вне интервала	Всего
Эмпирические частоты	45,03	49,97	95
Теоретические частоты	79,23	15,77	95

Проверка по критерию хи-квадрат с помощью функции ХИ2.ТЕСТ(), встроенной в табличный процессор Excel, дает уровень значимости результата $p = 4 \cdot 10^{-21}$. Проверка «вручную» с различными поправками на непрерывность дает величину другого порядка. Однако в любом случае уровень значимости результата $p < 0,001$, что говорит о том, что различие теоретического и эмпирического распределений в высшей степени достоверно, откуда вытекает, что воспроизводимость по данному показателю отсутствует.

Значимость. Исследователи также сообщают, что проверка по критерию Макнимара дает $\chi^2(1) = 59,1$ [22, р. аас4716-4], что соответствует уровню значимости $p = 1,5 \cdot 10^{-14} < 0,001$. То есть воспроизводимость по данному показателю также отсутствует.

Величины эффекта. Подсчитывается количество пар, у которых оригинальная величина эффекта больше, чем воспроизведенная. Если истинные величины эффекта в оригинальных и повторных исследованиях равны, то благодаря случайному отклонению наблюдаемая величина эффекта оригинального исследования будет превышать наблюдаемую величину эффекта повторного исследования с вероятностью 0,5. Из 99 пар исследований, для которых были известны величины эффекта и оригинального, и повторного исследования, 82 пары (82,8%) показали большее значение величины эффекта в оригинальном исследовании [22, р. аас4716-5]. Используя двусторонний биномиальный критерий и табличный процессор Excel, записываем следующую функцию,

вычисляющую уровень значимости результата: $2 \cdot (1 - \text{БИНОМ.РАСП}(82 - 1; 99; 0,5; \text{ИСТИНА}))$, которая возвращает значение $p = 2,2 \cdot 10^{-11} < 0,001$. Отличие эмпирического распределения от теоретического (50%; 50%) является достоверным. Следовательно, воспроизводимость по данному показателю отсутствует.

5. Примеры метаанализа результатов экспериментов по дистанционному видению

Почему в настоящем разделе использован метаанализ? На то есть две причины, одна из которых состоит в следующем. Первые три группы показателей, подробно рассмотренные в разделе 2, подходят для ситуации, когда имеются две выборки равного объема: одна представляет оригинальные исследования, а другая – повторные исследования. Однако эти группы показателей, как видно, не подходят в ситуации, когда у нас имеется несколько исследований, одно из которых можно условно считать оригинальным, а остальные – повторяющими его. Именно такой случай и рассматривается в настоящем разделе.

Двусторонний или односторонний критерий? Как известно, односторонние критерии рекомендуется выбирать в случае, когда у исследователя имеются веские основания полагать, что отклонения возможны лишь в одну сторону, или если нас интересуют лишь отклонения в одну сторону, а отклонения в другую сторону не имеют практического значения [14, с. 17]. Имеют ли отрицательные отклонения величины эффекта какое-либо значение для выявления дистанционного видения и возможны ли они? В работе [10, с. 112] показано, что отклонения в обе стороны имеют значение в контексте методологии времен Д.Б. Райна, когда рассматривался ряд серий испытаний, не связанных одна с другой. При этом показательным было даже сделанное шутки ради отрицательное отклонение в серии, если только оно значимо. Однако в контексте методологии, рассмотренной в работе [9], целесообразность учета отрицательных отклонений является весьма спорной. Во-первых, пока неизвестно, можно ли получить значимое отрицательное отклонение, если перципиент будет намеренно измышлять непохожие образы, которые впоследствии будут обработаны методом слепого судейства. Во-вторых, отрицательные отклонения величины эффекта, имеющиеся в малом числе серий, будут лишь уменьшать положительное значение взвешенного среднего (3), (5), увеличивая тем самым вероятность принятия нулевой гипотезы

($\mu = 0$, или $\theta = 0$). Таким образом, возможные эпизодические попытки произвести отрицательную величину эффекта не дадут отрицательного синтетического эффекта. Поэтому нет смысла учитывать отрицательное направление введением двустороннего критерия. В-третьих, односторонний критерий, как известно [4, с. 122], мощнее двустороннего.

