

**Динамика показателей массового взаимного оценивания
студенческих творческих работ (пир-асессмента)**

Михаил Петрович Карпенко,

доктор технических наук, профессор,

президент

*Частная образовательная организация «Современная гуманитарная
академия»,*

Басов Вадим Александрович,

кандидат физико-математических наук,

директор Дидакто-технологического управления

Библиотека информационно-образовательных ресурсов «Умней»

Установлено, что наиболее эффективным является воспитательный процесс, проводимый в деятельностной форме. Если пир-асессмент включен в учебный процесс, он становится деятельностью. При анализе результатов пир-асессмента интерес представляет не только качество работы студентов в роли ассессоров, но его динамика. Для обработки статистики выставляемых пир-отметок был определен комплекс методов обработки и показателей-критериев. Проведен анализ динамики показателей работы студентов в качестве ассессоров, что позволило продемонстрировать результаты воспитательного воздействия ассессмента. Наблюдается положительный эффект реализации дидактики ассессмента, студенты вырабатывают более устойчивые оценочные шкалы.

Ключевые слова: ассессмент, электронное обучение, воспитательное воздействие, отметки, стандартное отклонение.

В западных университетах взаимное оценивание студенческих творческих работ (эссе, рефератов) имеет солидную историю и насчитывает

более 50 лет. Во множестве источников декларируется положительное влияние пир-асессмента на учебный процесс [Agler, Miao, Moore, Zabrusky, 2004. С. 453–4612, Sadler, Good, 2006. С. 1–31, Topping, 2009. С. 20–27]. Более того, ряд исследователей отмечает, что оценивание студенческих работ сверстниками- студентами по качеству не уступает преподавательским оценкам [Mann, 2012. С. 4–8, Orpen, 1982. С. 567–572, Li, Xiong, Zang, Kornhaber, Lyu, Chung, Suen, 2016. С. 245-264]. Однако, в силу ряда причин, среди которых имеется и сопротивление преподавателей, пир-асессмент не становится массовым явлением и по масштабам не выходит за рамки экспериментов, проводимых отдельными преподавателями.

Одним из существенных различий западной и российской (а ранее - советской) парадигм образовательного процесса является большое внимание, уделяемое российской парадигмой его воспитательной составляющей. И действительно, воспитание и социализация студентов в российских вузах проводится, но в основном в вербальных формах. Тогда как еще со времен педагога Макаренко А.С. стало классическим понимание, что наиболее эффективным является воспитательный процесс, проводимый в деятельностиной форме. Делом студентов является учеба, пир-асессмент с обратной связью включен в учебный процесс, поэтому он является деятельностью и является наиболее эффективным средством воспитательного воздействия на студентов.

Обратной связью является расчет числовых показателей асессмента с доведением их до сведения студентов-асессоров с соответствующими оценками и критикой. Заметим, что современное информационное оснащение образовательного процесса позволяет с помощью соответствующей программной платформы сделать асессмент массовым, автоматизируя все фазы его организации, проведения, обработки и хранения результатов.

В Современной гуманитарной академии (г. Москва) массовый асессмент, вписанный в учебный процесс, реализуемый в электронной информационно-образовательной среде (ЭИОС) начал проводиться с 2014

года и на студентах других вузов проводится до настоящего времени. В этом масштабном эксперименте интерес представляет не только числовые показатели, позволяющие оценить творческие работы студентов и качество их работы в виде ассессоров, но и динамика этих показателей, так как именно динамика может продемонстрировать результаты воспитательного воздействия. Заметим, что наш эксперимент – уникальное явление – он в числовом виде представляет результат воспитания, полученный в предметной области.

С целью обеспечения сбора данных для исследования, в проводимые в ЭИОС Ровеб (Роботизированной образовательной веб – среде) творческие занятия был включен этап пир-ассесмента, в рамках которого обучающимся предлагался для анализа набор творческих работ, выполненных их «коллегами» - другими обучающимися. Этап пир-ассесмента предшествовал выполнению самого занятия каждым ассессором. Работы предоставлялись на условиях анонимности, т.е. обучающийся ничего не знал об авторе работы, не был с ним знаком. Каждую работу из набора пир-ассессор должен был оценить по 5-ти балльной шкале: (2-5) с точностью до 0,1 балла. Выставление отметок производилось также анонимно. Автор работы не имел доступа к информации о том, кто являлся ассессором его работ. Каждый обучающийся в ЭИОС Ровеб имел уникальный идентификатор, позволяющий администрации проводить анализ образующегося массива данных ассесмента, используя возможности IT – платформы.

