

УДК 371.315

ВЫБОР ИНСТРУМЕНТАРИЯ ДЛЯ ОБОСНОВАНИЯ НЕЗАВИСИМОСТИ ЭКСПЕРТНЫХ ОЦЕНОК КАЧЕСТВА РАЗРАБОТКИ ЭЛЕКТРОННЫХ ОБРАЗОВАТЕЛЬНЫХ РЕСУРСОВ*

С.Н. ЛАРИН,**кандидат технических наук,
старший научный сотрудник
E-mail: sergey77707@rambler.ru****Центральный экономико-математический институт РАН,
Москва, Российская Федерация****Е.В. ЖИЛЯКОВА,****кандидат экономических наук,
советник федерального государственного бюджетного учреждения
«Российский гуманитарный научный фонд»****E-mail: lenag@rfh.ru****Москва, Российская Федерация**

Предмет/тема. Следствием модернизации сферы образования в России стало активное внедрение в процесс обучения электронных образовательных ресурсов (ЭОР) в качестве инновационных составляющих для формирования и разработки современных образовательных технологий (СОТ). Однако далеко не все ЭОР по своим качественным характеристикам полностью отвечают требованиям ФГОС и других нормативных документов. Предметом статьи является экспертная оценка качества разработки ЭОР.

Цели/задачи. В этой связи основная задача данной статьи заключается не только в выборе инструментария для экспертной оценки качества разработки ЭОР, но и методов для обоснования независимости экспертных оценок. Экспертная оценка качества разработки ЭОР является специальным исследованием, для проведения которого должны использоваться современный инструментарий, позволяющий получать научно обоснованные и проверяемые результаты.

* Статья подготовлена при финансовой поддержке Российского гуманитарного научного фонда, проект № 13-06-00006а «Методология экспертной оценки качества электронных образовательных ресурсов».

Методология. Для решения указанной задачи в работе использовались основные положения теории экспертных оценок, структурированное представление дидактического контента средствами ЭОР, инструментов математической статистики, распределение независимых случайных величин и непараметрические методы для оценки тесноты связи, а также ранговые коэффициенты Спирмена и Кендалла.

Результаты. В ходе проведения исследований установлено, что экспертная функция является одной из важнейших и общепризнанных в педагогической практике. Решена проблема научной обоснованности независимости экспертных оценок и выбора соответствующего инструментария для экспертной оценки качества разработки ЭОР. Подтверждены гипотезы о независимости оценок экспертов и слабой зависимости между изучаемыми мнениями экспертов.

Выводы/значимость. В эпоху информатизации роль экспертизы существенным образом меняется. Из элемента обоснования управленческого решения она превращается в важный и значимый социальный институт оценки качества и начинает играть существенную роль в жизнедеятельности общества,

в том числе и в сфере образования. Описанные в статье методы и математический аппарат могут использоваться для экспертной оценки качества разработки ЭОР.

Ключевые слова: сфера образования, электронные образовательные ресурсы, дидактический контент, экспертная оценка, независимость экспертных оценок, аппарат математической статистики, непараметрические методы, оценка тесноты связи оценок.

Введение

Основной тенденцией модернизации сферы образования и науки в России в последнее время стало интенсивное внедрение информационно-коммуникационных технологий (ИКТ) в практику обучения и подготовки будущих специалистов. Это создает основу для вхождения России в мировое информационное пространство и стимулирует развитие горизонтальных связей в научно-образовательной сфере. Развитие фундаментальной науки и высшего образования соответствует приоритетам и основным требованиям Федеральных государственных образовательных стандартов (ФГОС) последнего поколения. Следствием этого становится активное использование в процессе обучения электронных образовательных ресурсов (ЭОР) и разнообразного программного обеспечения (ПО) в качестве инновационных составляющих для формирования информационных образовательных сред (ИОС) и разработки на этой основе современных образовательных технологий (СОТ)¹ [17, 18, 20].

Сегодня известно достаточно большое число разработанных и применяемых в учебных заведениях (УЗ) ЭОР. Однако далеко не все они полностью отвечают требованиям ФГОС и других нормативных документов, раскрывающих информационные особенности структурирования дидактического контента. В этой связи как никогда актуальной становится задача не только экспертной оценки качества разработки ЭОР, но и выбора инструментария для обоснования независимости экспертных оценок² [5, 9, 11, 14].

¹ Ларин С.Н., Жилиякова Е.В. Методы и приемы проведения независимой экспертизы // Вестник Воронежского государственного университета. Сер. Экономика и управление. Т. 2. Воронеж: Изд-во Воронежского гос. ун-та, 2009. С. 108–117.

² Ларин С.Н., Жилиякова Е.В. Использование инструментария математической статистики для обоснования эффективности независимой экспертизы инициативных исследовательских проектов // Вестник Воронежского государственного университета. Сер. Экономика и управление. Т. 2. Воронеж: Изд-во Воронежского гос. ун-та, 2011. С. 122–131.

Необходимость проведения экспертной оценки качества разработки ЭОР в настоящее время не вызывает сомнений. Более того, экспертная функция не только в науке, но и в педагогической практике является одной из важнейших и общепризнанных. Повышение роли и значения экспертной оценки качества разработки ЭОР в образовательной сфере объективно определяется двумя взаимосвязанными факторами. *Во-первых*, это демократизация управления УЗ, которая проявляется в расширении участия общества в процессах выработки и принятия соответствующих управленческих решений. *Во-вторых*, повышение рисков принятия неадекватных управленческих решений, которые могут привести к негативным социальным последствиям для развития УЗ.

