

## О НЕКОТОРЫХ МЕТОДИЧЕСКИХ АСПЕКТАХ ИССЛЕДОВАНИЯ ИНДИВИДУАЛЬНОСТИ ЧЕЛОВЕКА

**А.Ю. Калугин**

*Пермский государственный гуманитарно-педагогический университет, г. Пермь, Россия*

В российской психологии активно изучается индивидуальность человека, представленная на многих уровнях организации (от биологических до социальных). При этом зачастую используется корреляционный дизайн, позволяющий обнаружить связи между разноуровневыми свойствами индивидуальности. В публикации рассмотрены некоторые методические аспекты проведения корреляционного анализа, обсуждаются проблемы и возможные пути их решения. В частности, затронута проблема нелинейных зависимостей (параболической, гиперболической и т. п.), которые не обнаруживаются обычными корреляционными методами, однако могут быть выявлены с помощью методов нелинейных корреляций (расчет индекса корреляции, корреляционного отношения, максимального информационного коэффициента, корреляции дистанций, максимальной корреляции, метода «частных моментов» и т. п.). Обсуждается требование обязательности визуализации выявляемой взаимосвязи переменных (диаграммы рассеяния), позволяющее обнаружить скрытую структуру в данных, например, наличие подгрупп в выборке. Особое внимание уделено известной, но практически не рассматриваемой в отечественной психологии т. н. коррекции корреляции для ограниченного диапазона и связанным с этим трудностям. Показано, что изучение многочисленных попарных корреляций между разноуровневыми свойствами индивидуальности требует поправки на множественные сравнения, которая зачастую не проводится исследователями, создавая предпосылки для ложных выводов по результатам исследований. Обособленно рассматривается вопрос о робастных статистических методах исследования, в частности, о перестановочных тестах и бутстрепе, которые сочетают в себе устойчивость к помехам и высокую мощность. В заключение приводится обсуждение проблемы полноты представления результатов исследования и мнение о современных дискуссиях об уровне значимости, величине эффекта и доверительных интервалах, воспроизводимости психологических исследований и метааналитическом подходе. Представлен вывод об отсутствии в настоящее время однозначного решения проблемы проверки нулевой гипотезы (NHST) на фоне существенной критики роли уровня значимости и предлагаемых ему на смену интервальных оценок и величины эффекта. В качестве одного из выходов в сложившейся ситуации предложено представлять наиболее полную информацию о результатах исследования, включая указание на точный уровень значимости, величин доверительных интервалов, эффекта и т. п., которые в дальнейшем могут быть использованы в метаанализе, позволяя перейти к новому уровню научного обобщения.

*Ключевые слова:* индивидуальность, корреляционный анализ, нелинейные зависимости, коррекция корреляции для ограниченного диапазона, множественные сравнения.

В отечественной психологии существует несколько подходов к изучению индивидуальности человека (Ананьев, 2001; Мерлин, 1986; Русалов, 2012). Несмотря на некоторые разногласия, порой достаточно серьезные, эти подходы схожи тем, что предлагают изучать человека целостно, начиная с его биологических основ и заканчивая высшими социальными и духовными проявлениями. В послед-

ние десятилетия к проблеме всестороннего изучения индивидуальности человека обращаются идеологи теории черт, понимая важность учета как биологических, так и социальных детерминант личности (McAdams, 2006; McCrae, Costa, 2008).

Операционализация любой разработанной оригинальной теории – одна из важнейших задач, стоящих перед ее создателем, его учени-

ками и последователями. Поэтому все психологические школы, в той или иной мере занимающиеся разработкой новых теоретических построений (в том числе и научные сообщества, изучающие феномен индивидуальности), широко используют математический аппарат. Любая попытка изучения человека целостно требует включения в системный анализ множества переменных, отражающих разные стороны и уровни индивидуальности. Одним из аспектов такого отражения является выявление взаимного влияния переменных. В рамках такой парадигмы главными методами изучения многочисленных взаимосвязей становятся корреляционный и факторный анализы. Однако при сборе статистических данных и их обработке важно соблюдать ряд требований, игнорирование или нарушение которых может привести в последующем к серьезным искажениям. Существует множество пособий и статей, в которых указаны допущения применимости тех или иных статистических процедур, включая корреляционный и факторный анализы (Гржибовский, 2017; Наследов, 2008).

В настоящей публикации приводится упоминание лишь на некоторые из них, а также освещаются редко обсуждаемые проблемы применения корреляционного дизайна, рассматриваемые в контексте изучения индивидуальности.

*Проблема нелинейных зависимостей.* Не всегда связь между двумя свойствами характеризуется линейностью (формулировочнее – линейной зависимостью, чаще в понимании – прямой линейной зависимостью)<sup>1</sup>. Поскольку при разработке многих дизайнов исследования (и исследованиях индивидуальности) часто используются корреляционные дизайны, поэтому важно учитывать возможную нелинейность, чтобы не упустить значимые взаимосвязи (Калугин, 2018а, 2018б).

Более того, некоторые связи могут быть ошибочно приняты за несущественные, когда их характер изменился с линейного на нелинейный! Рассмотрим два примера. На рис. 1а представлена положительная прямолинейная зависимость, которая хорошо описывается с помощью  $r$  Пирсона ( $r = 0,94$ ,  $p < 0,001$ ) или  $\rho$  Спирмена ( $\rho = 0,95$ ,  $p < 0,001$ ). Однако при наличии нелинейной зависимости, например, гиперболической (рис. 1б) или параболической

(рис. 1с), корреляционный анализ как по Пирсону, так и Спирмену, не улавливая закономерности, занижает значение при гиперболическом варианте:  $r = -0,57$ ,  $p < 0,001$  и  $\rho = -0,55$ ;  $p < 0,001$  соответственно. В случае параболической зависимости эта тенденция становится еще очевиднее:  $r = -0,01$ ;  $p = 0,822$  и  $\rho = 0,001$ ;  $p = 0,980$  соответственно. Однако в ряде случаев зависимости при нелинейной связи лучше выявляются с помощью непараметрических критериев Спирмена и Кендалла, чем с помощью критерия Пирсона (Шитиков, Розенберг, 2013, с. 87).

Традиционными способами обнаружения нелинейной связи между переменными являются определение «индекса корреляции» и «корреляционного отношения» (Ферстер, Ренц, 1983). Нелинейные модели широко используются при регрессионном анализе, машинном обучении. Однако в случае с корреляционным анализом нелинейность крайне ограничено является предметом отдельного исследования и, скорее, она рассматривается в контексте проверки допущения применимости процедур линейного корреляционного анализа.

