

# Муниципальное образование: ИННОВАЦИИ И ЭКСПЕРИМЕНТ

Журнал для администрации управлений образования и руководителей образовательных учреждений

№ 5 • 2016

ПРЕДСЕДАТЕЛЬ РЕДАКЦИОННОЙ КОЛЛЕГИИ —

**Ломов Станислав Петрович**, доктор педагогических наук, профессор,  
академик-секретарь Российской академии образования, г. Москва

Журнал индексируется в ERIH PLUS®, 2016

«The journal is indexed by ERIH PLUS», 2016

## Редакционная коллегия

**Анисимов Олег Сергеевич** — доктор психологических наук, профессор, профессор кафедры акмеологии и психологии профессиональной деятельности Российской академии народного хозяйства и госслужбы, г. Москва, РФ.

**Артамонова Екатерина Иосифовна** — доктор педагогических наук, зав. кафедрой педагогики пединститута, президент Международной академии наук педагогического образования, г. Москва, РФ.

**Видмарович Наталия Петровна** — доктор филологических наук (PhD), профессор, доцент кафедры педагогики и андрагогики факультет философии Университет г. Белграда, Сербия.

**Днепров Сергей Антонович** — доктор педагогических наук, профессор, заведующий кафедрой РГГПУ, г. Екатеринбург, РФ.

**Зеленин Александр Васильевич** — доктор филологических наук (PhD), доктор философии в области лингвистики, Университет города Тампере, Финляндия.

**Козлова Светлана Акимовна** — доктор педагогических наук, профессор, профессор кафедры дошкольного образования МГПУ, главный редактор журнала «Дошкольник», г. Москва, РФ.

**Леванова Елена Александровна** — доктор педагогических наук, профессор, заведующая кафедрой социальной педагогики и психологии Московского педагогического государственного университета, г. Москва, РФ.

**Микулац Ирина** — кандидат педагогических наук (PhD), преподаватель, Университет города Пула, Хорватия (University of Pula, Croatia).

**Никандров Николай Дмитриевич** — доктор педагогических наук, профессор, Академик РАО, член Президиума РАО, Российская академия образования, г. Москва, РФ.

**Новикова Галина Павловна** — доктор педагогических наук, профессор, ректор Института развития образовательных систем, ученый секретарь отделения педагогики и методики дошкольного и начального образования МАНПО, г. Москва, РФ.

**Новоселова Светлана Юрьевна** — доктор педагогических наук, первый проректор по образовательной и научной деятельности ФГАОУ ДПО «Академия повышения квалификации и профессиональной переподготовки работников образования», г. Москва, РФ.

**Петерсон Людмила Георгиевна** — доктор педагогических наук, профессор, профессор ФГАОУ ДПО «Академия повышения квалификации и профессиональной переподготовки работников образования», г. Москва, РФ.

**Просвиркин Владимир Николаевич** — доктор педагогических наук, директор государственного бюджетного общеобразовательного учреждения г. Москвы «Центр образования «Школа здоровья» № 1679», г. Москва, РФ.

**Сумнительный Константин Евгеньевич** — доктор педагогических наук, профессор кафедры общего образования Московского института открытого образования, г. Москва, РФ.

**Сухова Елена Ивановна** — доктор педагогических наук, профессор, заведующий кафедрой дошкольного образования МГПУ, главный редактор журнала «Воспитание дошкольников», г. Москва, РФ.

**Темина Светлана Юрьевна** — доктор педагогических наук, профессор кафедры педагогики ОАНО ВО «Московский психолого-социальный университет», г. Москва, РФ.

**Хебиб Эмина** — доктор наук, (PhD), Associate Professor, Department for Pedagogy and Andragogy Faculty of Philosophy University of Belgrade, Serbia, профессор, философский факультет Загребского университета.

**Шудегов Виктор Евграфович** — доктор физико-математических наук, депутат, заместитель Председателя Комитета Государственной Думы по образованию, Федеральное собрание Российской Федерации. Государственная дума (6-й созыв: 2011–2016 гг.)

## Главный редактор

Алла Сиденко

## Зам. главного редактора

Елена Сиденко

## Ответственный секретарь

Андрей Бахтин

\* \* \* \* \*

Компьютерная вёрстка

Е. Конобеева

ISSN 2306-8329

www.in-exp.ru

Зарегистрирован Комитетом Российской Федерации по печати.  
Свидетельство о регистрации средств массовой информации  
ПИ № ФС77-28487 от 01.07.2007 г.

Двухлетний импакт-фактор  
РИНЦ — 0,938 (2015 год)

Место в рейтинге SCIENCE INDEX за 2013 г.  
по тематике «Народное образование. Педагогика» — 47.

Издание включено ВАК в перечень рецензируемых научных журналов и изданий, выпускаемых в Российской Федерации, в которых должны быть опубликованы основные результаты диссертаций на соискание ученой степени кандидата и доктора наук

Точка зрения редакции может не совпадать с точкой зрения авторов.

Ответственность за достоверность информации, содержащейся в публикуемых материалах, несут авторы.

За авторство публикуемых материалов редакция ответственности не несет.



«Журнал издается при информационной поддержке ФГАОУ ДПО «Академия повышения квалификации и профессиональной переподготовки работников образования»

Учредители журнала:

ООО «Инновации и эксперимент в образовании»

Адрес редакции:

140135, Московская обл., п. Никоновское,  
до востребования Сиденко А.С.

Электронный адрес редакции: sidenko@in-exp.ru

Подписано в печать 14.10.2016. Формат бумаги 60x84/8.  
Печать офсетная. Бумага офсетная. Тираж 500 экз.

Отпечатано в АО Первая Образцовая типография  
Филиал Чеховский Печатный Двор  
142300, Московская область, г. Чехов, ул. Полиграфистов, д. 1

© Сиденко А.С., Сиденко Е.А.

## Авторы номера

- Абрамов Руслан Агарунович** — доктор экономических наук, профессор, заведующий кафедрой государственного и муниципального управления ФБГОУ ВО «Российский экономический университет имени Г. В. Плеханова», г. Москва. E-mail: oef08@mail.ru
- Акимова Анжелика Александровна** — директор МБОУ «Средняя общеобразовательная школа № 11» городского округа Первоуральск Свердловской области, г. Первоуральск. E-mail: anz170468@mail.ru
- Веселова Валерия Валентиновна** — кандидат педагогических наук, старший научный сотрудник Московского психолого-социального университета, г. Москва. E-mail: veselovavv@yandex.ru
- Вострикова Мария Владимировна** — соискатель по общеинститутской кафедре теории и истории педагогики Института педагогики и психологии образования Московского городского педагогического университета, г. Москва. E-mail: mv.vostrikova@gmail.com
- Дьякова Ирина Викторовна** — старший преподаватель кафедры иностранных языков Военной академии воздушно-космической обороны имени Маршала Советского Союза Г. К. Жукова г. Твери. E-mail: I1ira11@mail.ru
- Загвязинский Владимир Ильич** — академик Российской академии образования, доктор педагогических наук, профессор, заведующий академической кафедрой методологии и теории социально-педагогических исследований Тюменского государственного университета, Тюмень. E-mail: education@utmn.ru
- Зерчанинова Татьяна Евгеньевна** — кандидат социологических наук, доцент, заведующий кафедрой регионального и муниципального управления ФГОУ ВПО «Российская академия народного хозяйства и государственной службы при Президенте Российской Федерации», Уральский институт управления — филиал, г. Екатеринбург. E-mail: Tatiana\_Z@ui.ranepa.ru
- Колмакова Ксения Сергеевна** — учитель музыки, МБОУ СОШ № 3, город Реж. E-mail: eger\_88.88@mail.ru
- Кольбе Алиса Сергеевна** — магистрант, Национальный исследовательский университет «Высшая школа экономики», Москва. E-mail: kolbe.alice@gmail.com
- Королев Денис Александрович** — кандидат технических наук, доцент, Национальный исследовательский университет «Высшая школа экономики», Москва. E-mail: dkorolev@hse.ru
- Лурье Леонид Израилевич** — доктор педагогических наук, профессор, директор общеобразовательного лицея № 1 г. Перми; профессор кафедры математического моделирования систем и процессов Пермского национального исследовательского политехнического университета; заведующий кафедрой теории и методики профессионального образования Пермского военного института внутренних войск МВД РФ, Пермь. E-mail: lourieleonid@gmail.com
- Паволоцкий Александр Владимирович** — кандидат педагогических наук, Национальный исследовательский университет «Высшая школа экономики», Москва. E-mail: apavolotsky@hse.ru
- Пыканов Игорь Владимирович** — соискатель по общеинститутской кафедре теории и истории педагогики, Институт педагогики и психологии образования Московского городского педагогического университета, г. Москва. E-mail: pykanov@gmail.com
- Степанов Александр Владимирович** — кандидат педагогических наук, доцент кафедры дизайна интерьера Российского государственного профессионально-педагогического университета, Екатеринбург. E-mail: S49@list.ru
- Степанова Татьяна Михайловна** — кандидат педагогических наук, доцент кафедры культурологии и дизайна Уральского федерального университета, Екатеринбург. E-mail: S49@list.ru
- Трофимова Елена Давидовна** — кандидат педагогических наук, доцент, Нижнетагильский государственный социально-педагогический институт, Нижний Тагил. E-mail: Akselen7025@ya.ru
- Федоров Владимир Анатольевич** — доктор педагогических наук, директор научно-образовательного Центра профессионально-педагогического образования Российского государственного профессионально-педагогического университета, Екатеринбург. E-mail: Fedorov1950@gmail.com
- Филатова Марина Николаевна** — кандидат педагогических наук, доцент кафедры педагогики внеурочной деятельности, ГАОУ ВО «Московский институт открытого образования», г. Москва. E-mail: filatova\_mn@mail.ru

