

ОБЩАЯ ПСИХОЛОГИЯ И ПСИХОЛОГИЯ ЛИЧНОСТИ

УДК 159.9.072.59

ПСИХОМЕТРИЧЕСКОЕ МОДЕЛИРОВАНИЕ: СИСТЕМАТИЧЕСКАЯ КРИТИКА БАЗОВЫХ ДОПУЩЕНИЙ И НЕКОТОРЫЕ АЛЬТЕРНАТИВЫ¹

Ю.А. Тюменева^{1,2}

¹ ФГБОУ ВО «Государственный Институт Русского языка им. А.С. Пушкина»,
Москва, ул. Академика Волгина, 6

² Национальный исследовательский университет «Высшая школа экономики», Москва,
ул. Мясницкая, 20

Резюме

В связи с популярностью психометрики, в которой вероятностное моделирование занимает лидирующие позиции, представляется полезным представить критический взгляд на эти методы, который, к сожалению, не обсуждается широко, особенно в русскоязычной литературе. Эта работа – систематический обзор растущей в последние годы критики в адрес нескольких допущений психометрического моделирования: 1) существуют ненаблюдаемые личностные черты и способности, которые влияют на выполнение заданий теста; 2) черты имеют стохастическую природу и, как следствие, для их изучения необходимы вероятностные модели; 3) данные, собранные на агрегированном уровне, аккуратно репрезентируют индивидуальные характеристики; 4) как и в естественных науках, модели в психометрии помогают раскрыть психологическую реальность, которая не поддается прямому наблюдению; 5) если проверка согласия эмпирических данных с моделью дает положительный результат, то модель правильно описывает реальность. Эти предположения критикуются на основании кругового характера определений самих признаков; логических ошибок, встроенных в сами допущения; методологического смещения вариативности и случайности поведения; отсутствия причинно-следственной связи между межиндивидуальной вариативностью ответов и индивидуальными ответами; поверхностных аналогий с моделями в естественных науках, к которым относится психометрика, а также подмены научных задач инструментальными и прагматическими. Мы приходим к выводу, что моделирование в психометрике контрпродуктивно, если используется как метод исследования психологической реальности, стоящей за тестом. Обсуждаются некоторые альтернативные практики количественных исследований, например, проверка существования вариации на

¹ Работа выполнена с использованием средств государственного бюджета по госзада-
нию на 2020–2024 годы (№ г.р. 00) FZNM-2020-0005 «Трансформация когнитивной и
коммуникативной деятельности человека в условиях современной информационной
среды».

уровне индивида и экспериментальный поиск ее объяснения. Существуют и другие возможные альтернативы, например, сетевой подход к психологическим явлениям, фасетная теория или моделирование, ориентированное на наблюдение. Хотя подобные исследования гораздо сложнее реализовать, чем стандартные тесты «соответствия данных модели», вероятно, именно они могут обеспечить рост психологических знаний.

Ключевые слова: психометрика; моделирование; реализм; репрезентация; черта; латентный конструкт; допущения; измерения

Введение

Типичный образ действия в психометрике заключается в следующем: на выборке респондентов с помощью специальных процедур (тестов или наблюдений) провоцируют заданное поведение, скажем, решение математических задач. По определенным параметрам этого поведения (например, по количеству правильно решенных задач) генерируют числовые данные; данные агрегируют; анализируют и сравнивают с модельным их распределением. Делается вывод о структуре ненаблюдаемой черты, в данном случае – о математической способности, которая предположительно и вызывает наблюдаемое поведение (ответы по тесту). Затем в соответствии с выявленной структурой респондентам начисляются баллы, отражающие их позицию в континууме черты. Это основные шаги измерения в психологии, хотя техники получения тестового балла и его интерпретации различаются в разных психометрических подходах (например, классической теории тестирования (СТТ, classical test theory), измерении Раша (Rasch measurement), современной теории тестирования (IRT, item response theory) или структурном моделировании).

Усилия психометриков направлены на проверку и поиск модели с лучшими индексами согласия данных с моделью. Однако эта проверка никак не затрагивает допущений, на которых строится вся практика. Эти допущения вытекают одно из другого, но для ясности изложения можно разбить их на три крупных блока: 1) существует черта (или способность) как самостоятельная реальность, влияющая на индивидуальное выполнение заданий по тесту; 2) эта черта количественна, т.е. индивиды разнятся по количеству этой черты; 3) черты имеют стохастическую природу, поэтому для их изучения нужны вероятностные модели, подобные тем, которые используются в естественных науках. Поскольку психометрические практики полагаются на эти допущения, но не проверяют их, выводы относительно психологической реальности находятся в весьма уязвимой позиции.

На сегодняшний день критические точки зрения относительно обоснованности распространенных психометрических практик достаточно ясно оформлены в литературе, хотя и разрозненны. В данной работе ставится задача представить эту критику систематически. Здесь она будет изложена в части, касающейся двух допущений: о существовании ненаблюдаемых черт, определяющих выполнение заданий теста, и о возможности применять вероятностное моделирование для изучения черт. На непроверяемое

допущение о количественности черт мы просто указываем, чтобы ограничить объем этой статьи до приемлемого.

Допущение 1. Существуют ненаблюдаемые индивидуальные черты, которые преимущественно определяют выполнение заданий теста.

Идея о «чертах», которые влияют на поведение и могут быть измерены по этому поведению, нетривиальна. Например, биология оперирует понятиями, схожими с психологическими, такими как «агрессивность», для оценки которых использует, например, частоту проявления специфического поведения, демонстрируемого в эксперименте. Объяснение этому поведению биологи обычно ищут в особенностях организма (напр., генетических или физиологических), ситуативных факторах (например, наличие пищи), прошлом опыте (памяти, научения) или их взаимодействии. В психологии такие понятия, как «агрессивность», отсылают к психологическим чертам – ненаблюдаемым сущностям, которые гипотетически являются причинами специфического поведения.

Исторически представление о ненаблюдаемых чертах возникло как объяснение корреляции между решением когнитивных задач (positive manifold) некоей общей интеллектуальной способностью, «отвечающей» за все эти решения (Spearman, 1904). В начале XX в. идея о чертах (гипотетических конструктах) позволила либерализовать доминирующий тогда в психологии операционализм (MacCorquodale & Meehl, 1948). Однако вопрос о сути таких ненаблюдаемых черт решался по-прежнему в духе того же операционализма: «черта» определялась статистически как общий фактор, который объяснял дисперсию результатов выполнения тестовых заданий, так что взятие этого фактора «под контроль» сводило на нет корреляцию между заданиями. Задача обнаружения «черт» также решалась статистически – нужно было найти такие факторы в собранных данных, чтобы локальная независимость могла быть продемонстрирована. Последующая разработка техник анализа (факторного, структурного моделирования, IRT) позволила выделять такие латентные факторы практически на любых данных (непрерывных, дихотомических, политомических, категориальных).

