

# **КВАЛИМЕТРИЧЕСКИЙ ПОДХОД В ОБРАЗОВАНИИ**

УДК [375.5:005.6]:37.012

**Н. Е. Эрганова,  
Л. В. Колясникова,  
Е. В. Игонина**

## **ТЕОРЕТИЧЕСКИЕ И ПРАКТИЧЕСКИЕ АСПЕКТЫ ВЫЯВЛЕНИЯ И ОЦЕНИВАНИЯ КАЧЕСТВА ОБРАЗОВАНИЯ В ШКОЛЕ**

*Аннотация.* В статье рассматривается одна из основных проблем педагогической теории и образовательной практики последних десятилетий – измерение качества подготовки учащихся и функционирования учебных заведений. Авторы справедливо отмечают, что социальные и экономические тенденции, доминирующие в современном мире, в том числе и в России, актуализируют совершенствование культуры производства и потребления знаний, а значит, и поиск новых подходов к управлению качеством образования, которое становится сегодня объектом пристального внимания сфер науки, государственного влияния, а также реального сектора экономики. Решение оценки качества напрямую связано с распознаванием подлинных (научных) и поддельных (псевдонаучных и ненаучных) вариантов диагностики, которое особенно сложно осуществить, когда дело касается латентных величин. В качестве инструмента такого разграничения и определения интегрального показателя качества предлагается метрическая система Г. Раша, представляющая собой методику конструирования, апробации и анализа диагностических материалов с помощью специальных компьютерных программ. Дан краткий теоретический обзор потенциала и достоинств метрической системы Раша, ценность которой состоит прежде всего в том, что она позволяет ответить на один из ключевых вопросов моделирования качества – относительно пригодности и полезности тех или иных показателей к проводимому исследованию.

*Ключевые слова:* измерение качества, качество образования в школе, интегральный показатель качества образования, метрическая система Г. Раша, латентная величина, индикаторная переменная.

*Abstract.* The paper considers one of the main theoretical and practical pedagogical problems of education quality assessment. The quality measurement depends on successful identification of genuine (scientific) and false diagnostic methods; the process becomes more complicated in case of latent variables. As a solution, the authors recommend the Rasch measurement model for identifying an integral indicator of education quality. The method in question involves designing, approbation and analysis of diagnostic materials, as well as mathematical and statistical data processing based on specialized computer software. The paper describes the advantages and theoretical potential of the Rasch method, and emphasizes its capacity for solving the key problem of quality modeling – i.e. suitability and utility of the indicator variables for the given research.

*Keywords:* quality measurement, quality of school education, integral indicator of education quality, the Rasch model, latent variable, indicator variable.

Под качеством образования в настоящее время понимается степень соответствия достигаемых результатов обучения официально утвержденным нормативным требованиям, социальным и личностным ожиданиям субъектов образовательной деятельности – иначе говоря, уровень удовлетворения государственного, общественного и личного заказов. В связи с этим для регулирования качества образования требуется выявление и учет воздействий на него не только внутренних факторов образовательного пространства, но и многообразных факторов внешней по отношению к нему среды [6, 10, 16].

Социальные и экономические тенденции, доминирующие в современном мире, в том числе и в России, актуализируют совершенствование культуры производства и потребления знаний, а значит, и поиск новых подходов к управлению качеством образования, которое становится сегодня объектом пристального внимания сфер науки, государственного влияния, а также реального сектора экономики.

В свете происходящего обозначилась очевидная потребность в более точных, мобильных и объективных системах оценивания качества общего, среднего и высшего профессионального образования. Причем для обеспечения комплексного анализа всех аспектов процесса обучения и результатов образовательной деятельности необходимо создавать и совершенствовать как общенациональную, так и региональные системы, определяющие успешность функционирования образовательных структур. Чтобы выводы о состоянии образования были обоснованными и способствовали правильному выбору коррекционных мер по повышению эффективности подготовки обучающихся, при отборе способов и средств измерения качества следует ориентироваться на передовые, уже апробированные и показавшие свою перспективность модели управления и компьютерные технологии [14].

