

<https://doi.org/10.3176/hum.soc.sci.1974.2.06>

Т. КИТВЕЛЬ

ОБ ОДНОЙ ВОЗМОЖНОСТИ УПОРЯДОЧЕНИЯ ПСИХОЛОГИЧЕСКИХ ПРИЗНАКОВ

В социальной психологии применяются многие методы для шкалирования изучаемых признаков. Приемы, разработанные Терстоуном, Ликертом, Гуттманом, Кумбсом и другими, широко известны и часто используются как в социально-психологических, так и в социологических исследованиях. Однако большинство этих методов создано скорее для конструирования шкал равных интервалов для определенного признака, чем для упорядочения самих признаков. Последнее требует нахождения какого-то более общего признака или такого алгоритма, который имел бы смысл не только в формальном плане, но и в плане содержания. Найти какой-то латентный, более общий признак для всех изучаемых не всегда удается — это хорошо известно из опыта применения именно с такой целью факторного анализа. Этим обусловлено также стремление исследователей зачастую конструировать свои схемы расположения (т. е. структуры, иерархии и т. д.) различных изучаемых признаков более или менее интуитивно и оправдывать их на уровне «здравого смысла». Наиболее сложно создание схемы различных отношений (установок, ценностных ориентаций) человека. Здесь трудно, если вообще возможно, выявить, какое отношение возникло раньше, какое позже; какое отношение вызывает другое или входит в другое в качестве его компонента. Р. Штольберг (ГДР) в одном из наиболее подробных исследований удовлетворенности трудом дал схему факторов, влияющих на нее, в виде концентрических кругов: в центре находится содержание труда; в следующем кругу — организация его, управление и условия труда; в третьем — признание результатов, возможность повышения квалификации, участие в принятии решений, товарищеская помощь; в четвертом — непрерывность плановых задач, понимание общественной и производственной перспективы, общественной необходимости труда; в пятом — отдаленность места работы от места жительства и используемые транспортные средства, в шестом — степень загруженности домашним хозяйством и уходом за детьми; в седьмом — гармоничность брака.¹ Выбор последовательности факторов остается неясным и спорным.

Один из путей упорядочения социально-психологических признаков дает использование так наз. несимметричных коэффициентов корреляции, в частности коэффициента g Л. Гудмана.²

Формула для расчета этого коэффициента проста:

$$g_{xy} = \frac{\sum f_i - F_d}{N - F_d},$$

где F_d — модальная частота объектов по признаку y ; f_i — модальная частота признака y при фиксированных значениях признака x ; N — количество объектов.

¹ R. Stollberg, Arbeitszufriedenheit-theoretische und praktische Probleme. Berlin, 1968, стр. 60.

² Л. Э. Гудман обозначил этот коэффициент λ . Нами использовано обозначение из: A. Siciński, Współczynniki zależności zmiennych niemierzalnych. В кн.: Metody statystyczne w socjologii. Warszawa, 1968, стр. 239.

(x и y — дискретно распределенные признаки). Данный коэффициент показывает, насколько уменьшаются ошибки при прогнозировании модального значения y при знании его распределения в разрезах x в сравнении с положением, когда мы этого не знаем. Т. е. можно сказать, что коэффициент g_{xy} в определенной мере измеряет влияние x на y . Так же можно измерить и обратное, т. е. влияние y на x , которое не всегда равно влиянию x на y . Точнее сказать — g_{xy} показывает ценность признака x в прогнозировании значения y у случайно выбранного объекта, а g_{yx} показывает обратное — прогностическую ценность признака y в угадывании x . Если знание значения x позволяет нам точнее прогнозировать значение y , чем знание значения y — значение x , то вполне разумно считать признак x более ценным и центральным, чем y . Для выявления, какой именно признак — x или y позволяет точнее прогнозировать значение другого, надо рассчитать разницу $g_{xy} - g_{yx}$.³