Чтобы обосновать существование черты как ненаблюдаемой причины (латентная переменная) тех или иных ответов на тест, применяется такая логика:

– если задания a_1 , a_2 , a_3 требуют для своего выполнения способности A и

– если данные по заданиям находятся в согласии с моделью, где «латентная переменная» преимущественно объясняет дисперсию по a_1 , a_2 , a_3 , то A существует; по выполнению заданий a_1 , a_2 , a_3 можно судить об A .

Эта логика представлена во многих учебниках и методологических статьях (Campbell, 1960; Kane, 2006; Brown & Croudace, 2014; Mueller & Hancock, 2018).

Контраргумент 1. Формально-логическая ошибка этого рассуждения заключается в том, что подтверждение следствий не означает подтвержде-

ния причины, так как для одного и того же следствия может быть несколько причин. Факт различий людей по выполнению каких-то действий может объясняться и без привлечения специальной инстанции вроде «черты» или «способности» (например, предыдущим опытом, обучением, другими знаниями, возможно, какими-то нейрологическими особенностями, а также взаимовлиянием факторов друг на друга) (Costantini et al., 2015).

Контраргумент 2. В психологии черта, как правило, определяется диспозиционно, т.е. через отношение между ситуацией и действием человека в ней: математические способности – это способность успешно решать математические задачи; общительность – это тенденция к общению; тревожность – склонность тревожиться и т.п. Возможно, что нет другого способа определить психологическое (Perugini, Costantini, Hughes & De Houwer, 2016). Но одно дело, когда мы используем диспозиционное определение, чтобы *описать* какую-то регулярность в поведении, а другое – когда мы используем определение, чтобы *объяснить* эту регулярность. Если умение решать задачи назвать математической способностью, то в этом нет логической ошибки – здесь математическая способность – это обобщающий термин, наименование определенного поведения; термином удобно пользоваться, он делает коммуникацию эффективнее. Но когда мы используем наименование для *объяснения* этого же поведения – это логическая ошибка: невозможно, чтобы математическая способность была причиной успешности в решении задач, так как успешность в решении задач и есть (по определению) математическая способность.

Интересно, что круговой характер определения черты осознается и в «официальной» психологической литературе: «Мы не должны говорить, что высокий балл связан с высокой способностью, скорее наоборот. Мы говорим о высокой способности индивида, потому что его достижения по тесту получили высокий балл. «Способность» – это просто обобщающее утверждение, касающееся его действий» (Cureton, 1951, p. 641).

Однако в более современном, четвертом, издании *Educational measurement* (Kane, 2006) отмечается, что черта не только указывает на «регулярность в поведении» (сам факт регулярности может быть важен для науки), но и что черта независима от заданий теста: «... в большинстве случаев у нас есть некоторое ощущение составляющих черту процессов, поэтому интерпретация черты не зависит полностью от достижений в тестируемой области. По крайней мере, мы знаем, что мы ощущаем, когда злимся, и как мы бы подступились к ответу на вопросы по химии» (Kane, 2006, p. 30).

Если *Educational measurement*, официальное издание Американской психологической ассоциации и Американского Совета по образованию, ссылается на «некоторое ощущение» как свидетельство независимости черты от самого измерения (теста); если этому вопросу посвящено несколько строк в 700-страничном издании, то вопрос о существовании независимой от теста черты нельзя считать ни решенным, ни даже поставленным на повестку дня.

Неудивительно поэтому, что «черта» и «способности» безосновательно, но привычно интерпретируются как самостоятельная реальность, влияющая на поведение. Так, существование кругового объяснения используется в хорошо известной пятифакторной модели личности: «Предположим, некоторые виды птиц сбиваются в стаи, потому что они выработали механизм (или набор механизмов), который позволяет каждой птице среагировать на ее окружение таким способом, что они собираются в стаи. Мы могли бы назвать эти механизмы «стайным» инстинктом, существующим у одних видов, но не у других. Это означает, что есть смысл говорить, что врожденный стайный инстинкт – это причина стайного поведения, потому что, когда он отсутствует, стайного поведения не наблюдается» (Terraciano & McCrae, 2012, p. 449). Таким образом, сбивание в стаи предлагается называть инстинктом сбивания в стаи (flocking instinct), и этот же инстинкт будем считать *причиной* сбивания в стаи, на что справедливо указал S. Boag (2011).

Другой пример: способность к реляционному мышлению описывается как умение замечать аналогии, аномалии, парадоксы и противоположности (Alexander, Dumas, Grossnickle, List & Firetto, 2016). Однако затем авторы проверяют гипотезу, что «реляционное мышление может влиять на латентные факторы аналогии, аномалии, парадоксов и противоположностей» (Alexander et al., 2016, p. 139). Но если реляционное мышление состоит из этих умений, как же тогда оно влияет на них?

Вербально определять действие в форме существительного (суметь решить задачи → *умение* решать задачи; терпеть → проявить *терпеливость*) вовсе не означает иметь дело с сущностью, стоящей за действием (Matthews, Deary & Whiteman, 2003). Иначе, мы *конструируем* бесконечную череду новых и новых способностей (черт), «материализованных» с помощью «вербальной магии» (Boag, 2011).

Почему это важно в контексте психологических измерений? Для измерения нужно, чтобы измеряемое, как минимум, существовало и находилось вне измерителя (Borsboom et al., 2004). Если утверждается, что математические способности влияют на выполнение задач теста и что по тесту мы судим об этой способности, то математическая способность должна находиться вне задач теста и влиять на них. Однако если математическая способность определяется через умение решать задачи, тем самым она и есть решение задач, следовательно, невозможно измерить математическую способность через решение задач, потому что измеритель совпадет с предметом измерения (Boag, 2011; Cramer et al., 2012).

Встает, однако, вопрос: а как же действует логика измерения в других науках? В биологии, например, агрессивное поведение измеряется по каким-то поведенческим реакциям, и по ним можно судить о росте агрессии. Разница в том, что биологи не считают некую латентную черту – агрессивность – причиной агрессивного поведения. Агрессивность для них – это просто короткое название для определенного паттерна поведения. Поэтому они приписывают рост числа агрессивных проявлений не росту «агрессив-

ности», но, например, действию определенных гормонов. Здесь ненаблюдаемая причина (гормоны) и следствие (агрессивность) – не одно и то же: гормоны не определяются через агрессию, а агрессия не определяется через гормоны.

То же самое и в других науках. За изменением температуры стоит энергия движения молекул, за хрупкостью – особенности кристаллической решетки, за горючестью – специфика химических веществ, за регенерацией ткани – размножение и дифференцировка клеток.

