

Разработка подходов к оцениванию итогового сочинения образовательными организациями высшего образования

Полежаева Мария Викторовна

кандидат педагогических наук, начальник отдела научно-методической экспертизы и психометрических исследований ФГБНУ «ФИПИ»

Симкин Виктор Николаевич

кандидат педагогических наук, ведущий научный сотрудник отдела научно-методической экспертизы и психометрических исследований ФГБНУ «ФИПИ», kim@fipi.ru

Ключевые слова: задания с развернутым ответом, схема оценивания, итоговое сочинение, критерии, дескриптор, аналитическое и холистическое оценивание.

Итоговое сочинение (изложение) как допуск к государственной итоговой аттестации (ГИА) выпускников образовательных организаций, реализующих программы среднего общего образования, впервые было введено в 2014/2015 учебном году. Основная цель итогового сочинения – выявление у обучающихся умения мыслить, анализировать и доказывать свою позицию с опорой на самостоятельно выбранные произведения отечественной и мировой литературы.

В соответствии с Порядком приема в вузы¹ при приеме на обучение по программам бакалавриата и программам специалитета организация высшего образования может добавлять в общий балл ЕГЭ баллы за оценку, выставленную ею по результатам проверки итогового сочинения, являющегося условием допуска к ГИА.

Образовательная организация высшего образования вправе разработать собственные критерии оценивания итогового сочинения, в том числе на основе предложенных Федеральным институтом педагогических измерений. В этой статье рассмотрены подходы, которыми рекомендуется руководствоваться при этом.

Что такое схема оценивания?

Как правило, описательные схемы оценивания разрабатываются для руководства анализом продукта или процессов, характеризующих учебные достижения. Обычно они применяются там, где необходимо качественное суждение в оценивании широкого круга субъектов или действий. Оценивание письменного текста можно привести в качестве общего примера применения схем оценивания.

¹ Порядок приема на обучение по образовательным программам высшего образования – программам бакалавриата, программам специалитета, программам магистратуры. Утвержден приказом Минобрнауки России от 14.10.2015 № 1147 (ред. от 31.07.2017) // URL: <http://base.garant.ru/71238710/>

Суждения относительно качества определенного письменного текста могут варьироваться в зависимости от критериев, которых придерживается конкретный оценивающий. У одного оценивающего могут превалировать лингвистические характеристики текста, в то время как другой оценивающий будет уделять больше внимания степени убедительности аргументации. Высококачественное эссе, скорее всего, будет отличать комбинация этих и других факторов. Разработав предварительно схему для оценочного процесса, мы сможем понизить его субъективность.

Когда схемы оценивания являются приемлемым оценочным инструментом?

Схемы оценивания — одна из альтернатив, доступных для оценки развернутых ответов. Они основаны на дескриптивных шкалах и позволяют оценить степень, в которой выражено соответствие определенному критерию.

То, является ли применение схем оценивания обоснованным, зависит от цели оценивания. Оценочные схемы имеют как минимум два достоинства в оценочном процессе. Во-первых, они позволяют исследовать степень соответствия определенному критерию; во-вторых, они предоставляют авторам работ обратную связь, позволяющую им совершенствоваться на основе полученной информации. Если эти достоинства отвечают цели оценивания, тогда оценочные рубрики, вероятно, являются адекватным оценочным средством.

Какие бывают виды схем оценивания?

Применяются разные типы схем оценивания. То, какой тип приемлем в конкретной ситуации, также зависит от цели оценивания. Как правило, рассматриваются различия между схемами оценивания, с одной стороны, холистическими и аналитическими, с другой стороны, обобщенными и специализированными.

В начальной фазе разработки оценочной схемы разработчик должен определить, что станет оценочным критерием. Например, два фактора, которые можно рассматривать при оценке письменной работы, — это грамматическая корректность и степень убедительности аргументации.

Аналитическая оценочная рубрика, во многом напоминающая контрольный лист, позволяет отдельно оценивать каждый из этих факторов. Каждый критерий оценивается по отдельной дескриптивной шкале.

