

Terra Economicus, 2019, 17(3), 6-29
DOI: 10.23683/2073-6606-2019-17-3-6-29

УПРАВЛЕНИЕ БЕЗ ИЗМЕРЕНИЙ

Виталий Леонидович ТАМБОВЦЕВ,

доктор экономических наук, главный научный сотрудник,
МГУ им. М.В. Ломоносова,
г. Москва, Россия,
e-mail: tambovtsev@econ.msu.ru

Цитирование: Тамбовцев, В. Л. (2019). Управление без измерений // *Terra Economicus*, 17(3), 6–29. DOI: 10.23683/2073-6606-2019-17-3-6-29

Характерной чертой реформ, которые проведены за последние десятилетия в секторе публичных услуг во многих странах, является использование в управлении принципов, перенесенных из сферы бизнеса. Эти реформы известны под общим названием «Новый государственный менеджмент». Одним из их ключевых признаков является установление организациям, оказывающим публичные услуги, количественных заданий и увязка вознаграждений работникам с уровнями исполнения этих заданий. В мировой литературе за эти годы опубликованы результаты большого числа эмпирических исследований последствий таких реформ, демонстрирующих их негативное влияние на качество предоставляемых услуг и стимулы работников. Между тем ситуация не меняется, поскольку руководители ведомств утверждают, что без установления количественных заданий невозможно управлять развитием их отраслей. В статье анализируется корректность таких утверждений. Дается краткая характеристика положений современной (репрезентационной) теории измерения, показывается, что субъективные суждения являются не только «полноправной» разновидностью измерений, но и неотъемлемой частью любых ситуаций принятия решения. Обосновывается различие заданий в коммерческих организациях (фирмах) и некоммерческих организациях, оказывающих публичные услуги. Первые имеют естественную метрику – деньги, а их величины устанавливаются исходя из критерия максимизации прибыли фирмы. Вторые не имеют естественной метрики и ясных критериев установления, в силу чего являются результатами субъективных суждений руководителей отраслей, облеченными в цифровую форму, т.е. иллюзиями количеств. Работники, стремясь выполнить такие задания, фактически работают на показатель, что и приводит к негативным последствиям для качества предоставляемых публичных услуг.

Ключевые слова: измерения; шкалы; суждения; задания; работа на показатель

MANAGEMENT WITHOUT MEASUREMENT

Vitaliy L. TAMBOVTSEV,

Doctor of Economics, Senior Staff Scientist,
Lomonosov Moscow State University,
Moscow, Russia,
e-mail: tambovtsev@econ.msu.ru

Citation: Tambovtsev, V. L. (2019). Management without measurement. *Terra Economicus*, 17(3), 6–29. DOI: 10.23683/2073-6606-2019-17-3-6-29

The public service reforms implemented over the past decades in many countries have a common feature – these reforms, known under the umbrella term of New Public Management (NPM), adopted practices of the business management. NPM principles imply setting quantifiable targets in public organizations. A number of recent empirical studies have shown negative impact of these reforms on service quality and motivation in the public sector. Nevertheless, the situation does not change because the superior authorities argue that it is impossible to manage the sector development without quantified targets. The paper analyses this argumentation. The paper deals with the critical points of contemporary (representative) measurement theory. As the author suggests, subjective judgment is not only a kind of measurement but is an indispensable part of any decision making. Differences between the targets in commercial organizations (firms) and those in public service sector are clarified. The former have a “natural” metric, that is, money, and its values depend on the firms’ profit maximization. The latter do not have “natural” metrics and specific criteria for measuring, so are merely the results of subjective judgments of the agencies’ heads, presented numerically. Virtually, those kinds of targets are not more than illusion of quantities. Striving to achieve those targets, the public servants actually engage in gaming, and this drives down the quality of the public services.

Keywords: measurement; scales; judgement; target; gaming

JEL classifications: C19, D91, H83, M10

1. «Нельзя управлять тем, что невозможно измерить»¹

Утверждение, ставшее заголовком этого раздела и кочующее из учебника в учебник, на первый взгляд представляется совершенно бесспорным. Действительно, как можно управлять, например, самолетом, если не измерены как минимум его скорость, высота над землей и направление движения? Как управлять фирмой, если не измерены ее входящие и исходящие денежные потоки, суммарные объемы заказов на будущее, запасы продукции, сырья и материалов и т.п.? Такого рода риторические вопросы можно ставить сколь угодно долго, и ответы будут очевидны: без измерения действительное управление невозможно!

Однако можно поставить и другие вопросы: что мы измеряем, когда принимаем решение создать семью и/или завести ребенка? Что измеряет молодой человек, решивший пойти учиться в университет, а не поступить на работу? Что измеряют люди, голосующие за одного кандидата, а не за другого? Это также риторические вопросы, но очевидные ответы на них совсем другие: никакого измерения здесь не происходит, хотя важность соответствующих решений для людей (по их последстви-

¹ Приписывается У. Хьюлетту, одному из основателей Hewlett-Packard.

ям) может быть куда выше, чем значимость для них управленческих решений, принимаемых на фирме.

Какой вывод можно сделать из сопоставления приведенных групп вопросов и ответов? С нашей точки зрения, он заключается в следующем: есть такие управленческие решения, качество которых без точных измерений *резко падает* (вплоть до возникновения невосполнимых ущербов), и есть такие управленческие решения, качество которых *не снижается* от отсутствия точных измерений. Не вдаваясь в дискуссию о том, что такое качество решения, отметим, что понимаем его здесь *ex post*, т.е. как получение в результате выполнения решений таких последствий, которые близки к ожидавшимся.

Л. Прусак в этой связи заметил: «Питер Дракер был прав, когда писал: “Что можно измерить, тем можно управлять”. Но почему это утверждение так часто становится ложным королларием типа “Что не может быть измерено, то не может хорошо управляться”?» (Prusak, 2010).

Ответ на этот вопрос, вероятно, связан с неразличением многими людьми смыслов этих двух утверждений, вследствие их незнакомства с формальной логикой и невниманием к логическим «тонкостям», однако *по последствиям* это смешение утверждений, приведенных Прусак, оказывается чрезвычайно значимым. Но, прежде чем перейти к их анализу, имеет смысл привести некоторые положения из теории измерений, поскольку именно на их основе только и можно обсуждать затронутые здесь проблемы.

2. «Измеряй измеримое и делай неизмеримое измеримым»²

Современная теория измерений³ восходит к основополагающим работам середины прошлого века (Stevens, 1946; Suppes & Zinnes, 1963; Суппес & Зиннес, 1967). Она трактует измерение как отображение одной системы в другую, где первая – это *объект* измерения, вещественный или не вещественный, а вторая – *шкала* измерения, т.е. знаковая система, включающая некоторые элементы (знаки разного типа) и отношения между ними. По характеру этих отношений принято различать четыре основных типа шкал: (а) номинальные, (б) порядковые, (в) разностей (интервалов) и (г) отношений, к которым часто добавляются (д) абсолютные шкалы⁴.

В номинальных шкалах единственным отношением между элементами является их эквивалентность (равенство) самим себе. Измерить в номинальной шкале – значит назвать измеряемый объект одним из имен, включенных в шкалу, иначе говоря, отнести объект к одному из классов, классифицировать его. Пример номинальной шкалы, хорошо знакомый экономистам, – классификатор JEL.

В порядковых шкалах между их элементами (градациями), кроме отношения эквивалентности, также определено отношение неравенства (больше, меньше, важнее, предпочтительнее и т.п.); иными словами, эти элементы ранжированы, и их часто называют баллами. Измерить объект в порядковой шкале – значит сопоставить с ним тот или иной балл. Баллы часто обозначают цифрами, однако с баллами нельзя проводить арифметические операции: как известно, двоечник + троечник отнюдь не равны отличнику...

В шкалах разностей (или интервальных) для их элементов определены не только отношение неравенства, но и операции сложения и вычитания. Примером может служить шкала интервалов времени, которые не только упорядочены по величине, но их можно также складывать и вычитать.

В шкалах отношений для элементов определены отношения эквивалентности и неравенства, а также все арифметические операции. Другими словами, эти эле-

² Галилео Галилей.

³ Именуемая также репрезентационной (чаще, но не вполне корректно – репрезентативной), поскольку трактует измерение как *представление (репрезентацию)* объектов измерения теми или иными знаками.

⁴ Сравнительно недавно в научный оборот введено представление о расплывчатых шкалах (fuzzy scales) (Finkelstein, 2003; Benoit & Foulloy, 2013).

менты – действительные числа, включая нуль (представляет отсутствие измеряемого свойства).

Наконец, абсолютные шкалы – это шкалы отношений, в которых существует однозначно определяемая (а не условная) единица измерения. Абсолютные шкалы обычно возникают для величин, являющихся отношением двух других, измеренных в одной и той же шкале отношений, например, для отдачи на инвестиции, где «естественной» единицей измерения выступает процент.

Процесс измерения как процесс отображения измеряемого объекта в ту или иную шкалу может осуществляться двумя разными способами: (1) с помощью специально созданного измерительного прибора и (2) непосредственно человеком («на глаз»). Они различаются по двум основным параметрам: издержкам и качеству, прежде всего – точности измерения. На первый взгляд, может показаться, что бесприборное измерение человеком всегда дешевле, но «зато» всегда менее качественно, однако это не так. Дело в том, что множества потенциальных объектов измерения (вещей, их свойств, отношений между ними, процессов и т.п.), для которых каждый из этих способов является наиболее подходящим, лишь *частично* пересекаются. Есть объекты измерения, для восприятия которых у человека либо просто нет соответствующих органов чувств, либо диапазон восприятия слишком узок для измерения интересующей величины. Примером могут служить радиоволны, магнитные поля и т.п. Одновременно есть объекты измерения, для восприятия информации о которых отсутствуют специально созданные приборы, хотя люди мгновенно и без особых издержек проводят измерение. Пример – определение того, можно или нельзя доверять какому-то индивиду, происходящее за очень короткое время визуального контакта (Todorov, Pakrashi & Oosterhof, 2009).

Результаты измерения вторым способом обычно называют *оценками*, а иногда, когда оценки даются опытными и знающими в соответствующей области людьми, – *экспертными оценками*. Особое название для результатов неприборного измерения подчеркивает, как правило, его предварительный, неточный, субъективный характер в сравнении с «настоящим» приборным измерением.