## Authors

- Abramov Ruslan A.** — Doctor of Economics, professor, head of state and municipal government FBGOU IN «Plekhanov Russian University of Economics», Moscow. E-mail: oef08@mail.ru
- Akimova Anzhelika A.** — School Director, Secondary school № 11, Pervouralsk. E-mail: anz170468@mail.ru
- Veselova Valery V.** — Ph.D., senior researcher of the Moscow Psychological and Social University, Moscow. E-mail: veselovavv@yandex.ru
- Vostrikova Maria V.** — competitor on the JINR Department of Theory and History of Pedagogy Institute of Pedagogy and Educational Psychology of the Moscow City Pedagogical University, Moscow. E-mail: mv.vostrikova@gmail.com
- Dyakova Irina V.** — senior teacher of the department of foreign languages Military Academy of Air-Space Defense after Marshal of the Soviet Union G. K. Zhoukov Tver Town. Email: *11ira11@mail.ru*
- Zagvyazinsky Vladimir I.** — Academician of the Russian Academy of Education, Doctor of Pedagogical Sciences, Professor, Head of the Academic Department for Methodology and Theory of Social and Pedagogical Research, Tyumen State University, Tyumen. E-mail: education@utmn.ru
- Zerchaninova Tatiana E.** — PhD in Sociology, Head of the Department of regional and municipal government, Russian Presidential Academy of National Economy and Public Administration, Ural institute of management — branch of RANEPa, Ekaterinburg. E-mail: Tatiana\_Z@ui.ranepa.ru
- Kolmakova Kseniya S.** — The teacher of music MBOU SOSh № 3, Rezh. E-mail: eger\_88.88@mail.ru
- Kolbe Alisa S.** — MSc student, National Research University «Higher School of Economics», Moscow. kolbe.alice@gmail.com
- Korolev Denis A.** — Ph.D., associate professor, National Research University «Higher School of Economics», Moscow. dkorolev@hse.ru
- Lourie Leonid I.** — Doctor of Pedagogical Sciences, Professor, Department of Mathematic Modeling of Systems and Processes, State National Research Polytechnic University of Perm; Head of the Department «Theory and Methodology of Professional Education», Perm Military Institute of Internal Troops of the Russian Federation; Principal of the Municipal Budgetary General Education Institution «Lyceum № 1» of Perm city, Perm. E-mail: lourieleonid@gmail.com
- Pavolotsky Aleksandr V.** — Ph.D., associate professor, National Research University «Higher School of Economics», Moscow. apavolotsky@hse.ru
- Pykanov Igor V.** — competitor on the JINR Department of Theory and History of Pedagogy, Institute of Pedagogy and Educational Psychology of the Moscow City Pedagogical University, Moscow. E-mail: pykanov@gmail.com
- Stepanov Aleksandr V.** — Candidate of Pedagogical Sciences, Associate Professor, Department of Interior Design, Institute of Arts, Russian State Vocational Pedagogical University, Yekaterinburg. E-mail: S49@list.ru
- Stepanova Tatyana M.** — Candidate of Pedagogical Sciences, Associate Professor, Department of Cultural Studies and Design, Ural Federal University, Yekaterinburg. E-mail: S49@list.ru
- Trofimova Elena D.** — the candidate of pedagogical Sciences, Nizhny Tagil State Socio- Pedagogical Institute, Nizhny Tagil. E-mail: Akselen7025@ya.ru
- Fedorov Vladimir A.** — Doctor of Pedagogical Sciences, Director of the Research and Education Centre of Vocational Pedagogical Education, Russian State Vocational Pedagogical University, Yekaterinburg. E-mail: Fedorov1950@gmail.com
- Filatova Marina N.** — PhD in pedagogy, Associate professor at the Chair of Pedagogy of Supplementary Education, State autonomous educational institution of higher education «Moscow Institute of Open Education», Moscow. E-mail: filatova\_mn@mail.ru

## В номере

### От теории к практике реализации ФГОС

- Е. Д. Трофимова*  
Роль музыки в формировании нравственных представлений младших школьников 6
- М. В. Вострикова*  
Особенности конструирования и реализации технологии педагогического взаимодействия школы и семьи на основе системно-деятельностного подхода 11
- И. В. Дьякова*  
Содержание устойчивого гармонично-развивающего пространства внеурочной деятельности младших школьников при обучении английскому языку 15

### Теория инновационной деятельности

- В. И. Загвязинский*  
О социальной значимости и востребованности педагогических исследований 20
- В. А. Федоров, А. В. Степанов, Т. М. Степанова*  
Новые организационные формы кафедральных структур современного вуза как условие совершенствования учебно-методической работы 27

### Инновации в образовательных учреждениях

- Л. И. Лурье*  
Почему выпускники педагогических вузов становятся менеджерами? 33
- Р. А. Абрамов*  
Аутсорсинговые модели участия населения в осуществлении местного самоуправления 42

- Д. А. Королев, А. С. Кольбе, А. В. Паволоцкий*  
Применение пиринговой оценки письменных работ в потоковых очных курсах 49

### Инновационные и исследовательские проекты, программы, подходы

- И. В. Пыканов*  
Функциональный подход к исследованию мотивационной управленческой деятельности руководителя школы 59
- А. А. Акимова, Т. Е. Зерчанинова*  
Исследование практики управления муниципальной системой образования города Первоуральск 64
- В. В. Веселова*  
Геймификация социума как одно из следствий развития информационно-коммуникационных технологий 71
- К. С. Колмакова*  
Создание авторской структурной модели урока музыки на основе современных педагогических технологий 75

### Современные подходы к качеству дополнительных общеобразовательных программ

- М. Н. Филатова*  
Современные подходы к разработке и оценке качества дополнительной общеобразовательной общеразвивающей программы 80

In the magazine

**From theory to practice, the implementation of the Federal State Educational Standard**

- E. D. Trofimova*  
Role of music in formation of the moral ideas younger schoolchildren 6
- M. V. Vostrikova*  
Features of the design and implementation of pedagogical technology of interaction of school and family-based system-activity approach 11
- I. V. Dyakova*  
The contents of the stable harmonically developing space of primary schoolchildren's extracurricular activities while teaching English 15

**The theory of innovation**

- V. I. Zagvyazinsky*  
Concerning the social importance and a demand of pedagogical researches 20
- V. A. Fedorov, A. V. Stepanov, T. M. Stepanova*  
Organizational forms of structures of the modern university as a condition for improvement of educational and methodical work 27

**Innovations in educational institutions**

- L. I. Lourie*  
Why graduates of pedagogical institutes become managers? 33
- R. A. Abramov*  
Outsourcing model of public participation in local government 42

- D. A. Korolev, A. S. Kolbe, A. V. Pavolotsky*  
Implementation of peer assessment in full-time courses 49

**Innovation and research projects, programs, approaches**

- I. V. Pykanov*  
Functional approach to the study of the motivational management activities of the head of school 59
- A. A. Akimova, T. E. Zerchaninova*  
Research of practice of management of a municipal education system of the city of Pervouralsk 64
- V. V. Veselova*  
Gamification society as a consequence of the development of information and communication technologies 71
- K. S. Kolmakova*  
Students' development of creative abilities at music lessons at school 75

**Modern approaches to quality more educational programs**

- Marina N. Filatova*  
Modern approaches to the development and quality assessment of enrichment programs in the general supplementary education 80

УДК 378.146, 378.147

**Королев Денис Александрович**

кандидат технических наук, доцент, Национальный исследовательский университет «Высшая школа экономики», Москва. E- mail: dkorolev@hse.ru

**Кольбе Алиса Сергеевна**

магистрант, Национальный исследовательский университет «Высшая школа экономики», Москва. E- mail: kolbe.alice@gmail.com

**Паволоцкий Александр Владимирович**

кандидат педагогических наук, Национальный исследовательский университет «Высшая школа экономики», Москва. E- mail: apavolotsky@hse.ru

## Применение пиринговой оценки письменных работ в потоковых очных курсах

---

**Аннотация.** Исследование применимости пиринговой оценки письменных работ в очных курсах проводилось с целью установления достаточных условий для сокращения преподавательской нагрузки и интенсификации активного изучения предмета студентами. На основе существующих методик проведения пиринговой оценки в массовых онлайн курсах были созданы модели оценки и правила проведения работ, учитывающие специфику аудитории и формата обучения. Эксперименты проводились в течение двух лет и позволили получить массив данных для анализа. В результате были проанализированы применённые во время курсов методы оценивания, а также смоделированы результаты при применении других формул. Комбинирование формул расчета оценок авторов и рецензентов с особыми правилами проведения работ позволяет добиться надежных результатов, этот подход является новым в образовательной практике. Практическая значимость исследования заключается в экспериментально подтвержденном опыте и смоделированных на экспериментальных данных формулах автоматического расчета оценок.

