

Ю.А. ТЮМЕНЕВА<sup>1</sup>, К.П. ВЕРГЕЛЕС<sup>2</sup>

<sup>1</sup> НИУ «Высшая школа экономики».

101000, Москва, ул. Мясницкая, д. 20.

<sup>2</sup> ФГБОУ ВО «Государственный институт русского языка им. А.С. Пушкина».

117485, Москва, ул. Академика Волгина, д. 6.

## ИССЛЕДОВАНИЕ ТРАНЗИТИВНОСТИ САМООЦЕНОК МОТИВАЦИИ<sup>1</sup>

*Аннотация.* Измерение в социальных науках подразумевает, что измеряемый признак количественный, то есть можно не только упорядочить величины признака, но и разницу между упорядоченными величинами выразить с помощью некоторой единицы измерения. Однако необходимость проверять это базовое допущение зачастую игнорируется. В пользу игнорирования имеются свои доводы, но фундаментально оно отвлекает социальные науки от выполнения основной задачи — исследовать реальность. В данной работе проверялось одно из требований к порядковой структуре мотивов — требование транзитивности: если  $a > b$  и  $b > c$ , то  $a > c$ . Если транзитивность не соблюдается, то мотивы невозможно оценивать даже в порядковой шкале («больше — меньше», «сильнее — слабее»), не говоря уже об их количественном измерении, на которое претендуют все методики с ликертовскими шкалами. На выборке в 250 студентов показано, что около половины респондентов соблюдали транзитивность при упорядочивании величин мотивов (внутренних, внешних и социальных), что оправдывает использование порядковых шкал для оценки, по крайней мере, для указанных мотивов и для двух величин: «больше» и «меньше»; однако даже в этих случаях для обоснования использования шкал Ликерта требуется дальнейшая проверка допущений об аддитивности для измерения мотивов. Другая же часть респондентов (около 40%) не смогли либо различить, либо упорядочить мотивы, поэтому не только измерение, но даже упорядочивание мотивов в этих случаях невозможно. Сделан вывод, что

---

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

ошибка транзитивности связана с индивидуальными особенностями респондентов и требует дальнейшего изучения как систематическая.

*Ключевые слова:* измерение в социальных науках; количественная и порядковая структура; транзитивность; мотивация.

*Для цитирования:* Тюменева Ю.А., Вергелес К.П. Исследование транзитивности самооценок мотивации // Социологический журнал. 2021. Том 27. № 2. С. 8–24. DOI: 10.19181/socjour.2021.27.2.8083

Шкалы Ликерта широко используются в социальных науках, в том числе в социологии, для измерения индивидуальных признаков: установок, мотивов, отношений и т. д. Однако редко исследуются базовые математические допущения для применения этих шкал. В данной статье на примере измерения мотивации обучения показано, что для участников оказывается не всегда возможным упорядочить разные виды мотивов между собой. Это подрывает одно из основных допущений об измеримости индивидуальных признаков — об их порядковой структуре. Поскольку порядковая структура — необходимое условие для аддитивности, а не наоборот, требования порядка должны соблюдаться прежде требований аддитивности. Невозможность упорядочить мотивы автоматически означает, что мотивы не имеют аддитивной структуры и, значит, неизмеримы. Такого рода экспликации базовых допущений представляют методический интерес для всех социальных наук.

#### **Количественная структура: упорядоченность и аддитивность**

Измеримый признак должен иметь количественную структуру, то есть он должен быть, как минимум, упорядочиваемым и аддитивным. Аксиоматические условия количественной структуры известны со времен Евклида [3, с. 142–152], систематически были изложены математиком Отто Гёльдером [14, р. 238] и в связи с психологическими измерениями обсуждались Дж. Мичеллом [19, р. 52–53, 117, 134; 20, р. 52–54]. Говоря формально, величины  $a$ ,  $b$  и  $c$  упорядочены относительно друг друга, если они:

- 1) транзитивны, то есть для любых значений  $a$ ,  $b$  и  $c$ , если  $a > b$  и  $b > c$ , то  $a > c$ ;
- 2) асимметричны, то есть для любых значений  $a$  и  $b$ , если  $a > b$ , то не  $b > a$ ;
- 3) строго связаны, то есть для любых значений  $a$  и  $b$ , если  $a \neq b$ , то или  $a > b$ , или  $b > a$ .

Важно, что порядковый признак может не удовлетворять требованиям аддитивности, то есть не быть количественным. Для аддитивной (и тем самым количественной) структуры должны соблюдаться дополнительные требования:

4) разрешимости: для каждой пары величин  $a$  и  $b$  либо  $a = b$ , либо есть такая величина  $c$ , что  $a = b + c$  или  $b = a + c$ ;

5) позитивности:  $(a + b) > a$ ;

6) коммутативности:  $a + b = b + a$ ;

7) ассоциативности:  $a + (b + c) = (a + b) + c$ .

Отсюда следует, что аксиомы порядка должны соблюдаться прежде, чем аксиомы количества. Это означает, что признак может быть порядковым, то есть допускать измерения в порядковой шкале, но не быть аддитивным, то есть не допускать измерений в интервальной шкале. И если условия порядка не соблюдаются, то признак автоматически становится неизмеримым не только в порядковой, но и в интервальной шкале.

### **Игнорирование проверки допущений о количественной природе признака**

Измерение индивидуальных признаков в социальных науках с помощью тестов, опросников, шкал дает результаты в виде «количества» измеренного признака у человека, в виде, например, баллов. Количественная оценка признака подразумевает, что сам признак измерим, то есть имеет количественную структуру, а именно величины признака упорядочиваемы и аддитивны. Однако допущение о количественной структуре признака зачастую проверке не подвергается, а сами допущения не эксплицируются.

