
УПРАВЛЕНИЕ В ОБЛАСТИ ОБРАЗОВАНИЯ

Библиографический список

1. *Мировая* цифровая библиотека (WDL). – Режим доступа: <http://www.wdl.org/ru>, свободный. – Заглавие с экрана. – Яз. рус.
2. *Дополнительная* довузовская подготовка иностранных граждан. – Режим доступа: <http://www.russia.edu.ru/information/met/970/> № 1, свободный. – Заглавие с экрана. – Яз. рус.
3. *Заиграев, И.* Интернет-технологии для дистанционного образования. – Режим доступа: <http://cis.rudn.ru/document/show.action?document.id=782>, свободный. – Заглавие с экрана. – Яз. рус.
4. *Князев, Е. А.* Анализ российского образовательного рынка для международного сотрудничества / Е. А. Князев, Ю. А. Афанасьев, В. П. Савиных, В. А. Фукин, Н. Н. Машников // Университетское управление: практика и анализ. – 2004. – № 1.
5. *Кривошапова, Т. В.* Болонский процесс в Казахстане и России: параллели и парадоксы. – Режим доступа: <http://www.ia-centr.ru/expert/3956>, свободный. – Заглавие с экрана. – Яз. рус.
6. *Научная* электронная библиотека eLIBRARY.ru. – Режим доступа: <http://elibrary.ru/defaultx.asp>, свободный. – Заглавие с экрана. – Яз. рус.
7. *Сайт* сетевого университета СНГ. – Режим доступа: http://imp.rudn.ru/su_sng, свободный. – Заглавие с экрана. – Яз. рус.

УДК 51-7

МОДЕЛЬ КОМПЛЕКСНОЙ ОЦЕНКИ ЧИТАЕЛЬНОСТИ ТЕСТОВЫХ МАТЕРИАЛОВ

С.В. Окладникова

Разработана модель комплексной оценки читабельности тестовых вопросов, основанная на автоматизированном расчете количественных характеристик содержания тестовых вопросов. С учетом конструктивных особенностей тестовых заданий сформирована номенклатура единичных показателей, рассчитаны нормированные значения и весовые коэффициенты, построена шкала диапазонов для определения уровня читабельности, доказана адекватность разработанной модели.

Ключевые слова: электронное тестирование, читабельность, комплексный показатель, оценка качества, тестовые материалы, тестовые вопросы.

Key words: electronic testing, readability, composite index, quality assessment, test materials, test questions.

В настоящее время в вузах электронное тестирование как метод оценки уровня знаний учащихся приобретает все большую популярность. Это обусловлено изменением системы высшего образования в рамках реализации основных принципов Болонского процесса, в частности, с появлением новых подходов к оценке качества образовательного процесса в вузе.

В странах, использующих высокоэффективные технологии и методики диагностики качества образования, разработкой, проверкой и тиражированием контрольных тестовых материалов занимаются специальные организации, например, в США – Educational Testing Service ETS (www.ets.org), в Великобритании – The Qualifications and Curriculum Authority QCA (<http://www.qca.org.uk>), в России – Федеральный институт педагогических измерений (<http://www.fipi.ru>).

Несмотря на развитие государственной системы тестирования, в России на сегодняшний день большую часть используемых в учебном процессе тестовых материалов (ТМ) составляют материалы, разработанные преподавателями учебных заведений в рамках преподаваемых ими дисциплин. Популярность их использования обусловлена, в первую очередь,

адаптивностью по отношению к преподаваемому материалу и ориентацией на педагогические потребности небольшого количества учащихся и т.п. Содержание ТМ формируется авторами вручную и целиком зависит от их индивидуальных особенностей, например, стиля изложения, уровня компьютерной грамотности и т.п.