Относительно недавно появились новые способы исследования нелинейности с помощью «частных моментов» (Viola, Nawrocki, 2012), реализованные в пакете «NNS» в среде  $R^2$ ; с определением т. н. максимального информационного коэффициента (maximal information coefficient, MIC) (Reshef et al., 2011), реализованного в пакете «minerva»; методом корреляции дистанций (distance correlation) (Szekely et al., 2007), реализованной в пакете «energy», и др. Наилучшие результаты, согласно результатам сравнения нескольких вариантов нелинейных корреляций (Deebani, Kachouie, 2018), демонстрирует максимальная корреляция (maximal correlation), реализованная в пакете «асерак» (Breiman, Friedman, 1985). *Наличие нескольких групп в данных.* В целом к статистическому анализу следует подходить с особой тщательностью, так как в рамках корреляционного дизайна можно совершить ошибку, ориентируясь только на

<sup>2</sup> Свободная программная среда вычислений с открытым исходным кодом, а также язык программирования для статистической обработки данных и работы с графикой; здесь и далее – статистическое программное обеспечение для анализа данных и, фактически, стандарт для статистических программ (прим. ред.)

<sup>1</sup> Далее – линейная зависимость, линейность (прим. ред.).

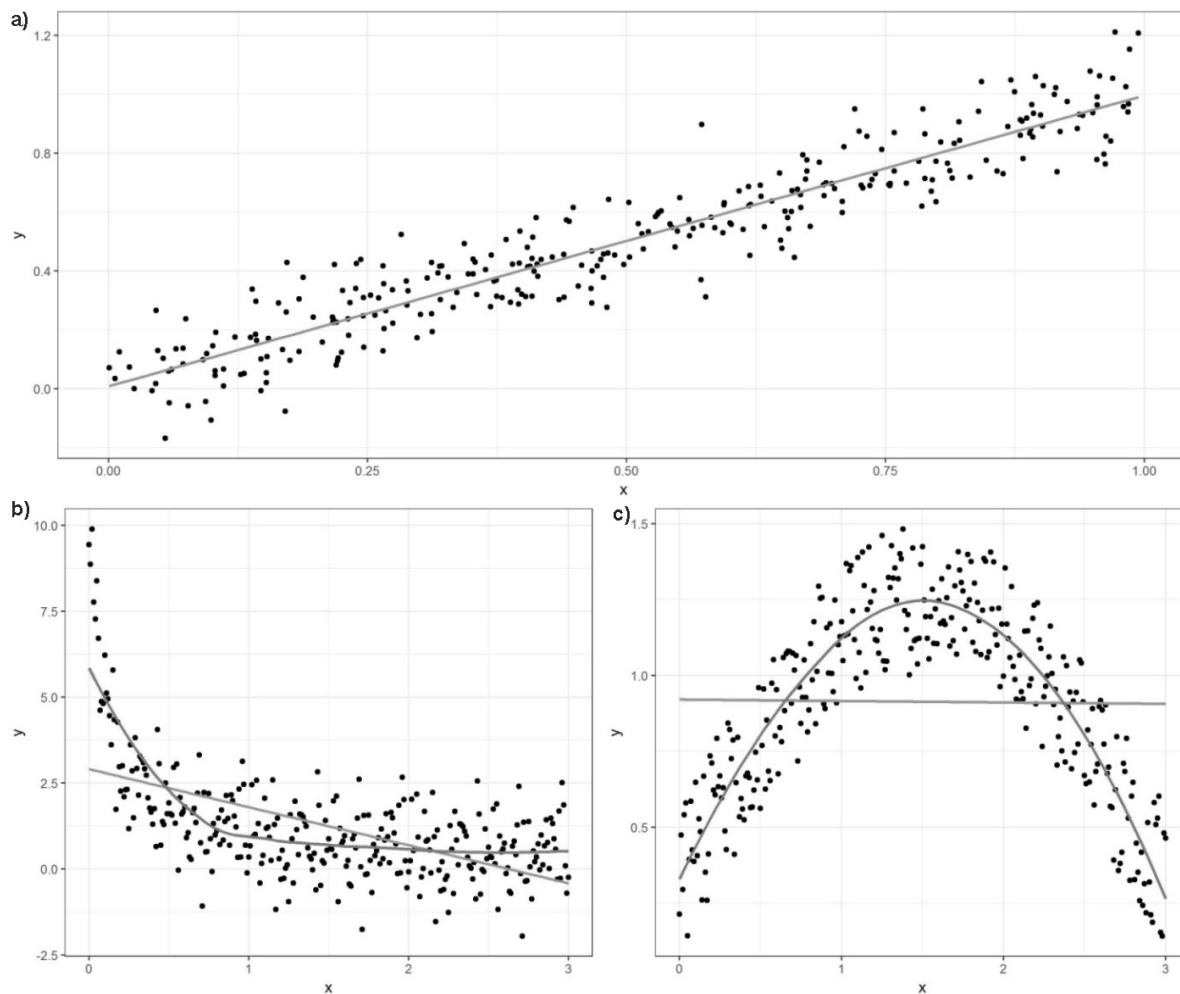


Рис. 1. Линейная (а), нелинейная гиперболическая (б) и параболическая (с) взаимосвязь: прямая линия – линия регрессии, отражающая линейную зависимость, кривая линия – линия регрессии, отражающая реальную зависимость в данных

Fig. 1. Rectilinear (a), nonlinear hyperbolic (b) and parabolic (c) relation: straight line – regression line, reflecting linear dependence, curved line – regression line, reflecting the real dependence in the data