Однако такое противопоставление во многом является *надуманным*. Дело в том, что приборное измерение – сравнение объекта с единицей той или иной шкалы отношений – в реальном мире происходит в физической среде, постоянно оказывающей случайные воздействия как на прибор, так и на объект измерения. Соответственно, результаты разных измерений объекта одним и тем же прибором каждый раз оказываются *разными*. «Это фундаментальное положение является законом природы. На основании громадного опыта практических измерений, накопленного к настоящему времени, может быть сформулировано следующее утверждение, называемое *основным постулатом метрологии*: отсчет является случайным числом. На этом постулате, который легко поддается проверке и остается справедливым в любых областях и видах измерений, основана вся метрология» (Шишкин, 1990: 43)⁵.

Поэтому представление о том, что результаты приборных измерений, в отличие от неприборных оценок, точны и объективны, является на деле заблуждением. Фактически вся классическая метрология, имеющая дело с приборами и шкалами отношений, посвящена разработке и обоснованию различных методов борьбы с ошибками.

Есть, однако, различие между приборным измерением и оценками (в том числе экспертными), которое значительно важнее, чем соотношение их точностей, но на которое не всегда обращают внимание. «Идеальная» система измерения должна обладать двумя свойствами: во-первых, она должна обеспечивать объективность, т.е. производить информацию, относящуюся только к объекту измерения, но не к его среде; во-вторых, результаты измерения должны иметь межсубъектный характер, т.е. произво-

⁵ См. также современные издания (Снежко, Захарова & Жирнова, 2008: 50; Родинков, Бокач & Булатов, 2010: 12 и др).

димая информация должна быть одной и той же для разных субъектов (Mari, 2005: 681).

Поскольку, как отмечено выше, на результат любого измерения (отсчет) влияет множество случайных факторов, будь то приборное или экспертное измерение, требование объективности может одинаково нарушаться для обоих типов систем измерения, как приборных, так и неприборных. Иначе обстоит дело с требованием межсубъектности. Поскольку приборы выдают информацию в шкалах отношений, то как единицы, так и результаты измерений в таких системах воспринимаются едино всеми индивидами: один сантиметр или метр значит для всех одно и то же⁶. Если в экспертной системе используются шкалы отношений (например, измерение длины на глаз), результат также оказывается межсубъектным, только менее точным («длина где-то около метра»). Однако, если экспертная система измерения использует порядковую или номинальную шкалу, межсубъектность может оказаться нарушенной, поскольку разные люди могут использовать разные «единицы измерения» (эталон, точки отсчета и т.п.) для оценки объекта измерения.

Поясним это на примере. Для оценки того, можно ли доверять какому-то человеку, разные индивиды опираются на несовпадающие *личные концепты доверия* (Тамбовцев, 2018), в силу чего результат оценки (скажем, «да, ему можно доверять») одним индивидом может не совпадать с результатом оценки этого человека другим индивидом.

Схожие ситуации возникают при измерении степени удовлетворенности людей качеством административных услуг, удовлетворенности студентов качеством преподавания, обобщенного доверия и т.п. Подчеркнем, что причина отсутствия межсубъектности во всех подобных случаях не в том, что люди по-разному понимают смысл задаваемых им вопросов, а в том, что у всех них разные «единицы измерения», или «эталон» оценки. Тот факт, что в соответствующих опросниках указывается «длина» шкалы и даются пояснения типа «1 – самый низкий уровень удовлетворенности, 10 – самый высокий», не устраняет различий в индивидуальных «длинах» единицы субъективного измерения, хотя и может несколько снизить их. Это означает, что проводить с результатами таких замеров арифметические операции (например, рассчитывать среднюю оценку преподавателя студентами) **нельзя**: это все равно что складывать 5 дюймов и 2 сантиметра и, поделив потом 7 пополам, получать ответ: 3,5 – правда, неизвестно чего...⁷

Означает ли сказанное, что упомянутые выше примеры показателей (удовлетворенности, обобщенного доверия и т.п.), получивших широкое распространение на практике, *в действительности не являются измерителями* того, что они призваны измерять? Да, означает. А являются в этой связи соответствующие оценки *вообще измерениями*? Да, являются – но измерением, которое может применяться (и применяется) исключительно в *личных* целях, в силу отсутствия межсубъектности производимой в его результате информации. Для того чтобы подобные оценки можно было (например) усреднять и высчитывать обобщенные оценки преподавателей, организаций, предоставляющих административные услуги, органов управления и т.п., необходимо прежде **как-то** добиться того, чтобы соответствующее свойство объекта измерения все «измерители» понимали одинаково, а главное – имели совпадающие (или очень близкие) «точки отсчета» и «единицы измерения». Без этого «усредненные» оценки просто лишены смысла (хотя «исходные данные» для них вполне осмысленны)⁸.

⁶ Если в стране используются другие единицы длины, например, дюймы и ярды, они пересчитываются в метры и сантиметры *единственным* образом.

⁷ Этот запрет не следует смешивать с запретом проводить арифметические операции с баллами как результатами неприборных измерений, осуществляющихся в *порядковых шкалах*: баллы внешне похожи на числа, но *числами не являются!*

⁸ «Принципиальная трудность состоит в том, что обычные инструменты статистики, нормально используемые для оценки точности измерения и его неопределенности, вероятно не будут работать надежно, если отношения между расстояниями различных частей шкал не полностью известны, что типично для порядковых и других качественных измерений» (Pendrill & Petersson, 2016: 1).

Естественным способом повышения уровня однородности оснований для оценки восприятий событий внешнего и внутреннего мира выступает коммуникация – *диалог*, обмен информацией в процессах «физического» взаимодействия индивидов, обеспечивающая взаимное обучение его участников. Диалог *не гарантирует* сближения «точек отсчета» оценок, но он является единственной надежной основой для такого сближения, для формирования взаимопонимания его участников. Альтернативой диалогу является индоктринация – *пропаганда*, т.е. попытка убедить одну группу индивидов (получателей сообщений) в том, что точка зрения другой группы (отправителей сообщений) «лучше», чем позиция первой группы. Мировой опыт показывает, что при использовании специальных «технологий убеждения», в частности, умелом использовании нарративов (de Graaf et al., 2012; Hoeken, Kolthoff & Sanders, 2016), на какое-то время можно внушить значительному числу людей ту или иную идею, однако если она не соответствует реальности или ее реализация приносит им прямой ущерб, то, как правило, в мировоззрении «распропагандированных» в результате естественного обучения происходят изменения. Тем самым в конечном счете (правдивый) диалог остается наиболее надежным способом обеспечения взаимопонимания людей.

Иногда стремление повысить объективность и межсубъектность результатов оценивания приобретает странные формы, основанные на непонимании существа проблемы. Так, при оценке научной значимости результатов исследований с содержательной точки зрения наиболее надежными являются мнения коллег, работающих в соответствующей области науки: ведь именно они понимают, что означает полученный кем-то результат для развития того или иного направления. Однако чиновники, призванные обеспечить организацию и финансирование исследований, полагая такие мнения «субъективными», считают более подходящими для «объективной» оценки показатели цитирования и производные от них (например, индекс Хирша). Но, как справедливо отмечают (Крупина & Клочков, 2014: 15), «когда рецензенты и редколлегии журналов принимают решения о публикации статей данного автора – это экспертные решения. И когда читатели научных журналов ссылаются на статьи данного автора, их решения также являются экспертными». Нужно отметить также, что выбор работ для цитирования определяется отнюдь не только их качеством (Bornmann & Daniel, 2008: 46–49). Поэтому показатели, которыми наука, с подачи бюрократов, «отчитывается перед обществом», фактически не отражают ее реальные достижения и влияние на общество и экономику (Glänzel, 2008).

Формирование оценок внешней среды и своего внутреннего состояния – деятельность, которой постоянно занят наш мозг. Ее непрерывность – основа выживания организма в природной и социальной среде, поэтому обработка сигналов, поступающих от органов чувств, и их сопоставление с различными «эталоном», характеризующими границы допустимого для того, чтобы выдать сигнал опасности в случае необходимости, происходят не только в течение активного периода суточного цикла, но и в состоянии сна. А. Рок приводит следующие слова М. Уилсона, нейроисследователя из Массачусетского технологического института: «Мозг постоянно оценивает новый опыт для того, чтобы определить, соответствует ли он ментальным моделям, построенным на основе предыдущего опыта, проверяя, насколько хорошо эти модели работают для предсказания новых событий и для ориентации решений» (Rock, 2004: 85).

Иными словами, *мозг постоянно занят измерениями*, поэтому максима Галилео Галилея, вынесенная в заголовок этого раздела статьи, может считаться таковой только в узком смысле измерения в *шкалах отношений*: ведь измерение в номинальных и порядковых шкалах *физически встроено* в наш мозг в ходе эволюции. Слабые стороны этих типов измерений (оценок) вкратце охарактеризованы выше, однако следует ли из них, что лозунг, распространенный среди датских менеджеров, занятых управлением качеством, «Одно измерение лучше, чем 1000 мнений» (Larsen, 2001) действительно имеет под собой основания? Для ответа на этот вопрос необхо-

димо рассмотреть роль измерений, количественных данных и расчетов в принятии управленческих решений.

3. «Что бы менеджер ни делал, он делает это путем принятия решений»⁹

Нормативная теория принятия решений, нацеленная на то, чтобы улучшить качество решений, предполагает проведение компьютерных расчетов на основе моделей той или иной степени сложности, требующих измерения различных переменных в порядковых шкалах, а идеально – в шкалах отношений. Дескриптивная теория принятия решений, анализирующая разнообразные *практики* их принятия, отмечает, что большая часть решений, от бытовых (Johnson & Puto, 1987; Reisen et al., 2008) до масштабных политических (Yetiv, 2013), принимается людьми на базе их субъективных интуитивных *суждений* (judgments) (Kahneman, 2003; Bazerman, 2010). Принятие решений лидерами бизнеса также не является исключением (Sinclair & Ashkanasy, 2002). К. Хэммонд, характеризуя суждения, утверждал, что им присущи неснижаемая неопределенность (irreducible uncertainty), неизбежная ошибочность (inevitable error) и неустраняемая несправедливость (unavoidable injustice) (Hammond, 1996). В решения, принимаемые людьми, в том числе в управленческие решения, свой вклад вносят также эмоциональные импульсы (Strack & Deutsch, 2004; Summers & Duxbury, 2012). Управленческим решениям присущи иллюзии, заблуждения и психологические уклоны (Barnes, 1984; Schwenk, 1985), часть из которых связана, согласно (Abatecola et al., 2018), с использованием эвристики (Tversky & Kahneman, 1974; Busenitz & Barney, 1994; Gigerenzer & Gaissmaier, 2011). Преодолению уклонов в принятии управленческих решений исследователи уделяют большое внимание (Keren, 1990; Morewedge et al., 2015).