Краткий исторический экскурс поможет объяснить, как получилось, что в социальных науках устоялась практика принимать допущение измеримости индивидуальных признаков без проверки. Хотя эта история не имеет прямой связи с эмпирической частью статьи, она показывает, что измеримость психологических и социальных феноменов (от ощущений до мотивов и ценностей) никогда не была доказана, но остается допущением с самого начала становления количественных социальных наук. Понимание этого факта должно помочь сформировать критическое отношение к привычным «количественным методам» и перенаправить активность исследователей с торопливых измерений социально-психологических признаков на тщательное их изучение.

Игнорирование необходимости базовых допущений об измеримости психологических феноменов стало и остается мейнстримом с момента появления эпохальной работы С. Стивенса [23]. В ней были артикулированы несколько идей, ключевых для современных количественных методов в социальных науках. Во-первых, термин «измерение» предлагалось теперь понимать в неклассическом его значении. Классическое понимание измерения, как оно принято, например, в самой развитой на сегодняшний день эмпирической науке — физике, — это обнаружение или оценка меры некоторой величины признака относительно какой-то другой его величины, взятой за единицу изме-

рения. Стивенс же, опираясь на репрезентативную теорию измерений (РТИ) [9; 22], предложил понимать под измерением «приписывание чисел». Во-вторых, в отличие от РТИ, концепция измерения Стивенса не предполагает существования отношений между объектами, которые были бы логически независимы от операций идентификации этих отношений. В согласии с операционализмом он понимает измерение не как отсылку к реальным объектам и не как абстракцию эмпирических отношений этих объектов, а как результат произвольной *операции*. Так, например, он пишет, что измерение в шкале отношений возможно там, где «существуют *операции* для определения равенности рангов, равенности интервалов, равенности отношений» [23, р. 679]. В-третьих, Стивенс расширил значение термина «измерение» от строго количественной оценки до любого суждения о признаке, включив в него классификации и упорядочивание.

Это переопределение ключевого для количественной науки понятия было предпринято С. Стивенсом, можно сказать, вынужденно. Его статья стала реакцией на выводы Британской ассоциации развития науки<sup>2</sup> о том, что психофизики (Г. Фехнер, Э. Вебер, Л. Терстоун, С. Стивенс и др.) не сумели продемонстрировать количественных свойств восприятия и ощущения человека [12; 13]. Эти выводы подытоживали семилетнюю работу состоящего преимущественно из физиков и математиков комитета ассоциации, который оценивал возможность (как утверждалось психофизиками) «количественной оценки сенсорных событий», то есть возможность измерять психологические феномены в духе классической физики: психофизики ставили задачу найти такое отношение (пропорцию, *ratio*) между различными количествами ощущения, которое можно было бы принять за единицу измерения ощущения. Это классическое, реалистское, понимание того, что представляет собой измерение, и психофизики того времени, все имеющие естественно-научную или математическую подготовку, использовали, конечно, его. Комитет, однако, счел процедуры, использованные психофизиками, и названные ими «измерения» не соответствующими их классическому значению, так как не обнаружил доказательств того, что для серии стимулов, каждый последующий из которых имел «едва заметное различие» от предыдущего, соответствующие приращения в интенсивности ощущения были равны между собой (цит. по: [19, р. 148]). *«Называние этих процедур измерением не добавляет ничего к его фактическому значению... Понятие измерения не имеет никакого мистического, внутренне присущего ему значения, которое, чудом не замеченное физиками, якобы предстоит открыть психологам. Это просто*

---

<sup>2</sup> Название ассоциации до 2009 г. В настоящее время это Британская научная ассоциация (British Association for the Advancement of Science). — *Прим. ред.*

*слово, конвенционально используемое для обозначения определенных идей. Его использование для обозначения других идей не расширяет, а разрушает его значение»* [13, р. 345]. Так что заключительная рекомендация звучала неутешительно для молодой психологии: предлагалось или отказаться от понятия «измерение», как не имеющего отношения к экспериментальным манипуляциям психофизиков и потому вводящего в заблуждение, или предоставить иные факты, которые можно было бы интерпретировать как измерение в «истинном смысле» [12].

Конечно, для психофизиков, прочно (хотя и ошибочно) убежденных примером физики, что изучать что-то можно, только измеряя, это, по сути, означало отказ в признании за психологией статуса научной дисциплины со всеми вытекающими отсюда имиджевыми, социальными и экономическими последствиями [11; 19]. Поэтому подход к измерению С. Стивенса, переопределяющий ключевой для науки термин, позволил изящно обойти проблему (не)измеримости психологических признаков и был быстро принят психологами. Он магическим образом уничтожил все логические противоречия и эффективно сформировал сегодняшний мейнстрим количественной психологии, где любой признак может быть измерен, если следовать правилам приписывания чисел.

В самом допущении, что измеряемые феномены (мотивы, установки, ценности, предпочтения и пр.) имеют количественную природу, логических противоречий нет. Возможно ли упорядочить величины некоторого признака и установить константную единицу измерения отношений между ними — это вопрос эмпирической проверки: если признак ее проходит, значит, он измерим. Отсюда ясно, почему доминирующую практику производства количественных «моделей», обходящуюся без проверки базовых допущений измеримости содержащихся в них переменных, Дж. Мичелл называет «патологической» [16; 17]. Статистики согласия, которыми часто обосновывают «модели» феноменов, в принципе, не способны проверить структуру измеряемого признака, поскольку сами основаны на допущении о его количественной природе [9; 22].