Кажется, что уместно возразить: ведь задачи теста лишь частично покрывают область возможного проявления способности, так что способность не тождественна какому-то набору действий, оцениваемых тестом, следовательно, она может влиять на эти и другие действия, а значит, может быть тестом измерена. Такая позиция и заявляется в основных учебниках и справочниках по психологическим измерениям: поведение, оцениваемое тестом, рассматривается как выборка от всех возможных проявлений черты (Анастаси, Урбина, 2007; Kane, 2006), отсюда – задача валидизировать тест, связав его с другими проявлениями этой же черты (например, через критериальную валидизацию).

Проблема этого возражения в том, что причина не может совпадать со следствиями ни частично, ни полностью, так что неважно, является тестовое поведение лишь частью «черты» или полностью ее покрывает. Если способность определяется через действия, то никакие из этих действий не являются ее «индикаторами», потому что нет никакой отдельной «ее»; есть состав деятельности, названный нами «способность».

Допущение 2. Ненаблюдаемые черты имеют стохастическую природу, поэтому для их изучения нужны вероятностные модели.

Суть допущения в том, что подобно моделированию в физике или биологии современное вероятностное психометрическое моделирование *исследует* психологическую реальность, которая недоступна прямой оценке. Черта понимается как «латентная переменная», которая предположительно скрывается под «наблюдаемыми (манифестируемыми) переменными» и «объясняет» их.

Через психометрическое моделирование пытаются статистически подтвердить существование латентной черты, проверяя согласие данных модели. При этом анализируются отклонения наблюдаемого ответа от математического ожидания согласно выбранной модели, и дисперсия, не объясненная данной моделью, разбивается на отдельные компоненты (Linacre, 1998). Если распределение дисперсии по этим компонентам равномерно, то принято считать, что никаких других «латентных» переменных, помимо модельных, нет.

Аргументы, которые звучат в поддержку этого допущения, следующие:

1. Подобно стохастической природе многих физических объектов, природа психологических свойств сложна и изменчива, поэтому для ее исследования больше подходят вероятностные методы.

2. Как и в естественных науках, модели в психометрике помогают выявить недоступную для прямого наблюдения психологическую реальность.

3. Если проверка согласия эмпирических данных с моделью дает положительный результат, значит, модель правильно описывает реальность.

Контраргумент 1. Во-первых, даже в первом приближении сложность поведения и чувств не означает их случайности. Если читатель дочитал данный текст до этого места, можно ли считать это результатом случайности? Если бы человек был столь же непредсказуемым существом, как его рассматривают в психометрике¹, то едва ли мы добирались бы на работу каждое утро, а вечером возвращались бы домой. Но даже если элемент случайности в человеческом поведении присутствует, то нет никакого понимания (ни теории, ни гипотезы), как эта случайность связывается с ответами на вопросы теста.

Более того, если разделять понятия случайности и изменчивости, а также допустить изменчивость индивидуального поведения, то к отклонениям человеческих реакций и поступков от прогнозируемых можно относиться не как к «ошибке измерения», а как к новым исследовательским вопросам и разработке новых методов изучения индивида. Например, теория перцептивного контроля (Forssell & Powers, 2009) интерпретирует изменчивость индивидуальных отклонений как ответ перцептивной системы на меняющиеся стимулы. Это, с одной стороны, делает изменчивость существенным компонентом адаптивной биологической системы, а с другой – открывает совершенно иные способы исследования индивидуального поведения. Или, например, подход Observation Oriented Modeling (Grice, 2011; Grice et al., 2016) позволяет получить информацию об индивидуальных паттернах поведения, которые пропускаются стандартными статистическими методами, поскольку последние основаны на усреднении агрегированных данных.

Контраргумент 2. Введение вероятностных моделей для измерения психологических феноменов не имеет под собой эмпирических оснований, в отличие от вероятностных моделей многих физических феноменов вроде тех, которые изучает квантовая физика, стохастическая природа которых была экспериментально обоснована. В той же физике описан класс объектов, хотя и ненаблюдаемых, который должен в соответствии с имеющимися знаниями и теорией вести себя стохастическим образом. В соответствии с этими же знаниями и теорией может быть разработан дизайн специального детектора для этих ненаблюдаемых объектов. Более того, в квантовой механике, где неопределенность занимает центральное место в определении текущего состояния вещества, переменные, составляющую систему, сами по себе измеримы как количественные переменные. Например, кван-

¹ Показательно в этой связи высказывание Георга Раша: «Что касается человеческих существ и их действий, представляется бесполезным делом конструировать модели, полезные для предсказания в отдельных случаях. Наоборот, человеческое существо действительно кажется довольно случайным не меньше, чем радиоактивная эмиссия» (Rasch, 1980, p. 11).

товые вычисления с использованием кубитов основаны на точном количественном измерении абсолютной температуры и количественно измеримых компонентах электрической активности (Vion et al., 2002).

Коротко говоря, в естественных науках модели применяются только в отношении объектов и процессов, где уже существуют количественные теории. Что еще важно, в саму модель включаются коэффициенты и параметры, рассчитанные экспериментально *до* моделирования, то есть, в модель на входе поступают результаты предварительных измерений (Thijssen, 2007; Flamm, 1998); более того, модель может быть сверена с реальными событиями для прояснения ее репрезентационных свойств. Так функционируют вероятностные математические модели климатических изменений, эволюционных процессов, эпидемиологических прогнозов и пр.

Физические модели выявляющего характера, которые тоже имеют много общего с целями «латентного моделирования», также включают результаты предварительных измерений в качестве параметров и констант. Так, мы можем ожидать сигнала детектора металла в присутствии металла, которого мы не видим (и в этом смысле латентного), основываясь на: 1) точном описании специфического класса элементов, металлов, и их свойств; 2) специально сконструированном устройстве, которое приспособлено вступать во взаимодействие с металлами, строго предсказанным образом; 3) заранее существующей логической схеме вывода о наличии / отсутствии сигнала (например, сигнал звучит, если и только если присутствует металл); 4) заранее известных и обоснованных ограничениях работоспособности устройства (расстояние до металла, электромагнитные помехи и пр.). Эта общая логика выявления скрытых объектов, как видим, глубоко укоренена в предварительных экспериментах и теориях об исследуемом предмете.

Образ действия психометрика не имеет с ничем общего с моделированием в естественных науках. Для психологических феноменов их стохастическая природа только предполагается, причем только на том основании, что детерминистские модели не могут удовлетворительно объяснять данные¹. В психометрике никогда не был описан класс объектов реальности, хотя бы и ненаблюдаемых, для которых модель могла бы сыграть роль детектора (Maraun & Gabriel, 2013) или репрезентации. «Латентные переменные», с которыми имеет дело психометрика, не имеют отношения к какому бы то ни было другому классу составляющих реального мира; они имеют дело со случайным переменным и их распределением. Но переменная – это просто функция (или параметр функции); она не принадлежит к составляющим реальности и поэтому не относится ни к наблюдаемым, ни к ненаблюдаемым; переменные не принадлежат и к составляющим языка –

¹ Кроме того, за счет необъясненной дисперсии в результатах тестирования, то есть за счет т.н. «случайной ошибки измерения», респонденты оцениваются в количественной шкале, а не просто упорядочиваются. Ситуация, когда улучшение шкалы от порядковой до интервальной происходит за счет наличия ошибки измерения, получило название «психометрический парадокс»: безошибочное измерение ухудшило бы «точность» оценки, превратив интервальную шкалу просто в порядковую (Michell, 2014).