На то, что качество измерений в последние десятилетия является одной из главных проблем педагогической теории и образовательной практики, разрешение которой напрямую связано с распознаванием подлинных (научных) и поддельных (псевдонаучных и ненаучных) вариантов диагностики, указывает, в частности, В. С. Аванесов [2]. Подлинная педагогическая диагностика, с его точки зрения, представляет собой «процесс определения меры интересующего латентного свойства ... на интервальной шкале посредством качественного теста, состоящего из системы заданий равномерно возрастающей трудности, позволяющего получать педагогически целесообразные результаты, отвечающие критериям надежности, валидности, объективности и эффективности» [1, с. 170].

Использование В. С. Аванесовым и другими авторами понятия «латентное свойство» (а также синонимичных ему «латентный признак», «латентная переменная», «латентное качество» и т. д.) связано с тем, что в области образования большая часть подвергаемых оцениванию характеристик может быть выявлена только опосредованно, через совокупность особым образом регистрируемых проявлений или показателей. Качество образования в целом относится к *латентным* (от лат. «latens» – скрывающийся), т. е.

скрытым для непосредственного наблюдения и не подлежащим прямым числовым подсчетам категориям, для диагностики которых не существует готовых определенных единиц измерения. Для анализа и построения прогнозов развития подобных категорий в научных исследованиях все чаще предлагается применять теорию соответствия наблюдаемых проявлений измеряемого качества однопараметрической модели Г. Раша [8, 12, 17–18]. Далее мы будем называть ее *метрической системой Г. Раша*. Рассмотрим ее сущность и достоинства относительно выявления качества функционирования и результативности деятельности общеобразовательной школы.

Разработка диагностических средств предполагает «теоретизацию» [4], «концептуализацию» [3], «операционализацию» [5] или формирование «операционального определения» [8] измеряемого или оцениваемого качества. Вне зависимости от различий наименований содержание данных процедур – это многостороннее исследование диагностируемого свойства с конструированием системы понятийных или эмпирических *индикаторов* (заданий, вопросов, показателей, пунктов анкеты и т. д.), которые должны подтверждать наличие или отсутствие у объекта признаков, данное свойство отражающих. Средство диагностики, таким образом, – это целостная совокупность индикаторных переменных, на основе которых осуществляются фиксация и последующее преобразование проявлений диагностируемого качества в количественные показатели, являющиеся итогами произведенного измерения или оценивания [11]. Для обозначения сформированной в ходе исследования совокупности индикаторов иногда используется понятие «*интегральный показатель измеряемого качества*».

Своеобразным инструментом разработки средств диагностики латентных величин и определения интегрального показателя измеряемого качества является метрическая система Г. Раша. Она представляет собой методику конструирования, апробации и анализа диагностических материалов с помощью специальных компьютерных программ (RUMM, Winsteps, MINISTEP и др.). Одно из

достоинств системы состоит в том, что при сопряженном вычислении трудности индикаторных переменных они тем не менее рассматриваются как независимые показатели: количественная характеристика исследуемого свойства у каждого объекта не определяется количественной характеристикой трудности предъявляемых к нему требований. Более того, трудность индикаторных переменных не ставится в зависимость от объемов и состава экспериментальной выборки.

Согласно логике модели Г. Раша, определение результата деятельности  $i$ -й школы по  $j$ -й индикаторной переменной осуществляется исходя из взаимодействия двух факторов: уровня развития латентной переменной (качества образования)  $\theta_i$  и уровня трудности индикатора  $\beta_j$ . Вероятность выполнения школой предъявляемого к ней требования (в виде соблюдения того или иного показателя ее деятельности) возрастает при увеличении уровня качества предоставляемого учреждением образования  $\theta_i$  и/или снижении трудности данного требования  $\beta_j$ . Данное условие может быть представлено формулой:

$$P\{X_{ij} = 1 | \theta_i, \beta_j\} = \varphi(\theta_i - \beta_j) = \frac{e^{\theta_i - \beta_j}}{1 + e^{\theta_i - \beta_j}} \quad (1)$$

где  $P(X_{ij})$  – вероятность выполнения  $i$ -й школой требования, зафиксированного в  $j$ -й индикаторной переменной [9, с. 16].

Из формулы (1) видно, что результат деятельности школы зависит от одного параметра – разности  $\theta_i - \beta_j$  (не случайно модель Г. Раша называется «однопараметрической»); остальные факторы, которые могут влиять на выполнение выдвигаемого требования, считаются вторичными.