Попытки актуализировать проблему для широкого профессионального сообщества и необходимость ее решения уже на протяжении полувека предпринимаются небольшой группой исследователей за рубежом [10; 16–18; 19; 21] и в России [1; 4; 6; 7]. Например, экспериментально проверялась порядковая структура «нейротизма» — одного из факторов «большой пятерки» личностных черт [21], обычно измеряемых в ликертовской шкале NEO (Neuroticism-Extroversion-Openness Inventory [15]) как аддитивные. В эксперименте С. Моррис с коллегами [21] респонденты попарно сравнивали себя и нескольких хорошо знакомых им людей по выраженности «нейротизма» с помощью формулировок стандартного опросника. Было показано, что основное

для порядковой структуры требование транзитивности нарушается в 12–25% случаев, а потому признак (точнее, его репрезентация в самооценке) не имеет даже порядковой структуры, не говоря уже об аддитивной. Однако в целом эти усилия не вызывают большого интереса у исследователей, что поразительно, если учесть фундаментальность проблемы измеримости для всех количественно ориентированных социальных наук. Поэтому нам кажется важным продолжить проверку измеримости индивидуальных признаков через эмпирическое тестирование аксиом, прежде всего о порядковой структуре признаков.

### **Проверка транзитивности оценок мотивации**

В социально-психологических исследованиях распространены многомерные методики, оценивающие сравнительную выраженность у человека тех или иных «мотивационных факторов», которыми человек руководствуется в своем поведении. В качестве таких методик можно привести «Шкалы академической мотивации» [2] или «Опросник профессиональной мотивации» [5]. Эти методики предлагают респондентам оценить на ликертовской шкале выраженность у себя различных мотивов. В зависимости от цели исследования индивидуальные самооценки по каждому виду мотивации усредняются или взвешиваются, и определяется доминирующий вид мотивации у респондента; затем данные агрегируются, чтобы оценить межгрупповые различия в выраженности мотивов, связи с жизненными событиями, учебными достижениями и пр.

Такое «измерение» видов мотивации предполагает, в частности:

- идентифицируемость видов мотивации как различных;
- сходство структуры разных видов мотивации;
- экстенсивную изменчивость величин разных видов мотивации без изменений в качестве;
- универсальную способность респондентов оценивать степень своего согласия с утверждением по шкале Ликерта, одинаково и однозначно интерпретировать ее деления и пр.

В данной работе тестируется одно из допущений количественного измерения мотивации, а именно что мотивы сравнимы между собой по величине. Это допущение касается порядковой структуры мотивов, оно предшествует более сильному допущению об их аддитивности. Если допущение о порядковой структуре подтвердится, это не приведет автоматически к подтверждению аддитивности, так как последняя должна удовлетворять большему количеству условий. Но если порядковая структура не подтвердится, то это автоматически опровергнет и аддитивность мотивации, указывая на ее неизмеримость.

Как было указано выше, порядковая структура должна удовлетворять требованию транзитивности: если  $a > b$  и  $b > c$ , то  $a > c$  (для

строого порядка, и если  $a \geq b$  и  $b \geq c$ , то  $a \geq c$  для нестрогого порядка). В нашем случае если мотивы доступны для упорядочивания, то есть если (а) они отличаются друг от друга, (б) имеют порядковую структуру, и (в) респонденты способны эту структуру различать, то можно ожидать выполнения требования транзитивности. Так, если мотиву А приписывается большая величина (степень, или уровень выраженности), чем мотиву В, а мотиву В — большая величина, чем мотиву С, то величина мотива А должна быть больше, чем величина мотива С. И если вдруг окажется, что  $A > B$  и  $B > C$ , но при этом  $A < C$ , то это будет являться нарушением транзитивности, а значит, одно, два или все три допущения об упорядоченности мотивов неверны: либо мотивы неразличимы между собой, либо они не имеют порядковой структуры, либо респонденты не могут их упорядочить. В любом случае это означает, что для оценки мотивации нельзя корректно использовать не только количественную (интервальную), но даже порядковую шкалу.

#### **Эмпирическая проверка транзитивности самооценок мотивов**

Исследовались три вида академической мотивации: внутренняя, связанная с интересом к учебе; внешняя, связанная с необходимостью получить диплом и найти работу; и социальная, связанная с притягательностью общения и социальных отношений в университете. Выбор мотивов был связан с их относительно легкой интуитивной различимостью. Это было важно, так как возможные нарушения транзитивности при упорядочивании мотивов должны объясняться сравнимостью величин мотивов, а не их плохой различимостью. Поэтому мы не стали брать мотивы, популярные у исследователей, но плохо различимые — такие, как мотивация достижения (стремление добиваться максимально высоких результатов и испытывать удовольствие в процессе решения трудных задач) или саморазвития (стремление к развитию своих способностей и достижению компетентности).

Для каждого мотива были сформулированы утверждения, подобные тем, что обычно используются в распространенных методиках измерения мотивации, и затем респондентов просили выбирать из попарно сгруппированных утверждений (в каждой паре они соответствовали разным мотивам) более значимое в качестве фактора, мотивирующего их учебу. Проверялась способность респондентов упорядочить утверждения так, чтобы между разными мотивами сохранялась транзитивность отношений. Если респонденты способны количественно оценивать мотивы (что является условием для измерения с помощью ликертовских шкал), тогда они непременно должны быть способны делать транзитивные парные сравнения, упорядочивая субъективную важность разных мотивов по отношению к одной и той же ситуации.