Качество разработки электронных образовательных ресурсов и проведение его экспертных оценок

В традиционном понимании экспертная оценка качества разработки того или иного продукта представляет собой всесторонний анализ группой специалистов (экспертов) его основных свойств и рабочих характеристик на предмет их соответствия установленным требованиям или стандартам качества. Для проведения такой экспертизы привлекаемые специалисты должны обладать исчерпывающими познаниями в конкретной предметной области. Отсюда можно заключить, что экспертная оценка качества является специальным исследованием, для проведения которого должны использоваться современные научные достижения, позволяющие получать достоверные, научно обоснованные и проверяемые результаты по конкретным свойствам и рабочим характеристикам того или иного продукта. В эпоху информатизации роль экспертизы существенным образом меняется. Из элемента обоснования управленческого решения она превращается в важный и значимый социальный институт оценки качества и начинает играть существенную роль в жизнедеятельности общества, в том числе и в сфере образования³ [12, 13, 16, 19].

При рассмотрении роли экспертной оценки качества разработки ЭОР возникают две взаимосвязанные проблемы: это формы институционализации

³ Грачев Г.В., Дрындин А.В., Расторгуев С.П. Автоматизированное рабочее место эксперта информационных материалов: основные направления разработки // Вопросы общей и социальной психологии: сборник научных статей. Саратов: Научная книга, 2006.

этого вида экспертной деятельности и ее адекватность тем задачам, которые она должна решать. Причем адекватность решения задач экспертизы как социального института, с одной стороны, зависит от научной обоснованности независимости экспертных оценок и выбора соответствующего инструментария (как общего, отражающего специфику проведения экспертизы вообще, так и специального — отражающего специфику экспертной оценки качества тех или иных объектов), с другой стороны — от адекватной ее организации в пределах сферы образования (институционализации экспертной деятельности).

Выполнение требования научной обоснованности независимости экспертных оценок и признания результатов экспертизы предполагает использование специалистами в конкретном виде экспертной деятельности научно обоснованного инструментария. При этом его понятийный аппарат, методы, правила и подходы к проведению экспертизы в конкретной предметной области должны обеспечить достижение повторяемости результатов экспертной оценки в одинаковых условиях даже при их использовании разными экспертами.

Для экспертной оценки качества разработки ЭОР должно публично оформиться экспертное сообщество, которое владеет научно обоснованным инструментарием проведения экспертизы. Экспертным сообществом должен быть публично представлен и убедительно показан (продемонстрирован) надежный инструментарий проведения экспертизы и научно обоснованные критерии независимости экспертных оценок⁴ [4, 12]. Здесь необходимо отметить, что потребности общества существенно опережают существующий в настоящее время в научном сообществе инструментарий проведения экспертизы и обоснования независимости экспертных оценок.

Основной задачей экспертной оценки качества разработки ЭОР является анализ его научного уровня, соответствия основным свойствам и рабочим характеристикам, возможностей его применения в сфере образования. Затем следует выработка рекомендаций о целесообразности внедрения ЭОР. Проведение исследования, основанного на использовании специальных знаний, и выдача соот-

ветствующего заключения экспертом в письменной форме осуществляется с соблюдением определенных правил его подготовки и оформления. При этом на эффективность экспертизы влияет не только степень независимости, но и **уровень компетентности специалистов-экспертов** [20].

Покажем возможности применения аппарата математической статистики для анализа результатов экспертной оценки качества разработки ЭОР и обоснования независимости оценок экспертов.

Обозначим символом n число ЭОР, подлежащих экспертной оценке качества разработки. Основное соотношение модели оценки качества разработки ЭОР будет иметь следующий вид:

$$\eta = \alpha \xi + \varepsilon, \quad (1)$$

где η — оценка качества ξ -го ЭОР;

α — коэффициент, который отражает совокупность основных свойств и рабочих характеристик ЭОР, а также компетентность эксперта в конкретной предметной области знаний;

ε — ошибка эксперта в оценке ЭОР.

Происхождение ошибки эксперта в оценке ЭОР имеет двойственный характер. *Во-первых*, ошибка может появиться из-за недостаточного уровня компетентности эксперта в конкретной предметной области знаний; *во-вторых*, из-за невнимательного изучения им совокупности основных свойств и рабочих характеристик ЭОР, которое должно дать представление о месте анализируемого ЭОР среди аналогичных разработок, его научной новизне, практической значимости, объеме представленного дидактического контента и используемых для этого методов и средств.

Для простоты и наглядности изложения материала допустим, что качество любого ЭОР будет оцениваться по 5-балльной шкале: «1» — крайне низкое; «2» — неудовлетворительное; «3» — удовлетворительное; «4» — хорошее; «5» — высокое. Если в модель (1) вместо каждого ξ -го ЭОР подставить оценки качества по 5-балльной шкале, то в идеальном случае получится, что $\alpha = 1$ и $\varepsilon = 0$. Однако практический опыт свидетельствует о том, что такое, в принципе, невозможно, главным образом в силу того, что параметр α остается неизвестным [13, 14].

Исходя из этого, становится ясно, что известные экспертные оценки представляют собой случайные величины, которые определенным образом зависят от числа анализируемых ЭОР разного качества q_i ($i = 1, 2, \dots, n$), доля которых изначально также неизвестна. Для частичного погашения влияния

⁴ Грачев Г.В., Дрындин А.В., Расторгуев С.П. Автоматизированное рабочее место эксперта информационных материалов: основные направления разработки // Вопросы общей и социальной психологии: сборник научных статей. Саратов: Научная книга, 2006.

ошибок необходимо использовать несколько экспертов при соблюдении условия независимости их оценок.