---

**Ключевые слова:**

кросс-рецензирование; пиринг; автоматизация проверки; очное обучение в вузе; оценивание работ студентов; электронная поддержка очных курсов

**Korolev Denis A.**

Ph.D., associate professor, National Research University «Higher School of Economics», Moscow. dkorolev@hse.ru

**Kolbe Alisa S.**

MSc student, National Research University «Higher School of Economics», Moscow. kolbe.alice@gmail.com

**Pavolotsky Aleksandr V.**

Ph.D., associate professor, National Research University «Higher School of Economics», Moscow. apavolotsky@hse.ru

## Implementation of peer assessment in full-time courses

---

**Abstract.** Research of applicability of peer assessment of written works in full-time courses are conducted in order to establish sufficient conditions for the reduction of teaching load and intensification of active learning by students. Basing on existing methodologies for peer evaluation of mass online courses were created assessment models and rules, specific to audience and training format. For two years of experiments real-life raw evaluation data for analysis and calculated results were gathered. The applied assessment methods were analyzed as well as the simulated results for other grading calculation formulas. Combining formulas for assessment of the authors and peers with special rules of work allow to achieve reliable results, this is a new approach to educational practice. The practical significance of the study is the experimentally proven experience and formulas modeled on the experimental data, that allow to calculate the grades without teacher.

---

**Keywords:**

peer assessment; automated assessment; classroom studies; assessment of assignments; electronic support for classroom courses

В 2011 году в Российских вузах прекратился набор на специалитет по большинству специальностей, произошел переход на «уровневую систему». Вместе с этим переходом произошли реформы, напрямую не связанные с Болонским процессом: произошло существенное укрупнение потоков. Эти реформы наложились на организационные преобразования — укрупнение вузов и оптимизацию их структуры. В этой статье рассматривается пример перехода к новой реальности на примере курса «Компьютерная графика», ранее читавшегося в МИЭМ на одной из кафедр. В результате преобразований в 2013–2014 годах курс сократился с 17 до восьми недель, а поток увеличился с 50 до 120–180 студентов, позже этот курс и вовсе был исключен из расписания. За два переходных года были опробованы способы компенсации потери контакта со студентами и сохранения их вовлеченности в учебный процесс.

Основным опасением при существенном сокращении объема и длительности курса и исключении из учебного плана курсовой работы была потеря персонального общения со студентами. На потоке в 50 человек за полгода преподаватель имеет возможность отслеживать работу каждого и быть в курсе затруднений и успехов отдельных студентов, вникать в суть их курсовых работ. Но в новых условиях пришлось искать другие способы взаимодействия со студентами, переключая основную рутину на автоматизацию и на них самих [2]. Особое внимание при этом уделялось тому, чтобы содержание курса не страдало и контроль знаний студентов не стал поверхностным [3]. Более того, в курсе заявлялось, что преподаватель имеет минимальное влияние на выставленные оценки, поскольку они преимущественно выставляются автоматически или путем пиринговой проверки работ самими студентами [1].

Пиринговая проверка, а также оценка, рецензирование (от англ. peer assessment<sup>1</sup>, evaluation, review) — это способ проверки и оценивания письменных работ, когда работа одного автора проверяется несколькими независимыми рецензентами, обычно из одного с автором круга. В данном случае проверку осуществляли студенты того же потока. Как правило, для повышения объективности оценки используется «слепой» или «двойной слепой» метод проверки, когда автор не знает, кто проверяет его работы и, во втором случае, когда и проверяющий не знает, кто автор данной работы [6]. Такой способ проверки или рецензирования активно используется в научном мире при рецензировании публикаций перед печатью [8], а также в синхронных массовых онлайн курсах, когда невозможно проверить задание автоматически, а ручная проверка при аудитории в несколько тысяч человек просто невозможна [9].

## Обзор

Использование пиринга позволяет не только сократить время, потраченное преподавателем на проверку работ, но и изменить сам подход к обучению. По наблюдениям [18], в среднем на прочтение и комментирование студенческой работы преподаватель тратит порядка 20–40 минут, а в случае с массовым образованием количество работ и, соответственно, время труда пропорционально возрастают. Автор отмечает, что задания с peer review развивают у студентов навык чтения научной литературы и писательский навык — студентам приходится писать научные работы на доступном языке, а также учит составлять конструктивную критику, в том числе отрицательную, что обычно вызывает сложности у студентов.

Комментарии к работе очень важны для дальнейшего совершенствования. Kwangsu Cho and Charles MacArthur в своем исследовании [5] сравнивают перепроверку работ авторами (студентами) после проверки рецензентами в лице студентов и экспертной проверки. Эксперимент показал, что после оценки экспертом (преподавателем) авторы (студенты) вносят минимальные поправки в свои работы, а в случае с получением комментариев от рецензентов-товарищей (студентов) в работы вносились достаточно объемные исправления. С другой стороны, эти изменения были ориентированы на объем, а не новое содержание, а после экспертной проверки авторы добавляли новую информацию.

Raoul A Mulder, Jon M Pearce, Chi Baik изучали студенческое восприятие до и после участия в peer review [13]. В данном исследовании многие студенты имели высокие ожидания до участия и для большинства ожидания оправдались. Студенты считают, что подобный подход полезным, особенно выделяют процесс рецензирования. Больше половины студентов отметили, что рецензии помогают улучшить работы.

Для peer review как формы учебной активности важной проблемой является студенческая мотивация [7]. Yanqing Wang, Yaowen Liang, Luning Liu and Ying Liu разработали свою систему «EduPCR4» для проведения peer review программного кода [19]. Система предлагает три типа баллов для мотивации студентов: баллы, начисляемые за предоставление работы в срок, качественные баллы, зависящие от задания, и бонусные баллы, которые могут быть как положительными, так и отрицательными — зависит от единства мнения рецензентов. Данная модель мотивирует студентов к участию как автором, так и рецензентом. Система «EduPCR4» учитывает также три случая, препятствующие истинным результатам проверки:

1. Продвинутая программа: автор программы имеет компетенцию выше большинства проверяющих, поэтому вторые могут неправильно понять оцениваемую работу.
2. Некомпетентный проверяющий: студент плохо успевает по предмету, поэтому его оценкам не всегда можно доверять.
3. Недобросовестные проверки: некоторые студенты выставляют оценки безотносительно работы (например, максимальный или минимальный балл).

<sup>1</sup> Cornell University, Center for Teaching Excellence. Peer assessment <https://www.cte.cornell.edu/teaching-ideas/assessing-student-learning/peer-assessment.html>

Тем не менее, «EduPCR4» остается уязвимой: если все студенты будут выставлять оценки недобросовестно, получить оценки и выявить радикальных проверяющих не получится.

Обработка рецензий и расчет итоговой оценки является ключевой задачей, связанной с peer review. Hoi K. Suen рассмотрел распространенные методы обработки результатов peer review [17]. Calibrated Peer Review (CPR) — подход, зависящий от реальной успеваемости рецензента, что позволяет задать весовой коэффициент его рецензии. Bayesian post hoc stabilization построен на Байесовских моделях. Однако данный метод не учитывает систематические ошибки. Последний подход широко применяется на популярном онлайн-ресурсе Coursera.org. В своем исследовании [15] они представляют три версии модели расчета оценки, однако методы отличаются несущественно и во всех за основу взято значение медианы. Последний из рассмотренных подходов — Credibility index, который является доработанной версией CPR. Если CPR направлен на выявление некомпетентности проверяющего, то Credibility index учитывает точность рецензента, его последовательность и переносимость между масштабом и заданием. Данный подход находится в режиме тестирования и разработки.