Отсутствие у многих преподавателей опыта по разработке ТМ и соответствующих знаний в области тестологии приводит к нарушению основных требований, предъявляемых к структуре, содержанию и оформлению тестовых заданий (ТЗ). Результаты исследований, проведенных в области качества ТМ, показали, что в среднем отсеивается 47 % разработанных ТЗ [5]. Исследование, проведенное в Астраханском государственном университете, позволило выявить основные причины, снижающие качество ТМ и разбить их на группы:

- нарушение методических требований, предъявляемых к содержанию ТЗ – 37 %;
- нарушение технологических требований, предъявляемых к форме ТЗ – 32 %;
- нарушение технических требований оформления ТМ – 16 %;
- нарушение принципов отбора содержания – 9 %;
- наличие грамматических ошибок – 6%.

Анализ графика показывает, что 48 % ошибок, допускаемых авторами, вызвано нарушением технологических требований, в частности избыточностью слов в формулировках ТЗ, которая в процессе тестирования значительно снижает уровень понимания у испытуемых поставленного в задании вопроса, и нарушением технических правил оформления ТЗ, существующих в эксплуатируемой автоматизированной системе тестирования (АСТ). Трудности, связанные с применением существующих методов оценки качества ТМ в вузе, обусловлены рядом факторов:

- необходимостью наличия большого объема статистической выборки результатов тестирования, достаточного для получения достоверных данных анализа, что для ряда дисциплин в силу небольшого количества часов, отведенных на их изучение, или в силу малочисленности групп учащихся возможно только в течение нескольких семестров;
- длительный период накопления статистических данных может привести к потере актуальности ТМ, так как для некоторых дисциплин (например, юридического, экономического направлений, информационных и телекоммуникационных технологий) характерно быстрое изменение содержания предметной области;
- обработка статистических данных предполагает эмпирическую проверку свойств ТМ на основе сложного математического аппарата, и ее выполнение крайне затруднительно без привлечения специальных программных продуктов;
- в силу специфики представления результатов анализа для их правильной интерпретации преподавателю необходимы дополнительные знания в области тестологии и знание самих методов статистической обработки.

В результате возникает потребность в разработке методов автоматизированной оценки качества ТЗ до начал их практического использования, которые позволили бы повысить как качество самих тестовых заданий, так и эффективность процесса создания ТМ. Одним из возможных способов решения поставленной задачи является оценка читабельности содержания ТМ, основанная на анализе количественных показателей формулировок ТЗ, характеризующих стиль изложения материала и степень его понимания целевой аудиторией.

Понятие качества тестовых материалов. Качество ТМ – многогранно. Для его описания используют сложную систему комплексных и единичных показателей (рис. 1). Основными являются валидность и надежность [1, 2]. В качестве критериев валидности рассматривают формальные, оценивающие технологичность представления содержания тестовых вопросов, и содержательные, оценивающие правильность отображения в них учебного материала. Надежность является комплексной характеристикой, главным образом зависящей от метрических свойств отдельных ТЗ, значения которых рассчитываются путем статисти-

ческой обработки сырых баллов, полученных по результатам предварительного тестирования испытуемых.



Рис. 1. Структура показателей качества тестовых материалов

Анализ существующих АСТ показал, что подсистемы анализа качества ТМ реализованы лишь у 30 % [4]. Большую часть из них составляют подсистемы, основанные на статистической обработке результатов тестирования испытуемых, при этом отсутствует простой и удобный для пользователя интерфейс, позволяющий быстро интерпретировать полученные результаты анализа. Распространенные на российском рынке программы оценки качества ТМ (RUMM 2020, WINSTEPS и др.) ориентированы на пользователей, имеющих навыки вычислительной работы на компьютере и знания в области статистического анализа. В силу своей высокой стоимости эти программы используются в специализированных центрах тестирования.

Существующие методы оценки читабельности текстов. Основным компонентом содержания ТЗ, как правило, является текст. В мировой практике в полиграфии, программировании, лингвистике и Public Relation при оценке качества текста используется понятие *readability* (в переводе с английского – *удобочитаемость, читабельность*). При этом под удобочитаемостью рассматривают свойство текста при чтении не вызывать повышенного утомления. В качестве факторов, влияющих на удобочитаемость текста, выделяют тип шрифта, его размер, цвет и т.п. Под читабельностью понимают возможность свободного прочтения текста человеком без чрезмерного напряжения, т.е. используются ли в тексте сложные термины, существует ли перегруженность причастными и деепричастными оборотами и т.п.