Расчет вероятности выполнения  $i$ -й школой требования, которое представлено  $j$ -й индикаторной переменной, производится, как отмечалось выше, посредством специальных компьютерных программ. По итогам данного вычисления выстраивается график трансформаций каждого из индикаторов (рис. 1). В данном графике на оси ординат отражается степень вероятности выполнения

школьным учреждением требования, закрепленного в индикаторной переменной, изменяющейся в интервале от 0,0 до 1,0; на оси абсцисс – разность между уровнем качества образования в школе  $\theta_i$  и уровнем трудности индикатора  $\beta_j$  при 0, фиксирующем равенство данных характеристик. На рис. 1 видно, что в последнем случае вероятность выполнения школой требования оказывается равной 0,5 ( $P(X_{ij}) = 0,5$ ); при положительной разнице ( $\theta_i - \beta_j$ ) она превышает это значение ( $P(X_{ij}) > 0,5$ ), а при отрицательной – не превосходит его ( $P(X_{ij}) < 0,5$ ).

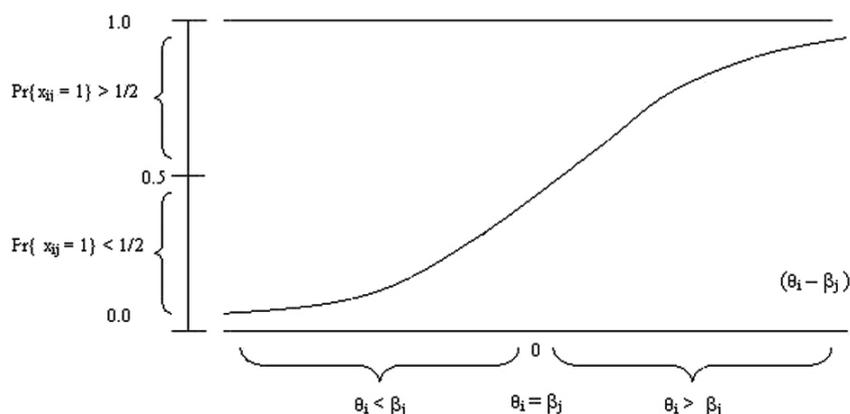


Рис. 1. График вероятности выполнения предъявляемого требования к деятельности школы в зависимости от разности между уровнями качества образования и трудности применяемого индикатора

Представленная формулой (1) модель Г. Раша, подвергнутая ряду математических преобразований, позволяет сравнить вероятность выполнения школами выдвигаемых к качеству деятельности требований независимо от трудности таковых, а трудность требований – не беря во внимание уровень качества образования в каждой из школ.

Кроме того, благодаря используемой в модели специальной единице измерения – логиту – появляется возможность представить оба показателя на одной интервальной линейной шкале. Логит был введен Георгом Рашем как общая мера латентных величин

в шкале натуральных логарифмов, он показывает приращение на единицу логарифма доли вероятности выполнения/ невыполнения требования, зафиксированного в индикаторе [8].

В технологическом смысле ценность метрической системы, разработанной Г. Рашем, состоит в том, что она разрешает преобразовать исходные «тестовые» результаты диагностики, которые согласно положениям современной теории тестирования не могут рассматриваться как соответствующие основным критериям оценивания, в интервальную шкалу натуральных логарифмов и обеспечить обоснование объективного педагогического измерения латентных величин.

Трансформация исходных результатов измерения в логиты уровня качества образования в школе и логиты уровня трудности индикаторных переменных производится как расчет их «логарифмической оценки»  $\theta_i$  и  $\beta_j$ :

$$\theta_i = \ln \frac{p_i}{q_i}, \quad (2)$$

$$\beta_j = \ln \frac{q_j}{p_j}, \quad (3)$$

где  $p_i$  – доля выполненных  $i$ -й школой требований;

$q_i$  – доля невыполненных  $i$ -й школой требований;

$p_j$  – доля школ, выполнивших  $j$ -е требование;

$q_j$  – доля школ, не выполнивших  $j$ -е требование [7, с. 90–91].

Затем производится стандартизация шкал исходных логитов сопоставимыми значениями средних арифметических и стандартных отклонений, в результате чего возникает соизмеримость обеих переменных величин.