Эксперимент проводился в 2014 - 2018гг. включительно. Всего в эксперименте приняли участие 8750 студентов, использовавших для обучения ЭИОС Ровеб. Образовательные программы были построены на принципах активного применения дидактики ассесмента. Для оценки влияния длительности обучения 5-ти летняя выборка была разделена на 10 частей по числу семестров обучения.

Таблица 1 Распределение обучающихся по семестрам обучения

Номер семестра обучения	Количество
1	828
2	1018
3	1253
4	1206
5	1762
6	710
7	1173
8	403
9	168
10	229
ИТОГО	8750

Для обработки накопленной статистики выставяемых обучающимся пир-отметок был определен комплекс методов обработки и показателей-критериев.

По отношению к работе, выдаваемой на ассесмент, было определено два количественных показателя: это диссектная отметка a_d (от лат. dissector — исследуемый, исследователь), в качестве которой по очереди бралась отметка, выставленная каждым исследуемым пир-ассессором, и трасовая отметка a_t (от англ. trust — доверие) – средняя отметка среди выставленных другими пир – ассессорами данной работы.

Таким образом, каждая i – тая творческая работа обучающихся имела два количественно определяемых показателя-критерия: $a_{d(i)}$ и $a_{t(i)}$.

Для каждого i -го пир-ассесмента, проведенного j -м обучающимся ассессором определялось:

- отклонение диссектной отметки от трасовой по формуле:

$$\delta_d(i) = a_d - a_t \quad (1)$$

- удельное отклонение диссектной отметки по формуле:

$$\omega_d(i) = \frac{\delta_d(i)}{a_d(i)} \quad (2)$$

После чего, для каждого j -ого обучающегося определялось:

- стандартное отклонение диссектных отметок $\delta_d(j)$ по формуле:

$$\sigma_d(j) = \sqrt{\frac{\sum(\delta_d(i))^2}{N}}, \quad (3)$$

где N – количество диссектных отметок;

$\delta_d(i)$ – по формуле (1);

- среднее удельное отклонение диссектной отметки $\overline{\omega}_d(j)$ по формуле:

$$\overline{\omega}_d(j) = \frac{1}{N} \sum \omega_d(i), \quad (4)$$

где N – по формуле (3);

$\omega_d(i)$ – по формуле (2);

- поправочный препозиционный коэффициент $K(j)$ по формуле:

$$K(j) = 1 - \overline{\omega}_d(j), \quad (5)$$

где $\overline{\omega}_d(j)$ – по формуле (4).

Стандартное отклонение диссектной отметки j -го асессора $\sigma_d(j)$ принималось в качестве его интегрального показателя $Q_{\text{инт.}}$. Этот показатель-критерий, отражает множественное действие факторов, влияющих на отклонения выставяемых асессором отметок от средних. Условно все действующие на асессора факторы можно разделить на две группы: одна представляет системные факторы, являющиеся характеристикой асессора как личности, другая - факторы, являющиеся следствием набора обстоятельств, так или иначе влияющих на принятие решения асессора, а степень такого влияния является другой характеристикой. Если системные факторы выражаются в систематическом завышении или занижении выставяемой отметки, то случайные факторы обеспечивают диссипацию (от лат. *dissipatio* — рассеяние) итогового результата.

Очевидно, что теоретическому исправлению в итоговом результате асессмента поддаются только системные факторы. Для этого вводится второй

важный препозиционный показатель-критерий: поправочный препозиционный коэффициент $K(j)$. Условно по его величине можно разделить всех обучающихся две группы:

$K(j) < 1$ оверстэйторы, систематически завышающие отметки ассессмента ;

$K(j) > 1$ лоустэйторы, систематически занижающие отметки ассессмента .

После определения для j -го ассессора поправочного препозиционного коэффициента $K(j)$ методика предусматривает исправление результатов, путем вычисления поправленных диссектных отметок $a'_d(j)$ для каждого i -го ассессмента по формуле:

$$a'_d(i) = K(j)a_d(i) , \quad (6)$$

где $K(j)$ – по формуле (5);

$a_d(i)$ – первоначальная i – тая диссектная отметка.

