

Афони́на М. В., кандидат педагогических наук, доцент кафедры теоретических основ информатики
Гонштейн М. А., студент института физико-математического образования
Алтайский государственный педагогический университет
г. Барнаул

АНАЛИЗ СОГЛАСОВАННОСТИ РАБОТЫ ЭКСПЕРТОВ ПРИ ПРОВЕРКЕ ЗАДАНИЙ С РАЗВЕРНУТЫМ ОТВЕТОМ ЕГЭ ПО ИНФОРМАТИКЕ И ИКТ

Критериально-ориентированный подход – способ, позволяющий оценить уровень усвоения учебного материала. Критериальное оценивание может быть использовано педагогами на в разных формах контроля достижений учащихся (рубежном, текущем, итоговом и пр.). В государственной итоговой аттестации учащихся по образовательным программам среднего общего образования в форме государственного экзамена (ЕГЭ) этот подход позволяет с помощью единых контрольно-измерительных материалов (КИМ), требований к выпускникам и критериев оценки выполнения заданий проверить и сопоставить их знания и подготовленность по предмету Информатика и ИКТ.

Такая форма проведения выпускного экзамена выпускника является одной из основных среди прочих. Кроме того, анализ результатов ЕГЭ позволяет сделать выводы о качестве работы учителей и педагогических коллективов в целом. По итогам ЕГЭ выполняется также анализ работы предметных комиссий, так как сложности и вопросы, возникшие на этапе проверки результатов экзамена, рассматриваются как основание для совершенствования КИМ и критериев оценивания заданий с развернутым ответом. Экзаменационная работа по информатике состоит из двух частей. Часть 1 содержит 23 задания с кратким ответом по всем основным разделам курса информатики. Она проверяется автоматически, без привлечения экспертов.

Проверку ответов на задания части 2 осуществляют эксперты, в соответствии с методическими рекомендациями по оцениванию заданий с развернутыми ответами и критериями оценивания, подготовленными федеральным институтом педагогических измерений (ФИПИ) [1, с. 22]. При этом критериальное оценивание результатов ЕГЭ используется как для тестовой части КИМ, так и для оценки заданий с развернутым ответом.

Каждую работу проверяют независимо друг от друга два эксперта. Если проверяя работу, эксперты оценили одно и то же задание с разницей в один балл, то в зачет компьютер выберет большее значение из двух, поставленных экспертами. Например, если один эксперт поставил 1 балл, а второй эксперт – 0 баллов, то в протокол будет поставлен 1 балл. Но если разница будет в 2 балла и более, то проверку этого задания компьютер назначит третьему эксперту, и его оценка уже будет окончательной. Третий эксперт назначается из числа членов предметной комиссии, имеющих статус «ведущий эксперт» или «старший эксперт», ранее не проверявших данную экзаменационную работу.

На бланке-копии работы, назначенной на третью проверку, в регистрационной части для сведения третьего эксперта указываются баллы, выставленные двумя экспертами при первой и второй проверке. Третий эксперт проводит оценивание по тем критериям, которые указаны в бланке-протоколе. Позиции оценивания, которые третий эксперт не проверяет, в протоколе заполнены символами "X" – это означает, что в этой части работ существенных разногласий у первого и второго эксперта не возникло. Проверка развернутых ответов участников ЕГЭ считается завершенной, когда все работы будут проверены экспертами необходимое количество раз, а результаты оценивания проанализированы программным обеспечением, предоставленным региональному центру обработки информации (РЦОИ) федеральным центром тестирования [1, с. 31].

Все эксперты оценивают работы в соответствии с едиными критериями. Экспертам выдают пакеты с копиями работ для проверки, с содержанием верного ответа и указаниями по оцениванию.

Пример задания 25 [3, с.16]:

Дан целочисленный массив из 30 элементов. Элементы массива могут принимать значения от 150 до 200 – рост учащихся выпускного класса. В команду по автогонкам входят все учащиеся, чей рост не более 175 см. Гарантируется, что такие учащиеся в классе есть. Опишите на русском языке или на одном из языков программирования алгоритм, который находит и выводит рост самого высокого участника гоночной команды.