это не концепция, не понятие, не теория и не идея. Бессмысленно говорить, что «когда мы концептуализируем переменную как ненаблюдаемую... мы предполагаем, что размещение человека на этой переменной может быть выведено из данных» (Borsboom, 2008, р. 30), так как нельзя «вывести» местоположение человека на переменной; это просто знание значения, которое принимает эта переменная (Maraun & Gabriel, 2013).

Поскольку моделирование в психометрике применяется к плохо изученной реальности (способности, чертам, и поведению, связанному с ответами на вопросы теста), при этом не существует никаких количественных теорий об этой реальности, то *не существует заранее известного правила* для соотнесения «латентной» переменной с получаемым «баллом» по тесту. В отличие от моделей в естественных науках, в психометрике правила связывания входящих переменных с исходящими подбираются *в процессе и после* анализа, поэтому модель не выводит на латентные переменные, но конструирует их способом, наиболее оптимальным для предписанной модели и входящих переменных (Maraun & Halpin, 2008).

Кроме того, не вводятся ограничения адекватности модели: модель предположительно может адекватно описывать «структуру конструкта» для всех людей, независимо от возраста, рода деятельности, социального и культурного контекста, здоровья и т.д. Нет также исследований, из которых такие ограничения могли бы следовать. Так что, какой бы модели в итоге не соответствовали данные тестирования, никаких независимых от теста свидетельств в отношении этой модели нет (Uher, 2021). Но без такого рода ограничений, психометрические модели являются лишь симуляторами того моделирования, которое используется в естественных науках.

В качестве примера, можно рассмотреть исследование карьерной адаптивности (Кондратюк, Бурмистрова-Савенкова, Моросанова, 2021)¹. Авторы демонстрируют психометрические характеристики «Шкалы карьерно-адаптационных способностей», включая выявление ее латентной факторной структуры. Под карьерной адаптивностью понимается «совокупность поведения, компетенций и установок индивидов, направленных на «приспособление себя к подходящей работе» (т.е. «на развитие карьеры»)). Далее авторы указывают, что «индивидуальные различия в карьерной адаптивности объясняют то, как по-разному люди строят и развивают свою карьеру». Мы имеем здесь классическое круговое определение, когда собирательный термин «карьерная адаптивность» превращается в самостоятельную ненаблюдаемую сущность, влияющую на приспособление к работе (т.е. адаптация влияет на адаптацию). Одной из задач авторов стала «проверка соответствия структуры теста конструкту карьерной адаптивности». Но поскольку структура «конструкта карьерной адаптивности» неизвестна, никакой проверки соответствия быть, конечно, не может. Фактиче-

¹ На месте этого исследования могло быть любое среди десятков других – достаточно ввести в Google Scholar ключевые слова «психометрика», «латентный», «конструкт», «модель».

ски, были проверены разные факторные решения на предмет их согласия с данными (которые сами пришли из опросника с неизвестными репрезентационными свойствами), взята наилучшая факторная модель (интерпретируемая в желательном направлении), и был сделан вывод о воспроизводимости конструктора с прогнозируемой структурой. Ясно, что какие бы результаты ни были получены, отклонить гипотезу о существовании «конструктора» как влияющую на ответы сущности при таком подходе невозможно в принципе, а отсутствие предварительных ограничений в отношении этой сущности позволяет «воспроизвести» любую структуру.

Вышесказанное объясняет, почему проверка согласия эмпирических данных с моделью, не означает проверки адекватности модели в отношении реальности. Проблема возникает не с самой проверкой, а со «склеиванием» статистических выводов о распределении ответов с выводами о «черте» (Maraun & Gabriel, 2013; Nowland, Beath & Boag, 2019).

Контраргумент 3. Психометрическое моделирование в подавляющем большинстве случаев использует агрегированные данные, т.е. дает информацию о латентных конструктах на уровне популяции. Но агрегированные данные страдают неустраняемыми недостатками при попытке их интерпретации на индивидуальном уровне, т.е. уровне, основном для психологии. Прежде всего, феномены, появляющиеся на уровне популяции, обладают свойством «эмерджентности», т.е. характеризуют только тот уровень системы, на котором появляются (как, например, фенотип появляется только на уровне целого организма, но отсутствует на клеточном уровне). Однако эта информация не обязательно будет адекватно описывать индивида – главный предмет интереса психологии. Было показано, например, что пятифакторная структура личности не воспроизводится на индивидуальном уровне, как и факторы интеллекта (Borsboom, Mellenbergh & Heerden, 2003; Borkenau & Ostendorf, 1998; Franić, Borsboom, Dolan & Boomsma, 2014).

Что происходит на агрегированном уровне – это оценка межиндивидуальной вариации в ответах, поэтому выведенные из таких корреляций факторы в принципе не в состоянии (каузально) объяснить индивидуальные ответы. Допустим, на уровне популяции мы получаем значительную корреляцию между успешностью выполнения математических задач теста, которые формируют «латентный конструкт» – математическую способность. Каково место этого конструкта в выполнении теста отдельным респондентом? Наиболее релевантные объяснения индивидуальных достижений возможны, например, через отсылку к предварительному опыту в решении таких задач, что в свою очередь отсылает к долговременной памяти и мышлению по аналогии; или, например, можно объяснить решение задачи хорошей рабочей памятью и вниманием, возможно, навыками счета, а также знанием некоторых эвристик, например, раскладыванием задачи на компоненты и пр. (см.: Пойа, 1959). Как в эти процессы вписывается «математическая способность»? Более того, если бы мы хотели проверить стабильность этой предполагаемой «черты» через серию повторных проверок одного и того же индивида, то мы немедленно бы столкну-

лись с отсутствием каких-либо корреляций, поскольку как раз в случае стабильной черты технически необходимая для корреляции вариативность была бы минимальна.

Все это означает, что для действительного объяснения происходящего требуется как минимум связать межиндивидуальные латентные конструкции с индивидуальными процессами. К сожалению, работы в этом направлении довольно редки в силу их гораздо большей затратности (временной, организационной и финансовой).

Кроме того, попытка использовать большие выборки якобы для минимизации ошибки измерения несколько лукава, поскольку само количественное измерение в психометрических моделях (в частности, в модели Раша) обязано исключительно ошибочному компоненту. «Количественная» шкала без ошибки измерения становится обычной порядковой шкалой Гутмана (Michell, 2014).