Далее предполагается, что в нашей экспертизе задействованы 2 эксперта, причем оценки друг друга им неизвестны. Это обстоятельство позволяет предположить, что при любых значениях параметра α , неизвестных значениях оценок эксперта совокупности основных свойств и рабочих характеристик ЭОР x и произвольных ошибок эксперта ε у него формируется оценка качества разработки ЭОР, которую он высказывает в форме одной из пяти ранее уже упомянутых оценок по 5-балльной шкале. В результате он получает некую округленную величину, ошибка округления которой остается неизвестной. Если допустить, что значения параметра α у разных экспертов одинаковы, то можно сформировать модели оценки качества разработки для любого ЭОР экспертами x_i и y_i :

$$x_i = \alpha q_i + \varepsilon_{i1} \quad (i = 1, 2, \dots, n), \quad (2)$$

$$y_i = \alpha q_i + \varepsilon_{i2} \quad (i = 1, 2, \dots, n), \quad (3)$$

где для 1-го и 2-го экспертов получаем модели получения оценок качества разработки любого ЭОР (2) и (3) соответственно⁵ [2]. При этом оценки экспертов независимы от порядка предъявления им любого числа ЭОР.

Для проведения дальнейших исследований сформулируем еще ряд важных предположений. Первое из них заключается в том, что ошибки разных экспертов при оценке разных ЭОР ε_{i1} и ε_{i2} распределены одинаково и независимы друг от друга. Однако это предположение приводит к появлению двух новых задач.

Суть *первой* задачи — необходимость проверки независимости оценок экспертов при любом, причем заранее неизвестном распределении значений показателей, отражающих основные свойства и рабочие характеристики качества анализируемых ЭОР. Кроме того, это предположение требует проверки того, насколько близки распределения оценок, сделанные разными экспертами, поскольку фактически они представляют собой сумму двух независимых случайных величин, распределение которых во многом совпадает. Первая из них $\alpha \xi$ зависит от распределения оценок экспертами качества анали-

зируемых ЭОР, а вторая представляет собой ошибки экспертов с одинаковым распределением, которое только искажает распределение самих оценок.

Сущность *второй* задачи — необходимость выявления совпадения неизвестных распределений оценок экспертов, для которых известно, что они одинаковы, и ничего больше.

Указанные задачи помогают решить, казалось бы, парадоксальную ситуацию. Ведь, с одной стороны, при независимых оценках двух экспертов 1-й из них не знает, какие оценки качества разработки по тем или иным ЭОР выставил 2-й эксперт. Но, с другой стороны, оценки этих экспертов x_i и y_i зависят от качества разработки q_i одного и того же ЭОР. Такая ситуация является достаточно типичной для сферы экономики, а теперь, как видим, и для сферы образования.

Этот парадокс приводит к появлению *третьей* задачи, суть которой заключается в необходимости проверки статистической зависимости оценок экспертами ЭОР при их неизвестном распределении.

Как видно, распределение статистической зависимости оценок экспертами ЭОР в каждой из трех задач остается неизвестным, поэтому и в методах, используемых для решения этих задач, не могут применяться распределения. Исходя из этого заключения, рассмотрим возможность применения критерия согласия χ^2 и ранговых коэффициентов корреляции для решения указанных задач [6–8, 10].

Предположим, что в результате экспертизы были получены экспертные оценки качества разработки ЭОР, которые приведены в табл. 1. Для наглядности и простоты доказательства предположим, что в экспертизе принимали участие два эксперта: 1-й и 2-й. При оценке качества разработки ЭОР эксперты могут выставить только одну оценку, исходя из приведенной 5-балльной шкалы: от 1 до 5. Во всех последующих таблицах указаны только оценки экспертов, сделанные ими в рамках этой шкалы. Далее выясним, зависит ли оценка качества

Таблица 1

Оценки качества разработки ЭОР экспертов (по шкале)

Номера предложений 2-го эксперта	Номера предложений 1-го эксперта				
	5	4	3	2	1
5	17	8	3	3	–
4	11	7	4	2	–
3	5	7	7	1	–
2	1	4	3	2	–
1	–	–	–	1	–

⁵ Ларин С.Н., Жиликова Е.В. Использование инструментария математической статистики для обоснования эффективности независимой экспертизы инициативных исследовательских проектов // Вестник Воронежского государственного университета. Сер. Экономика и управление. Т. 2. Воронеж: Изд-во Воронежского гос. ун-та, 2011. С. 122–131.

Таблица 2

Оценки двух экспертов качества разработки ЭОР (по шкале)

Номера предложений 2-го эксперта	Номера предложений 1-го эксперта				Итого у 2-го эксперта...
	5	4	3	2	
5	17	8	3	3	31
4	11	7	4	2	24
3	5	7	7	1	20
2	1	4	3	3	11
<i>Итого у 1-го эксперта...</i>	34	26	17	9	86

разработки ЭОР, поставленная 1-м экспертом, от оценки, полученной по результатам экспертизы 2-м экспертом (не получается ли ситуация, когда оценка по следующей экспертизе зависит от влияния оценки, полученной по результатам предыдущей?).

Чтобы проверить, независимы ли оценки и одинаковы ли эксперты, воспользуемся статистическим критерием согласия χ^2 , который справедлив при любом распределении предложений экспертов [10]. Для проверки зависимости оценок от качества разработки ЭОР применим коэффициенты корреляции, которые независимы от распределения.

Поскольку 1-й эксперт ни разу не поставил оценку 1 — «крайне низкое» ни одному ЭОР, то аналогично можно поступить и с оценками 2-го эксперта, объединив оценку 1 — «крайне низкое» в один ряд со следующей по порядку оценкой. Поэтому табл. 1 модифицируется в табл. 2.

Несмотря на то, что итоговая строка оценок 1-го эксперта не совпадает со столбцом итоговых оценок 2-го эксперта, распределения ЭОР с разными качественными характеристиками, по мнению каждого из экспертов, могут отличаться.