### Подход к построению учебного курса

Для краткого курса длительностью всего восемь недель на практике возможно провести всего две работы с пиринговой проверкой, учитывая, что на написание и на проверку работы даётся по одной неделе, после каждой работы нужно оставить время на разбор итогов (коротко на лекции и подробно — на сайте поддержки), чтобы в студенты учли свои ошибки в следующей работе или к зачёту. К тому же, это весьма ресурсоёмкое мероприятие и для студентов и для преподавателей. В такой ситуации хотелось избежать обычного для онлайн-курсов ground-truth тестирования, когда студентам даётся набор эталонных заранее проверенных экспертами работ и по результатам проверки этих работ студентами устанавливается, насколько оценка каждого студента близка к экспертной, далее их голоса учитываются с соответствующими поправочными коэффициентами. В данном случае была предпринята попытка использовать все два шанса для проведения эссе с пиринговой проверкой, чтобы студенты успели получить отклик на свои работы и написать ещё по одной с учётом полученного опыта.

Общий подход к построению курса и оценке работы студентов кратко описан на сайте поддержки курса<sup>1</sup>.

### Подход к оценке работ студентов в курсе

В реалиях очных курсов приходится признать, что мотивация студентов в общем случае имеет другой характер, чем у студентов онлайн-курсов, пришедших

изучать интересующую их дисциплину [12]. Как правило, значительная часть потока присутствует в учебном процессе номинально, в лучшем случае лишь выполняя требования, но говорить о тотальной вовлечённости студентов и заинтересованности темой курса, какая бы она ни была, невозможно. Это относится и к выборным предметам, и к общим потоковым курсам. Такой срез аудитории заставляет внимательнее смотреть на немногие мотивирующие инструменты, доступные преподавателю, поэтому остановимся подробнее на системе оценок в курсе, она важна, чтобы понять механизмы, использованные в организации пиринговой проверки работ.

Студенты имеют разные амбиции и разные интересы в своей работе, как и разные способности. Обычный учебный план предполагает одинаковую траекторию для всего потока, хотя очевидно, что из тех, кто в принципе намерен и способен погружаться в деятельность, одним ближе исследовательская работа, другим — практическая. Система оценок в курсе предполагала накопление баллов и различные учитываемые в накопленной оценке активности в сумме давали существенно больше баллов, чем максимально возможная оценка в ведомости. Студент мог сам выбирать, каким образом ему зарабатывать себе баллы, и мог посчитать, какие виды активности вероятнее принесут ему желаемые баллы, а чем ему заниматься неинтересно. При этом, компенсируя отсутствие «отрицательной оценки» в традиционном её понимании, здесь в шкале оценки отдельных видов работ могли присутствовать действительно отрицательные величины, то есть, начинающиеся с отрицательных чисел, и это наглядно демонстрировало влияние того или иного достижения или провала в каждом из заданий. Это не новый подход, он встречается в зарубежной практике, например, в ранее упомянутой работе [19], но в российских вузах он, если и практикуется, то на уровне инициатив преподавателей. Пиринговая проверка работ также не является распространенным инструментом в учебной практике в России, но отдельные преподаватели используют этот инструмент в своих курсах, например, в НИУ ВШЭ [16].

Оценки за работы и за их рецензирование начинались с отрицательной шкалы. Так, при общей шкале возможных баллов около 200 (100 баллов в курсе соответствовали максимальной оценке) письменная работа оценивалась в диапазоне 30 баллов, но от -9 до +20. Поначалу такой подход вызвал недоумение студентов, но «приложив» эту шкалу к обычной «школьной», они увидели, что «отрицательная» оценка находится там же, где и раньше — «двойка» из пятибалльной шкалы попадает в область отрицательных чисел, «тройка» начинает положительную часть шкалы. Аналогично, рецензирование работ оценивалось по пятибалльной шкале, но от «-2» до «+2». Все задания были необязательными, но оговаривалось, что «необязательно» не значит «бесплатно». Так, не написавшие эссе получали «0», а не выполнившие в срок рецензирование получали минимальный балл, то есть, «-2». Работы, уличённые в превышении допустимого уровня заимствований, автоматически получали минимальный балл, то есть, «-9» (рис. 1).

<sup>1</sup> «Успеваемость глазами преподавателя» — Компьютерная графика 2013. Сайт поддержки курса. [http://cg-2013.blogspot.ru/2013/10/blog-post\\_16.html](http://cg-2013.blogspot.ru/2013/10/blog-post_16.html)



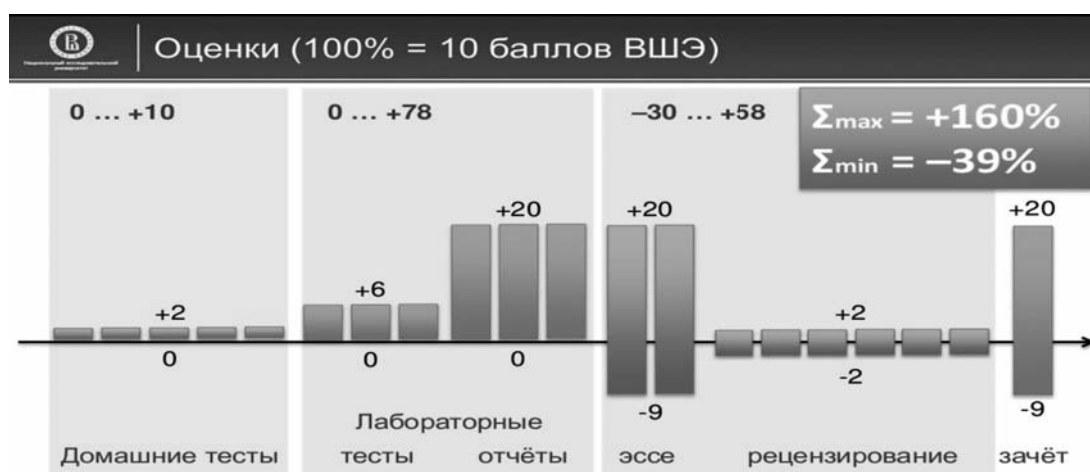


Рис. 1. Система оценок в курсе «Компьютерная графика» (2014). Слайд из презентации к вводной лекции.

Таким образом, «безопасная» модель поведения предполагала риски для тех, кто пишет эссе и стимулировала проверять чужие работы и это должно было обеспечить многократное превышение количества рецензентов над количеством авторов, что необходимо было для применения статистических методов при дальнейшей обработке результатов.

### Критерии оценки

В пиринговой оценке кроется риск разрыва уровня квалификации: ленивых студентов среди проверяющих может оказаться больше, чем креативных и энергичных. Но здесь вступает в силу одно из важных условий постановки работы с пиринговой проверкой: критерии оценки должны быть однозначны и понятны, легко идентифицируемы как проверяющим, так и контролирующим ход работы [14]. Среди таких критериев были и «контрольные вопросы», предполагающие численные ответы, например, «сколько источников указано в списке литературы» или «какой процент авторского текста показал анализатор».

Каждый студент в потоке получал один из вариантов тем, выбранный случайным образом из общего пула вариантов. К заданию прилагалась таблица критериев оценки<sup>1</sup> с детальным описанием не только самих критериев, но и необходимого содержания по каждому критерию, для получения того или иного балла. Критерии относились как к содержанию работы, так и к форме. Последнее было обусловлено отсутствием предварительной подготовки у студентов: большинство из них не могли с первого раза оформить документ по требованиям, что создавало бы им дополнительные препятствия при попытке подачи своих статей или тезисов в публикацию.

Пиринговая проверка обычно проводится вслепую. В проведенных экспериментах на второй год «слепым» был также и процесс распределения тем — каждый студент получал письмо с персональным номером выпавшей ему темы. Темы распределялись равномерно по общему числу студентов потока.

Таблица критериев оценки — это самый важный элемент в постановке работы над пиринговой проверкой. Если требования сформулированы чётко и заставляют обратить внимание автора (а позднее — рецензента) на различные стороны работы, то и сам процесс написания работы, удовлетворяющей требованиям, и процесс проверки таких работ становится полезным для студента, так как учит объективности и непредвзятости. Безусловно, на любое правило найдутся и те, кому интереснее найти незапрещённый обходной путь, но в данном случае это стимул к совершенствованию критериев по итогам каждой проведенной работы.

Добровольцы, решившие писать эссе, заполняли форму, в которой, помимо ссылки на файл и номера темы, указывали несколько численных характеристик своей работы, относящихся к библиографии — это была форма самоконтроля. Позднее студентам самим предлагалось указать процент авторского текста по данным eTXT. Предполагалось, что перед отправкой студент все-таки задумается, стоит ли посылать заведомо проблемную работу, если списал её. Показательно, что задача «укажите количество источников в работе» вызывала затруднения у студентов второго курса — показания рецензентов и авторов нередко отличались не только друг от друга, но и от фактического числа источников.