Распределение с учетом полученных значений всех индикаторных переменных по мере возрастания их трудности осуществляется как построение на одной плоскости всех их графиков с последующим анализом равномерности их расположения, что повышает точность измерения и дифференцирует достоверность его результатов. Для выявления дифференцирующей достоверности

дополнительно определяются индекс сепарабельности, критическое значение которого принимается равным 0,7. Возрастание меры трудности индикаторных переменных считается равномерным при отсутствии в нем различий больше 0,5 логит.

Проведение объективной диагностики качества образования в школе требует наличия целостного конструкта пригодного для данной цели измерительного инструмента, представленного совокупностью *совместимых* между собой индикаторных переменных. Совместимость индикаторов, сформированных с учетом одномерности диагностируемой латентной величины, и соответствие полученных результатов модели измерения выступают главными условиями эффективности проведенной диагностики. Соответствие индикаторов модели измерения наглядно отражается графически – в их характеристических кривых, конкретные примеры которых мы представим и прокомментируем ниже, пока же остановимся лишь на общих положениях их оформления.

Характеристическая кривая индикаторной переменной должна содержать информацию о коде индикатора (например, *Ю0001*), его названии (*показатель 1*), степени его соответствия модели измерения ( $\chi^2$ ) и объеме выборки испытуемых ( $N$ ). Степень соответствия индикатора модели измерения определяется по значению статистики  $\chi^2$ , рассчитываемому следующим образом: присвоенные школам оценки ранжируются, а сами школы распределяются по уровню качества образования на несколько примерно равных групп  $n$  (их количество зависит от целей исследования).

Для каждой из групп вычисляется среднее значение индикаторной переменной и на основе критерия  $\chi^2$  определяется степень близости экспериментальных точек характеристической кривой:

$$\chi^2 = \sum_{g=1}^n \frac{(T_{gj} - N_g \cdot P_{gj})^2}{N_g \cdot P_{gj}}, \quad (4)$$

где  $n$  – количество групп;

$g$  – номер группы;

$T_{gj}$  – число школ в  $g$ -й группе, выполнивших  $j$ -е требование;

$N_g$  – число школ в  $g$ -й группе;

$P_{gj}$  – вероятность выполнения  $j$ -го требования в  $g$ -й группе [9, с. 95].

Если значение  $\chi^2$  превышает установленный критический минимум 0,05, то индикаторная переменная признается адекватной модели измерения.

Совместимость всей совокупности индикаторов также определяется на основе  $\chi^2$  путем его деления на число степеней свободы с получением эмпирического уровня значимости, который должен быть выше 0,05. В таком случае совместимость индикаторных переменных является удовлетворительной и считается, что в анализируемом наборе отсутствуют «проблемные» варианты, разного рода дефекты (в содержании индикаторов, формах их представления и т. д.) и этот набор соответствует требованиям вероятностной однопараметрической модели Г. Раша.

Общая картина распределения индикаторных переменных по уровню их трудности и школ по уровню качества образования на линейной шкале логитов от наименьших их значений к наибольшим складывается при наложении двух гистограмм. Анализ их данных позволяет выявить границы диапазона варьирования каждой из количественных характеристик (согласно установленной норме он должен составлять не менее 6 логит), особенности распределения школ и взаимного расположения индикаторов, разницу между средними величинами (она не должна превышать 0,5 логит) и другие параметры. Забегая вперед, для наглядности приведем пример таких гистограмм, составленных на основе результатов предпринятого нами исследования (рис. 2).

Краткий теоретический обзор потенциала метрической системы Г. Раша позволяет сделать заключение, что на ее основе возможно описание широкого круга образовательных ситуаций. Какой-либо латентный показатель определяется операционально или с помощью индикаторных переменных, каждая из которых фактически характеризует один из аспектов, а вместе они составляют интегральный показатель исследуемой латентной характеристики.