Суждения, становящиеся основой решений, формируются также на основе интуиции (Hodgkinson et al., 2008), причем качество решений повышается, если она сочетается с аналитическим стилем мышления (Rusou et al., 2013; Abadie & Waroquier, 2019). Такая сочетаемость хорошо объясняется на основе теории дуальных процессов (Basel & Brühl, 2013) и может иметь, по мнению ряда ученых, эволюционные корни (Laurie & Rosati, 2015).

При всем том значении, которое играют суждения в процессах принятия управленческих решений, они – по общему мнению – были и остаются *неточными и субъективными*. Как же суждения могут играть столь важные роли в управлении? Частично мы ответили на этот вопрос выше: оценки – это разновидность измерения, т.е. части информационного обеспечения принятия решений. Здесь же подчеркнем, что это разновидность измерения, поставляющая в концентрированном виде информацию, которая не может быть произведена никаким приборным измерением (по крайней мере, при нынешнем уровне наших знаний).

П. Элбоу выделил три типа суждений: ранжирование (ranking), оценивание (evaluating) и одобрение (liking) (Elbow, 1993). *Ранжирование* – это представление обобщенного суждения об объектах (людях или процессах) в виде числа или оценки, позволяющих упорядочить эти объекты. *Оценивание* – это количественная или качественная характеристика каких-то свойств объекта. Наконец, *одобрение* – это выражение соответствия (или несоответствия в случае неодобрения) объекта предпочтениям субъекта, высказывающего суждение. Легко видеть, что ранжирование и одобрение – это те типы измерения, которые сегодня *не могут стать приборными*, поскольку они отражают *цели и интересы* субъектов оценки¹⁰. Вместе с тем они составляют важные, *неисключаемые* компоненты любого решения.

Важно также подчеркнуть, что *задания*, которые вполне могут иметь количественное измерение в виде тех или иных показателей и устанавливаются агентам «сверху»,

⁹ П. Драйкер (“Whatever a manager does he does through making decisions” – P.F. Drucker).

¹⁰ Более того, данные приборного измерения вполне могут быть объектами ранжирования и/или одобрения, в силу чего суждения можно считать (потенциальным) измерением «более высокого порядка», чем приборное измерение.

т.е. их принципами, *не заменяют* цели и интересы субъекта (агента). Ведь получив задание, агент (если он не полностью контролируем принципалом) имеет возможность выбирать, *как* ему выполнить задание. Работа на показатель (*gaming*), когда «буква» задания выполняется способом, который может иметь вредоносный характер для принципала и одновременно позволяет реализовать собственные цели агента, – типичная ситуация, возникающая в очерченных условиях.

Представления о *точности* суждений – в части столь важных для менеджмента *суждений о других людях* – прошли за последние полвека весьма интересную эволюцию. Д. Фандер отмечает, что в течение 1970–1980-х гг. преобладающей в исследовании социальных суждений была «парадигма ошибочности» («*error paradigm*»), ее девиз: «Беги от ошибок!» (*Shun error!*). «Поиск ошибок определил важный подход к изучению того, как люди оценивают другие личности, равно как и человеческих суждений, более широко» (Funder, 1995: 652). В качестве работ, характерных для этого подхода, он называет (Kahneman & Tversky, 1973; Ross, 1977; Nisbett & Ross, 1980).

Названные (и многие другие) работы, демонстрировавшие массовую неточность и приблизительность суждений о других людях, базировались на лабораторных экспериментах, так что именно эксперименты породили широко распространившуюся убежденность в преобладающей ошибочности суждений. В этой связи важную роль сыграла работа (Funder, 1987), в которой было показано, что ошибки в экспериментальных ситуациях (*errors*) вовсе не равнозначны заблуждениям в реальных жизненных ситуациях (*mistakes*): «одно и то же суждение, неверное по отношению к лабораторным стимулам, может быть правильным в более широком социальном контексте и отражать процессы, ведущие к точным суждениям в обычных обстоятельствах» (Funder, 1987: 76). Последующие исследования, выполненные в рамках «парадигмы точности» («*accuracy paradigm*»), продемонстрировали, что социальные суждения вовсе не так неизбежно ошибочны, как полагалось ранее. Прежде всего, в развитие своего подхода Д. Фандер предложил «реалистичную модель точности» суждений (*Realistic Accuracy Model, RAM*), которая связала точность социальных суждений с четырьмя основными факторами: характеристики задачи оценки; оцениваемые черты личности; количество и качество информации, на которой основывается суждение; опыт индивида, делающего суждение. Как отмечает Фандер, «людям обычно удается делать суждения о личности, которые достаточно точны, чтобы ориентироваться в сложном социальном мире» (Funder, 2012: 177). В русле этого подхода было показано, что интуитивные и аналитические суждения основаны на общих принципах, различаясь методами обработки имеющейся информации (Kruglanski & Gigerenzer, 2011), а вероятностные суждения оказались «на удивление рациональными» (*surprisingly rational*), если учитывать наличие шума при передаче и обработке сигналов (Costello & Watts, 2014). Более того, анализ продемонстрировал, что в ряде ситуаций интуитивные суждения дают более качественные решения, чем те, которым предшествует аналитическое размышление (Jussim, 2017; Maldonato et al., 2018), а свою недавнюю книгу, посвященную принятию решений в реальном мире, Г. Гигерензер назвал «Просто рациональные» (Gigerenzer, 2015).

Проблема точности суждений изучалась не только применительно к оценке черт личности, но и к более широкому кругу суждений, в частности к предсказаниям. Если в течение долгого времени распространенным было убеждение в том, что суждение – «враг точности, то сегодня суждения понимаются как неотъемлемый компонент предсказаний, и исследователи уделяют много внимания пониманию и улучшению их использования. Человеческие суждения могут обеспечить значительные улучшения точности предсказаний, но также могут быть предметом многих искажений» (Lawrence et al., 2006).

Таким образом, хотя неточность суждений (прежде всего, о других людях) оказалась сильно преувеличенной, сама по себе их качественность и неопределенность ни-

куда не исчезли. Как это сказывается на решениях, принимаемых на базе суждений, и последствиях таких решений?

Формируя суждения, люди *стремятся* к той степени их точности, которая позволяет принимать решения, последствия которых не причинили бы им вреда. Разумеется, это удастся не всегда, однако усилия, посвящаемые улучшению суждений и решений, обычно зависят от величины ожидаемых выгод и/или ущербов. Соблюдение такого соотношения – один из признаков *рациональности поведения* в условиях ограниченной информации и ограниченных возможностей ее обработки. Поэтому неточность – совершенно нормальное явление в подавляющем большинстве и суждений, и решений. Только там, где ожидаемые ущербы велики (тем более невозполнимы), люди уделяют значительное внимание анализу проблемной ситуации, измерению ее компонентов и т.п. Поэтому, например, сплетни и слухи являются вполне достаточной (по точности и надежности) информацией для принятия решений (Mitchell et al., 2016; Fonseca & Peters, 2018). Более того, исследователи предположили и нашли подтверждения тому, что неточная социальная информация может способствовать кооперации (Rauwolf et al., 2015; Rauwolf, 2016), хотя вопрос, безусловно, заслуживает дальнейшего изучения.

Формирование (более или менее) корректных суждений основано на присущей мозгу способности к статистическому (неявному, имплицитному) обучению, т.е. «способности с течением времени извлекать регулярности из окружающей среды» (Schariro & Turk-Browne, 2015: 501). Эта способность заключается в восприятии (подчас неявных, неосознаваемых) «сигналов» и «подсказок» (cue) из внешней среды (Sacco & Brown, 2018) и их обработке в определенных отделах мозга. Выявление регулярностей происходит независимо от отсутствия или наличия намерений их выявлять или повышенного внимания (awareness) к происходящему вокруг (Turk-Browne et al., 2009). Способность неявного обучения присуща всем людям, хотя и проявляется в неравной степени (Siegelman et al., 2016).

При этом эксперты не всегда имеют преимущество перед неспециалистами: обладая большими знаниями и опытом, они могут переносить в новые ситуации уже известные им подходы, что может вести к нелучшим решениям в нестандартных случаях (Heerkens et al., 2011). В пользу того, что в основе суждений лежат неосознанно выявленные регулярности (в частности связи между наблюдаемыми «подсказками» и обобщенными свойствами, такими как черты личности), свидетельствует факт, что стремление оправдать (объяснить и обосновать) суждения не повышает их точности (Hoffmann et al., 2017). Об этом же говорят эксперименты, показывавшие, что машинное обучение позволяет получить более точные суждения о чертах личности, чем те, которые формируются людьми (Youyou et al., 2015). Ведь специфика машинного обучения в том, что модель связи переменных вырабатывается компьютерной программой, а не задается заранее человеком, и именно так в мозгу выявляются регулярности. Подобным образом устраивается и обработка «больших данных» (Kosinski et al., 2016), так что машинное обучение следует путем, проложенным миллионы лет назад эволюцией (но, вероятно, с использованием более эффективных алгоритмов).

Таким образом, *человеческие суждения* (измерения в номинальной и порядковой шкалах) – *обязательный элемент управленческих решений*. Сводить последние исключительно к приборному измерению и шкалам отношений – значит выбрасывать из решений их «двигатели» – предпочтения, цели и намерения, без которых решений просто не бывает.

4. «Что измерено, то и важно»¹¹

Когда речь заходит об измерениях в организациях, возникают два тесно связанных вопроса: что измерять и для чего измерять? Ведь любое измерение (исключая интуитивные суждения) требует издержек – затрат труда, денег и других ресурсов, которые

¹¹ Г. Бивейн, К. Худ (“What’s measured is what matters”).