Как итог: поскольку ни на одном из этапов психометрического моделирования не вовлекаются данные, не зависящие от модели или от теста (который сам имеет неясное отношение к психологической реальности), то:

- выводы из моделирования в лучшем случае могут распространяться только на распределение ответов конкретных респондентов по конкретным заданиям теста;
- респонденты могут быть оценены только в отношении статистических фактов, проявившихся на агрегированных данных;
- психологическая реальность остается интактной.

Перспективы психометрического моделирования в психологии

Что означает вышесказанное для перспектив психометрического моделирования? Ответ зависит от статуса, в котором эти модели могут существовать как инструмент для решения прагматических задач (например, для предсказания или снижения размерности данных) и как научный метод исследования реальности.

С первых шагов психометрики ее практическая полезность для предсказания, отбора и распределения человеческих ресурсов (прежде всего в массовых и стандартизированных секторах экономики, таких как образование и вооруженные силы) стала тем обещанием, за счет которого психометрические методы оказались социально востребованной практикой. Хороший тест, т.е. тот, который в полной мере репрезентирует область целевых задач, аккуратно упорядочивает респондентов по выполнению заданий теста и максимизирует информацию о респонденте на фоне остальной выборки, такой тест безусловно важен. Пока методы психометрики производят «хорошие тесты», тем самым удовлетворяя определенную потребность общества, психометрика будет востребована и «практически полезна». Подчеркнем, что в этом статусе психометрическая модель не отвечает на вопросы о том, что именно является предметом ее оценки за пределами тестовых заданий. Еще на заре становления тестирования ученик Торндайка и ведущий

психолог-статистик США того времени Труман Ли Келли отметил: «Наши ментальные тесты измеряют что-то; нас может волновать, а может и не волновать, что именно, но это то, что нам выгодно измерять, поскольку это улучшает наши знания о том, что люди сделают в будущем. Измерительный прибор как мера того, что желательно измерить, стоит на первом месте, а то, мерой чего он является, – на втором» (Kelley, 1929, p. 86).

Поэтому вопреки распространенному мнению целью психометрического моделирования является не исследование реальности, как это происходит в естественных науках, а выравнивание данных и модели так, чтобы можно было доложить «тестовые баллы» соответственно достигнутой структуре. И надо отметить, что сейчас многие психометрические работы открыто позиционируются строго внутри статистических рамок и избегают эксплицитных утверждений о моделировании какой-либо сущности (черты или способности) за пределами теста, влияющей на ответы, и интерпретируют результаты только в отношении наиболее подходящего скоринга в соответствии с выявляемой структурой теста (напр., Артеменков, 2021).

Проблема с психометрическим моделированием возникает, когда оно используется в статусе полноценного исследовательского метода, т.е. с его помощью пытаются концептуализировать, «измерять» и «исследовать» психологические феномены. С учетом высказанных здесь ограничений стоит признать эту задачу, взятую на себя моделированием, обреченной на провал. Как уже сказано в тексте статьи, для того чтобы строить количественную модель феномена, нам нужны четкое описание класса таких феноменов, доказательства их количественной структуры, предварительные экспериментально полученные константы (параметры) модели, а также нам нужно знать границы применимости модели. Без этих ограничений психометрические модели будут лишь симуляторами моделирования, генерирующими данные с непонятными репрезентационными свойствами.

Фактически эксплуатация темы «полезности» для обоснования психометрики предполагает, что вопросы структурной связи между формальными латентными переменными и тем, что на самом деле они репрезентируют, уже решены. Предположим, что мы интерпретируем модели с латентными переменными как инструмент снижения размерности данных. Почему тогда появляются специфические факторные нагрузки, определенные какой-то формальной теорией латентных переменных? И есть ли какие-то другие законы или причины, почему мы должны использовать именно эти факторные нагрузки, например, для взвешивания? На самом деле эти нагрузки нечем обосновать, кроме как объяснением *ad hoc*. Любое обоснование, помимо структурного анализа, потребовало бы выйти за рамки модели. Поэтому оставаясь в рамках моделирования, никаких фундаментальных исследований реальности выполнить невозможно, как невозможно указать, каким образом и почему данные могут перестать согласовываться с моделью, а также с какой моделью они будут согласовываться в следующий раз (Michell, 2003a; Molenaar & von Eye, 1994).

Из всего вышесказанного следует, что, во-первых, какую бы реальность психометрика ни бралась оценивать, эта реальность должна сначала изучаться, а только потом моделироваться. Как указывал J. Michell (1999, 2003a, 2003b), сначала должна решаться научная задача (онтологическая), а только потом – инструментальная (эпистемологическая). Исследования же в психологии возможны и важны, но они не должны подменяться измерением и моделированием.

Приведем несколько примеров альтернативной исследовательской практики.

Прежде всего, что можно противопоставить допущению о существовании латентных черт? Хотя наш язык легко создает существительные из глаголов (настаивать – настойчивость, доминировать – доминирование, быть способным – способность), научному мышлению не обязательно поддаваться искушению создавать сущности с тем, чтобы потом измерять их. Даже если нет другого способа определить психологическое, кроме как через «тенденции», необходимо сначала продемонстрировать, как такие тенденции могут быть независимы от поведения, которое они якобы объясняют. С другой стороны, следует также признать, что участие в научном описании само по себе является необходимой частью научной задачи и что не все сторонники латентного подхода видят необходимость приписывать свойствам объяснительный статус. Например, в работе Gerber (2011) указываются примеры продуктивного выхода из кругового рассуждения, когда устойчивые поведенческие паттерны, оцениваемые опросником, объясняются не способностью, т.е. не статистическим «общим фактором», а предшествующим обучением, особенностями развития или восприятия. К примеру, такая черта, как «чувствительность к отвержению», определяется не через ожидание оказаться отвергнутым (что вело бы как раз к круговому определению), а через восприятие амбивалентных ситуаций и защитные реакции (Downey & Feldman, 1996). Это делает определение черты логически независимым от поведения, оцениваемого опросником: т.е. в опроснике спрашивается о «чувствительности к окружению», но сама эта чувствительность является не латентной «чертой», а следствием предшествующего опыта. Такая аккуратная позиция открывает дорогу и для теоретических разработок (теории привязанности, данного случая), и для эмпирических исследований (например, для изучения роли раннего отвержения, специфической активности мозга, предвещающей взаимодействие защитной реакцией и пр.)