Затем вычисляется отклонение $\delta'_d(i)$ поправленной диссектной отметки $a'_d(i)$ от трастовой a_t по формуле:

$$\delta'_d(i) = a'_d(i) - a_t \quad (7)$$

Производится перерасчет поправленного стандартного отклонения диссектных отметок по формуле:

$$\sigma'_d(j) = \sqrt{\frac{\sum(\delta'_d(i))^2}{N}},$$

где N – количество диссектных отметок (8)

Величина стандартного отклонения поправленных диссектных отметок $\sigma'_d(j)$ принимается в качестве третьего показателя-критерия j -го ассессора – ДИССИПАЦИОННОГО $q_{\text{дисс}}$.

Поскольку, как уже отмечалось выше, на интегральный показатель качества работы и профессионализма ассессоров оказывают совместное

действие диссипационный $q_{\text{дисс.}}$ и препозиционный показатели $q_{\text{преп.}}$, то допустимо определить:

$$Q_{\text{инт.}} = \sqrt{(q_{\text{дисс.}})^2 + (q_{\text{преп.}})^2} \quad (9)$$

Определение препозиционного показателя-критерия $q_{\text{преп.}}$ при этом производится по формуле:

$$q_{\text{преп.}} = \sqrt{(Q_{\text{инт.}})^2 - (q_{\text{дисс.}})^2} \quad (10)$$

От поправочного коэффициента K следует перейти к $K_{\text{преп.}}$ – коэффициенту препозиции, показателю преуменьшения или преувеличения выставляемых ассессором отметок: $K_{\text{преп.}} = \frac{1}{k}$ (11)

По результатам проведенных вычислений для каждого обучающегося – участника эксперимента произведено вычисление четырех показателей качества и профессионализма работы каждого ассессора, сведенные в таблицу 2.

Таблица 2 Показатели работы ассессора

Наименование показателя	Обозначение показателя	Значение показателя
коэффициент препозиции	$K_{\text{преп.}}$	показатель преуменьшения или преувеличения выставляемых ассессором отметок по сравнению со средними в результате действия внутренней оценочной шкалы
препозиционный показатель	$q_{\text{преп.}}$	показатель устойчивости внутренней оценочной шкалы.
диссипационный показатель	$q_{\text{дисс.}}$	показатель, являющийся мерой организованности\беспорядочности в результатах ассессмента (случайного рассеивания).

интегральный показатель	$Q_{\text{инт.}}$	показатель, отражающий множественное действие системных и случайных факторов, влияющих на отклонения выставляемых отметок
----------------------------	-------------------	---

Для каждой группы обучающихся данного семестра обучения произведен расчет средней величины каждого показателя (Таблица 3).

Таблица 3 Средние величины показателей работы ассессоров по семестрам эксперимента

НОМЕР СЕМЕСТРА	$\bar{K}_{\text{преп.}}$	$\bar{Q}_{\text{инт.}}$	$\bar{q}_{\text{дисс.}}$	$\bar{q}_{\text{преп.}}$
1	0,959	0,75	0,65	0,25
2	0,960	0,73	0,61	0,26
3	0,971	0,67	0,60	0,21
4	0,971	0,70	0,59	0,26
5	0,950	0,67	0,60	0,19
6	0,956	0,70	0,60	0,25
7	0,935	0,70	0,64	0,17
8	0,941	0,68	0,61	0,18
9	0,943	0,64	0,59	0,17
10	0,943	0,68	0,61	0,20

Для оценки динамики показателей работы ассессоров в процессе обучения построен ряд функциональных зависимостей в координатах: номер семестра – показатель (рис. 1-4).