## Метод

*Выборка.* В исследовании приняли участие 272 студента 1–3-го курса (средний возраст 20,4 года;  $sd = 2,1$ ), добровольные участники более широкого онлайн-исследования, цель которого не имела отношения к данной работе. После удаления из анализа случаев с нереалистично быстрым заполнением основной анкеты исследования и однотипными ответами (например, всегда выбирали среднее деление шкалы в опросниках анкеты) выборка составила 250 студентов.

*Инструмент и процедура.* Участников просили оценить важность различных мотивов для их учебы в университете по следующим формулировкам:

- внутренняя (internal) мотивация (интерес):
  - потому что мне интересно учиться ( $I_1$ );
  - потому что мне кажется важным то, что мы изучаем ( $I_2$ );
  - чтобы знать больше интересных для меня вещей ( $I_3$ );
- внешняя (external) мотивация (обязательства):
  - чтобы получить потом диплом ( $E_1$ );
  - чтобы потом получить работу ( $E_2$ );
  - чтобы меньше волноваться о своем будущем ( $E_3$ );
- социальная (social) мотивация:
  - чтобы встречаться с друзьями ( $S_1$ );
  - чтобы общаться с приятными мне людьми ( $S_2$ );
  - потому что здесь хорошая компания ( $S_3$ ).

Участники получили следующую инструкцию:

Вы увидите возможные причины того, почему люди ходят на учебу. Причины сгруппированы в пары. В каждой паре обе причины важны, но Вам нужно выбрать тот вариант, который лично для Вас более важен, чем другой вариант. Если обе причины важны одинаково и Вы не можете определить, что для Вас важнее, укажите вариант «Одинаково важно».

Пример задания.

Я хожу на занятия, прежде всего...

- 1) потому что мне кажется интересным то, что мы изучаем;
- 2) чтобы встретиться с друзьями;
- 3) одинаково важно.

В каждой паре сравнивались утверждения, соответствующие двум разным видам мотивации, так что каждый мотив  $I$  сравнивался с каждым мотивом  $E$  и с каждым мотивом  $S$ . Всего участники делали выбор более важного мотива в 27 парах утверждений (3·3·3).



### Результаты

В этом исследовании проверялась нестрогая транзитивность, то есть если мотив  $I$  был важнее или такой же по важности, как  $E$  ( $I \geq E$ ), а  $E$  важнее или такой же, как  $S$  ( $E \geq S$ ), то  $I$  мог быть только важнее или быть таким же, как  $S$  ( $I \geq S$ ). Если же мотив  $S$  оказывался в каком-то сравнении важнее, чем  $I$ , то это считалось нарушением транзитивности. Один триплет (три сравнения между мотивами) мог результативаться любым из восьми сочетаний:

- 1)  $I \geq E \geq S \geq S \geq I$ ;
- 2)  $I \geq E \geq S \geq I \geq S$ ;
- 3)  $I \geq E \geq S \geq E \geq I$ ;
- 4)  $I \geq E \geq S \geq E \geq S$ ;
- 5)  $E \geq I \geq S \geq S \geq I$ ;
- 6)  $E \geq I \geq S \geq I \geq S$ ;
- 7)  $E \geq I \geq S \geq E \geq I$ ;
- 8)  $E \geq I \geq S \geq E \geq S$ .

Два сочетания из восьми, а именно 1 и 8, свидетельствовали о нарушении транзитивности.

Анализ транзитивности был проведен 2 раза. Первый раз анализировались все возможные сочетания парных выборов между мотивами независимо от того, какое именно утверждение (из трех возможных) представляло каждый вид мотивации. То есть утверждения одного вида мотивации рассматривались как взаимозаменяемые. Здесь для каждого респондента был получен 81 триплет, где выбор любого утверждения  $I$  vs любого утверждения  $E$  соотносился с выбором любого  $E$  vs любого  $S$  и любого  $I$  vs любого  $S$ .

Второй раз анализ проводился только на тех триплетах, где каждый вид мотивации был представлен каким-то одним утверждением из трех. Например, был допустим триплет со сравнениями между  $I_1E_1$ ,  $E_1S_1$ ,  $I_1S_1$ , но не между  $I_1E_1$ ,  $E_1S_1$ ,  $I_1S_2$ . То есть сравнивались триплеты с фиксированными для каждого вида мотивации утверждениями. Это гарантировало, что нарушения транзитивности, если они происходили, не были вызваны семантическими различиями между утверждениями внутри одного вида мотивации. Для каждого респондента было проанализировано 27 триплетов с фиксированными утверждениями. Сравнение частоты нарушений транзитивности для 81 и для 27 триплетов могло показать, обоснованно ли считать психологически эквивалентными утверждения, относимые к одному какому-то виду мотивации.

### Анализ распределения нарушений по триплетам

Результаты показали, что и для 81 триплета, и для 27 триплетов в среднем фиксировалось около 5–6% нарушений транзитивности на триплет (табл. 1), то есть нарушения распределились относительно равномерно, и не было выявлено триплетов, которые бы «отвечали» за

большую часть ошибок: тест Q Кохрена был значим на уровне  $\alpha = 0,018$  для 81 триплета, чего с учетом множественности сравнений было недостаточно для отклонения гипотезы о равенстве дисперсий в триплетах, и незначим ( $\alpha = 0,252$ ) для 27 триплетов.