Холистическая схема может быть трансформирована и содержать отдельный набор критериев для оценки, например, убедительности аргументации. Такая трансформация приведет к появлению аналитической схемы с двумя факторами: качество письменной речи и убедительность аргументации. Каждый фактор будет оценен отдельным баллом. Часто числовые значения приписываются в оценке каждого критерия.

Как пояснялось выше, преимущество применения схемы оценивания, в отличие от просто взвешенного балла, заключается в том, что схемы содержат описание того, что ожидается от испытуемого на каждом уровне. Испытуемые могут использовать эту информацию для совершенствования своих достижений в будущем.

В некоторых случаях бывает невозможно разбить оценку на отдельные факторы. Когда наблюдается наложение наборов критериев для оценки разных факторов, холистическая схема оценивания может быть предпочтительней аналитической. В холистической схеме критерии рассматриваются в сочетании на единой дескриптивной шкале, то есть производится целостное оценивание по заданным признакам. Холистические оценочные схемы способствуют более широким суждениям относительно качества процесса или продукта.

Выбор применения аналитической схемы оценивания не исключает возможности присутствия холистического фактора. Холистическое суждение может быть встроено в аналитическую схему как одна из оцениваемых категорий. Одной из проблем здесь является то, что наложение критериев, установленных для холистического суждения, на другие оцениваемые факторы становится неизбежным. Когда одной из целей оценивания является присвоение общего балла, такое наложение должно быть внимательно отслежено и проконтролировано. Оценивающий должен определить, получают ли в результате наложения определенные критерии больший вес, чем предполагалось изначально. Другими словами, оценивающий должен проследить за тем, чтобы испытуемый не был непредумышленно излишне наказан за определенную ошибку.

Как разрабатываются схемы оценивания?

Первый шаг в разработке схемы оценивания представляет собой четкое определение качеств, которые должны быть продемонстрированы испытуемым для подтверждения достижения требуемого уровня умения в соответствующей области. Эти качества, выраженные в *критериях*, будут представлять их высший уровень.

Затем нужно будет принять решение относительно того, как лучше получить требуемую информацию – с помощью холистической или аналитической схемы. Если принимается решение в пользу аналитической схемы оценивания, тогда каждый критерий рассматривается отдельно и разрабатываются описания разных уровней на шкале. Эта процедура приводит к появлению отдельных описательных схем для каждого оцениваемого фактора. Для холистических оценочных рубрик рассматривается набор критериев для всех уровней единой шкалы, которая становится результатом процедуры разработки.

Следующий этап – разработка и написание дескрипторов для выбранного вида шкалы. Общие требования к дескрипторам можно представить следующим образом:

1. *Позитивность*. Желательно, чтобы дескриптор описывал то, что умеет делать участник, а не то, чего он не сделал, в чем ошибся.
2. *Определенность*. Дескриптор должен быть четким, не содержать неопределенность («Участник в той или иной форме...» – пример не вполне удачной формулировки).
3. *Ясность*. Дескриптор должен быть понятен тому, кому он предназначен.
4. *Краткость*. Дескриптор не должен содержать нерелевантную информацию.
5. *Независимость*. Каждый дескриптор должен быть понятен без ссылок на другие. К сожалению, достаточно распространенная ошибка – использование в описании одной характеристики ссылок на параметры, оцениваемые в этой же схеме отдельно.

После определения критериев высшего уровня внимание разработчика может быть переключено на определение минимально допустимого уровня. Какое решение будет представлять очень ограниченное, но минимально допустимое понимание оцениваемого концепта? Различие между критериями для высшего уровня решения и минимально допустимого с большой долей вероятности обозначит кри-

терии для среднего уровня. Такой подход приведет к трем уровням оценивания.

Если есть потребность в более широком разграничении, можно начать сравнение критериев для каждого из определенных оценочных уровней. Такая процедура сравнения может продолжаться до получения желаемого количества уровней оценки или до того момента, когда проведение различий между критериями становится невозможным. Если значимых различий между оцениваемыми категориями уже нельзя сделать, создание дополнительных категорий оценки следует прекратить. Лучше иметь меньшее количество значимых уровней категорий, чем большое количество таких, разницу между которыми очень трудно обозначить.