Существует целый ряд методов (табл. 1), с помощью которых можно оценить сложность написанного текста, где основными критериями выступают общее количество слов, средняя длина предложений и слов. Эти параметры легко поддаются количественному выражению и пригодны для автоматической оценки.

Рассмотренные модели ориентированы на оценку читабельности англоязычных текстов. Основные параметры русских и английских текстов не совпадают. Средняя длина слова английского словаря составляет 2,77 слога, а русского 3,03. Предложение, написанное на английском языке, в среднем в 1,43 раза больше, чем на русском. С учетом этого в исследовании, проведенном И.В. Оборневой [3], формула Флэша была скорректирована для оценки читабельности русскоязычных учебных текстов. Однако существует ряд факторов, которые делают невозможным применение данной модели для оценки читабельности ТМ. Во-первых, при расчете коэффициента читабельности объем анализируемой выборки текста должен быть не менее 100 слов, тогда как средняя длина формулировки отдельного ТЗ в среднем в 4 раза меньше. Во-вторых, структура ТЗ конструктивно отличается от структуры обычного текста наличием инструкции, вопроса и набора ответов.

Таблица 1

Существующие методы оценки читабельности текстов

Название метода	Используемые параметры текста	Математическая модель	Практическая применимость
Индекс туманности Ганнинга	Общее количество слов в тексте k ; количество предложений в тексте s ; средняя длина предложения w ; среднее количество многосложных слов (более трех слогов) l	$F_{Gunning} = \sum_{i=1}^s \frac{w_i}{s} + 0,4 \sum_{i=1}^s \frac{l_i}{w_i} = w + 0,4l$	Определение возраста целевой аудитории
Формула Флэша (тест FRES)	Общее количество слов в тексте k ; количество предложений в тексте s ; общее количество слогов в тексте f ; средняя длина предложения w ; средняя длина слова p (в слогах)	$F_{Flesch} = 206,835 - \left(\frac{k}{s} \cdot 1,015 + \frac{f}{k} \cdot 84,6\right) = 206,835 - (w \cdot 1,015 + p \cdot 84,6).$	Оценка читабельности текстов страховых договоров в США
Формула Флэша – Кинкэйда	Средняя длина предложения w ; средняя длина слова p	$F_{Flesch-Kincaid_grade} = w \cdot 0,39 + p \cdot 11,8 - 15,59.$ $F_{Flesch-Kincaid_age} = w \cdot 0,39 + p \cdot 11,8 - 10,59$	Определение уровня образования целевой аудитории в библиотеках при выборе книг и учебников
Индекс Колемана – Лиану	Общее количество символов в тексте x ; общее количество слов в тексте k ; количество предложений в тексте s ; средняя длина предложения w ; средняя длина слова P (в символах)	$F_{Coleman-Liau} = 5,89 \cdot \frac{x}{k} + 30 \cdot \frac{s}{k} - 15,8$	Оценка читабельности текстов
Формула Пауэрса – Самнера – Кеарла	Общее количество слогов в тексте f ; общее количество слов в тексте k ; количество предложений в тексте s ; средняя длина предложения w ; средняя длина слова p (в слогах)	$F_{Powers-Sumner-Kearl_grade} = w \cdot 0,0778 + f \cdot 0,0455 - 2,2029.$ $F_{Powers-Sumner-Kearl_age} = w \cdot 0,0778 + f \cdot 0,0455 - 2,7971.$	Оценка читабельности текстов, предназначенных для детей
Формула Маклаулина «SMOG»	Объем текста – 30 предложений, общее количество слов в тексте k ; количество многосложных слов в тексте L ; среднее количество многосложных слов в тексте l	$F_{SMOG_grade} = \sqrt{\frac{L}{s}} \sqrt{l} + 3$ $F_{SMOG_age} = \sqrt{\frac{L}{s}} \sqrt{l} + 8$	Прогнозирование 100 % понимания текста
Формула FORCAST	Общее количество слов в тексте k ; количество односложных слов в тексте b	$F_{FORCAST_grade} = 20 - k \cdot 0,0667 \cdot b$ $F_{FORCAST_age} = 25 - k \cdot 0,0667 \cdot b$	Оценка понятности технической документации в армии США, ориентирована только на взрослых читателей