Кроме формулировки задания приводится описание исходных данных на языках: Паскаль, Бейсик, СИ, Алгоритмический язык, Русский (естественный) язык, а также содержание верного ответа. Например, на языке паскаль содержание верного ответа выглядит так:

```
max:=150;
for i:=1 to N do
if (a[i]<=175) and (a[i]>max) then
  max:=a[i];
writeln(max);
```

Критерии содержат показатели выполненности задания, соответствие их конкретным баллам и комментарии к допустимым отклонениям, синтаксическим и алгоритмическим ошибкам.

Например, к заданию 25 даны указания для оценивания на 2, 1 и 0 возможных баллов:

На 2 балла задание может оцениваться, если учащимся «предложен правильный алгоритм, выдающий верное значение. При этом допускается запись алгоритма на другом языке, использующая аналогичные переменные. В случае, если язык программирования использует типизированные переменные, описания переменных должны быть аналогичны описаниям переменных на естественном языке. Использование не типизированных или необъявленных переменных возможно только в случае, если это допускается языком программирования, при этом количество переменных и их идентификаторы должны соответствовать условию задачи. В алгоритме, записанном на языке программирования, допускается наличие отдельных синтаксических ошибок, не искажающих замысла автора программы» [Там же].

Если не выполняются условия, перечисленные на 2 балла, то на 1 балл задание можно считать выполненным, когда предложено в целом верное решение, содержащее не более одной ошибки из числа следующих (если одинаковая ошибка повторяется несколько раз, она считается за одну ошибку):

1. Не инициализируется или неверно инициализируется переменная MAX.
2. В сравнении вместо знака «меньше или равно» используется знак «меньше».
3. Отсутствует вывод ответа.
4. Используется переменная, не объявленная в разделе описания переменных.
5. Не указано или неверно указано условие завершения цикла.
6. Индексная переменная в цикле не меняется (например, в цикле while).
7. Неверно расставлены операторные скобки.
8. В цикле происходит выход за границу массива.
9. В сложном логическом условии простые проверки верны, но условие в целом построено неверно (например, перепутаны операции «И» и «ИЛИ», неверно расставлены скобки в логическом выражении).

Если ошибок, перечисленных в п. 1–9, две или больше, или алгоритм сформулирован неверно, то эксперты выставляют 0 баллов [3,с.18] .

Несмотря на разработанные критерии оценивания, работа экспертов не всегда согласована. Причинами расхождений баллов являются:

1. Технические ошибки экспертов (выставление баллов не в ту строку протокола, оценивание по критериям другого задания и пр.).
2. Ошибки в предметной области (в заданиях допускается использование любого языка программирования и любые методы разработки алгоритма: от классического Бейсика с метками и оператором GOTO до C++, PHP и Perl; к эксперту может попасть вариант работы с применением библиотек нечасто используемых, что может вызвать сомнения эксперта по работе применяемых процедур).
3. Ошибки при интерпретации и применении критериев (например, ошибка 7 – «неверно расставленные операторные скобки»: если выпускник забыл написать последнюю операторную скобку «end.» – это одна ситуация, при исправлении которой принципиально ничего в алгоритме не изменится, другое дело, когда неправильно выделено тело цикла. Два случая не идентичны, и эксперты по разному подходят к применению критериев в таких ситуациях: один может допустить 1 балл, а другой – 0 при неверно организованном цикле. В таком случае обязательно необходимо согласование подходов в оценивании работ.).

Уменьшить рассогласованность помогут следующие меры:

1. Создание комфортных условий работы комиссии.
2. Разработка методических материалов.
3. Предоставление во время работы экспертам материалов с указаниями оценивания, бланков, схем.
4. Обучение экспертов.
5. Согласование критериев оценивания на разных уровнях

Важна аналитическая работа председателя комиссии (ПК) на протяжении периода работы ПК. Председатель и (или) его заместитель фиксирует у себя номер работ, вызвавшие наибольшие разногласия и вопросы экспертов при оценивании. Эти работы анализируются председателем ПК и передаются в ФИПИ. Далее такие работы используются для проведения обучения экспертов ПК следующих лет.