Теперь коснемся альтернатив устоявшимся способам психометрического моделирования. Как уже говорилось, вся вариативность, которую не удастся объяснить латентными переменными, обычно списывается на ошибку измерения, но ни источники такой вариативности, ни индивидуальные механизмы ее проявления вне «моделирования» почти не исследуются. Замечательны поэтому работы, которые проверяют существование вариативности на уровне индивида и ищут ее объяснения экспериментально. Так, например, используя непараметрическую альтернативу логистиче-

ской регрессии – «пороговый анализ» (Grice et al., 2016), который не «отменяет ошибки» через вычисление групповых средних (что бессмысленно, хотя бы потому что психометрические шкалы НЕ интервальные), можно представить ответы респондентов как специфические паттерны ответов. При этом каждый последующий ответ на утверждения опросника оценивается как дискретное состояние, связанное не только с текущим утверждением, но и с предыдущим. Действительно, сложно представить, что человек, оценивая утверждения в серии, полностью забывает свои предыдущие ответы (Lord & Novick, 1968). Такой подход позволяет вывести логику индивидуального поведения так, чтобы группа респондентов была как можно точнее разнесена по подгруппам с разными паттернами ответов (Grice, 2011, Grice et al., 2017). Такого рода методы не только дают больше информации о респондентах, чем стандартные психометрические измерения, но эта информация еще и индивидуализирована, и при этом не основана на непроверяемых допущениях.

Можно пойти с другой стороны и указать на области, где методы структурного моделирования могут широко и обоснованно применяться. Это области, имеющие дело с результатами истинных измерений, например, в экологии (Arhonditsis et al., 2006), эволюционной биологии (Pugesek, Tomer & Von Eye, 2003), генетике (Franić, Dolan, Borsboom & Boomsma, 2012). В науке о человеческом социальном поведении структурное моделирование и байесовский анализ довольно успешно применяются в исследованиях принятия решений и экономического поведения (Birnbaum, 2008; Cavagnaro & Davis-Stober, 2014). Здесь модели используют данные, полученные независимо от самих этих моделей, поэтому и результаты моделирования могут быть использованы для проверки модели в отношении реальности, которую модель должна репрезентировать.

Интересная перспектива просматривается в применении сетевого подхода (Haig & Borsboom, 2008; Schmittmann et al., 2013). Сетевой подход не делает допущений о латентной черте, а сосредотачивается на комплексе поведенческих событий, на том, как одно событие влечет за собой другое. Так, например, было показано, что симптоматика многих патопсихологических состояний (например, депрессии) может быть описана как несколько отдельных узлов событий, обуславливающих друг друга последовательно; что объясняет разнообразие клинической картины заболевания и его динамики (Borsboom, 2017). Однако хотя сетевой подход уходит от фантомных «латентных» конструкторов, в отсутствие выхода на реалистичную онтологию он остается только способом репрезентации *данных*, как указывают и сами приверженцы сетевого подхода.

В качестве примера «выхода на реалистичную онтологию», сошлемся на психометрическую модель, выстроенную в рамках теории Пиаже об этапах развития мышления, в частности, о возрастной динамике понимания пропорциональных отношений. Длительные наблюдения за действиями детей, эксперименты и анализ детских решений задач на «пропорции» показали, что если в задаче участвуют несколько размерностей (скажем, в

задаче с рычажными весами – вес груза и его расстояние от центра весов), то ребенок не сразу может взять в рассмотрение все размерности. Сначала он ориентируется только на наиболее заметный, фактор (например, вес), затем начинает учитывать влияние другого фактора (расстояние от грузика до центра весов), но решает задачу безошибочно, только если оба фактора действуют согласованно.

На финальной стадии ребенок может учесть все факторы, даже если их влияние не согласовано (например, соотнести веса грузов и их расстояние до центра) (Inhelder & Piaget, 1958; Siegler, 1981). Эта теория обеспечила разработку формальной модели, где этапы развития рассматриваются как латентные классы, а развитие – как упорядоченная серия дискретных переходов между этими классами (Jansen & Van der Maas, 1997). Частоты и характер ошибок на разных этапах, известные из экспериментальных данных, могли быть использованы как параметры и ограничения в психометрической модели для оценки ее адекватности.

Другие формализуемые альтернативы также существуют (см., например, фасетную теорию (Canter, 2012); моделирование, ориентированное на наблюдение (Grice, 2011)). Их обзор выходит за рамки данной работы, при этом их главные достоинства – отсутствие непроверенных допущений (например, латентные конструкторы, случайная ошибка, аддитивное и непрерывное распределение признака, одномерность и пр.), индивидуальный уровень анализа и внимание к процессам, определяющих ответы на задания и др. Хотя такого рода исследования гораздо сложнее реализовывать, чем стандартные проверки «моделей» тех или иных «конструкторов», именно такие исследования только и могут дать прирост психологического знания.

Заключение

Целью работы было собрать рассеянную в современной литературе критику психометрического моделирования и представить ее в систематизированном виде. Популярность психометрики, где вероятностное моделирование занимает ведущую позицию, быстро растет среди российских исследователей. Поэтому представляется полезным представить и критический взгляд на эти техники, который, к сожалению, в русскоязычной литературе никак не обсуждается. Автор надеется, что этот материал улучшит понимание ограничений этой методологии и усилит интерес к более глубокому поиску альтернативных способов познания психологической реальности.

Литература

- Анастаси, А., Урбина, С. (2007). *Психологическое тестирование* (А. А. Алексеев, пер. с англ.). СПб.: Питер.
- Артеменков, С. Л. (2021). Упорядоченные сети частных корреляций в психологических исследованиях. *Моделирование и анализ данных*, 11(2), 31–50. doi: 10.17759/mda.2021110202

Кондратюк, Н. Г., Бурмистрова-Савенкова, А. В., Моросанова, В. (2021). Шкала карьерно-адаптационных способностей М. Савикаса и Э. Порфели: психометрические характеристики русскоязычной версии на выборке подростков. *Вестник Российской государственной академии педагогических наук*, 18(3), 555–575. doi: 10.22363/2313-1683-2021-18-3-555-575

Ссылки на зарубежные источники см. в разделе *References* после англоязычного блока.

Поступила в редакцию 08.04.2022 г.; повторно 20.09.2022 г.; принята 18.10.2022 г.

Тюменева Юлия Алексеевна – ведущий научный сотрудник лаборатории когнитивных и лингвистических исследований Государственного института русского языка им. А.С. Пушкина; старший научный сотрудник Института образования Национального исследовательского университета «Высшая школа экономики», кандидат психологических наук.
E-mail: jutu@yandex.ru

For citation: Tyumeneva, Y. A. (2022). Psychometric Modelling: A Systematic Critique of Underlying Assumptions and Some Alternatives. *Sibirskiy Psikhologicheskij Zhurnal – Siberian journal of psychology*, 86, 6–25. In Russian. English Summary. doi: 10.17223/17267080/86/1

Psychometric Modelling: A Systematic Critique of Underlying Assumptions and Some Alternatives¹

Y.A. Tyumeneva^{1,2}

¹ *The Pushkin State Russian Language Institute, 6, Akademika Volgina Str., 117485, Moscow, Russian Federation.*

² *Institute of Education, HSE University, 20, Myasnitskaya Str., Moscow, 101000, Russian Federation.*