Исходя из приведенных выше аргументов, для дальнейшего анализа мы выбираем односторонний статистический критерий. Нулевая гипотеза H_0 будет гласить о том, что оцениваемая величина эффекта меньше либо равна нулю, а альтернативная гипотеза H_1 будет гласить о том, что оцениваемая величина эффекта больше нуля.

Метаанализ. Под эгидой Международной корпорации по внедрению научных достижений (SAIC) проведено 10 экспериментов в рамках программы по исследованию дистанционного видения. В шести из десяти указанных экспериментов (в экспериментах с номерами 1, 4, 5, 6, 9, 10) изучалось дистанционное видение. В работе [23] приведены данные по сериям испытаний в экспериментах 1, 9, 10, которых достаточно для проверки правильности статистического синтеза, осуществленного исследователями. Не имея возможности представить в настоящей статье подробный анализ, отметим только, что итоговые цифры по экспериментам 1, 9, 10 [23, р. 15] совпадают с результатами нашего синтеза.

Пример 1. В качестве первого примера выполним метаанализ результатов трех серий испытаний, составляющих эксперимент № 9. В таблице 2 тоном выделены ячейки, содержимое которых взято из работы [23, р. 27]. Как видно из таблицы, в эксперименте сделано три выборки: серия из 18 испытаний с перципиентом № 009, серия из 24 испытаний с перципиентом № 372 и серия из 28 испытаний с перципиентом № 389.

Таблица 2

Эксперимент № 9, проведенный под эгидой SAIC

Наблюдатель	N	Y	V	W	WY	WY^2
009	18	0,432	0,05556	18	7,776	3,359232
372	24	0,354	0,04167	24	8,496	3,007584
389	28	0,177	0,03571	28	4,956	0,877212
Итого	70			70	21,228	7,244028

Величины эффекта в сериях представлены в столбце Y . Дисперсия V в методе слепого судейства, как показано в работе [9, с. 91], вычисляется по формуле

$$V = \frac{1}{N},$$

где V – дисперсия наблюдаемой величины эффекта в исследовании; N – объем выборки (количество испытаний) в исследовании.

Веса исследований (W) в табл. 2 вычислены по формуле (2).

Проверка на гетерогенность. Подставляя в (1) итоговые результаты из табл. 2, получаем $Q=0,806485$. С учетом того, что $k=3$ (три серии испытаний), находим, что эмпирическое значение критерия $\chi^2 = 0,806$ при числе степеней свободы $df=2$ имеет уровень значимости $p_Q = 0,668$, который на порядок больше уровня значимости критерия 0,05 и в шесть раз больше уровня 0,1. Это означает «жизнеспособность нулевой гипотезы (о том, что дисперсия истинной величины эффекта в точности равна нулю)» (the viability of the null hypothesis (Is the true dispersion exactly zero)) [18, p. 121], поэтому мы будем считать, что величины эффекта в сериях испытаний являются гомогенными, и выбираем модель с фиксированным эффектом. Дополнительные вычисления по формулам (4) и (7) дают значения параметров $T^2 = 0$ и $F^2 = 0$, соответствующие модели с фиксированным эффектом и отсутствию гетерогенности. Проверка на гетерогенность завершена.

Выбор модели с фиксированным эффектом означает, что в трех сериях испытаний в рамках эксперимента № 9 получена воспроизводимость истинной величины эффекта.

По формуле (3) находим $M = 0,303257$ (оценка истинной величины эффекта θ). Вычисления по другим формулам, приведенным в [18, p. 66], дают

- оценку дисперсии суммарного эффекта $V_M = 0,014286$;
- оценку стандартного отклонения суммарного эффекта $SE_M = 0,119523$;
- верхнюю и нижнюю границы 95%-ного доверительного интервала для суммарного эффекта: $UL_M = 0,537522$; $LL_M = 0,068992$.