Рассмотрим **две гипотезы**. *Первая* гипотеза H_1 исходит из предположения о фактической независимости оценок экспертов. Поэтому ожидаемые частоты распределения при справедливости H_1 получаются в соответствии со статистическим критерием согласия χ^2 . Они приведены в табл. 3 и могут сравниваться с наблюдаемыми частотами

Таблица 3

Ожидаемые частоты распределения парных оценок экспертов

Оценки 2-го эксперта	Оценки 1-го эксперта			
	5	4	3	2
5	12,225	9,372	6,128	3,244
4	9,488	7,255	4,744	2,511
3	7,906	6,046	3,953	2,093
2	4,348	3,325	2,174	1,151

распределения, приведенными в табл. 2. *Конкурирующая* гипотеза H_2 состоит в том, что оценки экспертов считаются зависимыми между собой. В этом случае статистический критерий согласия обычно не формируется, ведь очень сложно сказать, какой будет эта зависимость.

Проверим первую гипотезу при помощи данных табл. 2 и 3. Для этого зададимся уровнем значимости $\alpha = 0,05$, определим число степеней свободы $m = (r - 1)(s - 1)$, а также расчетное значение статистического критерия согласия χ^2 . Поскольку в обеих таблицах число основных строк $r = 4$, а число основных столбцов $s = 4$, то расчетные значения числа степеней свободы m и статистического критерия согласия χ^2 составят соответственно:

$$m = (4 - 1)(4 - 1) = 9;$$

$$\chi^2 = 14,2573.$$

При заданном уровне значимости $\alpha = 0,05$ и расчетном значении числа степеней свободы $m = 9$ по специальным таблицам легко находим критическое значение статистического критерия согласия $\chi_{кр}^2(9; 0,05) = 16,92$. Как видим, расчетное значение этого критерия меньше, чем его критическое значение, то есть $\chi_p^2 = 14,25 \dots < \chi_{кр}^2(9; 0,05) = 16,92$. Это значит, что первая гипотеза H_1 подтверждается и можно утверждать, что оценки экспертов независимы даже в тех случаях, когда параметр α неизвестен, а распределение ошибок экспертов не совпадает.

Для того чтобы подтвердить предположение о том, что уровень знаний 1-го и 2-го экспертов в конкретной предметной области примерно одинаков, можно также использовать статистический критерий согласия χ^2 [10]. Это можно сделать, переформулировав гипотезу, причем учитывая, как было отмечено, что из одинаковости уровня знаний экспертов вытекает и одинаковое распределение оценок. Данное обстоятельство может только подтвердить (но не доказать) статистический критерий согласия χ^2 по следующим причинам: как правило, статистические методы чаще используются для опровержения чего-либо, а не для доказательства; одинаковое распределение оценок экспертов не подтверждает эквивалентность уровней их знаний в конкретной предметной области. Равенство распределений наглядно представлено в итоговых частотах табл. 2.

Проверка гипотезы о равенстве функций распределения основывается на наблюдениях, составленных из оценок n ЭОР каждым из двух экспертов. Данные для проверки равенства распределения оценок экспертов для данного случая приведены в табл. 4.

Таблица 4

Данные для проверки равенства распределений оценок экспертов

Номер эксперта	Число выставленных экспертами оценок				Количество наблюдений у каждого эксперта
	5	4	3	2	
1-й	34	26	17	9	86
2-й	31	24	20	11	86
Итого...	65	50	37	20	172

Таблица 5

Ожидаемые частоты появления оценок экспертов

Номер эксперта	Оценка каждого эксперта			
	5	4	3	2
1-й	32,5	25	18,5	10
2-й	32,5	25	18,5	10

Все приведенные в табл. 4 данные представляют собой выраженные через числа результаты наблюдений. Ожидаемые числовые результаты наблюдений выставленных экспертами оценок получаются после умножения на число наблюдений у любого эксперта (в данном случае их число одинаково и равно 86) частных от деления чисел последней строки на итоговое число (172). Результаты проведенных расчетов приведены в табл. 5.

Из табл. 4 и 5 следует, что расчетное значение статистического критерия согласия $\chi_p^2 = 0,8002$. Определим его критическое значение, допускающее ошибку распределения в 5% $\chi_{кр}^2(m; 0,05)$ с числом степеней свободы:

$$m = (s - 1)(r - 1) = (2 - 1)(4 - 1) = 3.$$

Оно будет равно $\chi_{кр}^2(3; 0,05) = 10,127$. Поскольку $\chi_p^2 = 0,8002 < \chi_{кр}^2(3; 0,05) = 10,127$, то подтверждается согласованность статистических данных со сделанными ранее предположениями. Следовательно, гипотеза об одинаковом уровне знаний экспертов также подтверждается.

Таким образом, расчетные значение статистического критерия согласия χ_p^2 согласуются с предположениями о независимости оценок экспертов и примерно одинаковом уровне их знаний в конкретной предметной области.

При экспертизе качества разработки ЭОР часто используются различные оценки, например, знаки отклонения рангов. Кроме того, для измерения взаимосвязи между отдельными признаками также используются оценки, не зависящие от распределения, или непараметрические коэффициенты связи. Эти коэффициенты определяются при условии, что анализируемые признаки подчиняются неизвест-

ным (возможно, даже различным) распределениям. Поэтому для условного обозначения признаков и оценки взаимосвязей между ними можно применять знаки отклонения, ранги и ранговые коэффициенты. При одинаковой количественной оценке всех значений исследуемого признака их ранг считается равным средней арифметической значений соответствующих номеров мест, которые определены экспертами. В этом случае приходится иметь дело с так называемыми связными рангами.