Таким образом, применение существующих моделей при оценке читабельности ТМ возможно только в случае их модификации с учетом конструктивных отличий ТЗ от текста, наполнения содержания ТЗ различными компонентами (формулы, графики и т.п.) и ориентацией на их использование русскоговорящей аудиторией.

Разработка модели оценки читабельности тестовых материалов. В результате опроса ведущих преподавателей Астраханского государственного университета и специалистов регионального ресурсного Центра дистанционного обучения методом комиссии была сформирована номенклатура, состоящая из 10 единичных показателей качества (ЕПК): C_1 – общее количество вариантов ответов, C_2 – количество дистракторов, C_3 – общее количество слов, C_4 – средняя длина слова, C_5 – количество предлогов, союзов, частиц, C_6 – количество знаков препинания, C_7 – количество иностранных символов (англ., греч. и т.п.), C_8 – количество математических символов, C_9 – количество цифр, C_{10} – количество формул [6].

Трудность в определении нормированных значений показателей $(C_j)_N$ вызвана отсутствием в настоящее время их стандартизованного или нормативного перечня. Наиболее приемлемым представляется подход, при котором значения нормирующих показателей $(C_j)_N$ определяются на основе вероятностного метода, предполагающего проведение статистической оценки выборочных значений C_j с учетом генеральной совокупности. Для решения задачи нормирования ЕПК в качестве опытных образцов были взяты тесты по различным дисциплинам, прошедшие статистический анализ качества и имеющие значение основного показателя качества – надежность, равное $\rho > 0,7$, что является допустимым для экзаменационных тестов. Средний объем накопленных статистических данных составил 310 человек/тестирований. При расчете показателя надежности (ρ) в качестве метода оценки надежности теста использовался метод расщепления, коэффициент надежности определялся по формуле Спирмена – Брауна [9].

В ходе предварительного анализа опытные тесты (T_s) были разделены на 2 группы. В первую группу вошли тесты, составленные по гуманитарным дисциплинам, в которых значения показателей C_8 и C_{10} равны 0. Во вторую группу вошли тесты, составленные по естественнонаучным и инженерно-техническим дисциплинам, в которых значения показателей C_8 и C_{10} не равны 0. Значения нормированных ЕПК были рассчитаны для каждой группы опытных тестов с погрешностью, не превышающей 3,68 %.

Определение весовых коэффициентов выполнялось на основе экспертного метода непосредственной численной оценки [8]. Цель экспертизы – численно определить, какой из ЕПК C_j при увеличении своего значения усложняет восприятие формулировок ТМ и с какой степенью. Для проведения экспертизы была предложена абсолютная шкала определения степени влияния ЕПК C_j на читабельность ТМ, позволяющая однозначно провести оценку объектов в баллах: 0 – не усложняет; 1 – усложняет очень слабо; 2 – усложняет слабо; 3 – усложняет средне; 4 – усложняет сильно; 5 – усложняет очень сильно. По результатам экспертизы были построены две матрицы: $K_1 = \{K_{ij} \mid i = 1 \dots 12, j = 1 \dots 8\}$ и $K_2 = \{K_{ij} \mid i = 1 \dots 12, j = 1 \dots 10\}$, в которых элементы K_{ij} принимали значения $0 < K_{ij} < 5$. Обработка мнений экспертов позволила выделить существенно значимые показатели ($C_1, C_2, C_3, C_4, C_7, C_{10}$), для которых выполняется условие: $K_j \geq 1/m$. С учетом рассчитанных нормированных значений ЕПК были сформированы новые наборы: $Q_1 = \{C_1, C_2, C_3, C_4, C_7\}$, $Q_2 = \{C_1, C_2, C_3, C_4\}$, $Q_3 = \{C_1, C_2, C_3, C_4, C_7, C_{10}\}$ и $Q_4 = \{C_1, C_2, C_3, C_4, C_{10}\}$, значения весовых коэффициентов для которых представлены в табл. 2. Согласованность мнений экспертов оценивалась на основе коэффициентов конкордации: $W_1 = 0,8$, $W_2 = 0,7$.