можно было бы использовать на другие цели, возможно, более значимые. Ответ на первый вопрос, если давать его в общем виде, достаточно очевиден: измерять следует то, что необходимо для управления деятельностью организации, включая то, что требуют от нее стейкхолдеры. Если спрос на информацию со стороны стейкхолдеров обусловлен их составом и действующими формальными и неформальными нормами (например, утвержденной методикой бухгалтерского учета, требованиями налоговых органов, законодательством об акционерных обществах, интересами инвесторов и т.п.), то внутренний спрос на измерения обусловлен особенностями технологии производственной деятельности и процессами менеджмента, действующими в организации.

Производящие организации (т.е. те, чья продукция не потребляется полностью их же членами) бывают, как известно, двух основных типов – коммерческие и некоммерческие¹². Продукция первых продается на рынках по складывающимся на них ценам¹³, вторых – распределяется по другим основаниям. В любой такой организации, членством более одного человека, необходимо обеспечить координацию действий участников для успешного производства продукции. Если число участников невелико и их права одинаковы, координация возможна в рамках прямого диалога; если же хотя бы одно из этих условий не выполняется, координация обычно обеспечивается механизмом предписания, или *иерархии*.

Как известно, иерархия подготавливает координацию действий следующим образом: на ее верху формируется план действий, определяющий целый ряд распределений ресурсов, потребляемых используемой в организации технологией, – материально-вещественных, энергетических, финансовых и, конечно, трудовых. Эти распределения должны быть внутренне согласованы в соответствии с нормами потребления ресурсов на единицу производимой продукции, обеспечивая каждое рабочее место соответствующим комплектом ресурсов. Каждый работник, занимающий то или иное рабочее место, получает *задание* по использованию ресурсов, имеющее характер или определенного *числа* (собрать за рабочее время *X* агрегатов стандартного качества, совершить *Y* операций по установленным правилам и т.п.), или некоторого *качественного требования* (собрать *как можно больше* агрегатов стандартного качества, совершить *как можно больше* операций по правилам и т.п.). В зависимости от уровня выполнения задания, измеряемого теми или иными *показателями*, работник получает определенное вознаграждение по заранее установленным правилам, которое и выступает стимулом к его трудовым усилиям.

Если бы *все* требования к работнику *исчерпывающе* выразались *одним* числом, очерченная схема работала бы весьма продуктивно (в пределах, определяемых качеством планов). Однако требования к работникам крайне редко (если вообще) могут быть выражены *единственным* требованием – числовым либо качественным. Как правило, эти требования *множественны*, что порождает практически непреодолимые трудности для вершук организации иерархий.

В начале 1990-х гг. Б. Олмстром и Р. Милгром строго доказали, что в ситуации множественных заданий, часть из которых выражается числом, а часть имеет качественный характер (причем вознаграждение завязано на уровне выполнения *количественных* показателей), работники, у которых под контролем сверху находится не весь процесс их работы, а только его *результаты*¹⁴, внимание и усилия направляют на выполнение количественных требований, пренебрегая исполнением качественных (Holmstrom & Milgrom, 1991). Иначе говоря, с точки зрения работника, в составе результатов его трудовых усилий важны только те, что *измеряются* и передаются его начальнику, поскольку последний на основе получаемых показателей достойно по-

¹² Так называемых гибридных организаций мы для простоты касаться не будем.

¹³ Случаи монопольных и/или регулируемых цен мы также не рассматриваем.

¹⁴ А это типичная ситуация, поскольку любой контроль требует издержек, и *полный контроль* требует *очень больших* издержек.

ощряет (или не поощряет) этого работника (Bevan & Hood, 2006). Поскольку задания работнику вовсе не заменяют для него его собственных целей, он организует свой труд так, чтобы в основном реализовать собственный интерес и лишь по остаточному принципу исполнить задание, причем не все задание, а только требуемые значения показателей, которые ему были установлены.

Нужно подчеркнуть, что Олмстром и Милгром не были *открывателями* этого явления, они только дали его научное объяснение, основанное на асимметрии информации. Само же это явление существовало (существует и будет существовать) во всех иерархиях, где отсутствует полный контроль действий работников, а задания имеют сложный характер, сочетая количественные и качественные компоненты. В плановых экономиках, таких как экономика СССР и его сателлитов, оно было известно как *работа на показатель*, в рыночных экономиках англоязычных стран – как *гейминг* (gaming) (Hood, 2006)¹⁵.

На практике преодоление негативных последствий работы на показатель шло двумя путями. Во-первых, путем подбора систем измерителей (показателей), максимально полно отражающих все значимые стороны требуемого результата труда работника, во-вторых, посредством расширения состава стимулов для его усилий, которые заинтересовывали бы его исполнять, и качественные, неизмеримые стороны и свойства результатов его усилий. Примером первого может служить сбалансированная система показателей (Kaplan & Norton, 1992), сбалансированная в том смысле, что отражает не только финансы фирмы, но и ее отношения со стейкхолдерами, примером второго – нашумевшие когда-то «кружки качества» (Tang et al., 1987).

Характеризуя первый путь, заметим, что он применим фактически только для *организаций*, а не для отдельных работников, поскольку (а) трудоемок и, соответственно, затратен; и (б) не способен решить проблему вычленения индивидуального вклада в случае взаимодополняемости усилий разных работников для получения результата, который только и может получить осмысленную стоимостную оценку на рынке. Второй путь также имеет естественные ограничения, поскольку внутренние мотивации плохо сочетаются с внешними, денежными (Benabou & Tirole, 2003). Тем не менее на этом пути заметны различные движения: идея «корпоративного гражданства» (Crane et al., 2008), «расширение возможностей» работников (Conger & Kanungo, 1988), формирование «психологического владения» (Pierce et al., 2001) и т.п. В целом же второй путь, снижающий «давление сверху» на работника «в обмен» на апелляцию к его внутренним мотивам, является наиболее продуктивным способом решения проблемы работы на показатель.

Этот путь, как представляется, выступает *единственным* способом ее решения в сфере производства такой продукции, как *услуги*, где возможности *прямого* приборного измерения их качества весьма ограничены, а использование косвенных (невалидных) измерений неизбежно *порождает* работу на показатель. Ведь если поощрение связано с такими показателями, то, достигнув их (но отнюдь не достигнув *действительного* качества), работник вправе требовать вознаграждения: показатели достигнуты! Поэтому, например, расширение возможностей работников весьма продуктивно проявляет себя именно в сфере услуг (Fulford & Enz, 1995).

5. «Если вы даете менеджеру количественное задание, он будет его выполнять, даже если в процессе выполнения должен будет разрушить компанию»¹⁶

Итак, работа на показатель является неизбежным следствием расхождения заданий и собственных целей работников. Будучи контрпродуктивной с точки зрения эффективности использования ресурсов (ведь ресурсы на «достижение показателей»

¹⁵ Одна из первых работ, посвященных анализу этого явления в рыночных экономиках, была опубликована в 1975 г. (Керт, 1975).

¹⁶ Эдвардс У. Деминг.

расходуются, а результат, *нужный потребителю*, не производится), в коммерческой сфере она достаточно быстро ограничила использование схем «управления по целям» и стимулирующих контрактов теми отраслями, где их выгоды превышают наносимый ими ущерб. Регулятором здесь выступил сам коммерческий характер организаций: ведь для них существовали и существуют *объективные измерители конечного результата*: прибыль, доходность, отдача на капитал или инвестиции и т.п. Именно забота о росте этих результатов определяет содержание внутрифирменных планов, включая количественные задания работникам: последние должны сделать столько продукции, чтобы эти результаты повысились (или хотя бы не снизились, имея в виду действия конкурентов). Конечно, расчет плановых заданий может быть ошибочным, но сам подход к их определению вполне понятен.

Названные схемы в конце 1980 – начале 1990-х гг. в целом ряде стран были перенесены в госсектор, в сферу оказания государственных услуг – муниципальное управление, здравоохранение, образование, социальную защиту, фундаментальную науку и т.п. (Gruening, 2001; Diefenbach, 2009), где объективные измерители конечных результатов, дающие основания для расчета заданий, *отсутствуют*. Обсуждение причин и динамики распространения этого подхода выходит за рамки задач этой статьи, отметим лишь, что попытки выявить позитивные последствия этого подхода оказывались в целом неудачными (Propper & Wilson, 2003; Verbeeten, 2008; Sanger, 2013; Dobija et al., 2019).

В качестве подтверждения продуктивности «управления по цифрам» в сфере государственных услуг можно встретить следующую логику ее оценки: после того как в отрасли ввели целевой показатель *П*, отчетные значения *П* повысились. Стало быть, установление целевых показателей положительно сказывается на деятельности отрасли (см., например, Bolli & Somogyi, 2011). Последователи такой логики то ли не видят, то ли не хотят видеть лукавый характер своих выводов: ведь, чтобы он был действительно обоснованным, нужно доказать, что именно *П* адекватно отражает *содержание* (смысл, качество, функцию и т.п.) *деятельности* организаций отрасли.

Но проблема в секторе госуслуг никогда не состояла в том, что организации не исполняют заданий, она заключалась и заключается в *качестве* этих услуг и *эффективности* их предоставления. А устанавливаемые задания – поскольку это количественные показатели – отражали и то и другое только *косвенно*, приводя к тому, что *задания исполнялись, но содержательные функции* организаций оставались обычно невыполненными. Поэтому неудивительно, что было опубликовано несколько работ, содержащих в той или иной форме в заголовке или тексте идиоматическое выражение «Hitting the target but missing the point», что в вольном переводе означает: «Задание выполнено, но задача провалена» (Radnor, 2008; Pugalis, 2013; Curtis, 2015; Vogsnes, 2018 и др.). Действительно, профессионалы (а в отраслях госсектора, где введено «управление по целям», работают именно профессионалы: медики, ученые, профессора), оценивая как количественные задания, так и непредвиденные (чиновниками, но не профессионалами) последствия их выполнения, уверенно говорят о негативных последствиях для качества.

Тем самым, на первый взгляд, *прямое, но качественное* измерение профессионалов оказывается в противоречии с *косвенным, но количественным* измерением чиновников. Однако так ли это?

Как мы подчеркнули в конце раздела 3, *субъективные качественные суждения* относительно целей и предпочтений – неотъемлемая часть любого решения. Точно такой же характер имеют *задания* в секторе госуслуг. Например, перед сектором высшего образования в Пакистане правительством поставлена задача: добиться того, чтобы не менее пяти университетов вошли в сотню лучших мировых вузов (Shoaiб & Mujtaba, 2018). Почему пять, а не четыре или шесть? Почему в сотню лучших, а не в первые 95 или 105? Ответы очевидны: это задание в *количественном отношении* имеет *совер-*

шенно произвольный характер, создавая видимость строгого измерения и будучи по существу субъективным качественным суждением.