Упорядочение признаков проводим именно на основе разниц $g_{xy} - g_{yx}$ для всех пар рассматриваемых признаков. В результате этого получаем симметричную матрицу, для упорядочения которой существует много способов. Метод Бергера, например, дает выигрышам (разницам с плюсовым знаком) количество «очков», равное количеству выигрышей у побежденного. В случае получения равного количества «очков» у двух или более признаков, процедуру следует повторить. Метод Бергера дает нам такую же последовательность, какая получается и при проекции этих признаков на первый собственный вектор матрицы выигрышей-проигрышей. При этом плюсовые значения $g_{xy} - g_{yx}$ обозначены 1, нулевые $1/2$ и отрицательные 0. При упорядочении признаков нас больше интересует последовательность как целая система, а не ее отдельные составляющие (разницы $g_{xy} - g_{yx}$), существенность отличия которых от нуля мы в таком случае отдельно проверять не будем. Когда речь идет об упорядочении таких признаков, как оценки, мнения и т. п., у психологов могут возникнуть возражения против метода Бергера. Во-первых, если предположить (как это, например, делается в теории когнитивного диссонанса), что человек попытается уменьшить количество противоречивых друг другу оценок, трудно поверить, чтобы при этом использовалась такая относительно сложная система взвешивания. Во-вторых, метод Бергера не дает нам наименее противоречивую последовательность, если под непротиворечивыми последовательностями понимать такие ряды признаков, результаты которых в таблице разниц $g_{xy} - g_{yx}$ имели бы максимальное количество одинаковых знаков в одной стороне от главной диагонали. Представляется, что такое понимание непротиворечивости больше отвечает психологическим теориям, чем метод Бергера.

Трудности, связанные с упорядочением оценок таким образом, чтобы в одну сторону от главной диагонали попало бы максимальное количество разниц с одинаковыми знаками, связаны с тем, что этому требованию отвечают многие различные последовательности и необходимо искать процедуру для выбора лучшей из них. Тогда из двух или более последовательностей следует выбрать ту, которая больше всего соответствует содержанию исследуемого явления, а при отсутствии оправданий по содержанию, можно вторичным критерием для выбора использовать метод Бергера.

Поиск последовательности, дающей наибольшее количество одинаковых знаков в одной стороне от главной диагонали, требует иногда применения метода «проб и ошибок», что при большом количестве признаков может значительно затруднить дело по сравнению с методом Бергера. Пользоваться методом Бергера или нет — это следует решать исходя из задач, поставленных исследователем. В примере (табл. 1 и 2) мы упорядочиваем признаки таким образом, чтобы получить наибольшее количество одинаковых знаков в одной стороне от главной диагонали.

Метод Бергера не дает точно такой же последовательности, как в случае, когда критерием упорядочения служит получение максимального количества одинаковых знаков в одной стороне от главной диагонали. Это можно проверить и по данным табл. 1. Достижение последней цели (увеличение количества одинаковых знаков в одной стороне от главной диагонали) заставляет оставить на первом месте признак с двумя

³ Мало вероятно, что направление более точного прогноза противоположно направлению причинной связи.

Таблица 1

Разница $g_{xy} - g_{yx}$ в группе опрошенных моложе 30 лет (в сотых долях)

Оценки	Признаки							
	1	2	3	4	5	6	7	8
1. Культурному уровню товарищей по работе	×	+02	+03	-03	00	-06	+03	+02
2. Состоянию здоровья	-02	×	00	+03	+07	+01	+05	+13
3. Своему культурному уровню	-03	00	×	+03	+03	+03	+03	+08
4. Зарплате	+03	-03	-03	×	+02	00	+03	+10
5. Успехам в труде	00	-07	-03	-02	×	+02	+02	+12
6. Жилищным условиям	+06	-01	-03	00	-02	×	+06	+05
7. Интересам к труду	-03	-05	-03	-03	-09	-06	×	+11
8. Отношениям в своей семье	-02	-13	-08	-10	-12	-05	-11	×