Принцип нумерации значений исследуемых признаков также лежит в основе использования непараметрических методов для изучения взаимосвязи между основными свойствами и рабочими характеристиками ЭОР. Матрицы, приведенные в табл. 1 и 2, позволяют восстановить все оценки 1-го и 2-го экспертов (x_p, y_i) по всем анализируемым ЭОР $i = 1, 2, \dots, n$ и сгруппировать их по парам. При этом сам порядок получения пар оценок экспертами качества разработки ЭОР уже не может быть восстановлен, но он в дальнейшем не играет существенной роли.

Для простоты получения значений рангов воспользуемся данными табл. 4. Если расположить оценки в порядке убывания их значений, то получим следующую картину для 1-го эксперта:

- оценки со значением 5 (их 34) займут места с 1-го по 34-е, а их средний ранг составит $(1 + 2 + \dots + 34) : 34 = 17,5$;
- оценки со значением 4 (их 26) займут места с 35-го по 60-е, а их средний ранг составит $(35 + \dots + 60) : 26 = 47,5$;
- оценки со значением 3 (их 17) займут места с 61-го по 77-е, а их средний ранг составит $(61 + \dots + 77) : 17 = 69$;
- оценки со значением 2 (их 9) займут места с 78-го по 86-е, а их средний ранг составит $(78 + \dots + 86) : 9 = 82$.

Аналогично для 2-го эксперта получим следующую картину:

- оценки со значением 5 (их 31) займут места с 1-го по 31-е, а их средний ранг составит 16;
- оценки со значением 4 (их 24) займут места с 32-го по 55-е, а их средний ранг составит 43,5;
- оценки со значением 3 (их 20) займут места с 56-го по 75-е, а их средний ранг составит 65,5;
- оценки со значением 2 (их 11) займут места с 76-го по 86-е, а их средний ранг составит 81.

Теперь можно восстановить пары оценок экспертов качества разработки для каждого ЭОР.

Для этого воспользуемся имеющимися данными следующим образом. Все получившиеся данные будут воспроизведены в табл. 6. Для простоты дальнейшего использования будем располагать уже полученные средние ранги в порядке убывания представленных 1-м экспертом оценок по всем ЭОР. Тогда будут получены (в соответствии с табл. 2) 17 наивысших оценок с номерами (5,5) у 1-го и 2-го экспертов, пары их средних рангов (17,5 и 16) представлены в табл. 6. Далее следуют 11 пар рангов (17,5 и 43,5), соответствующих оценкам (5,4) табл. 2. За ними следуют 5 пар рангов (17,5 и 65,5), соответствующих оценкам (5,3). Наконец, после одной (1) пары (17,5 и 81), соответствующей оценкам (5,2), появляются 8 пар рангов (47,5 и 16), соответствующих оценкам (4,5), далее идут 7 пар рангов (47,5 и 43,5), соответствующих оценкам (4,4) и т.д.

применяться для определения тесноты связи как между количественными, так и между качественными признаками при условии упорядоченного расположения их значений по степени убывания или возрастания признака. Это уже было проделано в табл. 6.

По данным табл. 6 (последняя строка) определим коэффициент Спирмена, т.е. зависимость между оценками 1-го и 2-го экспертов. Если не обращать внимания на связи между рангами, то сумма квадратов разностей (числитель) $\Sigma d^2 = 65\ 687$, знаменатель

$$n(n^2 - 1) / 6 = 105\ 995,$$

а сам коэффициент Спирмена будет равен $K_c = 0,3803$. Если учитывать связи рангов, то он будет равен $K_c = 0,3204$ из-за того, что поправки для первой строки табл. 4 по 1-му эксперту составят

$$T_x = 5\ 203 = 0,5 [34(34^2 - 1) + 26(26^2 - 1) + \dots + 9(9^2 - 1)],$$

а поправки для первой строки табл. 4 по 2-му эксперту будут равняться

$$T_y = 4\ 135 = 0,5 [31(31^2 - 1) + \dots + 11(11^2 - 1)].$$

Теперь числитель будет равен 30 970, а знаменатель — 96 651,1.

Поскольку связь между двумя величинами довольно слабая, то нужно убедиться в существенном ее отличии от 0. Следовательно, необходимо проверить значимость отличия от 0 коэффициента Спирмена при справедливости гипотезы об отсутствии связи между двумя исследуемыми величинами. Вычислим критическую статистику для этого случая:

$$K_{cp} = \rho_{x/y} \sqrt{\frac{n-2}{1-\rho_{x/y}^2}} = 0,3204 \sqrt{\frac{86-2}{1-0,1026}} = 2,66.$$

Поскольку случайная величина K_{cp} распределена примерно по закону Стьюдента с 84 степенями свободы, то из имеющихся данных получим $K_{скр} = 1,66$ для уровня значимости 5% (для 1% — 2,37). Отсюда следует, что расчетное значение $K_{cp} < K_{скр}$, что с очень большой уверенностью позволяет отвергнуть гипотезу об отсутствии связи. Следовательно, коэффициент Спирмена значим, а зависимость мнений экспертов, хотя и не очень тесная, но имеет место. Расчет коэффициента без учета связей даст такой же результат.

Для определения значений коэффициента Кендалла потребуется сделать несколько боль-

Таблица 6

Данные для проверки тесноты связи оценок экспертов

Число пар рангов	Значения пар рангов		Разность
17	17,5	16	1,5
11	17,5	43,5	-26
5	17,5	65,5	-48
1	17,5	81	-63,5
8	47,5	16	31,5
7	47,5	43,5	4
7	47,5	65,5	-18
4	47,5	81	-33,5
3	69	16	53
4	69	43,5	25,5
7	69	65,5	3,5
3	69	81	-12
3	82	16	66
2	82	43,5	38,5
1	82	65,5	16,5
3	82	81	1

Данные, представленные в табл. 4 и 6, позволяют получить всю необходимую информацию для проверки предположений о тесноте связи оценок экспертов, возникающей при просмотре ими одних и тех же ЭОР, в условиях, когда ничего не известно о распределении оценок самих экспертов. В таком случае проверка производится с помощью критериев, свободных от распределения, или непараметрических критериев.