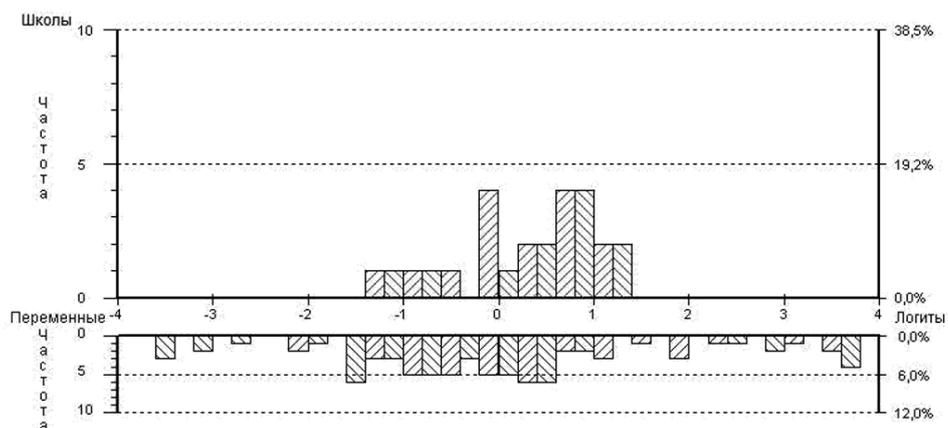


Рис. 2. Гистограммы распределения образовательных учреждений и индикаторных переменных на единой шкале логитов

Достоинства формирования интегральных показателей на основе модели измерения латентных переменных Раша состоят в следующем:

- снижается влияние фактора субъективности, поскольку нет необходимости в вынесении экспертных оценок;
- используется оценка совместимости разработанных индикаторов, т. е. определяется, действительно ли все показатели измеряют одну и ту же латентную переменную (в нашем случае качество образования в школе);
- латентная переменная измеряется по линейной шкале, что позволяет применять широкий спектр статистических процедур для решения задач мониторинга и сравнения объектов исследования;
- устанавливается зависимость, согласно которой увеличение числа привлекаемых индикаторных переменных повышает точность измерения латентной величины [8, 12].

В связи с тем, что качество образования – латентная категория, важно проверить, насколько каждый из показателей совместим с другими и с общей моделью качества образования в школе, т. е. в какой степени они определяют одну и ту же латентную величину. Если индикаторные переменные обладают свойством со-

вместимости, они приемлемы для измерения исследуемого качества. В противном случае, когда индикаторы характеризуют разные латентные величины, их набор нуждается в корректировке. Ценность метрической системы Раша состоит прежде всего в том, что она позволяет дать ответ на один из ключевых вопросов моделирования качества – относительно пригодности и полезности тех или иных показателей к проводимому исследованию [13].

*(продолжение в следующем номере)*

### **Литература**

1. Аванесов В. С. Основные направления развития педагогических измерений // Школьные технологии. 2012. № 1. С. 157–174.
2. Аванесов В. С. Проблема качества педагогических измерений // Педагогические измерения. 2004. № 2. С. 3–27.
3. Аванесов В. С. Проблема педагогического измерения латентных качеств // Педагогические измерения. № 3. 2010. С. 41–63.
4. Давлетов А. И., Серая А. А. Оценка эффективности изотерапии или метода психологической коррекции // Историческая и социально-образовательная мысль. 2010. № 1 (3). С. 86–94.
5. Звонников В. И. Современные средства оценивания результатов обучения: учебное пособие для студентов высших учебных заведений. Москва: Академия, 2007. 224 с.
6. Кальней В. А., Шишов С. Е. Технология мониторинга качества обучения в системе «учитель – ученик». Методическое пособие для учителя. Москва: Педагогическое общество России, 1999. 86 с.
7. Лепикова Н. В. Регулирование качества обучения студентов вуза с использованием информационных и коммуникационных технологий: дис. ... канд. пед. наук. Челябинск, 2007. 198 с.
8. Маслак А. А. Измерение латентных переменных в социально-экономических системах: теория и практика: монография. Славянск-на-Кубани: СГПИ, 2007. 424 с.
9. Осипов С. А. Разработка системы измерения латентных переменных на основе модели Раша для контроля уровня знаний обучаемых: дис. ... канд. тех. наук. Курск, 2003. 128 с.