Abstract

Due to the popularity of psychometrics, where probabilistic modelling takes a leading position, it seems useful to present a critical view of these techniques, which, unfortunately, is not extensively discussed, especially in the Russian-language literature. This paper is a systematic review of the growing criticism of several assumptions of psychometrics: 1) the existence of unobservable personality traits and abilities which are supposed to determine the performance of test items; 2) the stochastic nature of traits, and as a consequence, the need for probabilistic models to study them; 3) aggregate data represent individual characteristics; 4) as in the natural sciences, models in psychometrics help reveal psychological reality that is not directly observable; 5) if testing the agreement of empirical data with a model yields a positive result, then the model correctly describes reality. These assumptions are criticized on the basis of the circular nature of the definitions of the traits themselves, logical errors ingrained in the assumptions, the methodological mixture of variability and randomness regarding behavior, the lack of a causal link between inter-individual variation in responses and individual responses; superficial analogies with models in the natural science to which psychometrics refers, as well as the substitution of scientific tasks for instrumental and pragmatic

¹ The article was prepared as a part of the state assignment of Ministry of Education and Science of the Russian Federation for 2020–2024 (No. FZNM-2020-0005).

ones. We conclude that modelling in psychometrics is counterproductive if used as a method of exploring the psychological reality behind the test. Some alternative practices of quantitative research are discussed. For example, testing the existence of variation at the level of the individual and the experimental search for its explanation. There are other possible alternatives, such as, a network perspective on psychological phenomena; faceted theory; or observation-oriented modeling. Although this kind of research is much more difficult to implement than standard “goodness-of-fit” tests, it is probably this kind of research that can provide an increase in psychological knowledge.

Keywords: psychometrics; modelling; realism; representation; trait; latent construct; assumptions; measurement.

References

- Alexander, P.A., Dumas, D., Grossnickle, E.M., List, A., & Firetto, C.M. (2016). Measuring Relational Reasoning. *Journal of Experimental Education*, 84(1), 119–151. doi: 10.1080/00220973.2014.963216
- Anastazi, A., & Urbina, S. (2007). *Psikhologicheskoe testirovanie* [Psychological Testing]. (A. A. Alekseev, transl. from English). St. Petersburg: Piter.
- Arhonditsis, G.B., Stow, C.A., Steinberg, L.J., Kenney, M.A., Lathrop, R.C., McBride, S.J., & Reckhow, K.H. (2006). Exploring ecological patterns with structural equation modeling and Bayesian analysis. *Ecological Modelling*, 192(3–4), 385–409. doi: 10.1016/j.ecolmodel.2005.07.028
- Artemenkov, S.L. (2021). Uporyadochennyye seti chastnykh korrelyatsiy v psikhologicheskikh issledovaniyakh [Ordered networks of partial correlations in psychological research]. *Modelirovanie i analiz dannyykh*, 11(2), 31–50. doi: 10.17759/mda.2021110202
- Birnbaum, M.H. (2008). New paradoxes of risky decision making. *Psychological Review*, 115(2), 463–501. doi: 10.1037/0033-295X.115.2.463
- Boag, S. (2011). Explanation in personality psychology: “Verbal magic” and the five-factor model. *Philosophical Psychology*, 24(2), 223–243. doi: 10.1080/09515089.2010.548319
- Borsboom, D. (2008). Latent variable theory. *Measurement: Interdisciplinary Research & Perspective*, 6(1–2), 25–53. doi: 10.1080/15366360802035497
- Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, 16(1), 5–13. doi: 10.1002/wps.20375
- Borsboom, D., Mellenbergh, G. J., & Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, 110(2), 203. doi: 10.1037/0033-295X.110.2.203
- Borsboom, D., Mellenbergh, G., & Heerden, J. (2004). The Concept of Validity. *Psychological Review*, 111, 1061–71. doi: 10.1037/0033-295X.111.4.1061
- Brown, A., & Croudace, T.J. (2014). Scoring and estimating score precision using multidimensional IRT models. In S.P. Reise, D.A. Revicki (Eds.), *Handbook of Item Response Theory Modeling* (pp. 325–351). New York; London: Routledge.
- Campbell, D.T. (1960). Recommendations for APA test standards regarding construct, trait, or discriminant validity. *American Psychologist*, 15(8), 546–553. doi: 10.1037/h0048255
- Canter, D. (Ed.). (2012). *Facet Theory: Approaches to Social Research*. Springer Science & Business Media.
- Cavagnaro, D.R., & Davis-Stober, C.P. (2014). Transitive in our preferences, but transitive in different ways: an analysis of choice variability. *Decision*, 1, 102–122. <http://dx.doi.org/10.1037/dec0000011>
- Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Möttus, R., Waldorp, L.J., & Cramer, A.O. (2015). State of the aRt personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality*, 54, 13–29. doi: 10.1016/j.jrp.2014.07.003

- Cramer, A.O. J., Van Der Sluis, S., Noordhof, A., Wichers, M., Geschwind, N., Aggen, ... Borsboom, D. (2012). Measurable Like Temperature or Mereological like Flocking? on the Nature of Personality Traits. *European Journal of Personality*, 26(4), 451–459. doi: 10.1002/per.1879
- Cureton, E.E. (1951). Validity. In E.F. Lindquist (Ed.), *Educational Measurement* (pp. 621–694). Washington, DC: American Council on Education.
- Downey, G. & Feldman, S.I. (1996). Implications of rejection sensitivity for intimate relationships. *Journal of Personality and Social Psychology*, 70, 1327–1343. doi: 10.1037/0022-3514.70.6.1327
- Flamm, D. (1998). History and outlook of statistical physics. *arXiv preprint physics/9803005*. doi: 10.48550/arXiv.physics/9803005
- Forssell, D., & Powers, W. (2009). Perceptual Control Theory. *Living Control Systems Pub*, 72, 74–75.
- Franić, S., Borsboom, D., Dolan, C.V., & Boomsma, D.I. (2014). The big five personality traits: psychological entities or statistical constructs? *Behavior Genetics*, 44(6), 591–604. doi: 10.1007/s10519-013-9625-7
- Franić, S., Dolan, C.V., Borsboom, D., & Boomsma, D.I. (2012). Structural equation modeling in genetics. In R.H. Hoyle (Ed.), *Handbook of Structural Equation Modelling* (pp. 617–635). New York, NY: The Guilford Press.
- Gerber, J.P. (2011). Six solutions to the circular nature of trait psychology. In S. Boag & N. Tiliopoulos (Eds.), *Personality and individual differences: Theory, assessment and application* (pp. 297–306). New York: Nova.
- Grice, J.W. (2011). *Observation oriented modeling: Analysis of cause in the behavioral sciences*. New York, NY: Academic Press.
- Grice, J.W., Barrett, P.T., Cota, L., Taylor, Z., Felix, C., Garner, S., Medellin, E., & Vest, A. (2017). Four bad habits of modern psychologists. *Behavioral Sciences*, 7(3), Article 53. doi: 10.3390/bs7030053
- Grice, J.W., Cota, L.D., Barrett, P.T., Wuensch, K.L., & Poteat, M. (2016). A simple and transparent alternative to logistic regression. *Advances in Social Sciences Research Journal*, 3(7). doi: 10.14738/assrj.37.2125.
- Haig, B.D., & Borsboom, D. (2008). On the conceptual foundations of psychological measurement. *Measurement Interdisciplinary Research and Perspectives*, 6, 1–2. doi:10.1037/1040-3590.7.3.238
- Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York: Basic Books.
- Jansen, B.R., & van der Maas, H.L. (1997). Statistical test of the rule assessment methodology by latent class analysis. *Developmental Review*, 17(3), 321–357. doi: 10.1006/drev.1997.0437
- Kane, M. (2006). Validity. In R.L. Brennan (Ed.), *Educational Measurement* (pp. 17–64). Westport, CT: ACE/Praeger Publishers.
- Kelley, T.L. (1929). *Scientific Method*. Cleveland, OH: Ohio State University Press.
- Kondratyuk, N.G., Burmistrova-Savenkova, A.V., & Morosanova, V. (2021). A Psychometric Analysis of the Russian Career Adapt-Abilities Scale in High School Students. *Vestnik Rossiyskogo universiteta družby narodov. Seriya: Psikhologiya i pedagogika – RUDN Journal of Psychology and Pedagogics*. 18(3), 555–575. (In Russian). doi: 10.22363/2313-1683-2021-18-3-555-575
- Kyngdon, A. (2011). Psychological measurement needs units, ratios, and real quantities: A commentary on Humphry. *Measurement*, 9(1), 55–58. doi: 10.1080/15366367.2011.558791
- Laming, D.R.J. (1997). *The Measurement of Sensation*. Oxford, UK: Oxford University Press.
- Linacre, J.M. (1998). Structure in Rasch Residuals: Why Principal Components Analysis. *Rasch Measurement Transactions*, 12(2), 636. Retrieved from: <https://www.rasch.org/rmt/rmt122m.htm> (Accessed: 2 April 2021).