Таблица 2

Разница $g_{xy} - g_{yx}$ в группе опрошенных старше 30 лет

Оценки	Признаки							
	1	2	3	4	5	6	7	8
1. Зарплате	×	+01	00	+01	+07	+06	+14	+04
2. Состоянию здоровья	-01	×	+01	00	+02	+06	+07	+02
3. Своему культурному уровню	00	-01	×	+01	+04	+01	+08	+03
4. Успехам в труде	-01	00	-01	×	+01	+06	+02	+10
5. Жилищным условиям	-07	-02	-04	-01	×	+04	+03	+08
6. Интересам к труду	-06	-06	-01	-06	-02	×	+09	+11
7. Культурному уровню товарищей по работе	-14	-07	-08	-02	-03	-09	×	00
8. Отношениям в своей семье	-04	-02	-03	-10	-08	-11	00	×

отрицательными разностями $g_{xy} - g_{yx}$, так как любая его перестановка увеличила бы количество минусов в правой верхней части табл. 1.

Нулевые разности $g_{xy} - g_{yx}$ в какой-то мере затрудняют упорядочение признаков, если исследователь почему-то не желает воспользоваться методом Бергера. Как более «консервативный» метод можно рекомендовать причисление к нулевым разностям $g_{xy} - g_{yx}$ знака, уменьшающего количество (плюсов или минусов) тех знаков, которые в большинстве своем находятся в той же стороне от главной диагонали.

Остается проверить, насколько вероятной можно считать полученную последовательность признаков. Если бы признаки у объектов из генеральной совокупности были «равнозначными» в смысле их прогностической ценности, количество плюсов-минусов в одной стороне от главной диагонали было бы почти равным, а вероятность получения строгой последовательности (т. е. почти всех одинаковых знаков в одной стороне диагонали) очень малой. Но появление именно такой маловероятной картины дает основание отклонить гипотезу об отсутствии последовательности признаков.

Вероятность получения x плюсов в верхней правой части упорядоченной матрицы (где самые сильные в прогностическом смысле признаки стоят впереди остальных)

для n разниц $g_{xy} - g_{yx}$ равна $\frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$, где $n = (m-1)/2$,

m — количество признаков и p — вероятность успеха (в данном случае 0,5).⁴

⁴ Ф. Мостеллер, Р. Рурке, Дж. Томас, Вероятность. М., 1969, стр. 397.

Вероятность получения не менее чем r плюсов в верхней правой половине упорядоченной матрицы равна

$$P(x \geq r) = \sum_{x=r}^n \binom{n}{x} p^x (1-p)^{n-x}.$$

С помощью теста Вальда-Вольфовитца можно проверить значимость разного положения отдельных признаков (или их групп).⁵

Прежде чем проверить существенность различия упорядоченных последовательностей у двух групп исследованных, следует уточнить, какое именно различие нам надо знать. Можно проверить или просто различия двух последовательностей или же — в одинаковой ли мере стоящие впереди признаки «сильнее» стоящих позади них. Можно проверить и то и другое одновременно, т. е. одинаковы ли две последовательности и одинаков ли механизм их получения.

В первом случае рассчитываем значимость коэффициента ранговой корреляции τ Кендалла.

Для второго случая рассчитываем интегральную вероятность получения x_2 или более плюсов в n независимых испытаниях, если вероятность получения плюса в каждом случае равна $\frac{x_1}{n}$, где x_1 равен количеству плюсов в первой, взятой за основу сравнения последовательности, учитывая, что

$$b(x_2; n; p) = \frac{n!}{x_2!(n-x_2)!} \left(\frac{x_1}{n}\right)^{x_2} \left(1 - \frac{x_1}{n}\right)^{n-x_2}.$$

В третьем случае пользуемся той же формулой, только переставляем второй из сравниваемых рядов таким образом, чтобы он был сходен с первым, лежащим в основе сравнения рядом с плюсами в верхней правой части матрицы. Так как второй сравниваемый ряд переупорядочен, матрица уже не будет всегда иметь максимально возможное количество плюсов (x_2) в своей правой верхней части, а имеет их в равном или меньшем количестве x_{2a} т. е. $x_{2a} \leq x_2$. Величину x_{2a} следует поставить на место x_2 в приведенной выше формуле.