Каждую оценочную категорию следует определять, используя описание результатов деятельности, а не суждения о деятельности. Например, «текст испытуемого не содержит орфографических ошибок», предпочтительнее, чем «хорошая орфографическая грамотность». Слово «хорошая» требует от эксперта субъективного суждения, в то время как фраза «не содержит ошибок» позволяет сделать однозначный вывод. Формулировки рубрик требуют проверки привлечения группы экспертов для проверки. Различия в оценке, произошедшие вследствие нечетких формулировок, должны быть устранены.

Рекомендуемые критерии оценивания итогового сочинения организациями, реализующими образовательные программы высшего образования

Подходы к анализу результатов итогового сочинения были рассмотрены в статье Е.А. Зининой, Т.В. Браги². Большая работа над критериями оценивания сочинения по литературе идет в рамках разработки перспективной модели КИМ ЕГЭ по литературе³. При разработке подходов к оцениванию итогового сочинения исходили из того, что критерии должны согласовываться с принципом объективности оценивания сочинения. Итоговое сочинение,

² Зинина Е.А., Брага Т.В. Организационные и методические подходы к проведению и анализу итогового сочинения в субъектах Российской Федерации. // «Педагогические измерения» – 2017. – №2. – С. 86–90.

³ Зинин С.А. Итоговый экзамен по литературе и методика преподавания предмета: стратегия «обратной связи» // «Педагогические измерения» – 2017. – №1. – С. 27–31.

с одной стороны, носит надпредметный характер, то есть нацелено на проверку общих речевых компетенций обучающегося, выявление уровня его речевой культуры, оценку умения выпускника рассуждать по избранной теме, аргументировать свою позицию. С другой стороны, оно является литературоцентричным, так как содержит требование построения аргументации с обязательной опорой на литературный материал.

Для оценивания итогового сочинения предлагается использовать аналитическую шкалу с выделением следующих критериев:

- *Соответствие теме*
- *Аргументация. Привлечение литературного материала*
- *Композиция*
- *Качество речи*
- *Оригинальность сочинения*
- *Речевые нормы*
- *Орфографические нормы*
- *Пунктуационные нормы*
- *Грамматические нормы*
- *Фактическая точность в фоновом (не литературном) материале*

Холистическая интерпретация схемы говорит о том, что от испытуемых ожидается сочинение, отвечающее следующим требованиям:

«Участник раскрывает тему сочинения или в той или иной форме рассуждает на предложенную тему, выбрав убедительный путь её раскрытия (например, отвечает на вопрос, поставленный в теме, или размышляет над предложенной проблемой и т.п.), коммуникативный замысел сочинения выражен ясно. Участник при раскрытии темы сочинения строит рассуждение на основе произведения (произведений) отечественной или мировой литературы по собственному выбору, показывая разный уровень осмысления литературного материала: от элементов смыслового анализа (например, тематика, проблематика, сюжет, характеры и т.п.) до комплексного анализа художественного текста в единстве формы и содержания; выбор литературного произведения и аспекты его анализа оправданны с точки зрения темы;

допущено не более 1 фактической ошибки, связанной со знанием литературного материала (ошибка в написании автора и названия произведения, имен персонажей и топонимов произведения, в изложении сюжетной линии, литературных и исторических фактов и

т.п.). Сочинение отличается композиционной цельностью, логичностью изложения мыслей и соразмерностью частей, внутри смысловых частей нет нарушений последовательности и необоснованных повторов мысли. Сочинение характеризуется творческим, нестандартным подходом к раскрытию темы (присутствуют интересные мысли, или неожиданные и вместе с тем убедительные аргументы, или оригинальные наблюдения и проч.) или яркостью стиля. Речевых ошибок нет, или допущено 1–2 речевые ошибки. Орфографических ошибок нет, или допущена 1 ошибка. Пунктуационных ошибок нет, или допущена 1 ошибка. Грамматических ошибок нет, или допущена 1 грамматическая ошибка. Фактические ошибки отсутствуют»⁴.

Типичные ошибки при разработке критериев

В рамках заданной структуры вузы применяют (исходя из своих условий) варианты схем оценки сочинения, которые порой не вписываются в предложенные рекомендации. При корректировке рекомендуемых критериев или создании собственных критериев вузы вправе отражать свои специфические требования к абитуриентам через систему оценивания итоговых сочинений.