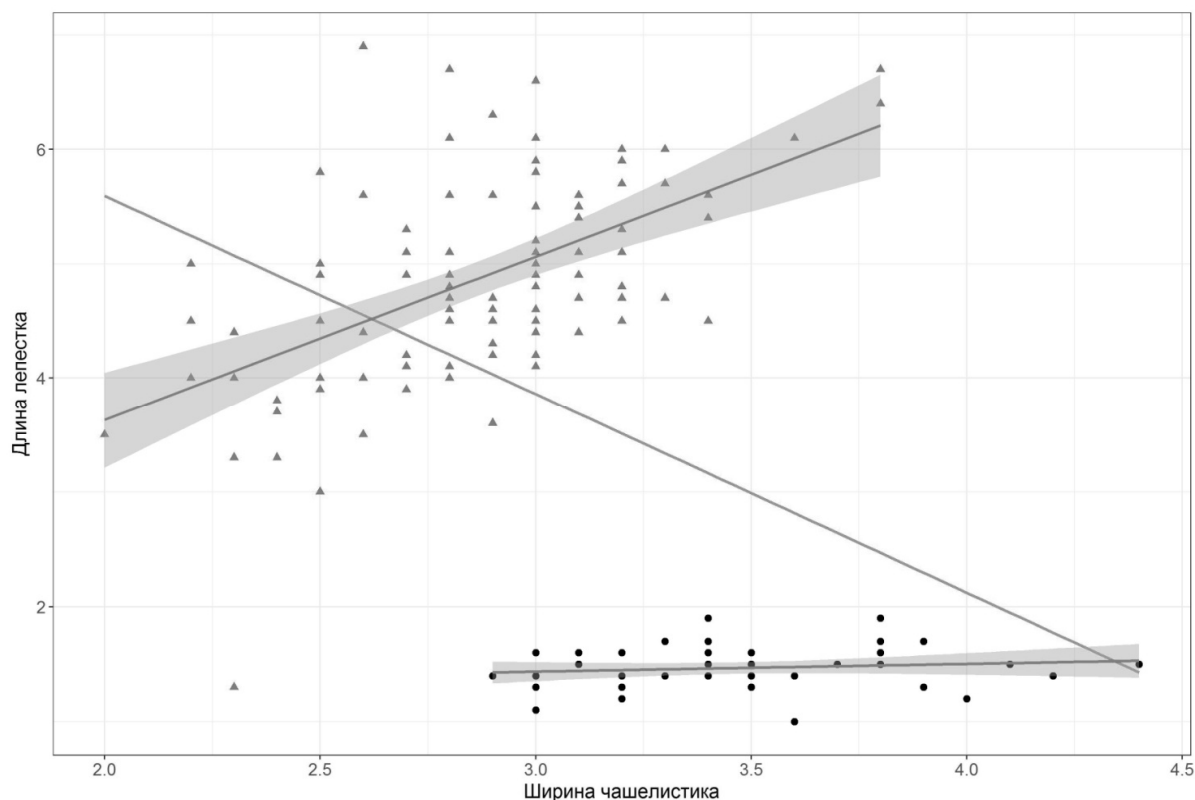
значение коэффициента корреляции. На рис. 2 представлена взаимосвязь ширины чашелистика и длины лепестка из классических данных Р. Фишера «Ирисы». Корреляционный анализ Пирсона обнаруживает значимую отрицательную взаимосвязь между данными ( $r = -0,43$ ;  $p < 0,001$ ), при этом при визуальном изучении диаграммы рассеяния в данных можно наблюдать явное присутствие двух групп. Корреляция по Пирсону, проведенная отдельно для каждой группы, дает следующие результаты:

- для большего размера группы:  $r = 0,54$ ;  $p < 0,001$ ;
- для меньшего размера группы:  $r = 0,13$ ;  $p = 0,359$ .

Таким образом, не только значимость корреляции меняется, но она даже меняет

знак на противоположный! Известное правило – «визуальное изучение диаграмм рассеяния» – часто не выполняется, и исследователь ориентируется только на табличные значения.

*Коррекция корреляции для ограниченного диапазона.* Также приведем пример проблемы, возникающей в случае разбиения результатов на три группы (высокие, средние, низкие) по какому-либо признаку, с дальнейшим включением такого признака в анализ (данные взяты из исследования (Калугин, 2015)). На рис. 3 представлена диаграмма рассеяния, демонстрирующая взаимосвязь между «эмоциональной реактивностью» и «выносливостью», измеренными по методике ФСВ-П Я. Стреляу. Общая выборка респондентов составила 287 человек. Далее с помощью кластерного анализа методом  $k$ -средних выборка



**Рис. 2.** Проблема наличия групп в исследуемых данных: треугольниками и точками изображены группы наблюдений, прямая линия – линия регрессии для анализа, проведенного на всей совокупности данных, прямые линии со шлейфом – линии регрессии для каждой совокупности отдельно, серый шлейф – 95 % доверительные интервалы

**Fig. 2.** The problem of the presence of groups in the studied data: the triangles and dots represent the groups of observations, the straight line is the regression line for the analysis performed on the entire data set, the straight lines with the loop are the regression lines for each set separately, the gray loop is 95% confidence intervals

была разделена на группы по критерию уровня значений переменной «эмоциональная реактивность» (ЭР): кластер с низкими показателями по ЭР включал 57 испытуемых, кластер со средними показателями ЭР – 137, кластер с высокими показателями ЭР – 93.

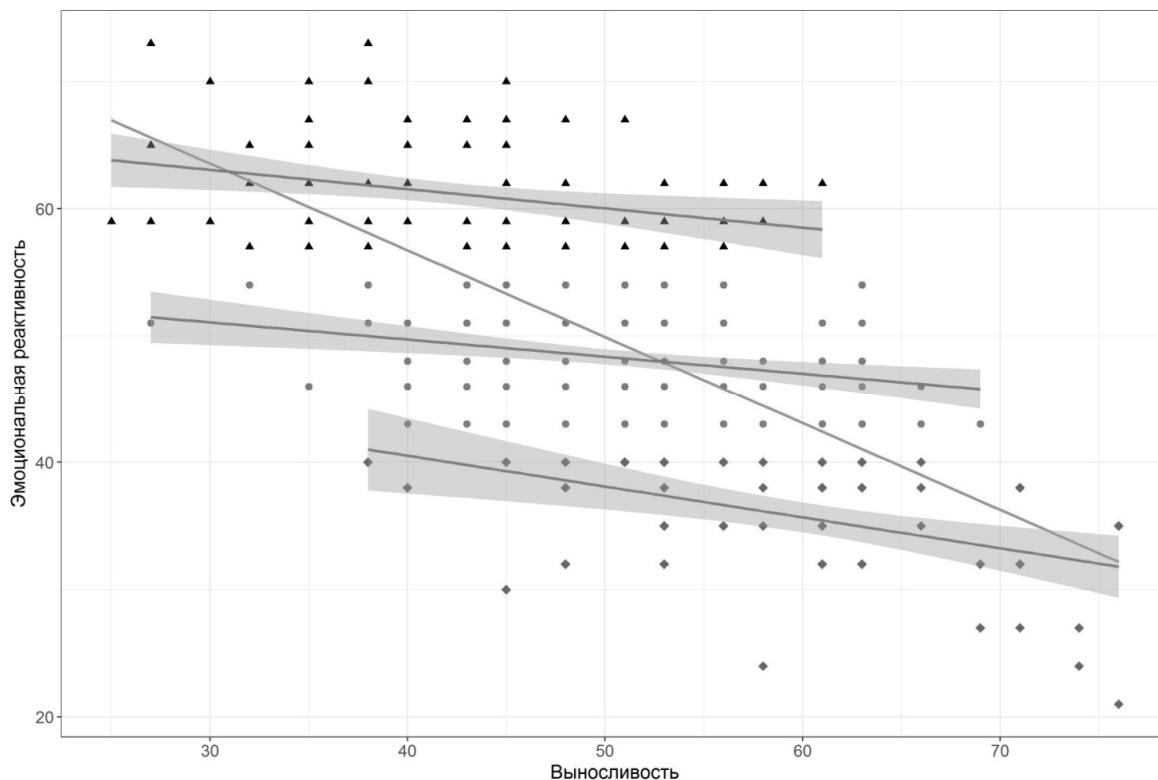
Исходя из анализа представленной на легенде рис. 3 информации, следует подчеркнуть, что линия регрессии в кластерах (подвыборках испытуемых) уже не так явно характеризует отрицательную корреляцию (в отличие от общегрупповых данных). Аргументами этому служат значения коэффициентов корреляции  $r$  Пирсона, средних арифметических  $M$  и стандартных отклонений  $SD$  для указанных групп:

- в общей выборке:  $r = -0,68$ ,  $p < 0,001$ ;  $M = 50$ ,  $SD = 10$ ;
- группа с низкими значениями ЭР:  $r = -0,44$ ;  $p < 0,001$ ;  $M = 35,63$ ;  $SD = 4,86$ ;
- группа со средними значениями ЭР:  $r = -0,28$ ;  $p = 0,001$ ;  $M = 48,23$ ;  $SD = 3,67$ ;

- группа с высокими значениями ЭР:  $r = -0,27$ ;  $p = 0,008$ ;  $M = 61,22$ ;  $SD = 4,14$ .

С математической позиции такие изменения обосновываются представлениями о корреляции как отношении ковариации к квадратному корню произведения двух дисперсий. Поэтому если дисперсия переменной будет искусственно ограничена (как в рассматриваемом случае), то корреляция будет уменьшена, несмотря на то, что наклон линии регрессии может остаться прежним (Revelle, 2018, с. 120). Следует отметить, что ограничение сразу по двум переменным приведет к еще большим искажениям результатов.

Отметим, что в психометрике разработаны методы коррекции такого искажения коэффициентов корреляции для ограниченного диапазона (correct correlations for restriction of range), также известные как «Thorndike's Case II» (в наше время существует несколько подходов к коррекции для разных случаев (Tran, 2011)).



**Рис. 3. Проблема ограниченного диапазона в корреляционном анализе: треугольниками, точками и ромбиками изображены выделенные кластеры, прямая линия – линия регрессии для анализа, проведенного на всей совокупности данных, прямые линии со шлейфом – линии регрессии для каждой совокупности отдельно, серый шлейф – 95 % доверительные интервалы**

**Fig. 3. The problem of limited range in the correlation analysis: the triangles, dots and diamonds represent the selected clusters, the straight line is the regression line for the analysis performed on the entire data set, the straight lines with the loop are the regression lines for each set separately, the gray loop is 95% confidence intervals**

В случае с «Thorndike's Case II» важно, чтобы переменные были линейно связаны. Также не следует включать в анализ статистику, сильно коррелирующую с переменной, на основании которой было произведено деление на группы, так как она косвенно тоже подвергается «ограничению диапазона» (Hunter, Schmidt, 2004).

Необходимые функции есть в пакетах «psych», «psychometric» и «selection» в среде R. Однако рассчитать скорректированные коэффициенты можно и «вручную» по следующей формуле (приводится по: Tran, 2011, с. 5):

$$R = \frac{r * S/s}{\sqrt{1 - r^2 + r^2 * S^2/s^2}}, \quad (1)$$

где  $r$  – корреляция в ограниченной выборке;  $S$  – стандартное отклонение в неограниченной выборке;  $s$  – стандартное отклонение в ограниченной выборке.

Таким образом, скорректированные коэффициенты корреляции для рассматриваемых групп на примере рис. 3 следующие:

- группа с низкими значениями ЭР:  $r = -0,71$ ;  $p < 0,001$ ;
- группа со средними значениями ЭР:  $r = -0,62$ ;  $p < 0,001$ ;
- группа с высокими значениями ЭР:  $r = -0,56$ ;  $p < 0,001$ .

Эти данные уже не столь сильно отличаются от значений корреляции, вычисленной для всей выборки респондентов ( $r = -0,68$ ,  $p < 0,001$ ).

Описанный случай не означает, что разделять массивы данных на подвыборки не стоит, но при этом важно учитывать возможные искажения и либо корректировать их, либо не включать в последующий анализ переменную, на основе которой было осуществлено деление на группы, а также сильно коррелирующие с ней переменные.

*Уровень значимости и величина эффекта.* В ряде случаев при разбиении на группы может наблюдаться исчезновение ранее выявленной значимой корреляции между переменными, что, например, в рамках теории интегральной индивидуальности В.С. Мерлина

(1986) может быть проинтерпретировано как разрушение связи между свойствами ввиду формирования индивидуального стиля. Однако исчезновение значимости связи может быть обусловлено иными причинами. Уровень значимости напрямую связан с объемом выборки, поэтому снижение размера выборки из-за разделения ее на подгруппы может привести к «незначимому» коэффициенту корреляции. За это и за некоторые другие проблемные моменты показатель значимости не раз критиковался (Сивуха, Козьяк, 2009; Cohen, 1994; Cumming, 2014) в основном сторонниками представлений о большей информативности величины эффекта и интервальных оценок (доверительных интервалов). Об этих подходах и их критике можно подробнее прочитать в ряде публикаций (Сивуха, Козьяк, 2009; Корнеев, 2016). Поэтому имеет смысл указывать помимо уровня значимости (желательно точного, а не округленного до одного из трех значений: 0,05; 0,01 и 0,001 – см., например, Publication Manual..., 2010) также доверительные интервалы и значения коэффициента величины эффекта (меры эффекта  $d$ ) Коэна (или его аналоги) (Tellez et al., 2015). Любая традиционная статистическая величина ( $t$ ,  $r$ ,  $F$ ,  $\chi^2$  и др.) может быть преобразована в  $d$ , например, с использованием соответствующих функций из библиотек «effsize», «psych» и др. в R либо с помощью онлайн-калькуляторов (например, [https://www.psychometrica.de/effect\\_size.html](https://www.psychometrica.de/effect_size.html)).

Помимо всего вышеописанного следует помнить, что сравнение коэффициентов корреляции разных групп «на глазок» неправомерно: например, разница между корреляциями  $r_1 = 0,5$  и  $r_2 = 0,3$  может быть значимой при размерах выборок  $n_1 = 250$  и  $n_2 = 100$  соответственно, однако она будет незначимой при размерах выборок  $n_1 = 200$  и  $n_2 = 100$ . Для таких целей существуют специализированные функции в пакете «psych» в R и в пакете Statistica корпорации StatSoft.