Между тем для тех, кому подобные задания устанавливаются, они выливаются в череду унижений, издевательств и оправданий (Heffernan & Heffernan, 2018), при том что сами рейтинги университетов весьма ненадежны, неустойчивы и зависят от произвольно устанавливаемых весов отдельных показателей (Rauhvargers, 2014) и почти не учитываются абитуриентами при выборе будущего места обучения (Higher Education Research Institute, 2007). Важно также подчеркнуть, что чиновники, устанавливающие подобные задания, обычно не несут какой-либо ответственности за их исполнение, включая предоставление должного ресурсного обеспечения, перенося всю ответственность вниз. Не несут они ответственности и за обоснованность заданий, никак не объясняя, почему задания именно таковы по величине.

6. «Поднято ярости масс – 3»¹⁷

Произвольность подавляющего большинства заданий в отраслях госуслуг, управление которыми перестроено в духе Нового государственного менеджмента, и их числовой характер, создающий иллюзию измерения, не оказывали бы негативного влияния на функционирование организаций этих отраслей, если бы такие задания имели характер *общих ориентиров*, задающих направление усилий руководителей и работников организаций. Тогда отчетные данные говорили бы о том, соответствуют ли усилия установленным ориентирам, «верной ли дорогой» идет организация и т.п., а руководители имели бы все основания для проведения аналитической работы внутри организации для того, чтобы улучшить ее соответствие этим ориентирам.

Казалось бы, что мешает наладить такую работу в иных условиях, когда неисполнение задания выступает основой для наказаний? Аналитическая работа и выработка улучшающих мер могут привести к позитивным изменениям спустя достаточно длительный срок, в то время как наказания могут последовать значительно раньше, в силу чего усилия главным образом направляются на работу на показатель. В таких условиях интерес к аналитическому использованию отчетной информации обусловливается преимущественно наличием у руководителей организаций *мотивации служения обществу* (public service motivation) (Moynihan & Pandey, 2010). Не следует также забывать и о том, что разные задачи требуют и разных измерителей одних и тех же процессов (Behn, 2003). Для того чтобы данные приборных измерений и оценки позволяли улучшить результаты, они должны позволять связывать ресурсы и результаты, т.е. отражать компоненты технологии производства результатов. Только на такой основе можно улучшать соответствие ресурсных компонентов по их количеству и качеству, выявлять «узкие места» и оптимизировать расходы ресурсов. При этом нужно четко понимать, что технологии оказания значительной части госуслуг, за исключением самых простых, типа выдачи справок, включают значительные компоненты неопределенности (услуги образования, здравоохранения), так что точные расчеты расходования ресурсов здесь невозможны (если, конечно, ориентироваться на получение *содержательных* результатов, а не на выполнение произвольно установленных «количественных» заданий).

Возвращаясь к названию этой статьи и исходя из проведенного анализа, можно сделать следующие выводы. Если понимать под измерением только приборное измерение в шкалах отношений, то управление без такого измерения не только возможно, но и неизбежно, поскольку практически весь «целевой блок» решений основан на оценках, представленных в номинальных и порядковых шкалах. Появление в составе целей *настоящих чисел*, а не качественных и порядковых оценок, выраженных цифрами, но не являющихся числами, возможно только на базе *обосновывающих расчетов*. Такие расчеты должны показать, что именно данное число является наиболее адекватным выражением предпочтений субъекта решения.

¹⁷ И. Ильф и Е. Петров

Если же понимать под измерением то, что изучается в репрезентационной теории измерений, то управление без измерений, конечно же, невозможно. Однако в рамках этой теории нет дискриминации неприборных измерений, так что *суждения* трактуются так же, как и измерения градусником или вольтметром. К сожалению, эта теория неизвестна подавляющему большинству чиновников, и среди них преобладает противопоставление «настоящего» измерения, которое выражается цифрой¹⁸, и «субъективных оценок»¹⁹. Первые считаются «хорошими», «правильными», а вторые – «неправильными», от которых следует избавляться и не использовать в «серьезных документах».

Практическим следствием этих заблуждений является то, что чиновники предпочитают использовать «строгие количественные» (а часто на самом деле цифровые, а не числовые, количественные), но невалидные показатели для *оправдания* управленческих решений, которые в действительности принимаются на основе субъективных оценок. Использование цифр выступает при этом как способ *снятия с себя ответственности* за (возможно) ошибочное по существу решение.

Таким образом, некорректное понимание измерения и его роли в управлении, помноженное на присущее людям доверие к числам и цифрам, оказывает в отраслях социально важных государственных услуг негативное влияние как на качество самих услуг, так и на работников, предоставляющих эти услуги.

ЛИТЕРАТУРА

- Крупина, С. М., & Ключков, В. В. (2014). Перспективы российской фундаментальной науки в условиях институциональных реформ: моделирование и качественные выводы, с. 11–24 / В: *Материалы 17-х Друкеровских чтений «Инновационные перспективы России и мира: теория и моделирование» (Москва, 2014)*. Новочеркасск: ЮРГТУ (НПИ).
- Родинков, О. В., Бокач, Н. А., & Булатов, А. В. (2010). *Основы метрологии физико-химических измерений и химического анализа: Учебно-методическое пособие*. СПб.: ВВМ.
- Снежко, А. А., Захарова, Н. В., & Жирнова, Е. А. (2008). *Общая теория измерений: Учеб. пособие*. Красноярск: Сибирский государственный аэрокосмический университет.
- Тамбовцев, В. Л. (2018). Категория доверия в исследованиях менеджмента // *Российский журнал менеджмента*, 16(4), 577–600.
- Шишкин, И. Ф. (1990). *Метрология, стандартизация и управление качеством: Учеб. для вузов*. М.: Изд-во стандартов.
- Abadie, M., & Waroquier, L. (2019). Evaluating the Benefits of Conscious and Unconscious Thought in Complex Decision Making // *Policy Insights from the Behavioral and Brain Sciences*, 6(1), 72–78.
- Abatecola, G., Caputo, A., & Cristofaro, M. (2018). Reviewing cognitive distortions in managerial decision making: Toward an integrative co-evolutionary framework // *Journal of Management Development*, 37(5), 409–424.
- Baesler, E. J. (1997). Persuasive effects of story and statistical evidence // *Argumentation & Advocacy*, 33(4), 170–175.
- Barnes, J. H. (1984). Cognitive biases and their impact on strategic planning // *Strategic Management Journal*, 5(2), 129–137.
- Basel, J. S., & Brühl, R. (2013). Rationality and dual process models of reasoning in managerial cognition and decision making // *European Management Journal*, 31(6), 745–754.

¹⁸ Понимание различий между *цифрами* (знаками определенной формы) и *числами* (абстрактными объектами, с которыми можно осуществлять определенные операции) также не является всеобщим знанием среди чиновников

¹⁹ В этом чиновники, впрочем, не отличаются от других: людям вообще свойственно доверие цифрам и числам (Porter, 1995), а их присутствие повышает убедительность новостей (Koetsenruijter, 2011), нарративов (Baesler, 1997), политических деклараций (Blastland & Dilnot, 2009) и т.п.

- Bazerman, M. (2010). *Judgment in managerial decision making*. New York: Jon Wiley and Sons.
- Behn, R. D. (2003). Why Measure Performance? Different Purposes Require Different Measures // *Public Administration Review*, 63(5), 586–606.
- Benabou, R., & Tirole, J. (2003). Intrinsic and extrinsic motivation // *Review of Economic Studies*, 70(3), 489–520.
- Benoit, E., & Foulloy, L. (2013). The role of fuzzy scales in measurement theory // *Measurement*, 46(8), 2921–2926.
- Bevan, G., & Hood, C. (2006). What's measured is what matters: targets and gaming in the English public health care system // *Public Administration*, 84(3), 517–538.
- Blastland, M., & Dilnot, A. (2009). *The Numbers Game: The Commonsense Guide to Understanding Numbers in the News, in Politics, and in Life*. London: Penguin.
- Bogsnes, B. (2018). Hitting the Target but Missing the Point // *Controlling & Management Review*, 62(5), 8–13.
- Bolli, T., & Somogyi, F. (2011). Do competitively acquired funds induce universities to increase productivity? // *Research Policy*, 40(1), 136–147.
- Bornmann, L., & Daniel, H.-D. (2008). What do citation counts measure? A review of studies on citing behavior // *Journal of Documentation*, 64(1), 45–80. DOI: 10.1108/00220410810844150.
- Busenitz, L. W., & Barney, J. B. (1994). Biases and heuristics in strategic decision making: Differences between entrepreneurs and managers in large organizations // *Academy of Management Proceedings*, (1), 85–89.
- Conger, J. A., & Kanungo, R. N. (1988). The Empowerment Process: Integrating Theory and Practice // *Academy of Management Review*, 13(3), 471–482.
- Costello, F., & Watts, P. (2014). Surprisingly Rational: Probability Theory Plus Noise Explains Biases in Judgment // *Psychological Review*, 121(3), 463–480.
- Crane, A., Matten, D., & Moon, J. (2008). The Emergence of Corporate Citizenship: Historical Development and Alternative Perspectives, pp. 25–49 / In: A. Scherer & G. Palazzo (eds.) *Handbook of Research on Global Corporate Citizenship*. Cheltenham: Edward Elgar.
- Curtis, I. (2015). *The use of targets in policing*. (https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/466058/ReviewTargets2015.pdf).
- de Graaf, A., Hoeken, H., Sanders, J., & Beentjes, J. W. J. (2012). Identification as a Mechanism of Narrative Persuasion // *Communication Research*, 39(6), 802–823.
- Diefenbach, T. (2009). New public management in public sector organizations: the dark sides of managerialistic 'enlightenment' // *Public Administration*, 87(4), 892–909.
- Dobija, D., Górska, A. M., Grossi, G., & Strzelczyk, W. (2019). Rational and symbolic uses of performance measurement: Experiences from Polish universities // *Accounting, Auditing & Accountability Journal*, 32(3), 750–781. DOI: 10.1108/AAAJ-08-2017-3106.
- Elbow, P. (1993). Ranking, Evaluating, and Liking: Sorting out Three Forms of Judgment // *College English*, 55(2), 187–206.
- Finkelstein, L. (2003). Widely, strongly and weakly defined measurement // *Measurement*, 34(1), 39–48. DOI: 10.1016/S0263-2241(03)00018-6.
- Fonseca, M. A., & Peters, K. (2018). Will any gossip do? Gossip does not need to be perfectly accurate to promote trust // *Games and Economic Behavior*, 107, 253–281.
- Fulford, M. D., & Enz, C. A. (1995). The Impact of Empowerment on Service Employees // *Journal of Managerial Issues*, 7(2), 161–175.
- Funder, D. C. (1987). Errors and Mistakes: Evaluating the Accuracy of Social Judgment // *Psychological Bulletin*, 101(1), 75–90.