Таблица 1

**Доля нарушений транзитивности на триплет, %**

| Кол-во триплетов в анализе | Минимум | Максимум | Медиана | Среднее |
|----------------------------|---------|----------|---------|---------|
| 81                         | 2,9     | 9,2      | 6,0     | 5,9     |
| 27                         | 2,8     | 9,2      | 4,8     | 5,0     |

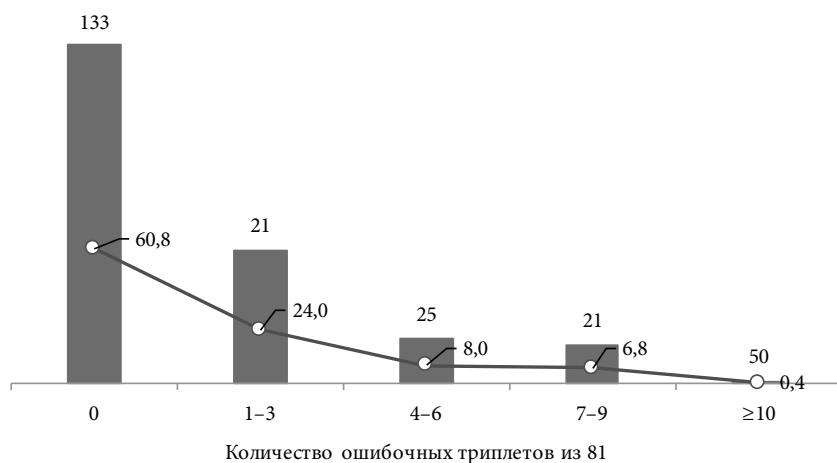
Анализ по респондентам ( $n = 250$ ) показал, что более половины участников смогли упорядочить важность мотивации без нарушения транзитивности, но четверть всех респондентов сделали около 80–90% всех нарушений транзитивности, так что распределились нарушения очень неравномерно (табл. 2).

Таблица 2

**Сводка по индивидуальным нарушениям транзитивности, %**

| Кол-во триплетов | Среднее | Вариация | Медиана | Случаев без нарушений | Среднее в случаях с нарушениями | Доля ошибок от общего их числа, совершенных в 25% выборки с максимальным количеством нарушений |
|------------------|---------|----------|---------|-----------------------|---------------------------------|--|
| 81               | 6       | 0–40     | 0       | 53                    | 13                              | 80   |
| 27               | 5       | 0–45     | 0       | 60                    | 13                              | 90   |

На рисунке 1 показано распределение нарушений более детально. Для 81 триплета, где утверждения были взаимозаменяемыми в рамках своего мотива, у 53% респондентов не было найдено ни одного нарушения транзитивности, у 8% найдено не более трех нарушений, но у каждого пятого 10 триплетов или более были «ошибочными» (рис. 1а). Для 27 триплетов с фиксированными утверждениями у 60% респондентов не было найдено ни одного нарушения транзитивности, у 24% было не более трех «ошибочных» триплетов, и только у одного человека более 10 триплетов были «ошибочными» (рис. 1б). Так что, хотя неравномерность нарушений транзитивности по выборке респондентов была найдена для обоих составов триплетов, транзитивность в триплетах с фиксированными утверждениями сохранялась чаще, чем в триплетах со взаимозаменяемыми утверждениями.



**Рис. 1. Распределение пропорций индивидуальных нарушений транзитивности:**

- а) для 81 триплета со взаимозаменяемыми утверждениями;**  
**б) для 27 триплетов с фиксированными утверждениями**

*Примечание:* Подписи столбцов в гистограммах показывают абсолютное количество респондентов с той или иной частотой нарушений транзитивности.

### Обсуждение

Целью этой работы было эмпирически проверить допущение о порядковой структуре мотивов — необходимое, хотя и недостаточное условие для ее количественной оценки — путем тестирования транзитивности отношений между их величинами.

Значительная часть респондентов (50–60%) безошибочно упорядочили величины исследуемых здесь видов мотивации по отношению друг к другу. Это говорит о том, что внутренняя, внешняя и социальная мотивации могут различаться и упорядочиваться по своей субъективно воспринимаемой величине. Следовательно, эти виды мотивации имеют порядковую структуру и могут оцениваться порядковой шкалой, по крайней мере, с двумя уровнями — «больше» и «меньше».

Оптимистичный вывод, однако, сразу же сталкивается с серьезным ограничением в виде значительной части респондентов, которые не могли различить и/или упорядочить эти мотивы, по крайней мере, в ряде случаев. Об этом свидетельствует, во-первых, тот факт, что за 80–90% всех ошибок «отвечало» 25% респондентов, а во-вторых, что среди респондентов, допустивших ошибки транзитивности (таких было 40–50% выборки), около 13% триплетов были ошибочными. В этом смысле 5–6% нарушений транзитивности, полученных в среднем по всей выборке, обязаны именно «безошибочным» респондентам, но никак не характеризуют группу, которая не смогла упорядочить

мотивы. Поэтому имеет смысл интерпретировать данные по нарушениям отдельно для группы респондентов, не сумевших упорядочить мотивы. Можем ли мы считать эти 13% нарушений случайными ошибками, связанными, скажем, с «невнимательностью» 40–50% респондентов?

Ошибки в упорядочивании, как и все другие ошибки, могут быть отнесены либо к случайным, либо к систематическим. Для интерпретации частоты случайных ошибок нужно задать какие-то осмысленные границы, в которых эмпирическая частота может быть оценена. Частота нарушений транзитивности будет наибольшей в случаях принципиальной невозможности не только оценить величину признака, но и упорядочить его разные величины; например, нельзя упорядочить тон голоса по ширине или романскую группу языков — по возрастной «романизму».