*Проблема множественных сравнений* (multiple comparisons, multiple testing problem). Традиционно при исследовании индивидуальности человека в корреляционном анализе или анализе различий используются десятки переменных, что приводит к опасности переоценить количество значимых связей или значимых различий (в большом объеме данных обязательно обнаружатся значимые взаимосвязи или различия).

Приведем пример частой ошибки: использование  $t$ -критерия Стьюдента для случая трех групп вместо применения однофакторного дисперсионного анализа. Исследуются различия по какой-либо переменной в трех группах (А, В и С). Для этого проводится три сравнения с помощью  $t$ -критерия: А-В, В-С и А-С. Предположим, что для каждого сравнения обнаружено значимое различие на уровне  $p < 0,05$ . Однако вероятность ошибки в этом случае рассчитывается по следующей формуле (приводится по: Bretz et al., 2010, p. 1):

$$P = 1 - (1 - \alpha)^m, \quad (2)$$

где  $\alpha$  – уровень значимости, на который ориентируется исследователь (обычно 0,05);

$m$  – количество тестирований нулевых гипотез;

$P$  – вероятность совершить хотя бы одну ошибку I рода.

Таким образом, для трех сравнений вероятность ошибиться составит:  $P = 1 - (1 - 0,05)^3 = 0,14$  (а вовсе не 0,05, как ожидает исследователь).

До настоящего времени разработано множество поправок для уровня значимости: поправка Бонферрони, поправка Холма, поправка Бенджамини-Хохберга и др. (Bretz et al., 2010). Корректировка уровня значимости на множественную проверку гипотез позволит повысить качество получаемых значимых взаимосвязей или различий, а также повышает уверенность в достоверности научных результатов. В случае использования среды R для получения таких поправок применяются соответствующие функции как базового пакета «stats», так и других пакетов, в том числе специализированных («multcomp» и др.).

*Перестановочные тесты и бутстреп.* Тестирование гипотез с помощью параметрических методов ( $t$ -критерий Стьюдента, корреляционный анализ Пирсона,  $F$ -критерий Фишера для однофакторного дисперсионного анализа и т. д.) требует соблюдения целого ряда допущений. Среди основных из них – необходимость обеспечения соответствия эмпирического распределения теоретическому нормальному, чего в конкретных случаях не всегда удается добиться. В случае асимметричного распределения и/или нарушения иных допущений обычно рекомендуют использовать непараметрические аналоги ( $U$ -критерий Манна – Уитни, корреляционный анализ Спирмена,  $H$ -критерий Краскела – Уоллиса и т. п.). Последние характеризуются

меньшей мощностью по сравнению с параметрическими. В последнее время при решении задач исследования все чаще предпочтение отдается т.н. робастным техникам<sup>3</sup>, которые, с одной стороны, устойчивы к помехам, а с другой – позволяют получить достаточно надежные результаты, в ряде случаев превосходящие по мощности параметрические тесты. К таким относятся, например, перестановочные тесты (permutation tests), которые позволяют получить эмпирическое распределение критерия и на этой основе рассчитать его уровень значимости (Кабаков, 2016; Шитиков, Розенберг, 2013). Как показывают сравнения мощности классических тестов и перестановочных, последние демонстрируют существенно большую мощность (Мелас и др., 2016). В рамках парадигмы рандомизированного подхода широко используется бутстреп<sup>4</sup> (bootstrap) и вычисление на его основе 95 % доверительных интервалов. В среде R функции для перестановочных тестов можно найти в пакетах «coin» и «lmPerm», функции для бутстрэпа – в пакете «boot».

*Проблема воспроизводимости.* В последние годы активно обсуждается проблема воспроизводимости психологических исследований (Open Science Collaboration..., 2015). Эта проблема напрямую касается и исследования индивидуальности: зачастую в рамках единой исследовательской программы используются одни и те же методы для изучения разноуровневых свойств индивидуальности. Таким образом, на протяжении десятилетий накапливается богатый эмпирический материал, доступный изучению и сравнению. Однако использование этих данных для проверки вос-

производимости и проведения метаанализа сопряжено с существенными трудностями: в связи с отсутствием четких требований к предоставлению материала, к степени его открытости результаты исследований зачастую подаются в усеченном виде (незначимые коэффициенты в корреляционном и низкие нагрузки в факторном анализе не приводятся, многочисленные взаимосвязи визуализируются с помощью корреляционных графов, а величина статистики при этом не указывается и т. п.), описательные статистики часто носят поверхностный характер и т. д.

В мировой науке наметился переход к новой парадигме, требующей «метааналитического мышления», максимально подробного представления полученных в исследовании данных; публикации, в том числе негативных результатов, наличия, в том числе, высокой статистической культуры исследователя (Корнеев и др., 2016; Корнилов, Корнилова, 2010; Корнилова, 2010; Сивуха, Козьяк, 2009; Cumming, 2012; Kline, 2013). Вероятно, требуется серьезная организационная и просветительская работа, связанная с выработкой четких требований к представлению результатов исследований (прежде всего – статистических) и доведением этих требований до научной общественности.

### Выводы

В настоящей публикации были рассмотрены некоторые теоретические и методические проблемы, стоящие перед исследователем феномена индивидуальности человека: учет характера корреляционных связей (линейность – нелинейность, внутренняя структура данных); специфика разбиения на подгруппы; проблема множественных сравнений; робастные и сходные с ними новые статистические процедуры; проблема воспроизводимости. Отдельный важный методический аспект – требование к полноте представления информации, в частности, приведение полных корреляционных таблиц и результатов факторного анализа, приведение точного уровня значимости, величин статистического эффекта и доверительных интервалов. Все это позволит использовать в дальнейшем метаанализ и перейти к новому уровню научного обобщения.

**Исследование выполнено за счет гранта Российского научного фонда (проект 18-18-00386), Институт психологии РАН.**

<sup>3</sup> Робастные техники или методы (англ. robustness, от robust – «сильный», «твёрдый», «устойчивый») – статистические методы, характеризующиеся устойчивостью к влиянию различного рода выбросов и помех на результат исследования, достигаемой путем выявления таких выбросов, снижения их влияния или исключения их из выборки (прим. ред.).

<sup>4</sup> Бутстрэп (англ. bootstrap) в статистике – практический метод исследования распределения статистик вероятностных распределений, основанный на многократной генерации выборок случайным методом (например, методом генерации случайных величин) с целью быстрого получения оценок основных их статистик (дисперсию, корреляцию и т. д., в том числе доверительные интервалы), используемых для построения сложных моделей (прим. ред.).