Так, некоторые вузы практикуют использование рекомендуемых критериев с изъятием какой-либо их части. Например, проводится оценивание содержания без учета грамотности или оценивание речевого оформления сочинения без оценки его содержания. Иногда вводятся специфические критерии, такие как: «Соблюдение этических норм»; «Филологическая компетентность», «Знание текстов художественного произведения и/или фактов истории, культуры».

Многие решения не вызывают возражений. Вместе с тем анализ вузовских критериев позволил выявить ряд типичных ошибок, которые существенно искажают компаративность представляемых результатов. Приведем ряд примеров.

1. *Повышение планки требований* (увеличивается минимальный объем сочинения, тре-

⁴ Письмо Росособнадзора от 12.10.2017 №10-718. Приложение 4. Критерии оценивания итогового сочинения (изложения). // URL: http://www.fipi.ru/sites/default/files/document/itog_soch/10-718_v_oiv_o_napraslenii_mr_po_sochineniyu_2017-2018_uchebnyy_god.pdf

буются ссылки на большее число литературных источников) или *добавление в оценивание требований*, выходящих за пределы инструкции по написанию сочинения.

Пример 1

Критерий «Культурологическая и/или филологическая компетентность». Обучающийся применяет термины и понятия литературоведения, культурологии, искусствоведения в качестве инструмента интерпретации и анализа – 1 балл.

2. *Неоправданное объединение рекомендованных критериев («Речевое оформление и грамотность», «Самостоятельность написания и соответствие теме», «Содержание и речь»).*

Пример 2

Оценка качества речи, грамотности сочинения: стилистическое, речевое оформление текста, степень свободного владения речью, соблюдение основных орфографических, пунктуационных и грамматических норм (в совокупности не более 15 ошибок) от 0 до 1 балла.

3. *Отсутствие некоторых важных для отбора абитуриентов критериев (оценивание по первым пяти критериям без проверки соблюдения речевых, орфографических, пунктуационных и грамматических норм. Возможно для некоторых специальностей это решение было бы оправдано, но в этом случае речь шла о приеме на профиль «Филологическое образование» (русский язык и литература) направления бакалавриата «Педагогическое образование».*

4. *Отсутствие «отсекающих» требований (при 0 баллов за соответствие теме сочинение оценивается по другим критериям).*

5. *Отсутствие необходимой дифференциации абитуриентов.*

Пример 3

Используется зачетная система: по 5 критериям ставятся не баллы, а зачет-незачет.

Пример 4

За сочинение ставится от 0 до 3 баллов, оценивание холистическое, отметка зависит от содержания и грамотности.

Пример 5

Сочинение оценивается по двум критериям: содержание и качество речи, грамотность: максимальный балл по каждому критерию – 1.

6. *Высокий уровень обобщения аспектов оценивания (1–0 баллов – за содержание; 1–0 – качество речи (вариант 1–0 баллов – за грамотность)).*

7. *Неоправданное присвоение критериям весовых коэффициентов (оценивание по критерию не 0–1–2, а 0–5–10, по другому критерию 0–5–10–15 и т.д.). Данный подход ведет к снижению надежности измерения.*

8. *Упрощение критериев («Глубина баллов» по всем критериям: 1–0).*

9. *Дробность оценивания (1 – 0,5 – 0 б).*

10. *Невнятность.*

Пример 6

1–2 ошибки – 5 баллов;

2–3 ошибки – 4 балла;

4–5 ошибок – 3 балла;

6–7 ошибок – 2 балла;

8–9 ошибок – 1 б. (0 баллов нет).

Пример 7

В сочинении сделано допустимое/недопустимое количество ошибок.