Преимущества и недостатки данного метода связаны прежде всего с преимуществами и недостатками лежащего в основе его коэффициента. Преимущество его в том, что данным коэффициентом можно пользоваться на всех уровнях измерения независимо от того, имеем ли мы дело со шкалой номинальной, порядковой, равных интервалов или отношений. Кроме того, он отражает криволинейные зависимости, и, наконец, простота вычислений является всегда ценным качеством, даже если в остальном методы равноценны.

Отрицательная сторона данного метода заключается в том, что при расчете коэффициентов не используется вся информация о распределении признаков между объектами. Как только модальные частоты объектов имеют одни и те же значения x и (или) y , коэффициенты g_{yx} и (или) g_{xy} равны нулю.

Особенно часто такое положение складывается, когда при изучении распределения оценок задается 2- или 3-балльная шкала типа «да — нет», «да — не знаю — нет» или «доволен — трудно сказать — недоволен». При 2- или 3-балльных шкалах для измерения признаков x и y больше вероятность (в сравнении со шкалами, имеющими больше градаций) того, что почти все объекты имеют одинаковое значение признака y в разрезах x (или наоборот). Малые значения g , которыми исследователь вынужден часто в таких случаях довольствоваться, еще не значат отсутствия взаимосвязи между признаками вообще, а являются показателями накопления большинства объектов в одном ряду или столбце таблицы. Вероятность такого накопления при более детальных

⁵ J. V. Bradley, Distribution free statistics. Englewood Cliffs, Prentice Hall, 1968, стр. 263.

шкалах измерения будет меньше, и соответственно можно ожидать получение отличающихся от нуля значений g_{xy} и g_{yx} .

При более длинных шкалах, с 7 или 10 баллами (градациями), опасение получения коэффициентов нулевых значений меньше, но при вторичных опросах часто бывает, что рассматриваемый объект оценивается уже не тем баллом, т. е. более длинные шкалы менее устойчивы. Поэтому исследователь должен всегда выбирать между более короткими и дающими относительно меньше информации и более длинными, но менее устойчивыми шкалами. Именно пятибалльная шкала оценок, нам кажется, отвечает обоим требованиям. Причем эти пять баллов не обязательно должны соответствовать по смыслу школьным оценкам успеваемости, они могут, например, быть истолкованы и следующим образом: 1 — намного ниже среднего, 2 — ниже среднего, 3 — среднее, 4 — выше среднего, 5 — намного выше среднего.

В экономических, медицинских и других исследованиях, где данные измерены по шкале равных интервалов или отношений (цена, вес, температура и т. п., а не оценки), можно пользоваться не коэффициентом g , а корреляционным отношением η .

Предлагаемым методом упорядочения признаков по их относительной prognostической ценности можно пользоваться не только в психологии, но и в экономических исследованиях (изучение спроса, интересов покупателей), квалиметрии (экспертные оценки), медицине (факторы, влияющие на состояние здоровья), физкультуре и спорте (взаимосвязь антропометрических данных с результатами) и т. д. Конечно, далеко не всегда удастся найти строгую последовательность признаков, но это уже не по вине метода. Признаки изучаемых объектов могут принадлежать многочисленным системам и цель данного метода — выявление только наличия одной из таких систем.

Рассмотрим на конкретном примере, как следует пользоваться описанным методом.

В «расположении» оценок различным сторонам жизни и трудовой деятельности можно предполагать определенную систему, так как маловероятно, чтобы оценки человеком какого-либо явления совсем не влияли на другие его оценки. Это, конечно, не значит, что возможны значительные корреляции между любыми видами оценок. Взаимосвязи оценок отражают взаимосвязи между оцениваемыми предметами, однако не прямо, а опосредованно свойствами человека как личности, прежде всего системой его отношений, мотивов поведения и преследуемых целей.