### Рецензенты

Большинство студентов из рассматриваемой выборки (второй курс бакалавриата инженерного факультета) не имели навыков написания академических текстов, равно как и навыков проверки чужих работ — в школе этому не учат, а в институте ко второму курсу заняться научной или педагогической деятельностью они ещё не успевают. Отсюда возникло опасение, что не только работы будут в среднем низкого качества, но, что важнее, их проверка окажется в руках столь же неквалифицированных рецензентов. Поскольку пиринговую проверку предполагалось использовать регулярно, то эту особенность аудитории нужно учитывать как данность. С другой стороны, учебный процесс и отдельные студенты в частности не должны страдать хотя бы от экспериментов преподава-

<sup>1</sup> Критерии оценки эссе — Компьютерная графика 2013. Сайт поддержки курса. [http://cg-2013.blogspot.ru/2013/09/blog-post\\_24.html](http://cg-2013.blogspot.ru/2013/09/blog-post_24.html)

телей. Поэтому, с одной стороны, в формулы расчета оценки работ закладывались механизмы дисквалификации неадекватных оценок и вводилась дифференцированная оценка работы самих рецензентов, а с другой стороны, проверка работ проводилась под контролем преподавателя и наиболее неоднозначные работы проходили ручную проверку, по результатам которой неадекватные рецензии аннулировались. Позже была применена формула более гибкой весовой дифференциации, позволявшая плавно градуировать вес рецензента и его персональную оценку в зависимости от удаленности от мнения большинства.

Требовалось обеспечить значительное превосходство числа рецензентов над числом авторов, поскольку одну работу должны проверить несколько человек и только в этом случае можно выявить возможные расхождения во мнениях. В самой системе оценок в курсе закладывались стимулирующие меры для рецензентов и преграждающие — для авторов работ. Формально, любой студент мог не писать и не проверять эти работы, если был уверен в успешности своих усилий в практической области (лабораторные работы давали достаточно высокий балл, если их выполнять добросовестно).

Здесь следует добавить, что позже, когда проводился разбор работ, в частности, аннулированных за плагиат, студенты подходили с претензией «я же потратил время и силы, почему я получил отрицательную оценку, хотя те, кто даже не пытался, остались при своём». Как видно, неотвратимость получения вознаграждения за любое действие или, по крайней мере, сохранения имеющегося положения прочно укоренилась в головах, что мешает студентам всерьёз воспринимать риски реальной жизни, в том числе — репутационные.

### Постановка эксперимента и используемые инструменты

Эксперимент проводился в 2013–14 и 2014–15 учебных годах. Основной «площадкой» был потоковый курс «Компьютерная графика» на втором курсе бакалавриата факультета Информационных технологий и вычислительной техники МИЭМ НИУ ВШЭ. Помимо этого, работы с пиринговой проверкой проводились на двух других курсах, один из них был элективным с аудиторией 30 человек, но основные результаты были получены по итогам двух потоковых курсов «Компьютерной графики».

В 2013 году курс впервые был переведен с кафедры (2 группы) на поток (5–7 групп) и сокращён вдвое (8 недель вместо 17), также пропала курсовая работа, составлявшая значимую часть в прежнем курсе. В новых условиях полагаться на ручную обработку студенческих работ было неразумно, поскольку весь курс проходил в сжатые сроки. Рабочий учебный план, в котором записана одна домашняя работа и экзамен, в формальном виде не обеспечивал погружения студентов в дисциплину, как и промежуточного контроля их успеваемости. Поэтому в курсе, помимо двух эссе и «домашней работы», заявленной в плане, были введены «домашние тесты», которые каждый мог проходить в любое время в пределах установленного диапазона времени (обычно — неделя), тесты на лаборатор-

ных работах, которые открывались для каждой группы в аудиторные часы. Вес таких тестов был выше, поскольку в них, как минимум, можно с большой вероятностью утверждать, что студент сам заполняет форму теста.

Наличие в вузе централизованно поддерживаемой LMS не решило задачу автоматизации работы в курсе — отчасти по техническим, отчасти по организационным причинам. Достаточно сказать, что технической гибкости не хватило даже для проведения тестов в подгруппах, не говоря о пиринговой проверке, которую эта система не поддерживает. Таким образом, инструментальные средства требовалось подобрать самостоятельно.

На этапе подготовки курса были изучены варианты существовавшего на тот момент программного обеспечения [20] и сервисов (например, iPeer<sup>1</sup>), но более детальное ознакомление с ними показало, что те немногие инструменты, которые удалось найти, были ориентированы на другой формат работы и не подходили, а для проведения исследований требовалась максимальная гибкость и возможность исправлять замеченные недоработки в сжатые сроки. В этой ситуации роль базы для экспериментов легла на сервисы Google Apps (тексты Google Documents, таблицы Google Spreadsheets и формы Google Forms), а также сервис Blogger для публикации материалов курса. Обратная связь от студентов принималась через формы, что позволяло надежно собирать их ответы в таблицы с точным указанием времени отправки. Таблицы, в свою очередь, позволяют автоматизировать обработку поступающей информации при помощи формул и скриптов. Результаты публиковались на сайте. Так из документооборота были исключены этапы переписки со студентами по электронной почте или в мессенджерах. Это отнюдь не исключало переписку для поддержки по содержательным или организационным вопросам, но все учебные транзакции стали проходить через формы и регистрироваться в таблицах, исключая человеческий фактор из обработки.

Далее, средствами Forms требовалось создать тесты, а средствами Spreadsheets — весь цикл подготовки и обработки информации пиринговой проверки. Для создания тестов Forms подходит ограниченно, но в данном случае тесты проводились больше для напоминания студентам о курсе, заставляя их вернуться к темам лекций после этих лекций, фактически, заменяя «повтор пройденного материала». Существенно сложнее оказалось реализовать при помощи таблиц весь цикл обработки пиринговой проверки эссе. В первый год применялась полуавтоматическая обработка данных: использовались лишь формулы в таблицах, что не позволяло работать с персональной рассылкой электронной почты, ограничивало обработку массивов данных. На этом этапе были сформулированы задачи для автоматизации обработки и выявлены недостатки алгоритмов расчета оценок. С другой стороны, табличное представление информации давало полную картину хода работы: все вычисления были детально видны на листах таблиц и любой студент, не согласившийся с оценкой, мог видеть, как она формировалась и что помешало ему получить желаемый балл.

<sup>1</sup> iPeer. Веб-приложение для проведения пиринговой оценки. <https://sourceforge.net/projects/ipeer/>

## Ход работы

В 2013 году темы эссе распределялись случайным образом по списку студентов и этот список публиковался на сайте поддержки курса, то есть, студенты знали, кому из их сокурсников какая тема досталась. Поскольку работа позиционировалась как добровольная, но сама эта активность была для студентов необычным нововведением, количество сданных работ примерно соответствовало ожидаемому (рассчитывалось, что работы напишут 20% от потока, а задания на их проверку выполнят 80%). Работы прислали 23 автора (19%), в то же время, специфику начисления баллов за рецензирование поняли не все и количество присланных рецензий оказалось ниже ожидаемого (69%). Тем не менее, заложенный в механизм оценки расчёт сработал — количество рецензий на одну работу было достаточным не только, чтобы увидеть разные мнения рецензентов, но и позволяло применять статистические методы для определения наиболее адекватных оценок.

Описанный ранее подход к оцениванию показал сильное «размытие» оценок<sup>1</sup>. И сильные, и слабые работы получали незначительно различающиеся баллы, большинство оценок, отклоняющихся к верхней или нижней границе шкалы, нивелировались оценками, попадающими в «безопасный» диапазон в середине шкалы (рис. 2).

Чтобы выявить недобросовестных рецензентов и показать студентам ориентиры на конкретных примерах, на сайте поддержки подробно разбирались все процессы оценивания и выборочного контроля<sup>2</sup>.

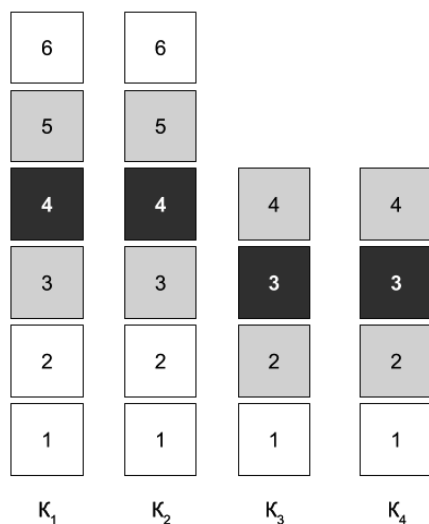


Рис. 2. «Безопасный диапазон» оценок по каждому из критериев  $K_1...K_4$ , в котором рецензент менее всего рискует выйти из доверительного интервала (закрашено светлым) или попасть в «дальнюю» часть весовой шкалы при оценке рецензий.