Частота случайных нарушений транзитивности будет наименьшей, если упорядочиваются субъективно различимые величины признаков с количественной структурой, такие как длина или вес. В этом случае частота нарушений будет определяться только возможностями респондента, например, различительными.

Случайная генерация выборов 27 попарных сравнений дала в среднем 25% триплетов с нарушением транзитивности (минимально 4%, максимально 48%), что и было верхним порогом для частоты случайных нарушений. В качестве нижней границы частоты случайных ответов была принята пропорция нарушений транзитивности, сделанных при попарном ранжировании испытуемыми собственных знакомых по объективным характеристикам рост и вес [23], когда в среднем около 4% сравнений нарушали транзитивность. Поскольку рост и вес имеют количественную структуру, только восприятие людей могло вызывать эти ошибки. Иными словами, даже если бы мотивы имели бесспорную количественную структуру, мы не могли бы избежать нарушений транзитивности примерно в 4% случаев. Поэтому, имея 4% как нижнюю границу и 25% как верхнюю, можно заключить, что 13% нарушений, допущенных в среднем респондентами, не сумевшими упорядочить мотивы, говорят о неслучайном происхождении этих нарушений и о том, что структура мотивов и/или восприятие респондента препятствует упорядочиванию.

Возможное объяснение заключается в том, что способность сравнивать выраженность мотивов связана с нечувствительностью некоторых респондентов к тонкой дифференциации субъективного опыта или с тем, что у этих респондентов структура мотивации (или восприятие собственной мотивации) не соотносится с заданными в нашем исследовании мотивами. В пользу последнего свидетельствует тот факт, что часть респондентов не показали нарушений транзитивности в триплетах с фиксированными утверждениями, а ошибались

в упорядочивании взаимозаменяемых мотивов. Если это так, то семантическая близость утверждений между собой внутри определенного вида мотивации не делает их психологически эквивалентными для части респондентов. Поскольку предположение о взаимозаменяемости стимулов в рамках какой-то шкалы (мотивации, ценности) — вообще частое допущение количественных методов, наши результаты в данной части заслуживают отдельного внимания.

В любом случае следует признать, что для существенной доли респондентов нельзя делать заключение о порядковой структуре мотивации, поэтому мы не можем ее оценить даже в порядковой шкале, не говоря уже об интервальной.

В данной работе мы не рассчитывали вскрыть природу ошибок транзитивности — это задача следующих исследований. Главной интенцией было обосновать необходимость проверки допущений, которые делаются при использовании ликертовских шкал, и исследовать структуру признаков, прежде чем приступать к их «измерению». При всех имиджевых и социально-экономических рисках, которые несет самокритическая проверка оснований привычных «измерительных» практик, это представляется единственным способом для психологии и социологии исследовать реальность, а не довольствоваться ее моделями неизвестной репрезентационной ценности.

## ЛИТЕРАТУРА

1. *Беребин М.А., Новохацки А.В.* О математической адекватности процедур психометрического нормирования результатов экспериментально-психологического исследования // Вестник Южно-Уральского государственного университета. Серия: Психология. 2010. № 27. С. 87–90.
2. *Гордеева Т.О., Сычев О.А., Осин Е.Н.* Опросник «Шкалы академической мотивации» // Психологический журнал. 2014. Т. 35. № 4. С. 96–107.
3. *Евклид.* Книга пятая // Начала / Пер. с греч. и коммент. Д.Д. Мордухай-Болтовского; При ред. участии М.Я. Выгодского и И.Н. Веселовского. [2-е изд., стер.]. М.; Л.: Гос. изд-во техн.-теорет. лит., 1950. С. 143–172.
4. *Руднев М.Г.* Инвариантность измерения базовых ценностей по методике Шварца среди русскоязычного населения четырех стран // Социология: 4М. 2013. № 37. С. 7–38.
5. *Осин Е.Н., Горбунова А.А. и др.* Профессиональная мотивация сотрудников российских предприятий: диагностика и связи с благополучием и успешностью деятельности // Организационная психология. 2017. Т. 7. № 4. С. 21–49.
6. *Пахомов А.П.* Проблема осмысленности психологических измерений // Психологический журнал. 2006. Т. 27. № 5. С. 75–82.
7. *Серенков П.С., Романчук В.М.* Качество как субъективно измеряемая величина // Приборы и методы измерений. 2019. Т. 10. № 1. С. 99–110. DOI: 10.21122/2220-9506-2019-10-1-99-110