Мы видим, что нулевая отметка не попадает в доверительный интервал суммарного эффекта, а это означает, что истинное значение вели-

чины эффекта (θ) отлично от нуля. И поскольку доверительный интервал лежит в положительной части оси, постольку мы отбрасываем нулевую гипотезу и принимаем альтернативную гипотезу $H_i: \theta > 0$.

Дополнительно по формулам из [18, р. 66] находим

- z-оценку суммарной величины эффекта $Z = 2,537231$;
- уровень значимости суммарной величины эффекта $p = 0,005587$.

Отсюда также видно, что полученная суммарная величина эффекта $M = 0,303$ является значимой ($p = 0,006 \leq 0,01$).

Полученные нами результаты $M = 0,303$ и $p = 0,006$ совпадают с результатами, приведенными в итоговой таблице в [23, р. 15].

Отсюда мы делаем следующие три вывода по данному эксперименту:

- 1) статистический синтез, осуществленный исследователями дистанционного видения правилен;
- 2) имеет место воспроизводимость истинной величины эффекта в трех сериях испытаний;
- 3) достоверно (при уровне значимости результата $p = 0,006$) принимается альтернативная гипотеза H_1 , гласящая о том, что оцениваемая величина эффекта больше нуля. И до тех пор, пока не доказано наличие сенсорных каналов передачи информации, мы вправе интерпретировать результат эксперимента как наличие дистанционного видения [11].

Аналогичные проверочные расчеты по экспериментам № 1 и 10 дают следующее: 1) величины эффекта в сериях являются гомогенными (имеется воспроизводимость величины эффекта); 2) значения суммарной величины эффекта и уровни значимости, полученные нами, совпадают с теми, которые приведены экспериментаторами (исходя из этого полагаем, что статистическая обработка была выполнена экспериментаторами правильно); 3) для эксперимента № 10 принимается альтернативная гипотеза $H_i: \theta > 0$ при уровне значимости результата $p = 9 \cdot 10^{-8}$.

Пример 2. Полагая, что в остальных трех экспериментах (с номерами 4, 5, 6) статистический синтез был осуществлен исследователями также правильно, проведем метаанализ результатов всех экспериментов по дистанционному видению, проведенных под эгидой SAIC.

В качестве исходных данных берем данные, представленные в трех первых столбцах табл. 3 (выделены тоном).

Таблица 3

Эксперименты по дистанционному видению,
проведенные под эгидой SAIC

Экспе- римент	N	Y	V	W	WY	WY^2	W^2	W^*	YW^*
1	200	0,124	0,005	200	24,8	3,075	40000	24,91	3,08852
4	40	-0,067	0,025	40	-2,68	0,18	1600	16,63	-1,1139
5	24	0,088	0,0417	24	2,112	0,186	576	13,02	1,1456
6	21	0,368	0,0476	21	7,728	2,844	441	12,08	4,44618
9	70	0,303	0,0143	70	21,21	6,427	4900	20,23	6,12933
10	90	0,55	0,0111	90	49,5	27,23	8100	21,62	11,8894
Итого				445	102,67	39,94	55617	108,5	25,5851

Проверка на гетерогенность. По указанным выше формулам находим $Q = 16,248$; $df = 5$; $p_Q = 0,006 \leq 0,01$. В соответствии с этими двумя показателями (Q, p_Q) имеется гетерогенность. Кроме того, по формулам (4), (7) и другим формулам из [18, р. 122–125; 3, с. 158–159] рассчитываем другие показатели гетерогенности:

- $T^2 = 0,035$ (оценку t^2);
- нижнюю и верхнюю границы доверительного интервала для дисперсии истинной величины эффекта: $LL_T^2 = 0,006$; $UL_T^2 = 0,104$;
- $T = 0,1875$ (оценку t);
- нижнюю и верхнюю границы доверительного интервала для стандартного отклонения истинной величины эффекта: $LL_T = 0,0772$; $UL_T = 0,3222$;
- $F^2 = 69\%$;
- нижнюю и верхнюю границы доверительного интервала: $LL_I^2 = 28\%$; $UL_I^2 = 87\%$.