Для теории и практики психологических наук большое значение имеет знание того, на основе каких оценок можно точнее прогнозировать остальные. И еще вопрос — если действительно систему оценок можно упорядочить с помощью данного метода, то какая группа исследуемых дает менее противоречивые последовательности — молодые или пожилые, менее или более образованные, мужчины и женщины и т. п.? Можно предположить, что возраст благоприятствует формированию системности в последовательности оценок, но так ли это?

В 1966 г. был проведен выборочный репрезентативный опрос взрослого населения ЭССР (старше 18 лет). Данные его положены в основу расчетов, приведенных в таблицах. Среди других вопросов опрашиваемый должен был по пятибалльной системе оценить свое здоровье, зарплату, интерес к выполняемой работе и т. д. Расположение этих оценок у опрошенных возрастом моложе и старше 30 лет дано в табл. 1 и 2. На основе признаков, предшествующих другим, можно точнее прогнозировать значение остальных признаков, чем наоборот.

Данные таблиц позволяют сделать выводы:

1. Система оценок молодежи больше зависит от оценок их товарищей (оценок товарищам).
2. Оценка зарплате занимает первое место у представителей старшего возраста — результат, логичность которого вполне понятна.
3. Оценка зарплате и собственному культурному уровню позволяет точнее прогнозировать оценку успехам в труде, чем наоборот. При этом эта тенденция сильнее в группе людей старше 30 лет. Самооценка (в данном случае оценка своему культурному уровню) служит отправной точкой для формирования не только оценки успехам в труде, но и интереса к работе. (Конечно, нельзя забывать, что мы далеко не всегда имеем

дело с прямолинейными и функциональными зависимостями между рассматриваемыми признаками типа «доволен зарплатой — доволен и содержанием труда»).

4. Оценка взаимоотношениям в семье выводится из всех остальных отношений. Хотя в социологической литературе часто подчеркивается важная роль влияния гармоничности внутрисемейных отношений на положительное отношение к труду, более правомерным представляется обратное — семейные отношения в какой-то мере даже больше служат «предохранительным» вентилем для «заземления» неполадок в отношении к самому себе, выполняемому труду, товарищам и т. п. Отсюда следует, что хорошая трудовая атмосфера, вероятно, нужнее для положительных взаимоотношений в семье, чем наоборот.

5. Оценка, данная товарищам по работе, которая была относительно независимым признаком у молодежи, оказалась на предпоследнем месте в группе старших. Из этого можно сделать вывод: возраст около 30 лет — это та граница, когда человек начинает ориентироваться исходя из «себя», а не из «других».

6. Вероятность получения 23 или более плюсовых разниц из 28 возможных (табл. 1) равняется почти нулю — менее 0,03. Менее 0,005 вероятность получения 33 или более плюсовых разниц (табл. 2). Эта величина настолько мала, что на основе ее можно отклонить гипотезу о случайной последовательности данных оценок⁶ в той совокупности, откуда взята выборка.

Полученные результаты подтверждаются работами советских психологов и философов (И. Божович, А. Неймарк, А. Ковалев, А. Спиркин и др.), в которых рассматриваются проблемы, связанные с формированием оценок себе и другим. Такие же не противоречивые в смысле содержания результаты получены в других разрезах (пол, образование и т. п.), что позволяет рекомендовать данный метод как для психологических и социологических, так и для близких им по содержанию исследований, где упорядочение признаков по их прогностической ценности поможет глубже вникнуть в изучаемое явление.

Институт истории
Академии наук Эстонской ССР

Поступила в редакцию
4/X 1973

⁶ Ф. Мостеллер, Р. Рурке, Дж. Томас, Вероятность, табл. 3.

T. KITVEL

ÜHESIT VÕIMALUSEST PSÜHHOLOOGILISTE TUNNUSTE JÄRJESTAMISEKS

Resümee

Artiklis vaadeldakse võimalust reastada psühholoogilisi tunnuseid nende vastastikuse prognostilise väärtuse järgi, mis on arvutatud L. A. Goodmani koefitsiendi λ (artiklis g) abil.