В процессе работы ПК председатель ПК запрашивает у руководителя РЦОИ статистическую информацию о ходе проверки развернутых ответов. При этом председателю ПК предоставляется информация о количестве проверенных работ один, два или три раза, о количестве работ, ожидающих первую, вторую или третью проверку. Кроме того, председатель ПК получает информацию о работе каждого эксперта ПК – количество проверенных работ, с кем в паре они проверялись, и сколько работ ушло на третью проверку. Результаты статистических отчетов председатель ПК использует для оптимизации работы ПК – приостановлении работы и проведении оперативного согласования, ограничении пакетов работ, назначаемых отдельным экспертам и пр.

Анализ работы экспертов может выполняться по следующим показателям:

- 1) количество проверенных работ (N);
- 2) количество или процент работ, отправленных на третью проверку (M) в сравнении с N;
- 3) количество или процент работ, получивших подтверждение оценки при третьей проверке (M+) в сравнении с N или с M;
- 4) количество или процент работ, не получивших подтверждение оценки при третьей проверке (M-) в сравнении с N или с M;
- 5) суммарное количество заниженных баллов по сравнению с 3-й проверкой (S<);
- 6) суммарное количество завышенных баллов по сравнению с 3-й проверкой (S>);
- 7) суммарное отклонение баллов от результатов третьей проверки (S=S<+ S>);
- 8) отношение суммарного отклонения баллов к количеству проверенных работ (T=S/N);
- 9) количество оцененных заданий в проверенных работах (K);

- 10) количество или процент заданий, совпавших с оценкой второго эксперта (K+);
- 11) количество заданий или процент, не совпавших с оценкой второго эксперта (K-);
- 12) накопительная сумма абсолютных разниц баллов со вторым экспертом (Z);
- 13) отношение накопительной суммы абсолютных разниц баллов со вторым экспертом к количеству проверенных работ или заданий ($P=Z/N$ или $P=Z/K$).

Количество проверенных работ показывает, насколько интенсивно эксперт выполняет проверку, но не отображает качество его работы. Чем больше проверенных работ у эксперта, тем больше вероятность появления разногласий, следовательно, растет и суммарное отклонение баллов. То же самое можно сказать и о девятом показателе. Второй показатель «количество или процент работ, отправленных на третью проверку» также не точно демонстрирует качество работы эксперта так как, одну работу проверяют два человека и это не значит, что они оба оценили неверно. Только после третьей проверки можно сказать, кто из экспертов был неправ, для чего и применяется третий показатель.

Четвертый показатель «количество или процент работ, не получивших подтверждение оценки при третьей проверке» демонстрирует, насколько эксперт отклоняется от критериев оценивания, то есть, чем больше этот показатель, тем ниже качество работы эксперта.

Пятый и шестой показатели позволяют судить о том, занижает или завышает, как правило, оценку эксперт при проверке. Данный критерий помогает в будущем продумать систему коррекции работы эксперта и повысить его квалификацию.

В большей степени помогает оценить качество работы эксперта следующий показатель – «суммарное отклонение баллов от результатов третьей проверки», но нельзя сравнивать работу экспертов по этому показателю. Если эксперт проверил мало работ, то и суммарное отклонение баллов в отличие от другого эксперта у него будет меньше. В этом случае помогает восьмой показатель – относительный.

Следующие (10-12) показатели позволяют более детально изучить работу эксперта при оценивании, но они не отображают ее интенсивность. Например, если эксперт проверил мало работ, то показатели совпадений и разногласий оценок со вторым экспертом будет небольшим. Для более точной оценки вводятся относительные показатели – 13.