Из числа непараметрических критериев для оценки тесноты связи наиболее часто используют ранговые коэффициент Спирмена (K_c) и коэффициент Кендалла (K_k) [6, 7]. Эти показатели могут

ших вычислений [7]. Исходные данные для этих вычислений представлены в табл. 6. Наибольшая проблема состоит в подсчете числа перестановок, которые требуются для приведения второго ряда из пар рангов в порядок, соответствующий первому. Дальнейшее гораздо проще, действие требует только расчетов, а не логики.

Второй ряд содержит четыре группы рангов, каждая из которых начинается с 16. Начнем перемещать ранг 16 в первую группу. Для перемещения каждого ранга 16 из второй группы его нужно поменять 1 раз (см. строку 1 табл. 6) с рангом 81 из первой группы. Затем 5 раз (см. строку 1 табл. 6) с рангом 65,5 из первой группы. Это потребует проведения как минимум шести ($1 + 5 = 6$) перестановок. Далее 11 раз с рангом 43,5. Итого получим $1 + 5 + 11 = 17$ перестановок для каждого ранга 16.

Всего ранг 16 во второй группе встречается 8 раз, т.е. для их перевода в первую группу нужно всего $136 (8 \times 17 = 136)$ перестановок. Каждый ранг 16 из третьей группы (их всего 3) нужно переставить 18 раз ($4 + 7 + 7 = 18$) с рангами 43,5, 65,5 и 81 из второй группы и 17 раз из первой группы. Всего $3 \times (4 + 7 + 7 + 17) = 3 \times 35 = 105$ перестановок. Наконец, каждый ранг 16 (их 3) из четвертой группы меняем 14 раз ($4 + 7 + 3 = 14$) с рангами 43,5, 65,5; затем 81, соответственно, из третьей; 35 раз — из второй и из первой групп. Ранг 16 из четвертой группы собран в одну за $3 \times (4 + 7 + 3 + 35) = 147$ перестановок. Всего потребуется $8 \times 17 + 3 \times 35 + 3 \times 49 = 388$ перестановок для того, чтобы ранги 16 были поставлены на первые места. Далее их считать уже не надо.

Аналогично поступаем и при перемещении ранга 43,5. Так как все ранги 16 уже переставлены, то каждый ранг 43,5 из второй группы (их 7) нужно $1 + 5 = 6$ раз переставить с рангами 81 и 65,5 из первой группы. Всего требуется $7 \times 6 = 42$ перестановки. Каждый ранг 43,5 из третьей группы (их 4) меняем с рангами 65,5 и 81 ($4 + 7 = 11$ раз, и 6 раз — из первой, т.е. 17 перестановками присоединяем к первой группе. Каждый из оставшихся двух рангов 43,5 из четвертой группы ($3 + 7 + 4 + 17 = 31$) перестановкой присоединяется к первой группе, а все ранги 43,5 перемешаются на свои места за $(7 \times 6 + 4 \times 17 + 2 \times 31) = 172$ перестановки. Далее они тоже не учитываются.

Точно так же ранги 65,5 можно упорядочить путем $7 \times 1 + 7 \times (4 + 1) + 1 \times (3 + 5) = 50$ перестановок. Ясно, что после этого все связные ранги 81 окажутся на своих местах.

Таким образом, получено число перестановок $Q = 388 + 172 + 50 = 610$. Из выражения $P + Q = 0,5n(n - 1)$ легко найти $P = 3\,655 - 610 = 3\,045$; $S = 3\,045 - 610 = 2\,445$ и вычислить коэффициент Кендалла $K_{ккы} = 0,666$. Если учесть связность рангов, то величина коэффициента Кендалла будет только больше. Но учитывать связи в данном случае не нужно, ведь существует некоторая объективная последовательность оценки качества q_i отдельных ЭОР, которую эксперты из-за округления не показали, но могли бы это сделать.

Хотя расчетное значение коэффициента Кендалла значительно выше расчетного значения коэффициента Спирмена, проверить его существование все-таки необходимо. Для $n = 86$ среднее квадратическое отклонение величины S равно $\sigma = 268,11$, расчетное значение $K_{кр} = 9,12 > K_{кр}$ для любого разумного уровня.

Заключение

Таким образом, полученные результаты свидетельствуют о малой зависимости между изучаемыми мнениями экспертов. Казалось бы, это дает все основания сделать вывод о том, что эксперты проводили свои оценки одинаково объективно и независимо друг от друга. Но правомерность такого вывода ставят под сомнение следующие обстоятельства.

Во-первых, при экспертизе качества разработки ЭОР обычно задействуют значительно больше двух экспертов. Поэтому в данном случае оценки 1-го и 2-го экспертов, по сути, представляют собой некую совокупность оценок разных экспертов, поставивших оценку одному из исследуемых ЭОР, однако среди всех экспертов, безусловно, найдутся такие пары, которые оценивали одни и те же ЭОР. Естественно, что их будет значительно меньше 86, как было рассмотрено, но и к ним применимы описанные ранее методы.

Во-вторых, определить уровень знаний каждого эксперта в конкретной предметной области по отношению к любому другому эксперту можно, применив теорию планирования эксперимента. Высокий уровень знаний эксперта имеет как положительные стороны, так и отрицательные. К положительной стороне можно отнести возможность отсеивать слабых и очень осторожных экспертов — распределение их оценок будет отличаться от распределений остальных. Отрицательная сторона заключается в том, что мнение одного сильного «идеального» эксперта, всегда правильно опреде-

ляющего качество разработки ЭОР, может потонуть среди мнений остальных, не столь сильных экспертов. Для обнаружения этого негативного эффекта применяется коллективная экспертиза.