- Funder, D. C. (1995). On the Accuracy of Personality Judgment: A Realistic Approach // *Psychological Review*, 102(4), 652–670.
- Funder, D. C. (2012). Accurate Personality Judgment // *Current Directions in Psychological Science*, 21(3), 177–182.
- Gigerenzer, G. (2015). *Simply rational: Decision making in the real world*. New York: Oxford University Press.
- Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making // *Annual Review of Psychology*, 62, 451–482.
- Glänzel, W. (2008). Seven Myths in Bibliometrics: About facts and fiction in quantitative science studies // *COLLNET Journal of Scientometrics and Information Management*, 2(1), 9–17.
- Gruening, G. (2001). Origin and theoretical basis of New Public Management // *International Public Management Journal*, 4(1), 1–25.
- Hammond, K. R. (1996). *Human judgment and social policy: Irreducible uncertainty, inevitable error, unavoidable injustice*. New York: Oxford University Press.
- Heerkens, H., Norde, C., & van der Heijden, B. (2011). Importance assessment of decision attributes: A qualitative study comparing experts and laypersons // *Management Decision*, 49(5), 748–761.
- Heffernan, T. A., & Heffernan, A. (2018). Language games: University responses to ranking metrics // *Higher Education Quarterly*, 72(1), 29–39.
- Higher Education Research Institute (2007). *College Rankings and College Choice: How Important Are College Rankings in Students' College Choice Process?* HERI Research Brief. August.
- Hodgkinson, G. P., Langan-Fox, J., & Sadler-Smith, E. (2008). Intuition: A fundamental bridging construct in the behavioural sciences // *British Journal of Psychology*, 99(1), 1–27.
- Hoeken, H., Kolthoff, M., & Sanders, J. (2016). Story Perspective and Character Similarity as Drivers of Identification and Narrative Persuasion // *Human Communication Research*, 42(2), 292–311. DOI: 10.1111/hcre.12076.
- Hoffmann, J. A., Gaissmaier, W., & von Helversen, B. (2017). Justifying the judgment process affects neither judgment accuracy, nor strategy use // *Judgment and Decision Making*, 12(6), 627–641.
- Holmstrom, B., & Milgrom, P. (1991). Multitask principal-agent analyses: incentive contracts, asset ownership, and job design // *Journal of Law, Economics, and Organization*, 7 (Spec. Is.), 24–52.
- Hood, C. (2006). Gaming in targetworld: The targets approach to managing British public services // *Public Administration Review*, 66(4), 515–521. DOI: 10.1017/S0140525X1500062X.
- Johnson, M. D., & Puto, C. P. (1987). A review of consumer judgment and choice, pp. 236–292 / In: M. J. Houston. *Review of marketing, 1987*. Chicago: American Marketing Association. (<https://scholarship.sha.cornell.edu/cgi/viewcontent.cgi?referer=https://scholar.google.com/&httpsredir=1&article=1434&context=articles>).
- Jussim, L. (2017). Précis of Social Perception and Social Reality: Why accuracy dominates bias and self-fulfilling prophecy // *Behavioral and Brain Sciences*, 40, e1, e1–e20. DOI: 10.1017/S0140525X1500062X.
- Kahneman, D. (2003). Maps of bounded rationality: a perspective on intuitive judgment and choice, pp. 449–89 / In: T. Frangsmyr (ed.) *Les Prix Nobel: The Nobel Prizes, 2002*. Stockholm: Nobel Found.
- Kahneman, D. X., & Tversky, A. (1973). On the psychology of prediction // *Psychological Review*, 80(4), 237–251.

- Kaplan, R. S., & Norton, D. P. (1992). The Balanced Scorecard: Measures that Drive Performance // *Harvard Business Review*, 70(1), 71–79.
- Keren, G. (1990). Cognitive aids and debiasing methods: Can cognitive pills cure cognitive ills? pp. 523–552 / In: J. P. Caverni, J. M. Fabre & M. Gonzales (eds.) *Cognitive biases*. New York: Elsevier.
- Kerr, S. (1975). On the folly of rewarding A, while hoping for B // *Academy of Management Journal*, 18(4), 769–783.
- Koetsenruijter, A. M. (2011). Using numbers in news increases story credibility // *Newspaper Research Journal*, 32(2), 74–82.
- Kosinski, M., Wang, Y., Lakkaraju, H., & Leskovec, J. (2016). Mining big data to extract patterns and predict real-life outcomes // *Psychological Methods*, 21(4), 493–506.
- Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and Deliberate Judgments Are Based on Common Principles // *Psychological Review*, 118(1), 97–109.
- Larsen, B. (2001). “One measurement is better than 1,000 opinions”: is it? // *Managerial Auditing Journal*, 16(2), 63–68. DOI: 10.1108/02686900110363618.
- Laurie, R. S. and Rosati, A. G. (2015). The Evolutionary Roots of Human Decision Making // *Annual Review of Psychology*, 66, 321–47.
- Lawrence, M., Goodwin, P., O’Connor, M., & Önkal, D. (2006). Judgmental forecasting: A review of progress over the last 25 years // *International Journal of Forecasting*, 22(3), 493–518.
- Maldonato, M., Dell’Orco, S., & Sperandio, R. (2018). When Intuitive Decisions Making, Based on Expertise, May Deliver Better Results than a Rational, Deliberate Approach, pp. 369–377 / In: A. Esposito, M. Faudez-Zanuy, F. Morabito & E. Pasero (eds.) *Multidisciplinary Approaches to Neural Computing*. Cham: Springer.
- Mari, L. P. (2005). Models of the Measurement Process, pp. 681–684 / In: P. Sydenham & R. Thorn (eds.) *Handbook of Measuring System Design*. Chichester: John Wiley & Sons.
- Mitchell, D., Bryson, J. J., Rauwolf, P., & Ingram, G. P. (2016). On the reliability of unreliable information: Gossip as cultural memory // *Interaction Studies*, 17(1), 1–25.
- Morewedge, C. K., Yoon, H., Scopelliti, I., Symborski, C. W., Korris, J. H., & Kassam, K. S. (2015). Debiasing Decisions: Improved Decision Making With a Single Training Intervention // *Policy Insights from the Behavioral and Brain Sciences*, 2(1), 129–140.
- Moynihan, D. P., & Pandey, S. K. (2010). The Big Question for Performance Management: Why Do Managers Use Performance Information? // *Journal of Public Administration Research and Theory*, 20(4), 849–866.
- Nisbett, R., & Ross, L. (1980). *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Pendril, L., & Petersson, N. (2016). Metrology of human-based and other qualitative measurements // *Measurement Science and Technology*, 27(9), article 094003. (<https://iopscience.iop.org/article/10.1088/0957-0233/27/9/094003/meta>).
- Pierce, J. L., Kostova, T., & Dirks, K. T. (2001). Toward a Theory of Psychological Ownership in Organizations // *Academy of Management Review*, 26(2), 298–310.
- Porter, T. M. (1995). *Trust in numbers: the pursuit of objectivity in science and public life*. Princeton: Princeton University Press.
- Propper, C., & Wilson, D. (2003). The Use and Usefulness of Performance Measures in the Public Sector // *Oxford Review of Economic Policy*, 19(2), 250–267.
- Prusak, L. (2010). What can’t be measured // *Harvard Business Review Blog*, 7 October. (<https://hbr.org/2010/10/what-cant-be-measured>).
- Pugalis, L. (2013). Hitting the target but missing the point: the case of area-based regeneration // *Community Development*, 44(5), 617–634.

- Radnor, Z. (2008). Hitting the Target and Missing the Point? Developing an Understanding of Organizational Gaming, pp. 94–105 / In: W. Van Dooren & S. Van de Walle (eds.) *Performance Information in the Public Sector*. London: Palgrave Macmillan.
- Rauhvargers, A. (2014). Where Are the Global Rankings Leading Us? An Analysis of Recent Methodological Changes and New Developments // *European Journal of Education*, 49(1), 29–44.
- Rauwolf, P. (2016). *Understanding the ubiquity of self-deception: the evolutionary utility of incorrect information*. PhD Thesis. University of Bath. (<https://core.ac.uk/download/pdf/42511326.pdf>).
- Rauwolf, P., Mitchell, D., & Bryson, J. J. (2015). Value homophily benefits cooperation but motivates employing incorrect social information // *Journal of Theoretical Biology*, 367, 246–261.
- Reisen, N., Hoffrage, U., & Mast, F. W. (2008). Identifying decision strategies in a consumer choice situation // *Judgment and Decision Making*, 3(8), 641–658.
- Rock, A. (2004). *The mind at night: The new science of how and why we dream*. Cambridge, MA: Basic Books.
- Ross, L. (1977). The intuitive psychologist and his shortcomings, pp. 174–214 / In: L. Berkowitz (ed.) *Advances in experimental social psychology*, vol. 10. New York: Academic Press.
- Rusou, Z., Zakay, D., & Usher, M. (2013). Pitting intuitive and analytical thinking against each other: The case of transitivity // *Psychonomic Bulletin & Review*, 20(3), 608–614.
- Sacco, D. F., & Brown, M. (2018). The face of personality: Adaptive inferences from facial cues are moderated by perceiver personality and motives // *Social and Personality Psychology Compass*, 12(8), e12410. DOI: 10.1111/spc3.12410.
- Sanger, M. B. (2013). Does measuring performance lead to better performance? // *Journal of Policy Analysis and Management*, 32(1), 185–203.
- Schapiro, A., & Turk-Browne, N. (2015). Statistical Learning, pp. 501–506 / In: A. W. Toga (ed.) *Brain Mapping: An Encyclopedic Reference*, vol. 3. Elsevier.
- Schwenk, C. R. (1985). Management illusions and biases: Their impact on strategic decisions // *Long Range Planning*, 18(5), 74–80.
- Shoaib, S., & Mujtaba, B. G. (2018). Perverse Incentives and Peccable Behavior in Professionals: A Qualitative Study of the Faculty // *Public Organization Review*, 18(4), 441–459.
- Siegelman, N., Bogaerts, L., & Frost, R. (2016). Measuring individual differences in statistical learning: Current pitfalls and possible solutions // *Behavior Research Methods*, 49(2), 418–432. DOI: 10.3758/s13428-016-0719-z.
- Sinclair, M., & Ashkanasy, N. M. (2002). Intuitive decision-making amongst leaders: More than just shooting from the hip // *Mt Eliza Business Review*, 5 (2), 32–40.
- Stevens, S. S. (1946). On the Theory of Scales of Measurement // *Science*, 103(2684), 677–680. DOI: 10.1126/science.103.2684.677.
- Strack, F., & Deutsch, R. (2004). Reflective and Impulsive Determinants of Social Behavior // *Personality and Social Psychology Review*, 8(3), 220–247.
- Summers, B., & Duxbury, D. (2012). Decision-dependent emotions and behavioral anomalies // *Organizational Behavior and Human Decision Processes*, 118(2), 226–238.
- Suppes, P., & Zinnes, J. (1963). Basic Measurement Theory, pp. 3–76 / In: D. Luce, R. R. Bush & E. Galanter (eds.) *Handbook of Mathematical Psychology*, vol. 1. Chichester: John Wiley & Sons. [Рус. изд.: Суппес, П., & Зиннес, Дж. (1967). Основы теории измерений, с. 9–110 / В кн.: Л. Д. Мешалкин (ред.) Психологические измерения. Москва: Мир.]