Поскольку нижние границы доверительных интервалов превышают ноль, постольку считаем, что T^2 и F^2 являются значимыми. Дисперсия истинной величины эффекта занимает существенную долю в общей дисперсии – в диапазоне приблизительно от 25% (низкая гетерогенность) до 75% и выше (высокая гетерогенность). Данные показатели также под-

тверждают гипотезу о том, что величины эффекта являются гетерогенными. *Величина эффекта не воспроизводится.*

Поэтому далее ведем расчет для модели со случайными эффектами по формуле (5) и другим формулам из [18, р. 65, 72–74]:

- взвешенного среднего (оценки величины μ) $M^* = 0,2359$;
- оценки дисперсии суммарного эффекта $V_M = 0,0092$;
- оценки стандартного отклонения суммарного эффекта $SE_M^* = 0,096$;
- нижней и верхней границ 95%-го доверительного интервала для суммарного эффекта: $LL_M = 0,477$; $UL_M^* = 0,424$.

Очевидно, что 95%-й доверительный интервал суммарного эффекта не покрывает нулевую отметку. Следовательно, истинное значение величины эффекта (μ) нельзя считать нулевым, а положительное значение взвешенного среднего при этом говорит об успешном результате. Принимаем альтернативную гипотезу $H_i: \mu > 0$.

Дополнительно по формулам из [18, р. 74] вычисляем

$$Z^* = 2,4565; p^* = 0,007 \leq 0,01.$$

Отсюда также видно, что результат является значимым. До тех пор, пока не доказано наличие сенсорных каналов передачи информации, мы вправе интерпретировать данный результат как наличие дистанционного видения.

Сводка результатов метаанализа приведена в табл. 4. Результаты экспериментов № 1 и № 10 нуждаются в пояснении. Например, в эксперименте № 10 величина $p(Q = 10,49; df=7)$ – уровень значимости результата Q -теста с семью степенями свободы – «всего» в 1,6 раза превышает уровень значимости критерия 0,1. Поэтому хотя формально Q -тест и показал гомогенность, но с учетом предостережения, приведенного выше, все же рассмотрим и другие два параметра гетерогенности: T^2 и I^2 . У этих двух параметров нижние границы доверительных интервалов упираются в нулевую отметку (расчетные значения отрицательные), что говорит о незначимости параметров T^2 и I^2 . Любой из этих двух параметров значим только в том случае, если выполняются два условия: 1) параметр больше нуля, 2) нижняя граница доверительного интервала превышает ноль (“If the lower limit exceeds zero” [18, р. 124, 125]). Таким обра-

Таблица 4

Результаты метаанализа результатов экспериментов
по дистанционному видению, проведенных под эгидой SAIC

Показатель	Значение	Границы 95%-го доверительного интервала	
		Нижняя	Верхняя
<i>Эксперимент № 1 (расчет не приведен)</i>			
Q	10,67		
Уровень значимости $p(Q = 10, 67; df = 9)$	0,30		
T_2	0,009	0	0,066
I_2	16	0	57
M	0,124	-0,015	0,262
V_M	0,005		
Уровень значимости суммарной величины эффекта (одностор.) p	0,040		
<i>Эксперимент № 9 (пример 1)</i>			
Q	0,806		
$p(Q = 0, 806; df = 2)$	0,668		
T_2	0		
I_2	0		
M	0,303	0,069	0,538
V_M	0,014		
P	0,006		
<i>Эксперимент № 10 (расчет не приведен)</i>			
Q	10,49		
$p(Q = 10, 49; df = 7)$	0,162		
T^2	0,045	0	0,213
I^2	33	0	70
M	0,55	0,34	0,76
V_M	0,011		
p	$9 \cdot 10^{-8}$		
<i>Все эксперименты по дистанционному видению (пример 2)</i>			
Q	16,25		
$p(Q = 16, 25; df = 5)$	0,006		
T^2	0,035	0,006	0,104
I^2	69	28	87
M	0,236	0,048	0,424
V_M^*	0,0092		
p^*	0,007		

зом, учитывая незначимость трех параметров гетерогенности: $p(Q = 10,49; df=7)$, T^2 и F^2 , – принимаем гипотезу о том, что во всех сериях испытаний, относящихся к эксперименту № 10, величины эффекта являются гомогенными. Иначе говоря, наблюдается воспроизводимость истинной величины эффекта в сериях эксперимента № 10. То же можно сказать и об эксперименте № 1.