Второй год эксперимента проходил на потоке 180 человек и для проведения курса на базе GoogleSpreadsheets

<sup>1</sup> «Эссе-1. Оценки работ». Компьютерная графика 2013. Сайт поддержки курса. <http://cg-2013.blogspot.ru/2013/10/1.html>

<sup>2</sup> «Как проверяются рецензии». Компьютерная графика 2013. Сайт поддержки курса. [http://cg-2013.blogspot.ru/2013/10/blog-post\\_6.html](http://cg-2013.blogspot.ru/2013/10/blog-post_6.html)

были созданы скрипты, что позволило вынести обработку данных из таблиц и использовать их только для сбора информации из форм и наглядного представления результатов обработки. Для наглядности также выводились некоторые промежуточные значения, что позволяло сохранить для студентов возможность видеть принцип формирования их оценки.

Автоматизация инструментальной части имела однозначно положительные последствия и для понимания студентами формирования их оценки, и для разгрузки преподавателя. В то же время подбор алгоритмов вычисления оценки стал задачей неоднозначной. С одной стороны, общие принципы расчета оценки пиринговым способом известны [11] из широко распространенных (на момент проведения эксперимента — недавно появившихся) массовых открытых онлайн курсов [15], с другой стороны, онлайн курсы имеют другой формат и другую аудиторию, обычно они проводят входное тестирование (ground truth) для определения весовых коэффициентов рецензентов [10]. Прямое копирование алгоритмов в данном случае недопустимо. Как было сказано ранее, мотивирующие приготовления были заложены ещё в системе оценок за курс, эту же специфику аудитории следовало заложить и в алгоритмы расчета оценок работ и рецензий.

Во избежание конфликтов и непонимания принципа расчета оценок, в первый год применялась простая формула: считалось среднее арифметическое по оценкам всех рецензентов данной работы по каждому критерию (их было четыре), после чего преподавателем вручную проверялись десять самых неоднозначных работ. Неоднозначность выявлялась по среднеквадратичному отклонению в оценках рецензентов. Далее оценки преподавателя (экспертные оценки) сравнивались с оценками каждого из рецензентов в установленном доверительном интервале (плюс-минус один балл по каждому критерию) и, если расхождения превышали допустимый порог, это отмечалось в таблице. Если рецензия имела более 50% критериев (в данном случае, три из четырех оценок выходили за доверительный диапазон) с отклоненной оценкой, то она аннулировалась, что исключало её из расчета общей оценки за работу и снимало все баллы, начисляемые рецензенту за проверку данной работы (рис. 3).

На рис. 3 показана проверка рецензий на три работы (фамилии рецензентов скрыты). В столбцах В–Е стоят оценки рецензентов, столбец F — номер проверяемой работы. Видно, что работа № 2 попала под контрольную проверку преподавателем и в столбцах М–Р стоят четыре оценки, с которыми сравниваются оценки каждого из экспертов. Из восьми экспертов двое разошлись во мнениях с преподавателем по трём из четырёх критериев более, чем на один балл (максимум баллов — шесть для первых двух оценок и четыре для последних двух). В столбцах Н–К показаны совпадения между оценками и видно, что две рецензии с тремя «0» были отклонены. Далее оценка за работу пересчитывалась с учетом оставшихся рецензий, а авторам отклоненных рецензий баллы за рецензирование не начислялись.

Также на рис. в столбце Q видна оценка аутентичности работ (в данном случае измерения делались преподавателем, а не рецензентами), и ясно, что первые две работы

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Фамилия	Полнота	Обосн.	Иллюст.	Верстка	Номер	Итог	Сравнение					Экспертная оценка				Аутентичность
2	Грановский	6	6	4	4	1	Ok										70
3	Зайцев	6	6	4	1	1	Ok										70
4	Ипатов	4	4	4	3	1	Ok										70
5	Лескин	6	4	4	2	1	Ok										70
6	Попов	4	4	3	2	1	Ok										70
7	Седов	4	3	3	2	1	Ok										70
8	Токарева	3	6	3	4	1	Ok										70
9	Чинков	4	5	4	4	1	Ok										70
10	Аниканов	5	6	3	3	2	1	0	0	1	1		3	4	4	4	71
11	Дерябина	5	5	4	4	2	1	0	1	1	1		3	4	4	4	71
12	Кольбе	3	3	3	2	2	1	1	1	1	0		3	4	4	4	71
13	Макаренко	5	5	3	3	2	1	0	1	1	1		3	4	4	4	71
14	Титова	4	4	3	3	2	1	1	1	1	1		3	4	4	4	71
15	Уланов	4	2	1	1	2	0	1	0	0	0		3	4	4	4	71
16	Федорченко	5	5	4	2	2	1	0	1	1	0		3	4	4	4	71
17	Чудовская	6	6	3	2	2	0	0	0	1	0		3	4	4	4	71
18	Абрамешин	4	5	3	2	3	Ok										85
19	Гречушников	4	4	4	2	3	Ok										85
20	Куриленков	4	4	2	2	3	Ok										85
21	Лапин	5	5	4	4	3	Ok										85
22	Полищук	2	3	3	3	3	Ok										85
23	Попов	5	4	3	2	3	Ok										85
24	Слободенюк	5	4	4	3	3	Ok										85
25	Чинков	4	4	3	3	3	Ok										85
26	Чудовская	5	5	3	3	3	Ok										85
27	Шашков	4	4	4	3	3	Ok										85

Рис. 3. Выборочный контроль рецензий и их сравнение с экспертной оценкой.

превысили допустимый уровень заимствований — они также отклоняются с выставлением минимального балла.

Несмотря на то, что на практике отклонено было всего восемь и шесть рецензий в первом и втором эссе соответственно, такие вычисления дают слишком дискретную реакцию на отклонения, что, в свою очередь, размывает фактическую оценку работ. Рецензенту достаточно поставить балл из «безопасного» диапазона и это гарантирует ему получение своих баллов (рис. 3). Шкалы критериев имели четыре и шесть градаций, доверительный интервал составляет три градации, то есть, несложно догадаться, что рецензенту нецелесообразно ставить крайние баллы, а внутри оставшихся лишь угадать по беглому взгляду на работу, в какую сторону может сместиться мнение преподавателя, не вникая в детали, что не соответствовало фактическим задачам всего мероприятия. Разумеется, среди студентов многие отнеслись к работе с интересом, тем более, что это для всех был первый опыт подобной деятельности, но алгоритм требовал совершенствования.

На второй год эксперимента было решено ввести непрерывную весовую шкалу для рецензий, по которой считать и весовой коэффициент оценки рецензента в групповой оценке, и оценку самой рецензии.

1. Для каждой компоненты оценки каждой работы определяется среднее арифметическое значение набора. Средние арифметические значения компонент становятся «эталонными» оценками.
4. Для каждой компоненты оценки каждой работы находится максимальное отклонение от эталонной соответствующей оценки. Это отклонение становится длиной шкалы.
5. Для каждой компоненты оценки каждого рецензента находится весовой коэффициент соответствия эталонной оценке, приведенный по шкале (от 0 до 1).

6. Итоговая оценка работы рассчитывается как сумма всех компонент оценки по всем рецензентам для данной работы, умноженных на соответствующий весовой коэффициент.

На рис. 4 виден пример расчета оценки для одного критерия из шести баллов. Закрашенные клеточки — это поступившие оценки рецензентов: «1» не поставил никто, «2» — один, «3» — двое и так далее. Средний балл в таком случае будет равен 4,5, там проходит ось мнения большинства. Длина шкалы — от оси до самой удаленной оценки (2), то есть, 2,5 балла. Вес оценок рецензентов будет падать пропорционально удалению от оси. Так, каждые 0,5 балла при линейном падении веса будут отнимать 20%. Любопытно, что при таком подсчете никто не получит полный балл, так как оценка имеет целочисленные значения и даже самые близкие к оси оценки получают вес 80%, а соседние с ними — по 40%. Компенсировать этот недостаток несложно, пропорционально «подтянув» после всех расчетов все значения, чтобы максимальные достигли 100%, но в эксперименте этого не делалось.