### Литература

1. Ананьев, Б.Г. Человек как предмет познания / Б.Г. Ананьев. – СПб.: Питер, 2001. – 288 с.
2. Гржибовский, А.М. Корреляционный анализ данных с использованием программного обеспечения Statistica и SPSS / А.М. Гржибовский, С.В. Иванов, М.А. Горбатова // Наука и здравоохранение. – 2017. – № 1. – С. 7–36.
3. Кабаков, Р.И. R в действии. Анализ и визуализация данных в программе R / Р.И. Кабаков. – М.: ДМК Пресс, 2016. – 588 с.
4. Калугин, А.Ю. Использование нейронных сетей в психологических исследованиях / А.Ю. Калугин // Искусственный интеллект в решении актуальных социальных и экономических проблем XXI века: сб. ст. по материалам Третьей всерос. науч.-практ. конф. (г. Пермь, 16–18 мая 2018 г.). – Пермь: ПГНИУ, 2018а. – С. 144–148.
5. Калугин, А.Ю. История и перспективы исследования интегральной индивидуальности в рамках системного подхода / А.Ю. Калугин // Вестник Пермского университета. Философия. Психология. Социология. – 2018б. – Вып. 2. – С. 252–263. DOI: 10.17072/2078-7898/2018-2-252-263.
6. Калугин, А.Ю. Ценностная направленность личности как выражение смыслообразующей активности: структура, динамика и функции в структуре интегральной индивидуальности (на материале студенчества): дис. ... канд. психол. наук / А.Ю. Калугин: – Пермь, 2015. – 179 с.
7. Корнеев, А.А. Критика методологии проверки нулевой гипотезы: ограничения и возможные пути выхода. Ч. I / А.А. Корнеев, Е.И. Рассказова, А.Н. Кричевец, А.Я. Койфман // Психологические исследования. – 2016. – Т. 9, № 45. – С. 1. – URL: <http://psystudy.ru/index.php/num/2016v9n45/1231-korneev45.html> (дата обращения: 02.09.2018).
8. Корнилов, С.А. Мета-аналитические исследования в психологии / С.А. Корнилов, Т.В. Корнилова // Психологический журнал. – 2010. – Т. 31, № 6. – С. 5–17.
9. Корнилова, Т.В. Основные тренды в развитии методов психологических исследований / Т.В. Корнилова // Экспериментальная психология в России: традиции и перспективы / под ред. В.А. Барабанщикова. – М.: Издво «Институт психологии РАН», 2010. – С. 42–46.
10. Мелас, В.Б. Численное сравнение перестановочных и классических методов проверки статистических гипотез / В.Б. Мелас, Д.И. Сальников, А.О. Гудулина // Вестник Санкт-Петербургского университета. Сер. 1. Математика. Механика. Астрономия. – 2016. – Т. 3(61), Вып. 3. – С. 415–423. DOI: 10.21638/11701/spbu01.2016.309.
11. Мерлин, В.С. Очерк теории интегрального исследования индивидуальности / В.С. Мерлин. – М.: Педагогика, 1986. – 256 с.
12. Наследов, А.Д. Математические методы психологического исследования. Анализ и интерпретация данных / А.Д. Наследов. – СПб.: Речь, 2008. – 392 с.
13. Русалов, В.М. Темперамент в структуре индивидуальности человека: Дифференциально-психофизиологические и психологические исследования / В.М. Русалов. – М.: Институт психологии РАН, 2012. – 528 с.
14. Сивуха, С.В. О реформе статистического вывода в психологии / С.В. Сивуха, А.А. Козьяк // Психология. Журнал Высшей школы экономики. – 2009. – Т. 6, № 4. – С. 66–86.
15. Ферстер, Э. Методы корреляционного и регрессионного анализа / Э. Ферстер, Б. Ренц. – М.: Финансы и статистика, 1983. – 304 с.
16. Шитиков, В.К. Рандомизация и бутстреп: статистический анализ в биологии и экологии с использованием R / В.К. Шитиков, Г.С. Розенберг. – Тольятти: Кассандра, 2013. – 314 с.
17. Breiman, L. Estimating optimal transformations for multiple regression and correlation / L. Breiman, J.H. Friedman // Journal of the American statistical Association. – 1985. – Vol. 80(391). – P. 580–598.
18. Bretz, F. Multiple comparisons using R / F. Bretz, T. Hothorn, P. Westfall. – Boca Raton: Chapman & Hall/CRC Press, 2010. – 205 p.
19. Cohen, J. The earth is round ( $p < .05$ ) / J. Cohen // American Psychologist. – 1994. – Vol. 49(12). – P. 997–1003. DOI: 10.1037/0003-066X.49.12.997.
20. Cumming, G. The new statistics: why and how / G. Cumming // Psychological Science. – 2014. – Vol. 25(1). – P. 7–29.
21. Cumming, G. Understanding the new statistics: effect sizes, confidence intervals, and meta-analysis / G. Cumming. – New York: Routledge, 2012. – 519 p.
22. Deebani, W. Ensemble Correlation Coefficient / W. Deebani, N.N. Kachouie



- // ISAIM. – 2018. – URL: [http://isaim2018.cs.virginia.edu/papers/ISAIM2018\\_Deebani\\_Kachouie.pdf](http://isaim2018.cs.virginia.edu/papers/ISAIM2018_Deebani_Kachouie.pdf) (дата обращения: 25.09.2018)
23. Hunter, J.E. *Methods of meta-analysis: Correcting error and bias in research findings. 2nd ed* / J.E. Hunter, F.L. Schmidt. – Thousand Oaks: Sage Publications, 2004. – 582 p.
24. Kline, R.B. *Beyond significance testing: statistics reform in the behavioral sciences* / R.B. Kline. – Washington: American Psychological Association, 2013. – 350 p.
25. McAdams, D.P. *A new Big Five: Fundamental principles for an integrative science of personality* / D.P. McAdams, J.L. Pals // *American Psychologist*. – 2006. – Vol. 61(3). – P. 204–217.
26. McCrae, R.R. *The five-factor theory of personality* / R.R. McCrae, Jr.P.T. Costa // *Handbook of personality: Theory and research. 3rd ed.* / Eds.: O.P. John, R.W. Robins, L.A. Pervin. – Guilford Press: New York, NY, USA, 2008. – P. 159–181.
27. Open Science Collaboration. *Estimating the reproducibility of psychological science* // *Science*. – 2015. – Vol. 349(6251). – P. 943. DOI: 10.1126/science.aac4716.
28. *Publication Manual of the American Psychological Association. 6th ed.* – Washington, DC: American Psychological Association, 2010. – 272 p.
29. Reshef, D.N. *Detecting novel associations in large data sets* / D.N. Reshef, Y.A. Reshef, H.K. Finucane et al. // *Science*. – 2011. – Vol. 334(6062). – P. 1518–1524. DOI: 10.1126/science.1205438
30. Revelle, W. *An introduction to psychometric theory with applications in R* / W. Revelle. – URL: <http://www.personality-project.org/r/book/chapter4.pdf> (дата обращения: 26.07.2018)
31. Szekely, G.J. *Measuring and testing dependence by correlation of distances* / G.J. Szekely, M.L. Rizzo, N.K. Bakirov // *The annals of statistics*. – 2007. – Vol. 35(6). – P. 2769–2794.
32. Tellez, A. *Effect size, confidence intervals and statistical power in psychological research* / A. Tellez, C.H. Garcia, V. Corral-Verdugo // *Psychology in Russia: State of the Art*. – 2015. – Vol. 8(3). – P. 27–46.
33. Tran, D.M. *Robustness of two formulas to correct Pearson correlation for restriction of range: Dissertation* / D.M. Tran. – Atlanta: Georgia State University, 2011. – URL: [http://scholarworks.gsu.edu/eps\\_diss/84](http://scholarworks.gsu.edu/eps_diss/84) (дата обращения: 15.06.2018)
34. Viole, F. *Deriving nonlinear correlation coefficients from partial moments* / F. Viole, D.N. Nawrocki // *SSRN Electronic Journal* (September 18, 2012). DOI: 10.2139/ssrn.2148522.