Выводы

1. Не существует стандартного инструмента для выявления воспроизводимости. На практике в психологии апробирован ряд показателей воспроизводимости, но еще не выработана какая-либо математическая формулировка (универсальная или неуниверсальная) интегрального критерия воспроизводимости, охватывающая все или определяющие показатели и дающая качественный ответ о наличии либо отсутствии воспроизводимости.

2. Если судить по отдельным показателям, то результаты апробации, упомянутой в п. 1, подтверждают положение философии науки, гласящее об ограниченном характере воспроизводимости в социально-гуманитарных науках, относимых к постнеклассической науке.

3. Ввиду существования нескольких моделей величины эффекта можно выделить три понимания требования *воспроизводимости величины эффекта*.

В первом понимании требуется идентичность (или воспроизводимость) значений *наблюдаемой* величины эффекта. Требование в таком виде эффективно лишь для класса явлений, у которых стохастическая составляющая пренебрежимо мала по сравнению с детерминированной. При этом можно считать, что и дисперсия истинной величины эффекта, и дисперсия наблюдаемой величины эффекта равны нулю.

Второе понимание требует воспроизводимости *истинной* величины эффекта. Наличие этой воспроизводимости проверяется с помощью Q -теста. Данное понимание применимо к классу явлений, отвечающему модели с фиксированным эффектом, в которой дисперсия истинной величины эффекта равна нулю, а дисперсия наблюдаемой величины эффекта больше нуля.

Следующий класс явлений отвечает модели со случайными эффектами, в которой отличны от нуля обе дисперсии: и дисперсия истинной величины эффекта, и дисперсия наблюдаемой величины эффекта. При

этом уже ничто не может быть воспроизведено в указанном выше понимании: ни наблюдаемая величина эффекта, ни истинная величина эффекта. В этом случае если и можно говорить о воспроизводимости, то в смысле ограниченного рассеивания значений величины эффекта, которое можно проверить, например, по следующей формуле: $T^2 < T_{ck}^2$, где T_{ck}^2 – критическое значение оценки дисперсии истинной величины эффекта. Это можно определить как третье понимание воспроизводимости. Как представляется, модель со случайными эффектами типична для постнеклассической науки.

Установка на воспроизводимость не должна противоречить интенции к истине. Если эффект не воспроизводится, из этого не следует, что эффект необходимо признать нулевым и что исследуемое явление не существует.

4. Метаанализ результатов экспериментов по дистанционному видению выявил ряд исследований, где наблюдается воспроизводимость истинной величины эффекта при ее ненулевом значении. Увеличение числа экспериментов, включаемых в метаанализ, демонстрирует ситуацию, в которой отсутствует воспроизводимость истинной величины эффекта, а для выяснения того, имеется ли воспроизводимость хотя бы в третьем понимании, требуется численное значение величины T_{ck}^2 . В то же время мы получаем точечные и интервальные оценки следующих показателей: общего среднего величины эффекта (математическое ожидание истинной величины эффекта), дисперсии истинной величины эффекта, стандартного отклонения истинной величины эффекта. Интервальная оценка и полученный уровень значимости свидетельствуют о ненулевой суммарной величине эффекта. До тех пор, пока не доказано наличие сенсорных каналов передачи информации, мы вправе интерпретировать данные результаты как наличие дистанционного видения.