В-третьих, не исключен вариант, когда к одному эксперту могут попасть преимущественно качественно разработанные ЭОР, а ко второму — откровенно слабые. Вполне естественно, что 2-й эксперт будет вынужден оценивать менее качественные ЭОР, чем 1-й эксперт. Для облегчения работы коллективного эксперта, который должен разграничить оценки экспертов в таком случае, можно сделать ранжирование слабых из самых качественно разработанных ЭОР, попавших на экспертизу. Но сделать такое упорядочение значительно труднее, чем упорядочить наиболее качественно разработанные ЭОР. Представляется, что для исследования согласованности оценок экспертов в этом случае лучше использовать модель с приоритетами. Однако описание технологии ее практического применения выходит за рамки данного исследования.

Список литературы

1. *Валеев Г.Х.* Экспертиза квалификационных научных исследований. М.: Логос, 2005. 112 с.
2. *Гусева А.И., Гаврилов С.И., Тихомирова А.Н.* Модель управления качеством информационно-образовательных ресурсов // Программные продукты и системы. 2010. № 1. С. 146–149.
3. *Евланов Л.Г., Кутузов В.А.* Экспертные оценки в управлении. М.: Экономика, 2001. 133 с.
4. *Жилякова Е.В.* Методы экспертной обработки по принципу согласованности с оценками большинства // Обзорные прикладной и промышленной математики. 2006. Т. 13. Вып. 3. С. 532–534.
5. *Иванов Д.А.* Экспертиза в образовании. М.: Изд. центр «Академия», 2008. 336 с.
6. *Кендал М.* Ранговые корреляции. М.: Статистика, 1975. 216 с.
7. *Кендал М., Стюарт А.* Статистические выводы и связи. М.: Наука, 1973. 899 с.
8. *Кобзарь А.И.* Прикладная математическая статистика. М.: Физматлит, 2006. 816 с.
9. *Косолапов Н.* Политика, экспертиза, общество: узлы взаимозависимости // Pro et Contra, 2003. Т. 8. № 26. С. 18–31.
10. *Крамер Г.* Математические методы статистики. М.: Мир, 1976. 648 с.
11. *Ларин С.Н., Жилякова Е.В.* Пути совершенствования механизмов независимой экспертизы и финансовой поддержки инициативных научных исследований // Национальные интересы: приоритеты и безопасность. 2011. № 35. С. 11–20.
12. *Ларин С.Н., Малков У.Х., Баранова Н.М.* Экспертная система оценки инновационных и технологических качеств электронных образовательных ресурсов // Национальные интересы: приоритеты и безопасность. 2014. № 25. С. 54–59.
13. *Михеев В.И.* Моделирование и методы теории измерений в педагогике. М.: УРСС, 2010. 224 с.
14. *Нестеров А.В.* Основы экспертно-исследовательской деятельности. М.: ВШЭ, 2009. 163 с.
15. *Орлов А.И.* Прикладная статистика. М.: Экзамен, 2006. 672 с.
16. Организация профессиональной и общественной экспертизы программ развития и образовательных программ учреждений общего образования: учебно-методический комплект / сост. и ред. Т.А. Мерцалова, С.Г. Косарецкий. М.: АСОУ, 2006. 192 с.
17. *Полат Е.С., Бухаркина М.Ю.* Современные педагогические и информационные технологии в системе образования: учебное пособие. М.: Академия, 2007. 368 с.
18. *Синаторов С.В.* Электронные образовательные ресурсы: классификация, требования к ним и принципы их разработки. URL: <http://www.sinato-rovsv.ru>.
19. Толковый словарь терминов понятийного аппарата информатизации образования. М.: ИИО РАО, 2006. 88 с.
20. *Филиппов А.Ф.* Участь эксперта // Отечественные записки. 2003. № 1. С. 7–15.
21. *Файзильберг Л.С.* Обучаемая система поддержки коллективного решения группы независимых экспертов // Управляющие системы и машины. 2003. № 4. С. 62–67.

SELECTION OF TOOLS TO SUBSTANTIATE THE INDEPENDENCE
 OF EXPERT ASSESSMENT OF QUALITY OF E-LEARNING RESOURCES

Sergei N. LARIN,
 Elena V. ZHILYAKOVA

Abstract

Importance Whereas the education industry is undergoing the improvement process in Russia, e-learning resources are being actively integrated into the learning and teaching process as innovative components needed to form and develop modern educational technologies. However, e-learning resources fail to entirely comply with the effective Federal State Standards on Education and other regulations in terms of quality. The article focuses on expert assessments of quality of e-learning resources.

Objectives In this respect, the main objective of the article is not only to select tools for expert assessment of e-learning resource quality, but also to substantiate the unbiased nature of expert assessments. Expert assessment of e-learning resource quality is a special-purpose study that requires modern tools to reach reasonable and verifiable scientific results.

Methods To solve the above objective, the researcher relies upon the principle provision of the expert assessment theory, structured presentation of the didactic contents using e-learning resources, tools of mathematical statistics, allocation of independent and random variables and nonparametric techniques to evaluate the extent of close relationship and Kendall and Spearman's rank correlation coefficients.

Results The research identifies that expert's function is one of the most important and generally accepted one in pedagogy. The research solves the matter of scientific substantiation and independence of expert assessment, and selection of relevant tools to carry out expert assessment of e-learning resource quality. The researchers also corroborate the hypotheses on independence of expert assessment and weak dependence of experts' opinions on each other.

Conclusions and Relevance In the age of information technologies, the role of expert assessment changes. Being an aspect for substantiating managerial decisions, it turns into an important and considerable social institution of quality assessment and starts playing a crucial role in social life, including education. The

methods and mathematical technique described in the article may be applied to perform expert assessment of e-learning resource quality.