**Калугин Алексей Юрьевич**, кандидат психологических наук, доцент кафедры практической психологии Пермского государственного гуманитарно-педагогического университета (Пермь), [kaluginau@yandex.ru](mailto:kaluginau@yandex.ru)

Поступила в редакцию 23 апреля 2019 г.

DOI: 10.14529/jpps190203

## ON SOME METHODOLOGICAL ASPECTS OF THE STUDY OF HUMAN INDIVIDUALITY

**A.Yu. Kalugin**, [kaluginau@yandex.ru](mailto:kaluginau@yandex.ru)

Perm State Humanitarian Pedagogical University, Perm, Russian Federation

Human individuality, presented on different levels (from biological to social ones), is of a high interest in Russian psychology, and the method of correlation design is widely used among researches, because it allows revealing relationships between multi-level properties of individuality. The present article examines several methodical aspects of the correlation analysis implementation, discussing problems and possible solutions. In particular, it considers the issue

of nonlinear dependencies (parabolic, hyperbolic etc.), which are impossible to reveal by common correlation methods, but which can be uncovered by using nonlinear correlations, such as correlation index, correlation ratio, maximal information coefficient, distance correlation, maximal correlation, “partial moments” method. Furthermore, it considers the necessity of visualizing variables correlation (scatterplots) that enables to reveal hidden data structures, for example, subgroups. Special attention is paid to correlations corrections for restriction of range and related difficulties that are well-known, but scarcely researched in Russian psychology. In process of investigating plentiful pairwise correlations between individuality properties on different levels it is important to consider an issue of multiple comparisons, which, however, is rarely taken into the account by researchers, leading to false results in many occasions. Moreover, the article examines robust statistical methods, particularly permutation tests and bootstrap. These methods combine robustness and high power. Finally, the study observes such issues as the completeness of results presentation and current debates about significance level, effect size and confidence intervals, reproducibility of psychological researches, and meta-analysis approach. Significance level has often been criticized; interval estimates and effect size were supposed to replace it. However, the problem of Null Hypothesis Significance Testing (NHST) has not been completely solved yet. A possible solution is presentation of complete data on research results including precise significance level, confidence intervals, effect size and etc. These estimations can be then applied in meta-analysis, which allows moving on to a new level of scientific generalizations.

*Keywords: individuality; correlation analysis; nonlinear dependencies; correct correlations for restriction of range; multiple comparisons.*

**The study was granted by the Russian Science Foundation (project No 18-18-00386), the Institute of Psychology of the Russian Academy of Sciences.**

### References

1. Anan'ev B.G. *Chelovek kak predmet poznaniya* [Man as the Object of Knowledge]. St. Petersburg, “Piter” Publ., 2001. 288 p.
2. Grzhibovskiy A.M. [Correlation Analysis of Data Using Statistica and SPSS Software]. *Nauka i Zdravooхранenie* [Science and Healthcare], 2017, no. 1, pp. 7–36. (in Russ.).
3. Kabacoff R.I. *R v deystvii. Analiz i vizualizaciya dannyh v programme R* [R in Action. Data Analysis and Graphics with R]. Moscow, DMK Press Publ., 2016. 588 p.
4. Kalugin A.Yu. [The Use of Neural Networks in Psychological Research]. *Iskusstvennyy intellekt v reshenii aktual'nyh social'nyh i ehkonomicheskikh problem XXI veka* [Artificial Intelligence in Solving Actual Social and Economic Problems of the XXI Century]. Perm', Perm State University Publ., 2018a, pp. 144–148. (in Russ.).
5. Kalugin A.Yu. [History and Prospects of Studying Integral Individuality within the System Approach]. *Vestnik Permskogo universiteta. Filosofiya. Psikhologiya. Sotsiologiya* [Perm University Herald. Series “Philosophy. Psychology. Sociology”]. 2018b, iss. 2, pp. 252–263. DOI: 10.17072/2078-7898/2018-2-252-263. (in Russ.).
6. Kalugin A.Yu. *Cennostnaya napravlennost' lichnosti kak vyrazhenie smysloobrazuyushchey aktivnosti: struktura, dinamika i funkcii v strukture integral'noy individual'nosti (na materiale studentchestva): diss... kand.psikhol. nauk.* [Value Orientation of the Personality as an Expression of Meaning-Making Activity: Structure, Dynamics and Function in the Structure of Integral Individuality (Based on the Study of Students). Cand. diss. (Psychology)]. Perm', 2015. 179 p.
7. Korneev A.A., Rasskazova E.I., Krichevets A.N., Koifman A.Ya. [Criticism of Null Hypothesis Significance Testing: Limitations and Possible Ways Out. Part I]. *Psichologicheskie issledovaniya* [Psychological Studies]. 2016, vol. 9(45), p. 1. Available at: <http://psystudy.ru/index.php/num/2016v9n45/1231-korneev45.html>. (accessed 2.09.2018). (in Russ.).
8. Kornilov S.A., Kornilova T.V. [Meta-Analytic Studies in Psychology]. *Psichologicheskiy zhurnal* [Psychological Journal]. 2010, vol. 31(6), pp. 5–17. (in Russ.).
9. Kornilova T.V. [Main Trends in the Development of Methods of Psychological Research]. *Ehksperimental'naya psikhologiya v Rossii: tradicii i perspektivy* [Experimental Psychology in Russia: Traditions and Perspectives]. Moscow, Institut psikhologii RAN Publ., 2010, pp. 42–46. (in Russ.).