Шкала в таком подходе — это максимальное расстояние от мнения большинства до самого отдаленного от него мнения отдельного рецензента, поэтому на рисунке 4 шкала уходит в область несуществующих значений оценки. Поскольку разные критерии имели разное количество баллов, то и предельная длина шкалы варьировалась. На практике она зависела ещё и от единодушия рецензентов — если все сошлись во мнениях и поставили одинаковую оценку, то единственный, поставивший иной балл, попадал на край шкалы и такая оценка обесценивалась. При таком подходе «удалённость» отдельно взятой оценки от «эталонной» (то есть, средней) определяла падение веса этой оценки в итоговой оценке, равно как и оценку за саму рецензию. Этот подход, конечно, не

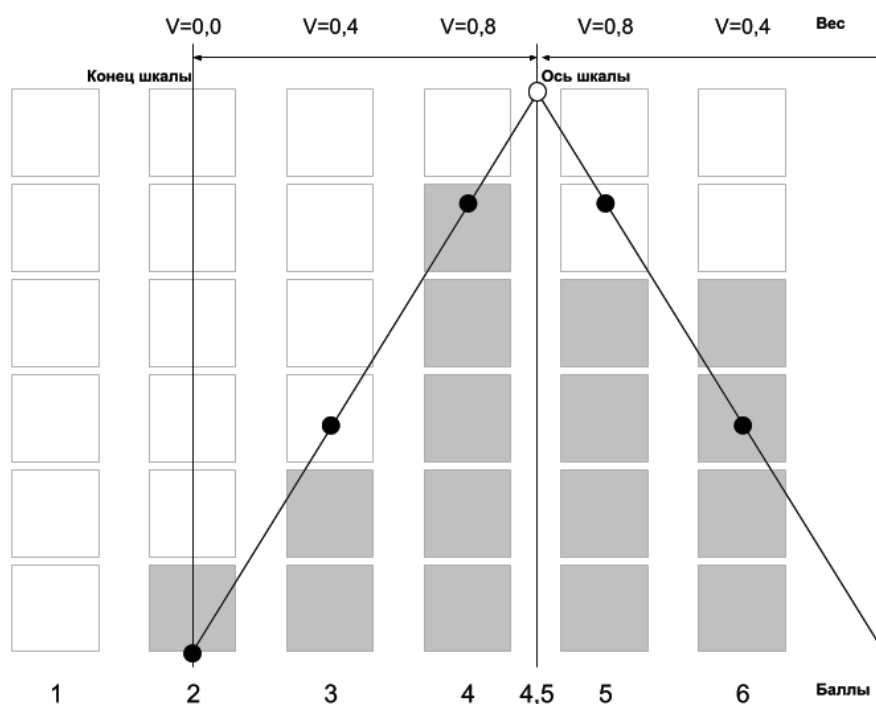


Рис. 4. Расчет весовых коэффициентов рецензий с линейной весовой шкалой (показан расчет по одному критерию с максимумом оценки 6 баллов).

гарантировал, что эталонной станет действительно адекватная оценка, но особенности демократического управления (то есть, приоритета мнения большинства) — не тема данной статьи.

На практике проводилась частичная проверка работ преподавателем и далее экспертные оценки прини-

мались за «эталонную», что смещало точку отсчета шкалы. В такой ситуации случалось, что немногие поставившие далекую от мнения большинства оценку, оказывались у основания шкалы, а большинство теряло вес в соответствии с удаленностью от нового центра истины (рис. 5).

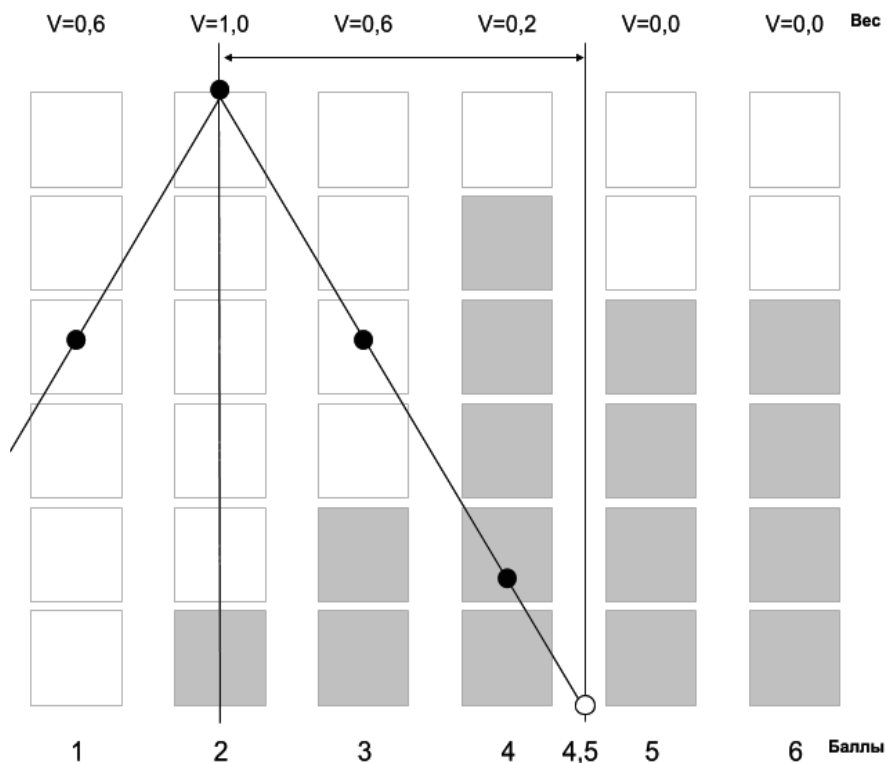


Рис. 5. Влияние смещения оси шкалы при экспертной оценке (на примере показан случай, когда эксперт выставил оценку «2», существенно отличающуюся от мнения большинства).

Определение шкалы по максимальной удаленности от усредненного мнения большинства (или от мнения эксперта) давало довольно большие диапазоны, что опять приводило к размытию оценки и сохраняло относительно высокий балл для рецензентов, выставляющих оценки «в безопасном диапазоне», если мнение эксперта не отличалось от мнения большинства радикально, как это показано на рис. 5, хотя и такой, правда, единичный случай был в 2014 году.

### Последующая работа

По итогам эксперимента были собраны все полученные от рецензентов оценки и этот массив данных стал основой для следующего исследования. Зная оценки всех участников эксперимента, включая преподавателя, можно построить формулы, которые дадут высокую степень корреляции с экспертной (преподавательской) оценкой. В данной статье мы не будем останавливаться на этом исследовании, кратко оно отражено в докладе [4].

### Результаты и выводы

1. Численные итоги активности студентов в ходе экспериментов 2013 и 2014 годов отражены в табл. 1. Здесь мы видим соотношение числа авторов и рецензентов как следствие мер по стимулированию рецензентов и демонстрации ответственности авторов за некачественную работу, заложены в системе оценки работ в курсе. Каждому студенту предлагалось проверить по три работы, поэтому число рецензий в среднем втрое больше числа рецензентов.

Таблица 1

Активность студентов в ходе экспериментов

	Эссе-1 2013	Эссе-2 2013	Эссе-1 2014	Эссе-2 2014
Эссе	23	12	8	12
Рецензии (всего)	249	270	363	453
Рецензенты	83	90	131	149

2. Даже детализированная по критериям оценки работ при пиринговой проверке склонна к смещению к «безопасному диапазону». В такой ситуации рецензент имеет меньше шансов выпасть из доверительного интервала (в первом запуске эксперимента) или на попасть на край весовой шкалы (во втором запуске). Решить это можно введением контрольных вопросов с однозначной оценкой, пересекающихся с оценкой по основным выбранным критериям. Например, если в шкале критериев приводятся ориентиры по количественным характеристикам (допустим, ссылкам на источники), то явное указание численных характеристик (допустим, 3 из 10 минимально необходимых) будет значить, что по соответствующему критерию («Обоснованность») высокий балл уже не может быть выставлен, поскольку в таблице критериев приводятся соответствующие численные ориентиры.

3. Описанные эксперименты привнесли новое содержание в преподавание курса и развили у студентов навыки оценивания работ по объективным критериям, а система

оценки дала им представление о том, как оцениваются их работы. Сами студенты в финальном опросе неоднократно отмечали, что за всё время учёбы в школе и институте они впервые поняли, как ставится оценка. В то же время, полностью решить задачу автоматизации проверки такой подход не позволил — преподавателю приходилось делать, пусть и частичную, проверку работ и вмешиваться в процесс расчета оценки, хоть и опосредованно.

4. По итогам проведенных экспериментов был получен массив оценок и набор анкет с отзывами студентов. Численные данные были обезличены для возможности публичного использования и послужили основным материалом для последующих исследований. В частности, путем моделирования различных рабочих ситуаций и наборов рецензентов были получены формулы расчета оценки для первой (заменяющего ground-truth этап) и последующих пиринговых проверок, а также вычислены желательные и минимальные соотношения количества рецензентов к количеству авторов.

5. Описанный метод и использовавшиеся инструменты применялись и на других учебных курсах. В частности, было проверено, что способ применим и в группах численностью 25–30 человек, если сохраняется допустимое отношение количества рецензентов к количеству авторов.