Литература

1. Вачков И.В., Вачкова С.Н. Воспроизводимость психологических экспериментов как проблема постнеклассической науки // Культурно-историческая психология. – 2016. – Т. 12, № 1. – С. 97–101. DOI: 10.17759/chp.2016120110.

2. Губанов Н.И., Губанов Н.Н., Волков А.Э. Особенности критериев научности в социально-гуманитарных науках // Гуманитарное образование в креативно-антропологическом измерении: Сб. науч. ст. 12-й Всерос. науч.-практ. конф., 19–20 нояб. 2015 г., Екатеринбург / Науч. ред. С.З. Гончаров; Рос. гос. проф.-пед. ун-т. – Екатеринбург: ИД «Ажур», 2015. – С. 15–23.

3. Дембицкий С. Оценка гетерогенности в метаанализе (на примере данных кросс-национальных исследований) // Социология: теория, методы, маркетинг. – 2013. – № 1. – С. 154–165.
4. Закс Л. Статистическое оценивание. – М.: Статистика, 1976. – 598 с.
5. Ильин В.В. Философия: Учебник: В 2 т. – Ростов-н/Д.: Феникс, 2006. – Т. 1. – 832 с.
6. Лекторский В.А. Научное и вненаучное мышление: скользящая граница // Научные и вненаучные формы мышления: Симпозиум (Москва, 4–9 апреля 1995 г.). Москва – Киль, 1996. – URL: <http://philosophy.ru/iphras/library/ruspaper/LEKTORS1.htm> (дата обращения: 29.05.2014).
7. Лейкевич Т.Г. Философия: Курс лекций. – М.: ИНФРА-М, 2000. – 240 с.
8. Метрология и электрорадиоизмерения в телекоммуникационных системах: Учеб. пособие / Под общ. ред. Б.Н. Тихонова. – М.: Горячая линия-Телеком, 2007. – 374 с.
9. Осипов В.Е. Дистанционное видение в свете двух признаков паранауки // Философия науки. – 2017. – № 3 (74). – С. 81–105.
10. Осипов В.Е. К вопросу о статистическом анализе условий и результатов экспериментов Д.Б. Райна // Философия науки. – 2016. – № 4 (71). – С. 91–114.
11. Осипов В.Е. Эмпирические основания дискуссии об экстрасенсорном восприятии // Эпистемология и философия науки. – 2014. – № 4 (42). – С. 171–191.
12. Рузавин Г.И. Методология научного познания: Учеб. пособие для вузов. – М.: ЮНИТИ-ДАНА, 2012. – 287 с.
13. Рузавин Г.И. Случайность // Философия: Энциклопедический словарь / Под ред. А.А. Ивина. – М.: Гардарики, 2004. – С. 774–775.
14. Смагунова А.Н., Козлов В.А. Примеры применения математической теории эксперимента в рентгенофлуоресцентном анализе. – Иркутск: Изд-во Иркут. ун-та, 1990. – 232 с.
15. Холтон Дж. Что такое «антинаука»? // Вопросы философии. – 1992. – № 2. – С. 26–58.
16. Черникова И.В. Природа науки и критерии научности // Гуманитарный вектор. – 2012. – № 3 (31). – С. 89–96.
17. Штанько В.И. Философия и методология науки: Учеб. пособие для аспирантов и магистрантов естественнонаучных и технических вузов. – Харьков: ХНУРЭ, 2002. – 292 с.
18. Borenstein M., Hedges L., Higgins J., Rothstein H. Introduction to Meta-Analysis. – Chichester: John Wiley & Sons, Ltd, 2009. – 421 p.
19. Cumming G. The new statistics: why and how // Psychological Science. – 2014. – Vol. 25 (1). – P. 7–29.
20. DerSimonian R., Laird N. A meta-analysis in clinical trials // Controlled Clinical Trials. – 1986. – Vol. 7. – P. 177–188.
21. Marker M. A Bayesian Approach to Meta-Regression: The Relationship Between Body Mass Index and All-Cause Mortality: A Dissertation ... Doctor of Philosophy. – Tallahassee: Florida State University, 2007. – 144 p.
22. Nosek B. et al. Estimating the reproducibility of psychological science // Science. – 2015. – Vol. 349, Iss. 6251. DOI: 10.1126/science.aac4716. – URL: <http://datacolada.org/wp-content/uploads/2016/03/5341-Nosek-et-al-Science-2015-Estimating-the-reproducibility-of-psychological-science.pdf> (дата обращения: 05.01.2018).
23. Uts J. An assessment of the evidence for psychic functioning // Journal of Scientific Exploration. – 1996. – Vol. 10, No. 1. – P. 3–30.
24. Viechtbauer W. Conducting meta-analyses in R with the metafor package // Journal of Statistical Software. – 2010. – Vol. 36, Iss. 3. – P. 1–48.