10. Melas V.B., Salnikov D.I., Gudulina A.O. [The Numerical Comparing of Classical and Permutation Methods of Statistical Hypothesis Testing]. *Vestnik Sankt-Peterburgskogo universiteta. Seriya 1. Matematika. Mekhanika. Astronomiya* [Vestnik of Saint Petersburg University. Series 1. Mathematics. Mechanics. Astronomy]. 2016, vol. 3(61), pp. 415–423. DOI: 10.21638/11701/spbu01.2016.309. (in Russ.)
11. Merlin V.S. *Ocherk teorii integral'nogo issledovaniya individual'nosti* [Outline of the Theory of Individuality Integral Research]. Moscow, "Pedagogika" Publ., 1986. 256 p.
12. Nasledov A.D. *Matematicheskie metody psikhologicheskogo issledovaniya. Analiz i interpretatsiya dannyh* [Mathematical Methods of Psychological Research. Data Analysis and Interpretation]. St. Petersburg, "Rech" Publ., 2008. 392 p.
13. Rusalov V.M. *Temperament v strukture individual'nosti cheloveka: Differentsial'no-psikhofiziologicheskie i psikhologicheskie issledovaniya* [Temperament in the Structure of Human Individuality: Psychophysiological Studies]. Moscow, Institut psikhologii RAN Publ., 2012. 528 p.
14. Sivuha S.V., Kozyak A.A. [On the Reform of Statistical Inference in Psychology]. *Psihologiya. ZHurnal Vysshey shkoly ehkonomiki* [Psychology. Journal of the Higher School of Economics]. 2009, vol. 6(4), pp. 66–86. (in Russ.).
15. Ferster Eh., Renc B. *Metody korrelyatsionnogo i regressionnogo analiza* [Methods of Correlation and Regression Analysis]. Moscow, Finansy i statistika Publ., 1983. 304 p.
16. Shitikov V.K., Rozenberg G.S. *Randomizatsiya i butstrep: statisticheskiy analiz v biologii i ehkologii s ispol'zovaniem R* [Randomization and Bootstrap: Statistical Analysis in Biology and Ecology Using R]. Tol'yatti, "Kassandra" Publ., 2013. 314 p.
17. Breiman L., Friedman J.H. Estimating Optimal Transformations for Multiple Regression and Correlation. *Journal of the American Statistical Association*. 1985, vol. 80(391), pp. 580–598. DOI: 10.1080/01621459.1985.10478157.
18. Bretz F., Hothorn T., Westfall P. *Multiple Comparisons Using R*. Boca Raton, Chapman & Hall/CRC Press, 2010. 205 p.
19. Cohen J. The Earth is Round ( $p < .05$ ). *American Psychologist*. 1994, vol. 49(12), pp. 997–1003. DOI: 10.1037/0003-066X.49.12.997.
20. Cumming G. *Understanding the New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis*. New York, Routledge Publ., 2012. 519 p. DOI: 10.4324/9780203807002.
21. Cumming G. The New Statistics: Why and How. *Psychological Science*. 2014, vol. 25(1), pp. 7–29. DOI: 10.1177/0956797613504966.
22. Deebani W., Kachouie N.N. Ensemble Correlation Coefficient. ISAIM, 2018. Available at: [http://isaim2018.cs.virginia.edu/papers/ISAIM2018\\_Deebani\\_Kachouie.pdf](http://isaim2018.cs.virginia.edu/papers/ISAIM2018_Deebani_Kachouie.pdf). (accessed 25.09.2018).
23. Hunter J.E., Schmidt F.L. *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*. 2nd ed. Thousand Oaks, Sage Publications, 2004. 582 p.
24. Kline R.B. *Beyond Significance Testing: Statistics Reform in the Behavioral Sciences*. Washington, American Psychological Association Publ., 2013. 350 p. DOI: 10.1037/14136-000.
25. McAdams D.P., Pals J.L. A New Big Five: Fundamental Principles for an Integrative Science of Personality. *American Psychologist*. 2006, vol. 61(3), pp. 204–217. DOI: 10.1037/0003-066X.61.3.204.
26. McCrae R.R., Costa Jr.P.T. The Five-Factor Theory of Personality. *Handbook of Personality: Theory and Research*. 3rd ed. New York, NY, USA, Guilford Press, 2008, pp. 159–181.
27. Open Science Collaboration. Estimating the Reproducibility of Psychological Science. *Science*. 2015, vol. 349(6251), pp. 943. DOI: 10.1126/science.aac4716.
28. *Publication Manual of the American Psychological Association*. 6th ed. Washington, DC, American Psychological Association Publ., 2010. 272 p.
29. Reshef D.N., Reshef Y.A., Finucane H.K., Grossman S.R., McVean G., Turnbaugh P.J., Lander E.S., Mitzenmacher M., Sabeti, P.C. *Detecting Novel Associations in Large Data Sets*. *Science*. 2011, vol. 334(6062), pp. 1518–1524. DOI: 10.1126/science.1205438
30. Revelle W. An Introduction to Psychometric Theory with Applications in R. Available at: <http://www.personality-project.org/r/book/chapter4.pdf>. (accessed 26.07.2018).

31. Szekely G.J., Rizzo M.L., Bakirov N.K. Measuring and Testing Dependence by Correlation of Distances. *The Annals of Statistics*. 2007, vol. 35(6), pp. 2769–2794. DOI: 10.1214/009053607000000505.

32. Tellez A., Garcia C.H., Corral-Verdugo V. Effect Size, Confidence Intervals and Statistical Power in Psychological Research. *Psychology in Russia: State of the Art*. 2015, vol. 8(3), pp. 27–46. DOI: 10.11621/pir.2015.0303.

33. Tran D.M. Robustness of Two Formulas to Correct Pearson Correlation for Restriction of Range: Dissertation. Atlanta, Georgia State University Publ., 2011. Available at: [http://scholarworks.gsu.edu/eps\\_diss/84](http://scholarworks.gsu.edu/eps_diss/84). (accessed 15.06.2018).

34. Viole F., Nawrocki D.N. Deriving Nonlinear Correlation Coefficients from Partial Moments. *SSRN Electronic Journal* (September 18, 2012). DOI: 10.2139/ssrn.2148522.

*Received 23 April 2019*

---

### ОБРАЗЕЦ ЦИТИРОВАНИЯ

Калугин, А.Ю. О некоторых методических аспектах исследования индивидуальности человека / А.Ю. Калугин // Психология. Психофизиология. – 2019. – Т. 12, № 2. – С. 29–40. DOI: 10.14529/jpps190203

### FOR CITATION

Kalugin A.Yu. On Some Methodological Aspects of the Study of Human Individuality. *Psychology. Psychophysiology*. 2019, vol. 12, no. 2, pp. 29–40. (in Russ.). DOI: 10.14529/jpps190203

---