- Tang, T. L. P., Tollison, P. S., & Whiteside, H. D. (1987). The effect of quality circle initiation on motivation to attend quality circle meetings and on task performance // *Personnel Psychology*, 40(4), 799–814.
- Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure // *Social Cognition*, 27(6), 813–833.
- Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural Evidence of Statistical Learning: Efficient Detection of Visual Regularities Without Awareness // *Journal of Cognitive Neuroscience*, 21(10), 1934–1945.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases // *Science*, 185(4157), 1124–1131.
- Verbeeten, F. H. M. (2008). Performance management practices in public sector organizations: Impact on performance // *Accounting, Auditing & Accountability Journal*, 21(3), 427–454.
- Yetiv, S. A. (2013). *National Security through a Cockeyed Lens: How Cognitive Bias Impacts U.S. Foreign Policy*. Baltimore, MD: Johns Hopkins University Press.
- Youyou, W., Kosinski, M., & Stillwell, D. J. (2015). Computer-based personality judgments are more accurate than those made by humans // *Proceedings of the National Academy of Sciences of the USA*, 112(4), 1036–1040. DOI: 10.1073/pnas.1418680112.

REFERENCES

- Abadie, M., & Waroquier, L. (2019). Evaluating the Benefits of Conscious and Unconscious Thought in Complex Decision Making. *Policy Insights from the Behavioral and Brain Sciences*, 6(1), 72–78.
- Abatecola, G., Caputo, A., & Cristofaro, M. (2018). Reviewing cognitive distortions in managerial decision making: Toward an integrative co-evolutionary framework. *Journal of Management Development*, 37(5), 409–424.
- Baessler, E. J. (1997). Persuasive effects of story and statistical evidence. *Argumentation & Advocacy*, 33(4), 170–175.
- Barnes, J. H. (1984). Cognitive biases and their impact on strategic planning. *Strategic Management Journal*, 5(2), 129–137.
- Basel, J. S., & Brühl, R. (2013). Rationality and dual process models of reasoning in managerial cognition and decision making. *European Management Journal*, 31(6), 745–754.
- Bazerman, M. (2010). *Judgment in managerial decision making*. New York: Jon Wiley and Sons.
- Behn, R. D. (2003). Why Measure Performance? Different Purposes Require Different Measures. *Public Administration Review*, 63(5), 586–606.
- Benabou, R., & Tirole, J. (2003). Intrinsic and extrinsic motivation. *Review of Economic Studies*, 70(3), 489–520.
- Benoit, E., & Foulloy, L. (2013). The role of fuzzy scales in measurement theory. *Measurement*, 46(8), 2921–2926.
- Bevan, G., & Hood, C. (2006). What's measured is what matters: targets and gaming in the English public health care system. *Public Administration*, 84(3), 517–538.
- Blastland, M., & Dilnot, A. (2009). *The Numbers Game: The Commonsense Guide to Understanding Numbers in the News, in Politics, and in Life*. London: Penguin.
- Bogsnes, B. (2018). Hitting the Target but Missing the Point. *Controlling & Management Review*, 62(5), 8–13.
- Bolli, T., & Somogyi, F. (2011). Do competitively acquired funds induce universities to increase productivity? *Research Policy*, 40(1), 136–147.

- Bornmann, L., & Daniel, H.-D. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, 64(1), 45–80. DOI: 10.1108/00220410810844150.
- Busenitz, L. W., & Barney, J. B. (1994). Biases and heuristics in strategic decision making: Differences between entrepreneurs and managers in large organizations. *Academy of Management Proceedings*, (1), 85–89.
- Conger, J. A., & Kanungo, R. N. (1988). The Empowerment Process: Integrating Theory and Practice. *Academy of Management Review*, 13(3), 471–482.
- Costello, F., & Watts, P. (2014). Surprisingly Rational: Probability Theory Plus Noise Explains Biases in Judgment. *Psychological Review*, 121(3), 463–480.
- Crane, A., Matten, D., & Moon, J. (2008). The Emergence of Corporate Citizenship: Historical Development and Alternative Perspectives, pp. 25–49 / In: A. Scherer & G. Palazzo (eds.) *Handbook of Research on Global Corporate Citizenship*. Cheltenham: Edward Elgar.
- Curtis, I. (2015). *The use of targets in policing*. (https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/466058/ReviewTargets2015.pdf).
- de Graaf, A., Hoeken, H., Sanders, J., & Beentjes, J. W. J. (2012). Identification as a Mechanism of Narrative Persuasion. *Communication Research*, 39(6), 802–823.
- Diefenbach, T. (2009). New public management in public sector organizations: the dark sides of managerialistic ‘enlightenment’. *Public Administration*, 87(4), 892–909.
- Dobija, D., Górska, A. M., Grossi, G., & Strzelczyk, W. (2019). Rational and symbolic uses of performance measurement: Experiences from Polish universities. *Accounting, Auditing & Accountability Journal*, 32(3), 750–781. DOI: 10.1108/AAAJ-08-2017-3106.
- Elbow, P. (1993). Ranking, Evaluating, and Liking: Sorting out Three Forms of Judgment. *College English*, 55(2), 187–206.
- Finkelstein, L. (2003). Widely, strongly and weakly defined measurement. *Measurement*, 34(1), 39–48. DOI: 10.1016/S0263-2241(03)00018-6.
- Fonseca, M. A., & Peters, K. (2018). Will any gossip do? Gossip does not need to be perfectly accurate to promote trust. *Games and Economic Behavior*, 107, 253–281.
- Fulford, M. D., & Enz, C. A. (1995). The Impact of Empowerment on Service Employees. *Journal of Managerial Issues*, 7(2), 161–175.
- Funder, D. C. (1987). Errors and Mistakes: Evaluating the Accuracy of Social Judgment. *Psychological Bulletin*, 101(1), 75–90.
- Funder, D. C. (1995). On the Accuracy of Personality Judgment: A Realistic Approach. *Psychological Review*, 102(4), 652–670.
- Funder, D. C. (2012). Accurate Personality Judgment. *Current Directions in Psychological Science*, 21(3), 177–182.
- Gigerenzer, G. (2015). *Simply rational: Decision making in the real world*. New York: Oxford University Press.
- Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, 62, 451–482.
- Glänzel, W. (2008). Seven Myths in Bibliometrics: About facts and fiction in quantitative science studies. *COLLNET Journal of Scientometrics and Information Management*, 2(1), 9–17.
- Gruening, G. (2001). Origin and theoretical basis of New Public Management. *International Public Management Journal*, 4(1), 1–25.
- Hammond, K. R. (1996). *Human judgment and social policy: Irreducible uncertainty, inevitable error, unavoidable injustice*. New York: Oxford University Press.

- Heerkens, H., Norde, C., & van der Heijden, B. (2011). Importance assessment of decision attributes: A qualitative study comparing experts and laypersons. *Management Decision*, 49(5), 748–761.
- Heffernan, T. A., & Heffernan, A. (2018). Language games: University responses to ranking metrics. *Higher Education Quarterly*, 72(1), 29–39.
- Higher Education Research Institute (2007). *College Rankings and College Choice: How Important Are College Rankings in Students' College Choice Process?* HERI Research Brief. August.
- Hodgkinson, G. P., Langan-Fox, J., & Sadler-Smith, E. (2008). Intuition: A fundamental bridging construct in the behavioural sciences. *British Journal of Psychology*, 99(1), 1–27.
- Hoeken, H., Kolthoff, M., & Sanders, J. (2016). Story Perspective and Character Similarity as Drivers of Identification and Narrative Persuasion. *Human Communication Research*, 42(2), 292–311. DOI: 10.1111/hcre.12076.
- Hoffmann, J. A., Gaissmaier, W., & von Helversen, B. (2017). Justifying the judgment process affects neither judgment accuracy, nor strategy use. *Judgment and Decision Making*, 12(6), 627–641.
- Holmstrom, B., & Milgrom, P. (1991). Multitask principal-agent analyses: incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization*, 7 (Spec. Is.), 24–52.
- Hood, C. (2006). Gaming in targetworld: The targets approach to managing British public services. *Public Administration Review*, 66(4), 515–521. DOI: 10.1017/S0140525X1500062X.
- Johnson, M. D., & Puto, C. P. (1987). A review of consumer judgment and choice, pp. 236–292 / In: M. J. Houston. *Review of marketing 1987*. Chicago: American Marketing Association. (<https://scholarship.sha.cornell.edu/cgi/viewcontent.cgi?referer=https://scholar.google.com/&httpsredir=1&article=1434&context=articles>).
- Jussim, L. (2017). Précis of Social Perception and Social Reality: Why accuracy dominates bias and self-fulfilling prophecy. *Behavioral and Brain Sciences*, 40, e1, e1-e20. DOI: 10.1017/S0140525X1500062X.
- Kahneman, D. (2003). Maps of bounded rationality: a perspective on intuitive judgment and choice, pp. 449–89 / In: T. Frangmyr (ed.) *Les Prix Nobel: The Nobel Prizes, 2002*. Stockholm: Nobel Found.
- Kahneman, D. X, & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80(4), 237–251.
- Kaplan, R. S., & Norton, D. P. (1992). The Balanced Scorecard: Measures that Drive Performance. *Harvard Business Review*, 70(1), 71–79.
- Keren, G. (1990). Cognitive aids and debiasing methods: Can cognitive pills cure cognitive ills? pp. 523–552 / In: J. P. Caverni, J. M. Fabre & M. Gonzales (eds.) *Cognitive biases*. New York: Elsevier.
- Kerr, S. (1975). On the folly of rewarding A, while hoping for B. *Academy of Management Journal*, 18(4), 769–783.
- Koetsenruijter, A. M. (2011). Using numbers in news increases story credibility. *Newspaper Research Journal*, 32(2), 74–82.
- Kosinski, M., Wang, Y., Lakkaraju, H., & Leskovec, J. (2016). Mining big data to extract patterns and predict real-life outcomes. *Psychological Methods*, 21(4), 493–506.
- Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and Deliberate Judgments Are Based on Common Principles. *Psychological Review*, 118(1), 97–109.