References

1. *Vachkov, I.V. & S.N. Vachkova.* (2016). Vosproizvodimost psikhologicheskikh eksperimentov kak problema postneklassicheskoy nauki [Reproducibility of psychological experiments as a problem of post-nonclassical science]. *Kulturno-istoricheskaya psikhologiya* [Cultural and Historical Psychology], Vol. 12, No. 1, 97–101. DOI:10.17759/chp.2016120110.
2. *Gubanov, N.I., N.N. Gubanov & A.E. Volkov.* (2015). Osobennosti kriteriev nauchnosti v sotsialno-gumanitarnykh naukakh [Features of scientific criteria in social and human sciences]. In: Goncharov, S.Z. (Ed.) *Gumanitarnoe obrazovanie v kreativno-antropologicheskom izmerenii*: Sb. nauch. st. 12-y Vseros. nauch.-prakt. konf., 19–20 noyab. 2015 g., Yekaterinburg [Education in the Humanities in Creative-anthropological Terms: Collection of Scientific Articles of the 12th All-Russian Scientific and Practical Conference, November 19–20, 2015, Yekaterinburg]. Yekaterinburg, Azhur Publ., 15–23.
3. Dembitskiy, S. (2013). Otsenka geterogenosti v metaanalize (na primere dannykh kross-natsionalnykh issledovaniy) [Evaluation of heterogeneity in meta-analysis (the case of cross-national research data)]. *Sotsiologiya: teoriya, metody, marketing* [Sociology: Theory, Methods, Marketing], 1, 154–165.
4. *Zachs, L.* (1976). *Statisticheskoe otsenivanie* [Statistical Estimation]. Moscow, Statistika Publ., 598. (In Russ.)
5. *Ilyin, V.V.* (2006). *Filosofiya: Uchebnik. V 2 t.* [Philosophy: Textbook. In 2 vols.], Vol. 1. Rostov-on-Don, Feniks Publ., 832.
6. *Lektorskiy, V.A.* (1996). Nauchnoe i vnauchnoe myshlenie: skolzyashchaya granitsa [Scientific and non-scientific thinking: a sliding boundary]. In: *Nauchnye i vnauchnye formy myshleniya: Simpozium* (Moskva, 4–9 aprelya 1995 g.) [Scientific and Non-scientific Ways of Thinking: Symposium (Moscow, April 4–9, 1995)]. Moscow & Kiel. Available at: <http://philosophy.ru/iphras/library/ruspaper/LEKTORS1.htm> (date of access 29.05.2014).
7. *Leshkevich, T.G.* (2000). *Filosofiya: kurs lektsiy* [Philosophy: Course of Lectures]. Moscow, INFRA-M Publ., 240.
8. *Tikhonov, B.N.* (Ed.) (2007). *Metrologiya i elektroradioizmereniya v telekommunikatsionnykh sistemakh: Uchebnoe posobie* [Metrology and Electric and Radio Measurements in Telecommunication Systems: Manual]. Moscow, Goryachaya liniya-Telekom Publ., 374.
9. *Osipov, V.E.* (2017). Distsantsionnoe videnie v svete dvukh priznakov paranauki [Remote viewing in the light of two parasceince signs]. *Filosofiya nauki* [Philosophy of Science], 3 (74), 81–105.
10. *Osipov, V.E.* (2016). K voprosu o statisticheskom analize usloviy i rezultatov eksperimentov D.B. Rayna [On the problem of statistical analysis of conditions and results of J.B. Rhine's experiments]. *Filosofiya nauki* [Philosophy of Science], 4 (71), 91–114.
11. *Osipov, V.E.* (2014). Empiricheskie osnovaniya diskussii ob ekstrasensornom vospriyatii [Empirical foundations of the discussion on extrasensory perception]. *Epistemologiya i filosofiya nauki* [Epistemology and Philosophy of Science], 4 (42), 171–191.
12. *Ruzavin, G.I.* (2012). *Metodologiya nauchnogo poznaniya: Ucheb. posobie dlya vuzov* [Methodology of Scientific Knowledge: Manual for High School]. Moscow, YUNITI-DANA Publ., 287.
13. *Ruzavin, G.I.* (2004). Sluchaynost [Randomness]. In: Ivin, A.A. (Ed.). *Filosofiya: Entsiklopedicheskii slovar* [Philosophy: Encyclopedic Dictionary]. Moscow, Gardariki Publ., 774.
14. *Smagunova, A.N. & V.A. Kozlov.* (1990). Primery primeneniya matematicheskoy teorii eksperimenta v rentgenofluorestsentnom analize [Examples of Application of the Mathematical Theory of Experiment in X-ray Fluorescence Analysis]. Irkutsk, Irkutsk University Publ., 232.