**Нормированные значения единичных показателей качества
тестовых материалов и значения их весовых коэффициентов**

Наборы ЕПК	Значения $(C_j)_N$ и (K_j) для ЕПК C_j									
	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}
$(C_j)_N$	5,7	3,99	24,18	3,28	2,5	2,51	3,68	—	5,54	—
Q_1	0,261	0,148	0,264	0,186	—	—	0,141	—	—	—
Q_2	0,304	0,172	0,307	0,217	—	—	—	—	—	—
$(C_j)_N$	6,59	4,86	27,5	3,51	3,13	2,49	3,64	8,24	9,1	3,4
Q_3	0,187	0,175	0,180	0,137	—	—	0,126	—	—	0,195
Q_4	0,214	0,200	0,206	0,157	—	—	—	—	—	0,223

Сформированная номенклатура характеризуется большим числом разнородных ЕПК, что обуславливает определенные трудности в нахождении их функциональной зависимости. Увеличение значения каждого из показателей C_j ухудшает качество формулировок ТЗ, т.е. делая их менее читабельными. Поэтому уравнение модели комплексной оценки читабельности ТМ было разработано на основе принципа среднего геометрического, а значение относительного единичного показателя рассчитывалось по формуле:

$$q_j = \frac{(C_j)_N}{C_j}$$

Комплексный показатель качества признается состоятельным, если его значения монотонно возрастают (или убывают) при улучшении качества продукции. Если в ТЗ значения показателей $C_j = (C_j)_N$, то значение комплексного показателя $Q = 1$, что соответствует его «идеальному» значению. Увеличение натуральных значений C_j повышает читабельность ТЗ, т.е. его содержание становится более трудным для восприятия, при этом значение комплексного показателя $Q < 1$. Уменьшение натуральных значений C_j наоборот делает содержание ТЗ более легким, а значение комплексного показателя $Q > 1$. С учетом этого для определения уровня читабельности ТЗ были определены следующие границы диапазонов комплексного показателя.

1. Читабельность ТМ соответствует установленному показателю качества и считается нормальной при выполнении условия:

$$Q = \prod_{j=1}^m q_j (C_j)^{K_j}, \text{ если } \forall C_j \quad (C_j)_N - \Delta_{C_j} \leq C_j \leq (C_j)_N + \Delta_{C_j} \quad (1)$$

2. При ограничении уровня читабельности сверху шкала градации определяется при выполнении условия:

$$Q_{\max} = \prod_{j=1}^m q_j (C_j)^{K_j}, \text{ если } \forall C_j \quad ((C_j)_N)_{\min} \leq C_j \leq (C_j)_N - \Delta_{C_j} \quad (2)$$

Уровень читабельности устанавливается легкий.

3. При ограничении уровня читабельности снизу шкала градации определяется при выполнении условия:

$$Q_{\min} = \prod_{j=1}^m q_j (C_j)^{K_j}, \text{ если } \forall C_j \quad (C_j)_N + \Delta_{C_j} \leq C_j \leq ((C_j)_N)_{\max} \quad (3)$$

Уровень читабельности устанавливается сложный.

4. Читабельность ТМ признается не соответствующей установленному показателю качества при условии:

$$Q = 0, \text{ если } \begin{cases} \forall C_j & C_j < ((C_j)_N)_{\min} \\ \forall C_j & C_j > ((C_j)_N)_{\max} \end{cases} \quad (4)$$

УПРАВЛЕНИЕ В ОБЛАСТИ ОБРАЗОВАНИЯ

В формулах (1) – (4) Δ_{c_j} – допустимое отклонение единичного показателя от нормированного значения. Рассчитанные на основе (1)–(4) значения границ диапазонов для каждого набора показателей представлены в табл. 3.