- Krupina, S. M., & Klochkov, V. V. (2014). Perspectives of Russian Basic Science under Institutional Reforms: Modelling and Qualitative Conclusions, pp. 11–24 / In: *Proceedings of 17th Drucker's Readings "Russia and World Innovation Perspectives: Theory and Modelling"* (Moscow, 2014). Novocherkassk: Platov South-Russian State Polytechnic University (NPI) Publ. (In Russian.)
- Larsen, B. (2001). "One measurement is better than 1,000 opinions": is it? *Managerial Auditing Journal*, 16(2), 63–68. DOI: 10.1108/02686900110363618.
- Laurie, R. S. and Rosati, A. G. (2015). The Evolutionary Roots of Human Decision Making. *Annual Review of Psychology*, 66, 321–47.
- Lawrence, M., Goodwin, P., O'Connor, M., & Önkal, D. (2006). Judgmental forecasting: A review of progress over the last 25 years. *International Journal of Forecasting*, 22(3), 493–518.
- Maldonato, M., Dell'Orco, S., & Sperandio, R. (2018). When Intuitive Decisions Making, Based on Expertise, May Deliver Better Results than a Rational, Deliberate Approach, pp. 369–377 / In: A. Esposito, M. Faudez-Zanuy, F. Morabito & E. Pasero (eds.) *Multidisciplinary Approaches to Neural Computing*. Cham: Springer.
- Mari, L. P. (2005). Models of the Measurement Process, pp. 681–684 / In: P. Sydenham & R. Thorn (eds.) *Handbook of Measuring System Design*. Chichester: John Wiley & Sons.
- Mitchell, D., Bryson, J. J., Rauwolf, P., & Ingram, G. P. (2016). On the reliability of unreliable information: Gossip as cultural memory. *Interaction Studies*, 17(1), 1–25.
- Morewedge, C. K., Yoon, H., Scopelliti, I., Symborski, C. W., Korris, J. H., & Kassam, K. S. (2015). Debiasing Decisions: Improved Decision Making With a Single Training Intervention. *Policy Insights from the Behavioral and Brain Sciences*, 2(1), 129–140.
- Moynihan, D. P., & Pandey, S. K. (2010). The Big Question for Performance Management: Why Do Managers Use Performance Information? *Journal of Public Administration Research and Theory*, 20(4), 849–866.
- Nisbett, R., & Ross, L. (1980). *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Pendrill, L., & Petersson, N. (2016). Metrology of human-based and other qualitative measurements. *Measurement Science and Technology*, 27(9), article 094003. (<https://iopscience.iop.org/article/10.1088/0957-0233/27/9/094003/meta>).
- Pierce, J. L., Kostova, T., & Dirks, K. T. (2001). Toward a Theory of Psychological Ownership in Organizations. *Academy of Management Review*, 26(2), 298–310.
- Porter, T. M. (1995). *Trust in numbers: the pursuit of objectivity in science and public life*. Princeton: Princeton University Press.
- Propper, C., & Wilson, D. (2003). The Use and Usefulness of Performance Measures in the Public Sector. *Oxford Review of Economic Policy*, 19(2), 250–267.
- Prusak, L. (2010). What can't be measured. *Harvard Business Review Blog*, 7 October. (<https://hbr.org/2010/10/what-cant-be-measured>).
- Pugalis, L. (2013). Hitting the target but missing the point: the case of area-based regeneration. *Community Development*, 44(5), 617–634.
- Radnor, Z. (2008). Hitting the Target and Missing the Point? Developing an Understanding of Organizational Gaming, pp. 94–105 / In: W. Van Dooren & S. Van de Walle (eds.) *Performance Information in the Public Sector*. London: Palgrave Macmillan.
- Rauhvargers, A. (2014). Where Are the Global Rankings Leading Us? An Analysis of Recent Methodological Changes and New Developments. *European Journal of Education*, 49(1), 29–44.
- Rauwolf, P. (2016). *Understanding the ubiquity of self-deception: the evolutionary utility of incorrect information*. PhD Thesis. University of Bath. (<https://core.ac.uk/download/pdf/42511326.pdf>).

- Rauwolf, P., Mitchell, D., & Bryson, J. J. (2015). Value homophily benefits cooperation but motivates employing incorrect social information. *Journal of Theoretical Biology*, 367, 246–261.
- Reisen, N., Hoffrage, U., & Mast, F. W. (2008). Identifying decision strategies in a consumer choice situation. *Judgment and Decision Making*, 3(8), 641–658.
- Rock, A. (2004). *The mind at night: The new science of how and why we dream*. Cambridge, MA: Basic Books.
- Rodinkov, O. V., Bokach, N. A., & Bulatov, A. V. (2010). *Basics of metrology for physical and chemical measurement and chemical analysis. A study guide*. Sankt-Petersburg: VVM Publ. (In Russian.)
- Ross, L. (1977). The intuitive psychologist and his shortcomings, pp. 174–214 / In: L. Berkowitz (ed.) *Advances in experimental social psychology*, vol. 10. New York: Academic Press.
- Rusou, Z., Zakay, D., & Usher, M. (2013). Pitting intuitive and analytical thinking against each other: The case of transitivity. *Psychonomic Bulletin & Review*, 20(3), 608–614.
- Sacco, D. F., & Brown, M. (2018). The face of personality: Adaptive inferences from facial cues are moderated by perceiver personality and motives. *Social and Personality Psychology Compass*, 12(8), e12410. DOI: 10.1111/spc3.12410.
- Sanger, M. B. (2013). Does measuring performance lead to better performance? *Journal of Policy Analysis and Management*, 32(1), 185–203.
- Schapiro, A., & Turk-Browne, N. (2015). Statistical Learning, pp. 501–506 / In: A. W. Toga (ed.) *Brain Mapping: An Encyclopedic Reference*, vol. 3. Elsevier.
- Schwenk, C. R. (1985). Management illusions and biases: Their impact on strategic decisions. *Long Range Planning*, 18(5), 74–80.
- Shishkin, I. F. (1990). *Metrology, standardization and quality management*. Moscow: Izdatelstvo Standartov Publ. (In Russian.)
- Shoab, S., & Mujtaba, B. G. (2018). Perverse Incentives and Peccable Behavior in Professionals: A Qualitative Study of the Faculty. *Public Organization Review*, 18(4), 441–459.
- Siegelman, N., Bogaerts, L., & Frost, R. (2016). Measuring individual differences in statistical learning: Current pitfalls and possible solutions. *Behavior Research Methods*, 49(2), 418–432. DOI: 10.3758/s13428-016-0719-z.
- Sinclair, M., & Ashkanasy, N. M. (2002). Intuitive decision-making amongst leaders: More than just shooting from the hip. *Mt Eliza Business Review*, 5 (2), 32–40.
- Snezhko, A. A., Zakharova, N. V., & Zhirnova, E. A. (2008). *General theory of measurement. A Study guide*. Krasnojarsk: Reshetnev Siberian State University of Science and Technology (Reshetnev University). (In Russian.)
- Stevens, S. S. (1946). On the Theory of Scales of Measurement. *Science*, 103(2684), 677–680. DOI: 10.1126/science.103.2684.677.
- Strack, F., & Deutsch, R. (2004). Reflective and Impulsive Determinants of Social Behavior. *Personality and Social Psychology Review*, 8(3), 220–247.
- Summers, B., & Duxbury, D. (2012). Decision-dependent emotions and behavioral anomalies. *Organizational Behavior and Human Decision Processes*, 118(2), 226–238.
- Suppes, P., & Zinnes, J. (1963). Basic Measurement Theory, pp. 3–76 / In: D. Luce, R. R. Bush & E. Galanter (eds.) *Handbook of Mathematical Psychology*, vol. 1. Chichester: John Wiley & Sons.
- Tambovtsev, V. L. (2018). A construct of trust in management studies. *Russian Management Journal*, 16(4), 577–600. (In Russian.)

- Tang, T. L. P., Tollison, P. S., & Whiteside, H. D. (1987). The effect of quality circle initiation on motivation to attend quality circle meetings and on task performance. *Personnel Psychology*, 40(4), 799–814.
- Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition*, 27(6), 813–833.
- Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural Evidence of Statistical Learning: Efficient Detection of Visual Regularities Without Awareness. *Journal of Cognitive Neuroscience*, 21(10), 1934–1945.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131.
- Verbeeten, F. H. M. (2008). Performance management practices in public sector organizations: Impact on performance. *Accounting, Auditing & Accountability Journal*, 21(3), 427–454.
- Yetiv, S. A. (2013). *National Security through a Cockeyed Lens: How Cognitive Bias Impacts U.S. Foreign Policy*. Baltimore, MD: Johns Hopkins University Press.
- Youyou, W., Kosinski, M., & Stillwell, D. J. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences of the USA*, 112(4), 1036–1040. DOI: 10.1073/pnas.1418680112.