- Lord, F.M., & Novick, M. (1968). *Statistical Theories of Mental Test Scores*. Reading, Mass.: Addison-Wesley.
- Maas, Van der H.L.J., Molenaar, D., Maris, G., Kievit, R.A., & Borsboom, D. (2011). Cognitive Psychology Meets Psychometric Theory: On the Relation Between Process Models for Decision Making and Latent Variable Models for Individual Differences. *Psychological Review*, 118(2), 339–356. doi: 10.1037/a0022749
- MacCorquodale, K., & Meehl, P.E. (1948). On a distinction between hypothetical constructs and intervening variables. *Psychological Review*, 55(2), 95–107. doi: 10.1037/h0056029
- Maraun, M.D., & Gabriel, S.M. (2013). Illegitimate concept equating in the partial fusion of construct validation theory and latent variable modeling. *New Ideas in Psychology*, 31(1), 32–42. doi: 10.1080/15366360802035596
- Maraun, M.D., & Halpin, P.F. (2008). Manifest and Latent Variates. *Measurement: Interdisciplinary Research and Perspectives*, 6, 113–117. doi: 10.1080/15366360802035596
- Matthews, G., Deary, I.J., & Whiteman, M.C. (2003). *Personality Traits*. Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9780511812736
- Michell, J. (1999). *Measurement in Psychology: A Critical History of a Methodological Concept*. Cambridge, UK: Cambridge University Press.
- Michell, J. (2003a). Epistemology of measurement: the relevance of its history for quantification in the social sciences. *Social Science Information*, 42, 515–534. doi: 10.1177/0539018403424004
- Michell, J. (2003b). Pragmatism, positivism and the quantitative imperative. *Theory & Psychology*, 13, 45–52. doi: 10.1177/0959354303013001761
- Michell, J. (2014). The Rasch paradox, conjoint measurement, and psychometrics: Response to Humphry and Sijtsma. *Theory & Psychology*, 24(1), 111–123. doi: 10.1177/0959354313517524
- Molenaar, P.C. M., & von Eye, A. (1994). On the arbitrary nature of latent variables. In A. von Eye & C.C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 226–242). Thousand Oaks CA: Sage Publications, Inc.
- Mueller, R.O., & Hancock, G.R. (2018). Structural equation modeling. In G.R. Hancock, L.M. Stapleton, R.O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences* (pp. 445–456). New York: Routledge. doi: 10.4324/9781315755649
- Nowland, T., Beath, A., & Boag, S. (2019). Objectivity, realism, and psychometrics. *Measurement: Journal of the International Measurement Confederation*, 145, 292–299. doi: 10.1016/j.measurement.2019.05.038
- Perugini, M., Costantini, G., Hughes, S., & De Houwer, J. (2016). A functional perspective on personality. *International Journal of Psychology*, 51(1), 33–39. doi: 10.1002/ijop.12175
- Pugesek, B.H., Tomer, A., & Von Eye, A. (2003). *Structural equation modeling: applications in ecological and evolutionary biology*. Cambridge, UK: Cambridge University Press.
- Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. University of Chicago Press.
- Schmittmann, V.D., Cramer, A.O. J., Waldorp, L.J., Epskamp, S., Kievit, R.A., & Borsboom, D. (2013). Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas in Psychology*, 31(1), 43–53. doi: 10.1016/j.newideapsych.2011.02.007
- Siegler, R.S. (1981). Developmental sequences within and between concepts. *Monographs of the Society for Research in Child Development*, 46(2), 1–74. doi: 10.2307/1165995
- Spearman, C. (1904). 'General intelligence,' objectively determined and measured. *The American Journal of Psychology*, 15(2), 201–293. doi: 10.1037/11491-006
- Terracciano, A., & McCrae, R.R. (2012). Why do (some) birds flock? Causality and the structure of characteristic adaptations. *European Journal of Personality*, 26, 449–450. doi: 10.1002/per.1878
- Thijssen, J. (2007). *Computational Physics*. Cambridge, UK: Cambridge University Press.

- Uher, J. (2021). Psychometrics is not measurement: Unraveling a fundamental misconception in quantitative psychology and the complex network of its underlying fallacies. *Journal of Theoretical and Philosophical Psychology*, 41(1), 58–84. doi: 10.1037/teo0000176
- Vion, D., Aassime, A., Cottet, A., Joyez, P., Pothier, H., Urbina, C. ... & Devoret, M.H. (2002). Manipulating the quantum state of an electrical circuit. *Science*, 296(5569), 886–889. doi: 10.1126/science.1069372

Received 08.04.2022; Revised 20.09.2022; Accepted 18.10.2022

Yulia A. Tyumeneva – Leading researcher, Language and Cognition Laboratory, The Pushkin State Russian Language Institute; Senior researcher, Institute of Education, HSE University, PhD.

Email: jutu@yandex.ru