Таблица 3

Границы диапазонов комплексного показателя читабельности

Наборы ЕПК				Градация оценки
Q_1	Q_2	Q_3	Q_4	
[0,6 : 0,96]	[0,62 : 0,96]	[0,65 : 0,97]	[0,65 : 0,97]	Сложный
[0,95 : 1,06]	[0,95 : 1,06]	[0,96 : 1,08]	[0,96 : 1,08]	Нормальный
[1,07 : 2,62]	[1,07 : 2,47]	[1,09 : 3,73]	[1,09 : 3,33]	Легкий

Достоверность всех полученных в ходе исследования статистических данных была подтверждена соответствующими критериями.

Оценка адекватности предложенной модели. В качестве основных требований, предъявляемых к математическим моделям, являются требования точности и адекватности. Анализ точности модели комплексной оценки читабельности (1)–(4) проведен путем сравнения результатов, полученных в ходе моделирования, и оценок экспертов, которые были приняты в качестве истинных. При планировании эксперимента в качестве опытных образцов были использованы ТЗ различного уровня трудности по различным дисциплинам. В роли экспертов выступили преподаватели, которым было предложено определить читабельность ТЗ по категориям: 1 – легкое, 2 – нормальное, 3 – трудное. Обработка результатов проведения эксперимента заключалась в подсчете количества ТЗ, отнесенных экспертами и моделью в разные группы уровней читабельности ТМ. По результатам проведенных исследований были сделаны выводы:

- 1) относительная максимальная погрешность изменяется в диапазоне $2,86 < E_m < 7,81$ %;
- 2) наибольшей точностью оценки читабельности ТМ обладает модель Q_2 , у которой максимальная относительная погрешность составляет $E_m = 4,55$ %, а наименьшей – модель Q_4 с погрешностью $E_m = 7,81$ %;
- 3) разработанные модели оценивают ТМ, имеющие уровень читабельности:
 - легкий – с погрешностью $2,86 < E_m < 6,25$ %;
 - нормальный – с погрешностью $4,23 < E_m < 7,81$ %;
 - трудный – с погрешностью $4,55 < E_m < 7,69$ %.

Оценка адекватности выполнялась на основе критериев Стьюдента (t) и Фишера (F), табличные значения которых были определены при доверительной вероятности $P = 0,95$. В результате анализа было установлено, что фактические значения по обоим критериям удовлетворяют условиям достоверности ($F < F_{st}$, $t < t_{st}$), следовательно, гипотеза о случайном расхождении средних значений и дисперсий в исследуемых выборках подтверждается, и полученные модели адекватны.

Практическое использование модели оценки читабельности тестовых материалов. На основе предложенной модели была разработана автоматизированная система тестирования АСТ-ДО [7]. Анализ эффективности от внедрения АСТ-ДО выполнялся в отношении числа отбракованных ТЗ и ошибок, допускаемых авторами при подготовке ТМ (рис. 2, 3).

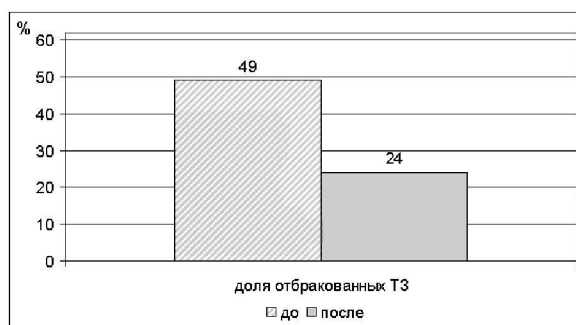


Рис. 2. Анализ эффективности процесса разработки и оценки качества ТМ с использованием АСТ-ДО в отношении числа отбракованных ТЗ