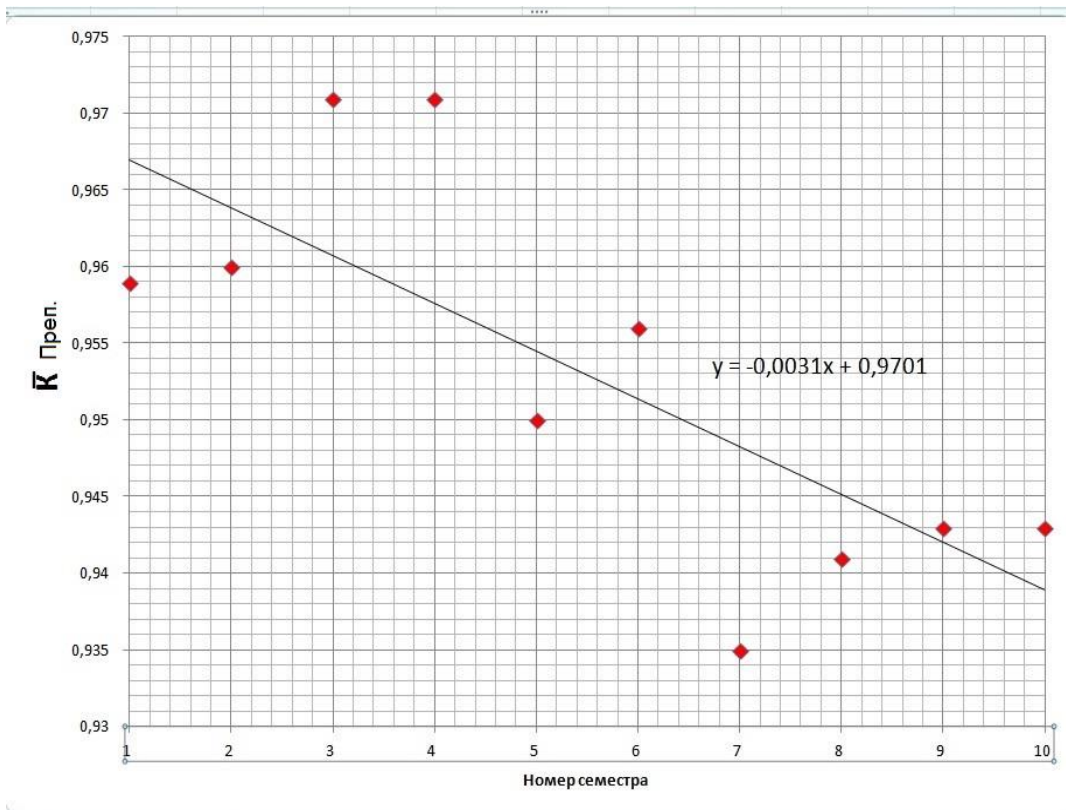


Рис. 1 Зависимость группового препозиционного показателя \bar{K} работы ассессора от продолжительности эксперимента

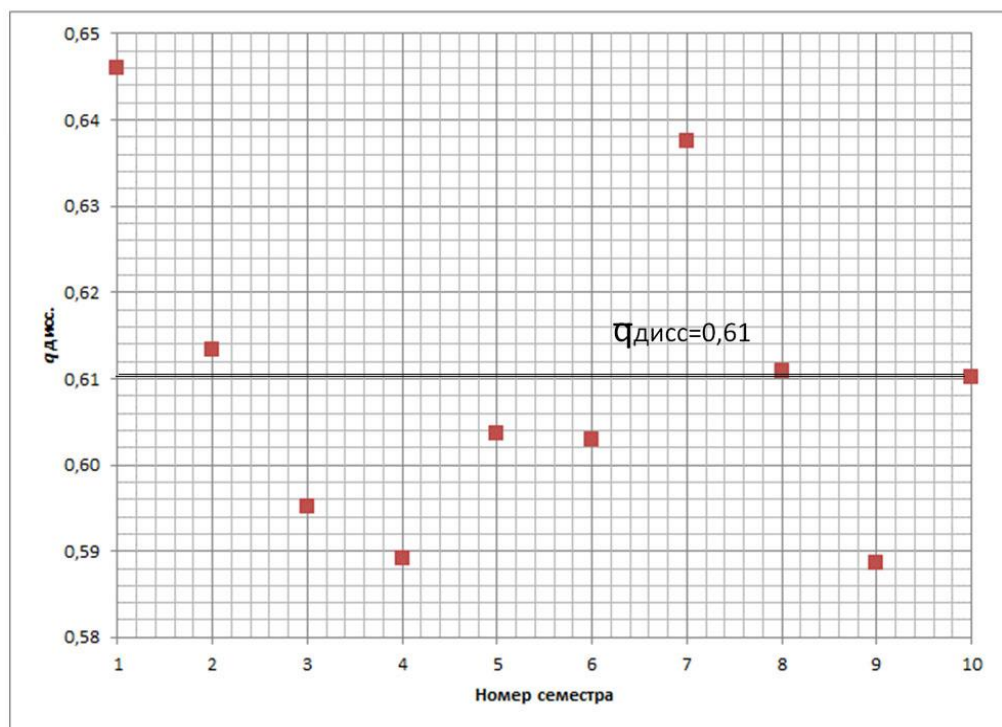


Рис. 2 Зависимость среднего группового диссипационного показателя $\bar{q}_{\text{дисс.}}$ работы ассессора от продолжительности эксперимента