8. *Barrett P.* The EFPA test-review model: When good intentions meet a methodological thought disorder // Behavioral Sciences. 2018. Vol. 8. No. 1 (5). P. 1–22. DOI: 10.3390/bs8010005
9. *Campbell N.R.* Physics, the Elements. London: Cambridge University Press, 1920. — 566 p.
10. *Cliff N.* Abstract Measurement Theory and the Revolution That Never Happened // Psychological Science. 1992. Vol. 3. No. 3. P. 186–190. DOI: 10.1111/j.1467-9280.1992.tb00024.x
11. *Danziger K.* Constructing the subject: Historical origins of psychological research. Cambridge: Cambridge University Press, 1994. — 254 p. DOI: 10.1017/s0033291700022480
12. *Ferguson A., Myers C.S., Bartlett R.J., et al.* Quantitative estimates of sensory events: Interim report of the committee appointed to consider and report upon the possibility of quantitative estimates of sensory events // British Association for the Advancement of Science. 1938. No. 108. P. 277–334. DOI: 10.1038/130334a0
13. *Ferguson A., Myers C.S., Bartlett R.J., et al.* Quantitative estimates of sensory events: Final report of the committee appointed to consider and report upon the possibility of quantitative estimates of sensory events // British Association for the Advancement of Science. 1940. No. 1. P. 331–349.
14. *Hölder O.* The Axioms of Quantity and the Theory of Measurement / Transl. from Germ. by J. Michell, C. Ernst // Journal of Mathematical Psychology. 1996. Vol. 40. No. 3. P. 235–252. DOI:10.1006/jmps.1996.0023
15. *McCrae R.R., Costa P.T.* Validation of the five-factor model of personality across instruments and observers // Journal of Personality and Social Psychology. 1987. Vol. 52. No. 1. P. 81–90. DOI:10.1037/0022-3514.52.1.81
16. *Michell J.* Normal science, pathological science and psychometrics // Theory & Psychology. 2000. Vol. 10. No. 5. P. 639–667. DOI:10.1177/0959354300105004
17. *Michell J.* Is Psychometrics Pathological Science? // Measurement: Interdisciplinary Research & Perspective. 2008. Vol. 6. No. 1–2. P. 7–24.
18. *Michell J.* Sensitivity of preferences and ratings to ordered metric structure in attitudes // Australian Journal of Psychology. 1998. Vol. 50. No. 3. P. 199–204. DOI:10.1080/00049539808258796
19. *Michell J.* Measurement in psychology: A critical history of a methodological concept. Cambridge: Cambridge University Press, 1999. — 268 p.
20. *Michell J.* An introduction to the logic of psychological measurement. New York: Psychology Press, 2014. — 190 p.
21. *Morris S.D., Grice J.W., Cox R.A.* Scale Imposition as Quantitative Alchemy: Studies on the Transitivity of Neuroticism Ratings // Basic and Applied Social Psychology. 2017. Vol. 39. Iss. 1. P. 1–18. DOI:10.1080/01973533.2016.1256288
22. *Russell B.* The principles of mathematics. London: Cambridge University Press, 1903. — 586 p.
23. *Stevens S.S.* On the theory of scales of measurement // Science. 1946. No. 103. P. 667–680.

**СВЕДЕНИЯ ОБ АВТОРАХ**

**Тюменева Юлия Алексеевна** — кандидат психологических наук, старший научный сотрудник, Институт образования, Национальный исследовательский университет «Высшая школа экономики».

**Телефон:** +7 (910) 407-51-40. **Электронная почта:** jutu@yandex.ru

**Вергелес Ксения Петровна** — младший научный сотрудник лаборатории когнитивных и лингвистических исследований, ФГБОУ ВО «Государственный институт русского языка им. А.С. Пушкина».

**Телефон:** +7 (985) 489-59-94. **Электронная почта:** vergeles.k.soc@gmail.com

Дата поступления: 21.10.2020.

---

**SOTSIOLOGICHESKIY ZHURNAL = SOCIOLOGICAL JOURNAL. 2021.**  
**VOL. 27. NO. 2. P. 8–24.** DOI: 10.19181/socjour.2021.27.2.8083

**YULIA A. TUMENEVA<sup>1</sup>, KSENIYA P. VERGELES<sup>2</sup>**

<sup>1</sup>International Laboratory for Evaluation of Practices and Innovations in Education, Institute of Education, National Research University Higher School of Economics. Myasnitskaya St. 20, 101000, Moscow, Russian Federation.

<sup>2</sup>Center for Cognition and Communication, Pushkin State Russian Language Institute. Acadimika Volgina St. 6, 117485, Moscow, Russian Federation.

**STUDYING THE TRANSITIVITY OF MOTIVATION RATINGS**

*Abstract.* Measurement in social sciences implies that the measured feature is quantitative, or in other words that it is possible not only to arrange the values of any given attribute, but also to express the difference between ordered magnitudes using a certain unit of measurement. However the need to verify this basic assumption is often ignored. And though there are a few possible excuses for this, but fundamentally this neglect distracts the social sciences from its main task of exploring reality. In this work, one of the requirements for the ordinal structure of motives was checked, namely the requirement of transitivity: if  $a > b$  and  $b > c$ , then  $a > c$ . If transitivity is not observed, then motives cannot be evaluated even on an ordinal scale (“more – less”, “stronger – weaker”), not to mention their quantitative measurement, which all methods that use Likert scales are supposedly tailored to. On a sample of 250 students, it was shown that about half of the respondents established transitivity when arranging their motives (internal, external and social ones), which justifies the use of ordinal scales for motivation assessment, at least for these motives and for two values: “more” and “less”; however, even in these cases, further validation of the assumptions about additivity when it comes to measuring motives is required to justify the use of Likert scales. The other part of the respondents (about 40%) could neither distinguish nor arrange their motives, therefore not only measuring, but even defining the order of their motives in these cases is impossible. It is concluded that the transitivity error is associated with the individual characteristics of the respondents and requires further study as a systematic error.

*Keywords:* measurement in social sciences; ordinal and quantitative structure; transitivity; motivation.