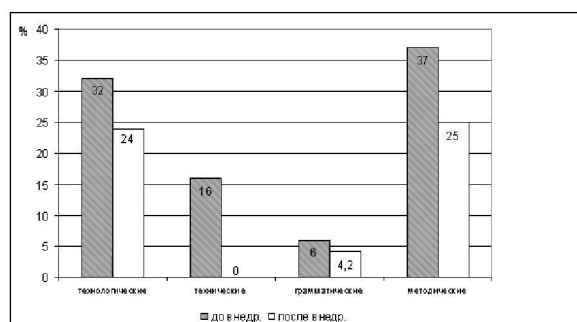


Рис. 3. Анализ эффективности процесса разработки и оценки качества ТМ с использованием АСТ-ДО в отношении ошибок, допускаемых авторами

Анализ эффективности в отношении числа отбракованных ТЗ показал, что в среднем их число сократилось в 2 раза. Анализ эффективности в отношении числа ошибок, допускаемых авторами, показал, что использование АСТ-ДО при подготовке ТМ позволило сократить в среднем число технологических ошибок в 1,33 раза, грамматических – в 1,5 раза, методических – в 1,48 раза и исключить технические ошибки.

В целом, предложенные модели комплексной оценки читабельности, реализованные в АСТ-ДО, позволяют повысить эффективность процесса разработки и оценки качества ТМ до начала их практического использования в результате обеспечения автора автоматизированной поддержкой в отношении контроля качества содержания ТМ в процессе их разработки.

Библиографический список

1. Анастаси, А. Психологическое тестирование / Анна Анастаси, Сюзен Урбина ; пер. с англ. под ред. А. А. Алексеева. – Изд. 7-е, перераб. и доп. – СПб. : Питер, 2007. – 102 с.
2. Гласс, Дж. Статистические методы в педагогике и психологии / Дж. Гласс, Дж. Стэнли. – М. : Прогресс, 1976. – 496 с.
3. Оборнева, И. В. Математическая модель оценки учебных текстов / И. В. Оборнева // Вестник МГПУ. – М. : МГПУ, 2005. – № 1 (4). – С. 141–147. – (Сер. Информатика и информатизация образования).
4. Окладникова, С. В. Автоматизированная система тестирования, расширенная подсистемой оценки качества объектов тестового материала до начала их практического использования / С. В. Окладникова // Образовательные технологии. – Воронеж : ВГПУ, 2007. – № 2. – С. 78–81.
5. Окладникова, С. В. Анализ ошибок формирования содержательной части тестовых заданий / С. В. Окладникова // Проблемы информатизации образования: региональный аспект : мат-лы Всерос. НПК (27–29 апреля 2006 г.). – Чебоксары : Изд-во Л. А. Наумова, 2006. – С. 243–246.
6. Окладникова, С. В. Количественная оценка содержания тестового задания для компьютерного тестирования / С. В. Окладникова // Изв. ВолгГТУ. – Волгоград : Изд-во ВолгГТУ, 2007. – № 9 (35), вып. 3. – С. 88–90. – (Сер. Актуальные проблемы управления, вычислительной техники и информатики в технических системах).
7. Окладникова, С. В. Свидетельство № 2007611447. «Автоматизированная система тестирования для дистанционного обучения на основе алгоритмов адаптивного управления АСТ_ДО» / С. В. Окладникова, М. В. Таюшев. – Заявл. 12.02.07 ; зарег. 06.04.07. – М. : Федер. ин-т промышленной собственности Федеральной службы по интеллектуальной собственности (Роспатент), 2007.
8. Окладникова, С. В. Формирование системы весовых коэффициентов интегрального критерия, оценивающего фактическую сложность формулировок тестовых заданий / С. В. Окладникова // Вестник компьютерных и информационных технологий. – М. : Машиностроение, 2009. – № 6. – С. 46–50.
9. Челышкова, М. Б. Теория и практика конструирования педагогических тестов : учеб. пос. для вузов / М. Б. Челышкова. – М. : Логос-М, 2002. – 432 с.