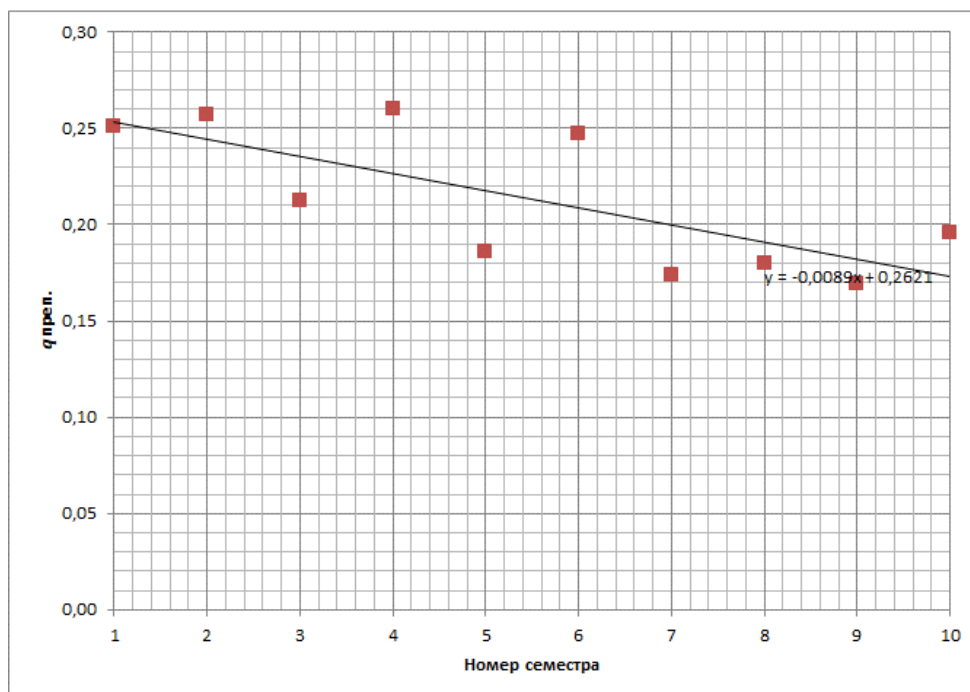


Рис. 3 Зависимость среднего группового препозиционного показателя $\bar{q}_{\text{преп.}}$ работы ассессора от продолжительности эксперимента

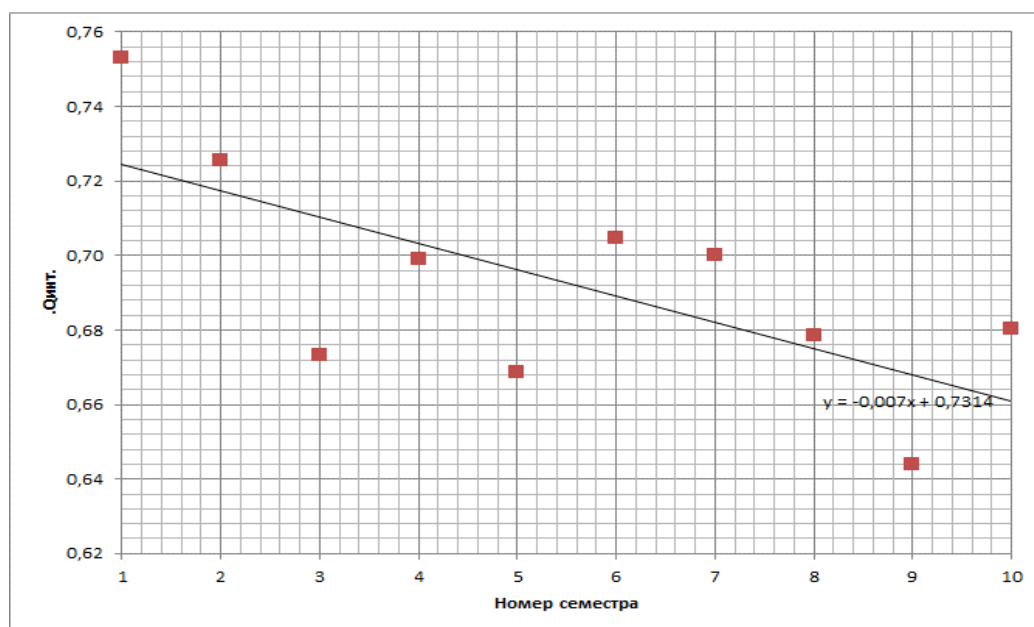


Рис. 4 Зависимость группового интегрального показателя $\bar{Q}_{\text{инт.}}$ работы ассессора от продолжительности эксперимента