15. *Holton, G.* (1992). Chto takoe "antinauka"? [What is "anti-science"?]. *Voprosy filosofii* [Problems of Philosophy], 2, 26–58. (In Russ.).
16. *Chernikova, I.V.* (2012). Priroda nauki i kriterii nauchnosti [The nature of science and criteria for being scientific]. *Gumanitarnyy vektor* [Humanitarian Vector], 3 (31), 89–96.
17. *Shtanko, V.I.* (2002). *Filosofiya i metodologiya nauki: Uchebnoe posobie dlya aspirantov i magistrantov estestvennonauchnykh i tekhnicheskikh vuzov* [Philosophy and Methodology of Science: Manual for Postgraduates and Undergraduates of Scientific and Technical Universities]. Kharkov, Kharkov National University of Radio Electronics Publ., 292.
18. *Borenstein, M., L. Hedges, J. Higgins & H. Rothstein.* (2009). *Introduction to Meta-Analysis*. Chichester: John Wiley & Sons, Ltd., 421.
19. *Cumming, G.* (2014). The new statistics: why and how. *Psychological Science*, 25 (1), 7–29.
20. *DerSimonian, R. & N. Laird.* (1986). A meta-analysis in clinical trials. *Controlled Clinical Trials*, 7, 177–188.
21. *Marker, M.* (2007). *A Bayesian Approach to Meta-Regression: The Relationship Between Body Mass Index and All-Cause Mortality: A Dissertation ... Doctor of Philosophy*. Tallahassee, Florida State University, 144.
22. *Nosek, B. et al.* (2015). Estimating the reproducibility of psychological science. *Science*, 28 Aug., Vol. 349, Iss. 6251. DOI: 10.1126/science.aac4716.
23. *Utts, J.* (1996). An assessment of the evidence for psychic functioning. *Journal of Scientific Exploration*, Vol. 10, No. 1, 3–30.
24. *Viechtbauer, W.* (2010). Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* 36, 1–48.

Сведения об авторе

Осипов Вадим Евгеньевич – старший преподаватель кафедры «Радиотехнические устройства и системы диагностики», Омский государственный технический университет (644050, Омск, просп. Мира, д. 11, e-mail: osvad@list.ru; osipov@omgtu.ru).

Information about the author

Osipov, Vadim Evgenievich – Senior Lecturer at the Chair of Radio Engineering Devices and Diagnostics Systems, Omsk State Technical University (11, Mira av., 644050, Omsk, e-mail: info@omgtu.ru; osvad@list.ru).

Дата поступления 20.08.2018