### Литература

1. Антипова Е. П., Левчук Н. Л. Формирование инновационного потенциала учащихся // Муниципальное образование: инновации и эксперимент. 2015. № 6. С. 53–58.
2. Гузанов Б. Н., Пустовалова Е. И., Выгузова Е. В. Самостоятельная работа студентов как основа формирования компетенций в условиях многоуровневой подготовки // Муниципальное образование: инновации и эксперимент. 2016. № 3. С. 65–73.
3. Дешко И. П., Кряженков К. Г., Тулинов С. В. Оценка результативности обучения в массовом открытом онлайн курсе по сетевым технологиям // Муниципальное образование: инновации и эксперимент. 2016. № 3. С. 10–15.
4. Кольбе А. С., Королев Д. А. Исследование применимости пирингового метода оценки в очных курсах // 14-ая международная научно-практическая конференция «Управление качеством». Тезисы докладов. М.: ПРОБЕЛ-2000, МАТИ, 2015. С. 219–229.
5. Cho K. et al. Student revision with peer and expert reviewing / Cho K., MacArthur C. // Learning And Instruction. 2010. Т. 20. № 4. С. 328–338.
6. dLu R. A comparison of anonymous versus identifiable e-peer review on college student writing performance and the extent of critical feedback / dLu R., Bol L. // Journal of Interactive Online Learning. 2007. Т. 6. № 2. С. 100–115.
7. Estévez-Ayres I. et al. An Algorithm for Peer Review Matching in Massive Courses for Minimising Students' Frustration / Estévez-Ayres I., García R. M. C., Fisteus J. A., Kloos C. D. // J. UCS. 2013. Т. 19. № 15. С. 2173–2197.
8. Hames I. Ethical Guidelines for Peer Reviewers // Committee on Publication Ethics (COPE). March, 2013. v.1 URL: [http://publicationethics.org/files/Peer%20review%20guidelines\\_0.pdf](http://publicationethics.org/files/Peer%20review%20guidelines_0.pdf) (Дата обращения: 15.09.2016).
9. Kloos C. D. et al. Mixing and blending MOOC Technologies with face-to-face pedagogies / C. Delgado Kloos, P.J. Muñoz-Merino,

- C. Alario-Hoyos, I. Estévez Ayres and C. Fernández-Panadero // 2015 IEEE Global Engineering Education Conference (EDUCON). IEEE, 2015. C. 967–971.
10. Kulkarni C. et al. Peer and self assessment in massive online classes / Kulkarni C., Wei K.P., Le H., Chia D., Papadopoulos K., Cheng J., Koller D., Klemmer S.R. // Design thinking research. Springer International Publishing, 2015. C. 131–168.
  11. Kulkarni C. et al. Scaling self and peer assessment to the global design classroom / Kulkarni C., Pang-Wei K., Le H., Chia D., Papadopoulos K., Koller D., Klemmer S.R. // Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013.
  12. Kuzminykh K. «Global Classroom» Experiment at Higher School of Economics: Who Takes MOOCs offered by Russian Universities? // Higher Education in Russia and Beyond. 2015. T. 3. № 1. C. 8–9.
  13. Mulder R. A. et al. Peer Review in Higher Education: Student Perceptions before and after Participation / Mulder R. A., Pearce J. M., Baik C. // Active Learning In Higher Education. 2014. № 15(2). C. 157–171.
  14. Norton L. S. Essay-writing: what really counts? // Higher Education. 1990. T. 20. № 4. C. 411–442.
  15. Piech C. et al. Tuned models of peer assessment in MOOCs / Piech C., Huang J., Chen Z., Do C., Ng A., Koller D. // arXiv preprint arXiv:1307.2579. 2013.
  16. Stognieva O. Implementing Peer Assessment in a Russian University ESP Classroom // Journal of Language and Education. 2015. T. 1. № 4. C. 63–73.
  17. Suen H. K. Peer assessment for massive open online courses (MOOCs) // The International Review of Research in Open and Distributed Learning. 2014. T. 15. № 3.
  18. Taylor S. M. Can Peer Review Help Johnny Write Better? Critiquing classmates' work helps students better diagnose their own writing problems, but many do not know how to comment constructively. Here are successful peer-review strategies // Education Digest. 2014. T. 80. № 4. C. 4–10.
  19. Wang Y. et al. A Motivation Model of Peer Assessment in Programming Language Learning / Wang Y., Liang Y., Liu L., Liu Y. // arXiv preprint arXiv:1401.6113. 2014.
  20. Williams R. Automated essay grading: An evaluation of four conceptual models // Expanding Horizons in Teaching and Learning. Proceedings of the 10th Annual Teaching Learning Forum. Perth, Australia: Curtin University of Technology, 2001. C. 7–9.
- Management». Abstracts. M.: PROBEL-2000, MATI, 2015. P. 219–229. (In Russian)
5. Cho K. et al. Student revision with peer and expert reviewing / Cho K., MacArthur C. // Learning And Instruction. 2010. V. 20. No. 4. P. 328–338.
  6. dLu R. A comparison of anonymous versus identifiable e-peer review on college student writing performance and the extent of critical feedback / dLu R., Bol L. // Journal of Interactive Online Learning. 2007. V. 6. No. 2. P. 100–115.
  7. Estévez-Ayres I. et al. An Algorithm for Peer Review Matching in Massive Courses for Minimising Students' Frustration / Estévez-Ayres I., García R. M. C., Fisteus J. A., Kloos C. D. // J. UCS. 2013. V. 19. No. 15. P. 2173–2197.
  8. Hames I. Ethical Guidelines for Peer Reviewers // Committee on Publication Ethics (COPE). March, 2013. v.1 URL: [http://publicationethics.org/files/Peer%20review%20guidelines\\_0.pdf](http://publicationethics.org/files/Peer%20review%20guidelines_0.pdf) (Дата обращения: 15.09.2016).
  9. Kloos C. D. et al. Mixing and blending MOOC Technologies with face-to-face pedagogies / C. Delgado Kloos, P. J. Muñoz-Merino, C. Alario-Hoyos, I. Estévez Ayres and C. Fernández-Panadero // 2015 IEEE Global Engineering Education Conference (EDUCON). IEEE, 2015. P. 967–971.
  10. Kulkarni C. et al. Peer and self assessment in massive online classes / Kulkarni C., Wei K.P., Le H., Chia D., Papadopoulos K., Cheng J., Koller D., Klemmer S.R. // Design thinking research. Springer International Publishing, 2015. P. 131–168.
  11. Kulkarni C. et al. Scaling self and peer assessment to the global design classroom / Kulkarni C., Pang-Wei K., Le H., Chia D., Papadopoulos K., Koller D., Klemmer S.R. // Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013.
  12. Kuzminykh K. «Global Classroom» Experiment at Higher School of Economics: Who Takes MOOCs offered by Russian Universities? // Higher Education in Russia and Beyond. 2015. V. 3. No. 1. P. 8–9.
  13. Mulder R. A. et al. Peer Review in Higher Education: Student Perceptions before and after Participation / Mulder R. A., Pearce J. M., Baik C. // Active Learning In Higher Education. 2014. No. 15(2). P. 157–171.
  14. Norton L. S. Essay-writing: what really counts? // Higher Education. 1990. V. 20. No. 4. P. 411–442.
  15. Piech C. et al. Tuned models of peer assessment in MOOCs / Piech C., Huang J., Chen Z., Do C., Ng A., Koller D. // arXiv preprint arXiv:1307.2579. 2013.
  16. Stognieva O. Implementing Peer Assessment in a Russian University ESP Classroom // Journal of Language and Education. 2015. V. 1. No. 4. P. 63–73.
  17. Suen H. K. Peer assessment for massive open online courses (MOOCs) // The International Review of Research in Open and Distributed Learning. 2014. V. 15. No. 3.
  18. Taylor S. M. Can Peer Review Help Johnny Write Better? Critiquing classmates' work helps students better diagnose their own writing problems, but many do not know how to comment constructively. Here are successful peer-review strategies // Education Digest. 2014. V. 80. No. 4. P. 4–10.
  19. Wang Y. et al. A Motivation Model of Peer Assessment in Programming Language Learning / Wang Y., Liang Y., Liu L., Liu Y. // arXiv preprint arXiv:1401.6113. 2014.
  20. Williams R. Automated essay grading: An evaluation of four conceptual models // Expanding Horizons in Teaching and Learning. Proceedings of the 10th Annual Teaching Learning Forum. Perth, Australia: Curtin University of Technology, 2001. P. 7–9.

## References

1. Antipova E. P., Liauchuk N. L. Formation of innovative potential of students // Municipal education: innovation and experiment. 2015. No. 6. P. 53–58. (In Russian)
2. Guzanov B. N., Pustovalova E. I., Vyguzova E. V. Independent work of students as a basis for the formation of competences in the conditions of multilevel preparation // Municipal education: innovation and experiment. 2016. No. 3. P. 65–73. (In Russian)
3. Dshko I. P., Kryazhenkov K. G., Tulinov S. V. Training impact assessment in massive open online course on network technologies // Municipal education: innovation and experiment. 2016. No. 3. P. 10–15. (In Russian)
4. Kolbe A. S., Korolev D. A. Investigation of the applicability of the method of assessment peering in classroom courses // 14th International Scientific and Practical Conference «Quality