Keywords: education, industry, e-learning, resource, didactic content, expert assessment, independence, technique, mathematical statistics, nonparametric methods, correlation ratio, assessment

References

1. Valeev G.Kh. *Ekspertiza kvalifikatsionnykh nauchnykh issledovaniy* [Examination of qualifying researches]. Moscow, Logos Publ., 2005, 112 p.
2. Guseva A.I., Gavrilov S.I., Tikhomirova A.N. Model' upravleniya kachestvom informatsionno-obrazovatel'nykh resursov [Information and educational resource quality management model]. *Programmnye produkty i sistemy = Software Products and Systems*, 2010, no. 1, pp. 146–149.
3. Evlanov L.G., Kutuzov V.A. *Ekspertnye otsenki v upravlenii* [Expert assessment in management]. Moscow, Ekonomika Publ., 2001, 133 p.
4. Zhilyakova E.V. Metody ekspertnoi obrabotki po printsipu soglasovannosti s otsenkami bol'shinstva [Methods for expert reviews by the principle of consistency with the majority estimates]. *Obozrenie prikladnoi i promyshlennoi matematiki = Review of Applied and Industrial Mathematics*, 2006, vol. 13, iss. 3, pp. 532–534.
5. Ivanov D.A. *Ekspertiza v obrazovanii* [Expert reviews in education]. Moscow, Akademiya Publ., 2008, 336 p.
6. Kendall M. *Rangovye korrelyatsii* [Rank Correlation Methods]. Moscow, Statistika Publ., 1975, 216 p.
7. Kendall M., Stuart A. *Statisticheskie vyvody i svyazi* [Statistical Conclusions and Relations]. Moscow, Nauka Publ., 1973, 899 p.
8. Kobzar' A.I. *Prikladnaya matematicheskaya statistika* [Applied mathematical statistics]. Moscow, Fizmatlit Publ., 2006, 816 p.

9. Kosolapov N. Politika, ekspertiza, obshchestvo: uzly vzaimozavisimosti [Policy, expertise, society: mutual dependence nodes]. *Pro et Contra*, 2003, vol. 8, no. 26, pp. 18–31.
10. Cramer H. *Matematicheskie metody statistiki* [Mathematical Methods of Statistics]. Moscow, Mir Publ., 1976, 648 p.
11. Larin S.N., Zhilyakova E.V. Puti sovershenstvovaniya mekhanizmov nezavisimoi ekspertizy i finansovoi podderzhki initsiativnykh nauchnykh issledovaniy [Ways of improving the mechanisms for independent examination of and financial support for special-purpose scientific researches]. *Natsional'nye interesy: priority i bezopasnost' = National Interests: Priorities and Security*, 2011, no. 35, pp. 11–20.
12. Larin S.N., Malkov U.Kh., Baranova N.M. Ekspertnaya sistema otsenki innovatsionnykh i tekhnologicheskikh kachestv elektronnykh obrazovatel'nykh resursov [Expert system for assessing innovation and technological qualities of electronic educational resources]. *Natsional'nye interesy: priority i bezopasnost' = National Interests: Priorities and Security*, 2014, no. 25, pp. 54–59.
13. Mikheev V.I. *Modelirovanie i metody teorii izmerenii v pedagogike* [Modeling and measurement theory in pedagogy]. Moscow, URSS Publ., 2010, 224 p.
14. Nesterov A.V. *Osnovy ekspertno-issledovatel'skoi deyatel'nosti* [Fundamentals of expert research]. Moscow, HSE Publ., 2009, 163 p.
15. Orlov A.I. *Prikladnaya statistika* [Applied statistics]. Moscow, Ekzamen Publ., 2006, 672 p.
16. *Organizatsiya professional'noi i obshchestvennoi ekspertizy programm razvitiya i obrazovatel'nykh programm uchrezhdenii obshchego obrazovaniya* [Organization of the professional and public evaluation of development programs and educational programs of basic education institutions]. Moscow, Academy of Public Administration Publ., 2006, 192 p.
17. Polat E.S., Bukharkina M.Yu. *Sovremennyye pedagogicheskie i informatsionnye tekhnologii v sisteme obrazovaniya* [Modern teaching and information technologies in the education industry]. Moscow, Akademiya Publ., 2007, 368 p.
18. Sinatorov S.V. *Elektronnye obrazovatel'nye resursy: klassifikatsiya, trebovaniya k nim i printsipy ikh razrabotki* [Electronic educational resources: classification, requirements and principles of their development]. Available at: <http://www.sinatorovsv.ru>. (In Russ.)
19. *Tolkovyy slovar' terminov ponyatiinogo apparata informatizatsii obrazovaniya* [Dictionary of terms relating to the conceptual framework of IT support in education]. Moscow, Institute of Information Technologies in Education of Russian Academy of Education Publ., 2006, 88 p.
20. Filippov A.F. Uchast' eksperta [The role of experts]. *Otechestvennye zapiski = Notes of the Fatherland*, 2003, no. 1, pp. 7–15.
21. Faizil'berg L.S. Obuchaemaya sistema podderzhki kollektivnogo resheniya gruppy nezavisimykh ekspertov [Trained system for supporting a collective solution of a team of independent experts]. *Upravlyayushchie sistemy i mashiny = Controlling Systems and Machines*, 2003, no. 4, pp. 62–67.

Sergei N. LARIN

Central Economics and Mathematics Institute of RAS, Moscow, Russian Federation
sergey77707@rambler.ru

Elena V. ZHILYAKOVA

Russian Foundation for Humanities, Moscow, Russian Federation
lenag@rffh.ru

Acknowledgments

The article was supported by the Russian Foundation for Humanities, project No. 13-06-00006a Methodology for Expert Assessment of Quality of Electronic Educational Resources.