10. Панасюк В. П. Школа и качество: выбор будущего. С-Петербург: КАРО, 2003. 384 с.
11. Поздняков С. А. Метод и алгоритмы измерения латентных переменных при управлении в образовательных системах: дис. ... канд. техн. наук. Курск, 2009. 168 с.
12. Разработка методики измерения качества образования в школе: методические рекомендации / А. А. Маслак, О. В. Леус, А. А. Данилов. Славянск-на-Кубани: СПИИ, 2009. 67 с.
13. Севрук А. И. Качество в образовании. Проблемы, модели, технологии. Пермь, 2000. 158 с.
14. Севрук А. И., Юнина Е. А. Мониторинг качества преподавания в школе: учебное пособие. Москва: Педагогическое общество России, 2003. 144 с.
15. Третьякова Т. В. Система оценки качества образования и ее построение в регионах с территориальными и национальными особенностями: автореферат диссертации на соискание ученой степени доктора педагогических наук. Якутск, 2010. 45 с.
16. Шишов С. Е., Кальней В. А. Мониторинг качества образования в школе. Москва: Педагогическое общество России, 1999. 354 с.
17. Andrich D. Rasch models for development. London: Sage Publications, 1988. 94 p.
18. Smith E. V., Smith M. S. Introduction to Rasch Measurement. Theory, models and applications. Maple Grove, Minnesota: JAM Press, 2004. 689 p.

### References

1. Avanesov V. S. Main directions of development of pedagogical measurements. *Shkol'nye tehnologii*. [School technologies]. 2012. № 1. P. 157–174. (In Russian)
2. Avanesov V. S. The problem of quality of pedagogical measurements. *Pedagogicheskie izmerenija*. [Pedagogical measurements]. 2004. № 2. P. 3–27. (In Russian)
3. Avanesov V. S. The problem of pedagogical measurement of latent qualities. *Pedagogicheskie izmerenija*. [Pedagogical measurements]. 2010. № 3. P. 41–63. (In Russian)

4. Davletov A. I., Seraya A. A. Assessment of efficiency of fine arts therapy or method of psychological correction. *Istoricheskaja i social'no-obrazovatel'naja mysl'*. [Historical, social and educational thought]. 2010. № 1 (3). P. 86–94. (In Russian)
5. Zvonnikov V. I. Modern means of estimation of results of training: manual for students of higher educational institutions. Moscow: *Akademija* [Academy]. 2007. 224 p. (In Russian)
6. Kalney V. A., Shishov S. E. Technology of monitoring of quality of education in system «teacher – student». Moscow: *Pedagogicheskoe obshhestvo Rossii*. [Pedagogical society of Russia]. 1999. 86 p. (In Russian)
7. Lepikova N. V. Regulation of quality of training of students of higher education institution with use of information and communication technologies. Cand. Diss Chelyabinsk. 2007. 198 p. (In Russian)
8. Maslak A. A. Measurement of latent variables in social and economic systems: theory and practice. Slavyansk-on-Kuban: SGPI [Publishing center of the SGPI], 2007. 424 p. (In Russian)
9. Osipov S. A. Creation of system of measurement of latent variables on the basis of G. Rash's model for control of level of trainees' knowledge. Cand. Diss. Kursk, 2003. 128 p. (In Russian)
10. Panasyuk V. P. School and quality: the choice of the future. SPb: KAPO [CARO], 2003. 384 p. (In Russian)
11. Pozdnyakov S. A. Method and algorithms of measurement of latent variables at management in educational systems. Cand. Diss. Kursk, 2009. 168 p. (In Russian)
12. Creation of measurement technique of quality of education at school: methodical recommendations / A. A. Maslak, O. V. Leus, A. A. Danilov. Slavyansk-on-Kuban: SGPI [Publishing center of the SGPI], 2009. 67 p. (In Russian)
13. Sevruk A. I. Quality in education. Problems, models, technologies. Perm, 2000. 158 p. (In Russian)
14. Sevruk A. I., Yunina E. A. Monitoring the quality of teaching at schools: tutorial. Moscow: *Pedagogicheskoe obshhestvo Rossii*. [Pedagogical society of Russia]. 2003. 144 p. (in Russian)

15. Tretyakova T. V. System of assessment of quality of education and its construction in the regions with the territorial and national features. Doct. diss. Yakutsk, 2010. (In Russian)

16. Shishov S. E., Kalney V. A. Monitoring of quality of education at school. Moscow: Pedagogicheskoe obshchestvo Rossii [Pedagogical society of Russia], 1999. 354 p. (In Russian)

17. Andrich D. Rasch models for development. London: Sage Publications, 1988. 94 p. (translated from English)

18. Smith E. V., Smith M. S. Introduction to Rasch Measurement. Theory, models and applications. Maple Grove, Minnesota: JAM Press, 2004. 689 p. (translated from English)