**For citation:** Tumeneva, Y.A., Vergeles, K.P. The transitivity of motivation ratings. *Sotsiologicheskii Zhurnal = Sociological Journal*. 2021. Vol. 27. No. 2. P. 8–24. DOI: 10.19181/socjour.2021.27.2.8083

## REFERENCES

1. Berebin M.A., Novohacki A.V. On the mathematical adequacy of the procedures for psychometric normalization of the results of experimental psychological research. *Vestnik Juzhno-Ural'skogo gosudarstvennogo universiteta. Psihologija*. 2010. No. 27. P. 87–90. (In Russ.)
2. Gordeeva T.O., Sychev O.A., Osin E.N. Questionnaire “Scales of academic motivation”. *Psihologicheskij zhurnal*. 2014. Vol. 35. No. 4. P. 96–107. (In Russ.)
3. Евклид. *Pyataya kniga. Nachala*. [Fifth book. Beginings.] Transl. from Greek and comment. by D.D. Mordukhai-Boltovskii, M.Ya. Vygodskii, I.N. Veselovskii. Moscow; Leningrad: Gos. izd-vo tehn.-teoret. lit. publ., 1950. P. 143–172. (In Russ.)
4. Rudnev M.G. Measurement Invariance of Basic Human Values Assessed with Schwartz Instrument among four Russian-Speaking Populations. *Sociologija: 4M*. 2013. No. 37. P. 7–38. (In Russ.)
5. Osin E.N., Gorbunova A.A., et al. Professional motivation of employees of Russian enterprises: diagnostics and connections with well-being and success of activity. *Organizacionnaja psihologija*. 2017. Vol. 7. No. 4. P. 21–49. (In Russ.)
6. Pahomov A.P. The problem of meaningfulness of psychological measurements. *Psihologicheskij zhurnal*. 2006. Vol. 27. No. 5. P. 75–82. (In Russ.)
7. Serenkov P.S., Romanchak V.M. Quality as a subjectively measured quantity. *Pribory i metody izmerenii*. 2019. Vol. 10. No. 1. P. 99–110. DOI: 10.21122/2220-9506-2019-10-1-99-110 (In Russ.)
8. Barrett P. The EFPA test-review model: When good intentions meet a methodological thought disorder. *Behavioral Sciences*. 2018. Vol. 8. No. 1 (5). P. 1–22. DOI: 10.3390/bs8010005
9. Campbell N.R. *Physics, the Elements*. L.: Cambridge University Press, 1920. 566 p.
10. Cliff N. Abstract Measurement Theory and the Revolution That Never Happened. *Psychological Science*. 1992. Vol. 3. No. 3. P. 186–190. DOI: 10.1111/j.1467-9280.1992.tb00024.x
11. Danziger K. *Constructing the subject: Historical origins of psychological research*. Cambridge: Cambridge University Press, 1994. 254 p. DOI: 10.1017/s0033291700022480
12. Ferguson A., Myers C.S., Bartlett R.J., et al. Quantitative estimates of sensory events: Interim report of the committee appointed to consider and report upon the possibility of quantitative estimates of sensory events. *British Association for the Advancement of Science*. 1938. No. 108. P. 277–334. DOI: 10.1038/130334a0
13. Ferguson A., Myers C.S., Bartlett R.J., et al. Quantitative estimates of sensory events: Final report of the committee appointed to consider and report upon the possibility of quantitative estimates of sensory events. *British Association for the Advancement of Science Advancement of Science*. 1940. No. 1. P. 331–349.
14. Hölder O. The Axioms of Quantity and the Theory of Measurement. Transl. from Germ. by J. Michell, C. Ernst. *Journal of Mathematical Psychology*. 1996. Vol. 40. No. 3. P. 235–252. DOI:10.1006/jmps.1996.0023
15. McCrae R.R., Costa P.T. Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*. 1987. Vol. 52. No. 1. P. 81–90. DOI:10.1037/0022-3514.52.1.81
16. Michell J. Normal science, pathological science and psychometrics. *Theory & Psychology*. 2000. Vol. 10. No. 5. P. 639–667. DOI:10.1177/0959354300105004
17. Michell J. Is Psychometrics Pathological Science? *Measurement: Interdisciplinary Research & Perspective*. 2008. Vol. 6. No. 1–2. P. 7–24.
18. Michell J. Sensitivity of preferences and ratings to ordered metric structure in attitudes. *Australian Journal of Psychology*. 1998. Vol. 50. No. 3. P. 199–204. DOI:10.1080/00049539808258796



19. Michell J. *Measurement in psychology: A critical history of a methodological concept*. Cambridge: Cambridge University Press, 1999. 268 p.
20. Michell J. *An introduction to the logic of psychological measurement*. N.Y.: Psychology Press, 2014. 190 p.
21. Morris S.D., Grice J.W., Cox R.A. Scale Imposition as Quantitative Alchemy: Studies on the Transitivity of Neuroticism Ratings. *Basic and Applied Social Psychology*. 2017. Vol. 39. Iss. 1. P. 1–18. DOI:10.1080/01973533.2016.1256288
22. Russell B. *The principles of mathematics*. L.: Cambridge University Press, 1903. 586 p.
23. Stevens S.S. On the theory of scales of measurement. *Science*. 1946. No. 103. P. 667–680.

*INFORMATION ABOUT THE AUTHORS*

**Yulia A. Tumeneva** — Candidate of Psychological Sciences, Associate Professor, Institute of Education / Department of Educational Programmes; Senior Research Fellow, International Laboratory for Evaluation of Practices and Innovations in Education, Institute of Education, National Research University Higher School of Economics. **Phone:** +7 (910) 407-51-40. **Email:** jutu@yandex.ru

**Kseniya P. Vergeles** — Junior Research Fellow, Center for Cognition and Communication, Pushkin State Russian Language Institute. **Phone:** +7 (985) 489-59-94. **Email:** vergeles.k.soc@gmail.com

Received: 21.10.2020.

---