11. *Громоздкость формулировок критериев.*

Пример 8

K2. Аргументация. Привлечение литературного материала

3 балла может быть выставлено, если есть прямой и счерпывающий ответ по теме, обнаружено отличное знание и глубокое понимание текста художественного произведения, а также умение пользоваться литературным материалом для раскрытия темы, давать оценку излагаемым фактам, композиционно стройно, логически последовательно и аргументировано излагать свои мысли, писать правильным и выразительным литературным языком. Орфографические, пунктуационные, речевые, грамматические и фактические ошибки либо отсутствуют вообще, либо обнаружены единичные нарушения норм письменной речи. <...>

2 балла ставится за сочинение, в котором дан *схематичный* ответ на тему сочинения, допущены *отклонения* от темы, *неточности* в изложении фактического материала, *нарушения* последовательности изложения мыслей. Обнаружено *не более 4* (суммарно) орфографических и пунктуационных ошибок. Присутствуют и другие виды ошибок.

1 балл ставится за сочинение, в котором преобладают *общие фразы*, не подтвержденные литературным материалом; изложение носит *трафаретный* характер или сводится к *простому пересказу* произведения (или учебника), имеются *серьезные нарушения* последовательности в выражении мыслей. Обнаружено *не более 8* (суммарно) орфографических и пунктуационных ошибок.

0 баллов ставится за сочинение, в котором *не понята и не раскрыта тема*, обнаруживается полное *незнание литературного материала*. Допущено *более 8* (суммарно) орфографических и пунктуационных ошибок. Присутствуют и другие виды ошибок.

Итак, на основе анализа типичных недостатков можно сформулировать следующие рекомендации вузам по составлению критериев:

- использовать аналитический подход к оцениванию, а не холистический;
- применять в качестве основы рекомендуемые критерии оценивания итогового сочинения для организаций, реализующих образовательные программы высшего образования;
- проанализировать, какие требования к абитуриентам предъявляются вузом, возможно, дифференцировать эти требования в зависимости от направлений подготовки;
- на основе проведенного анализа выбрать критерии оценивания сочинения;
- не использовать критерии, которые выходят за пределы инструкции, предложенной участникам, пишущим сочинение;
- не забывать о необходимости учитывать минимальный объем текста и задать ведущий(-е) критерий(-и), в случае получения 0 баллов по которому(ым) сочинение дальше не проверяется;
- не использовать громоздкие и нечеткие формулировки;
- максимальный балл по критериям сделать не больше 4, лучше 2–3 балла, не использовать весовые коэффициенты для критериев (0–5–10) и дробное оценивание (0–0,5–1);

- максимальный балл за сочинение сделать от 10 до 20 баллов с дальнейшим переводом в заданную шкалу.

Для оценки того, насколько ясной является определенная схема оценивания, могут быть использованы приведенные ниже вопросы.

1. Четко ли определены критерии оценки?

2. Достаточно ли ясны различия между критериями оценки и степенями выражения описываемых ими качеств?

3. Могут ли два независимых эксперта прийти к одинаковому результату оценивания одной работы на основе предложенной схемы?

Ответ «нет» на любой из этих вопросов должен привести к пересмотру соответствующих категорий схемы.

Другой способ уточнения схемы оценивания связан с использованием «якорных» работ. «Якорные работы» – это предварительно оцененные работы, которые иллюстрируют нюансы оценочных схем. Отдельный эксперт имеет возможность обратиться к якорной работе в любой момент процесса оценивания для установления различия между уровнями оценки.

После того как сделано все возможное для уточнения оценочных категорий, оценочные схемы и «якорные» работы могут быть представлены для апробации отдельной группе преподавателей. Любые расхождения в их оценках будут указывать на то, какие компоненты оценочной схемы нуждаются в дальнейшей доработке и разъяснениях. Необходимо обсуждать все расхождения в интерпретации дескрипторов и вносить необходимые уточнения. Хотя этот процесс времязатратен, он может значительно повысить надежность оценивания.

Кроме корректного составления схемы оценивания сочинения необходимо также учесть необходимость подготовки экспертов для проверки работ. Как показал опыт проведения ЕГЭ и подготовки экспертов по оцениванию ответов обучающихся на задания с развернутым ответом, учителя школ и преподаватели вузов не привыкли оценивать ответы экзаменуемых по определенным критериям. Даже при наличии качественных критериев остается элемент субъективности. Чтобы свести данный элемент к минимуму, необходимо проводить целенаправленную подготовку экспертов и предварительное согласование подходов к оцениванию.