Анализ полученных результатов показывает, что только диссипационный показатель работы ассессора не имеет хоть сколько-нибудь выраженного тренда по отношению к продолжительности эксперимента. Средний групповой препозиционный показатель характеризует данные группы как лоустейторов. Таким образом, обучающимся в среднем характерно занижать отметку ассессмента. При этом с ростом номера семестра обучения тенденция усиливается (Рис.1). Тренд показателя диссипации выражен горизонтальной линией (Рис.2). То есть мера случайного рассеивания отклонений отметок не имеет динамики, во времени не изменяется. Но величина разброса препозиционного показателя со временем снижается (Рис.3). В отношении интегрального показателя также наблюдается снижение величины разброса. (Рис.4).

Таким образом, мы можем констатировать положительный эффект ассессмента, студенты вырабатывают более устойчивые оценочные шкалы.

В рамках изучения закономерностей коллегиальных оценок студенческих работ в 2018 году в Современной Гуманитарной Академии был проведён эксперимент по исследованию расхождения отметок, выставяемых студенческим работам преподавателями [Карпенко, 2019. С. 25-30]. В эксперименте добровольно участвовали 38 профессиональных преподавателей высшей школы, все они имели педагогический стаж более 10 лет и учёную степень, 9 из них были докторами наук. Каждый выставял отметки 10-ти работам, так что каждой диссектной отметке противостояла средняя величина 9 отметок пир-ассессоров (тростовая отметка). Получен результат: математическое ожидание стандарта отклонения диссектной отметки от тростовой – 0,68 балла. Таким образом в роли пир-ассессоров студенты, у которых динамика стандарта отклонения выразилась трендом 0,724 балла в начале эксперимента и 0,662 балла – в конце, ни в чём не уступают опытным преподавателям. Более того, к 10-му семестру их оценочные шкалы были хоть и ненамного, но в среднем были качественно лучше, чем у преподавателей.

Можно сделать вывод о том, что по мере освоения образовательной программы, внутренние оценочные шкалы обучающихся становятся все более устойчивыми. Предложенные количественные меры оценки результатов асессмента могут быть с успехом применены при оценке воспитательного воздействия электронных информационных образовательных сред с помощью реализации дидактики асессмента.

Литература.

1. Карпенко М.П. Экспериментальное исследование расхождения отметок, выставляемых студенческим работам преподавателями. // Высшая школа, 2019, №10. С. 25 - 30.
2. Lin Agler, Lin Miao, DeWayne Moore, Karen M. Zabucky. Effects of Personality on Metacognitive Self-Assessments // College Student Journal. 2004. № 38.3. P. 453–461.
3. Mann B.L. Expert-Peer Online Assessment and Formula // The 10th Annual Hawaii International Conference on Education, Honolulu, Hawaii. January. 2012. P. 4–8.
4. Orpen C. Student versus lecturer assessment of learning // Higher Education. 1982. № 11. P. 567–572.
5. Philip M. Sadler, Eddie Good. The Impact of Self and Peer Grading on Student Learning // Educational Assessment. 2006, № 11.1. P. 1–31.
6. Topping K. J. Peer assessment // Theory into Practice, 2009. № 48(1). P. 20–27.

Сведения об авторах:

Михаил Петрович Карпенко,

доктор технических наук, профессор,
президент Частной образовательной организации
«Современная гуманитарная академия».

e-mail: vfokina@campus.muh.ru

Адрес: 109029, г. Москва, ул. Нижегородская, д. 32, стр. 4

Тел. (495) 727-12-95.

Басов Вадим Александрович,

кандидат физико-математических наук,
директор дидакто-технологического управления
Библиотеки информационно-образовательных ресурсов «Умней»,

e-mail: vbasov@ campus.muh.ru

Адрес: 109029, г. Москва, ул. Нижегородская, д. 32, стр. 4

Тел. +7(915) 224-92-95