Для дополнительного контроля качества работы экспертов вышестоящими службами используются следующие механизмы:

1. Для ведущих экспертов – анализ статистики удовлетворенных апелляций, анализ согласованности работы ПК;
2. Для старших экспертов – анализ статистики удовлетворенных апелляций, анализ согласованности работы экспертов ПК;
3. Для основных экспертов – анализ согласованности работы экспертов ПК.

Согласование критериев выполняется на разных уровнях:

- согласование критериев оценивания между председателями ПК регионов (на курсах повышения квалификации, конференциях и семинарах);
- предварительное согласование критериев оценивания экспертами (на курсах повышения квалификации и до начала проверок);
- оперативное согласование (во время проверок работ).

В соответствии с порядком привлечения экспертов к работе проводится испытание, в результате которого экспертам присваиваются статусы (ведущий, старший, основной). Показатели согласованности оценивания, должны соответствовать значениям, определенным органами исполнительной власти субъекта Российской Федерации, осуществляющими государственное управление в сфере образования для присвоения экспертам каждого из статусов. Рекомендуемые показатели согласованности и их значения по результатам квалификационных испытаний для присвоения каждого из статусов экспертам ПК по информатике и ИКТ [4, с. 26]:

- Ведущий эксперт – 5%;
- Старший эксперт – 7%;
- Основной эксперт – 10%.

Анализ проверки заданий с развернутым ответом по информатике показывает не только уровень согласованности работы каждой предметной комиссии, но и позволяет выявлять тех экспертов, которые, отклоняются от согласованной политики предметной комиссии. В составе предметной комиссии выявляются люди, у которых экспертная позиция – узкопрофессиональная предметная, а на проверке развернутых ответов такой позиции не должно быть. В этой ситуации, кем бы ни был проверяющий – он просто эксперт, который оценивает работы четко по критериям и не имеет право на собственное мнение при проверке развернутых ответов.

Анализ результатов оценивания экспертами развернутых ответов участников ЕГЭ проводится в отношении согласованности работы экспертов во время первой, второй и третьей проверок, данных о проверках по экзаменационным работам, результат по которым был изменен по решению конфликтной комиссии и результатов межрегиональной перекрестной проверки и перепроверки работ участников ЕГЭ. Вся полученная информация детализируется, внимательно отслеживается и передается руководителям системы образования и структурам, которые отвечают за подготовку экспертов.

Работа, основанная на анализе приведенных показателей, процедурах согласования, а так же включающая создание комфортных условий для работы экспертов позволяет повысить качество проверок.

Библиографический список

1. Об утверждении «Порядка проведения государственной итоговой аттестации по образовательным программам среднего общего образования»: Приказ Министерства образования и науки Российской Федерации от 9 января 2017 года № N 1400 (с изменениями на 9 января 2017 года) [Электронный ресурс] / подгот. АО «Кодекс» // АО «Кодекс». – М., 2017. – Режим доступа: <http://docs.cntd.ru/document/499071166>.
2. Унифицированные учебные материалы для подготовки председателей предметных комиссий ЕГЭ 2016 года: Информатика и ИКТ – Электронные текстовые данные. – (Москва, 2016.) – Режим доступа: http://fipi.ru/sites/default/files/document/1456847812/inf_dlya_predsedateley_pk.doc, – Загл. с экрана. – Для предметных комиссий субъектов РФ.
3. Методические материалы для председателей и членов региональных предметных комиссий по проверке выполнения заданий с развернутым ответом экзаменационных работ ЕГЭ 2016 года – Электронные текстовые данные. – (Москва, 2016.) – Режим доступа: http://fipi.ru/sites/default/files/document/1455276612/inf_met_rek_2016.doc. – Загл. с экрана. – Для предметных комиссий субъектов РФ.
4. Методические рекомендации по формированию и организации работы предметных комиссий субъекта Российской Федерации при проведении государственной итоговой аттестации по образовательным программам среднего общего образования – Электронные текстовые данные. – (Москва, 2016.) – Режим доступа: http://www.ege.edu.ru/common/upload/docs_new/MR_po_formirovaniu_i_organizatsii_raboty_PK_GIA-11_2017.docx. – Загл. с экрана. – Методические рекомендации.