Tunnuse X prognostilise väärtuse võib lugeda suuremaks tunnuse Y omast, kui $g_{xy} > g_{yx}$, s. t. tunnuse X väärtuste tundmine lubab suuremal määral hoiduda eksimisest tunnuse Y väärtuste ennustamisel kui tunnuse Y väärtuste tundmine seda võimaldab tunnuse X väärtuste prognoosimisel.

Esimeseks sammuks on seega vahede $g_{xy} - g_{yx}$ arvutamine kõigi tunnuste paaride jaoks. Järgnevalt kasutame tunnuste reastamiseks mõnda sellist meetodit, mis annab vaid üheselt määratud järjestuse (näit. Bergeri süsteemi). Esitatud näites taotlesime tunnuste reastust maksimaalse plusside arvuga $g_{xy} - g_{yx}$ ees seisvate märkide maatriksi paremal ülemisel poolel, kuna selline reastamise meetod tundus kõige paremini vastavat ainevallale.

Saadud järjestuse tõenäosust võib kontrollida eeldusel, et üldkogumis, millest objektid on juhusliku valiku abil võetud, on tunnused oma prognostiliselt väärtuselt võrdsed. Kui ülaltoodud meetodi järgi reastatud tunnuste «võitude-kaotuste» maatriksis on ühel

pool peadiagonaali tunduvalt rohkem ühesuguse märgiga vahesid $g_{xy} - g_{yx}$, siis võib binomiaaljaotuse tabelites leida saadud plusside-miinuste jaotumise tõenäosuse ning loobuda eeldusest tunnuste prognostilise võrdsuse kohta, kui r või rohkema plussi saamine n võimalikust ühel pool peadiagonaali jääb vähemaks kui 0,05.

Artiklis on näidatud võimalusi järjestuste erinevuse olulisuse kontrolliks. Tulemused, mida selle meetodi kasutamine andis küsitlute (1714 üle 18 aasta vanust küsitlute 1966. a. ENSV-s) hinnangute järjestamisel, annavad alust selle meetodi laiemaks kasutamiseks mitte ainult psühholoogidele, vaid ka sotsioloogidele ja majandusteadlastele.

Eesti NSV Teaduste Akadeemia
Ajaloos Instituut

Toimetusse saabunud
4. X 1973

T. KITVEL

ABOUT A METHOD FOR ARRANGING PSYCHOLOGICAL VARIABLES

Summary

In this paper the method of arranging psychological variables by their prognostical value (computed as a measure of association λ , proposed by L. A. Goodman, here referred to as g) is proposed.

The prognostical value of variable X is higher than the prognostical value of variable Y , if $g_{xy} > g_{yx}$, i.e. if knowing the value of variable Y allows to avoid more errors in prognosing the value of Y than knowing the value of Y does about X .

The method for arranging the variables consists of computing the differences of $g_{xy} - g_{yx}$ for all pairs of variables (X and Y). The next step is to arrange all of the variables by the method which gives only one possible ranking, for example, the method proposed by Berger. (In the example we maximized the sum of plus signs in the right-above side of table of differences $g_{xy} - g_{yx}$.)

The last process is to compute the probability of the resulting set of pluses and minuses in our tournament table with variables as competitors and $g_{xy} - g_{yx}$'s as differences in scores.

As a result we have an ordering for the variables and can compute the probability of such ordering under the hypothesis that the underlying ranking in general population is a random one. In the case of near-zero (under 0.05) probability for the resulting set of signs before $g_{xy} - g_{yx}$, the hypothesis about the random ranking of variables can be discarded. Distribution-free tests could be applied for testing the differences in location for a single variable or a set of variables. Also there are proposed significance tests for the differences between two sets of rankings for the same variables.

The results obtained in the arranging of judgements about some important aspects of everyday life are encouraging enough to propose the method given above as a convenient tool for psychological, sociological or economic studies.

Academy of Sciences of the Estonian SSR,
Institute of History

Received
Oct. 4, 1973