

# Оценочные шкалы в образовательной технологии

**Вячеслав ГУЗЕЕВ**, доктор педагогических наук, профессор, зав. кафедрой образовательных технологий АПК и ПРО

**Всякая оценка предполагает прежде всего диагностику состояния оцениваемого объекта. Оценка чего-либо связана с измерением определённых параметров. В широком смысле измерение — это поиск упорядоченной классификации.**

## Типология оценочных шкал

Какие методы измерения используются? Шкалирование, метод экспертных оценок, тестирование, контент-анализ. Рассмотрим виды шкал и особенности их применения.

Количественные шкалы могут быть непрерывными и дискретными. Однако непрерывные шкалы применяются очень редко, поэтому мы не станем их обсуждать. Кроме того, употребляемая нами терминология не всегда соответствует принятой в других источниках. В частности, порядковые дескриптивные шкалы часто называют *номинальными*.

## Количественные шкалы. Планируемые результаты обучения и образовательные стандарты

С точки зрения теории всякое измерение есть отображение множества измеряемых объектов в некоторое множество величин. Если в качестве множества величин выступает числовое множество, то говорят о количественной шкале. Таким образом, количественная шкала предназначена для представления оценки числом. Это, например, хорошо всем известная школьная система балльных оценок. Если начало отсчёта не связано с учеником, не зависит от него, а ученику лишь находится соответствующее место на шкале, то говорят об *абсолютной оценочной шкале*. Типичные инструменты образовательной технологии — тесты достижений или более привычные нам контрольные работы — нацелены как раз на идентификацию состояния ученика по абсолютной количественной шкале.

«Бывают такие случаи, когда мальчик или девочка проявили титанические, не свойственные для себя усилия над выполнением задания. И в подобной ситуации их результат может казаться всем ниже, чем результат отличника, обладающего превосходной памятью и способностью без особых усилий осваивать материал. Именно тут и возникает впечатление о несправедливости оценок учителя»<sup>1</sup>. Особенно остро встаёт эта проблема при обучении подростков. «Чувство самоидентичности у подростка получает новое психологическое качество — это непосредственное ощущение и рефлексия внутреннего опыта. Подросток становится скрытным и со взрослыми, и со сверстниками. Глубинное «Я» не может быть описано его носителем в рациональных терминах — всегда остаётся что-то, не поддающееся определению. Оно воспринимается подростком как интимная субстанция, таинство души. Подросток оберегает это таинство в себе — гарантию ощущения своей автономности, индивидуальности, самооценности. Риск самораскрытия состоит в потенциальной возможности других вторгнуться в глубинное «Я» и выражается в опасности внешней оценки»<sup>2</sup>. В таких ситуациях, возможно, лучше использовать относительную шкалу.

<sup>1</sup> Журавлёв В.И. Основы педагогической конфликтологии. М.: Российское педагогическое агентство, 1995. С. 66.

<sup>2</sup> Мастеров Б.М. Психология саморазвития: психотехника риска и правила безопасности. М.: Интерпракс, 1994. С. 50.

Относительная шкала предполагает сравнение текущего состояния ученика с его же состоянием некоторое время назад. Следовательно, эта шкала связана с самим измеряемым объектом и отражает его изменение, развитие. Однако при такой шкале нередко «...ученик,

получивший плохую отметку за свой ответ, может утверждать, что она низкая потому, что учитель его не любит. А учитель утверждает, что он ставит отметки соответственно предъявленному уровню знаний, независимо от отношений с учащимися. Иначе говоря, мы сталкиваемся с ситуацией, когда ученик воспринимает выставление отметки как средство реализации отношения к нему, тогда как учитель рассматривает выставление отметки сугубо в рамках деловых отношений»<sup>3</sup>.

<sup>3</sup> Костенчук И.А. Диалог в педагогическом общении. М.: Ин-т пед. новаций РАО, 1995. С. 21.

Абсолютная количественная шкала строится очень просто: умеет ученик безошибочно выполнять задания из некоего набора — такая-то оценка, из другого — иная. Так же может быть выстроена и относительная шкала: за определённый период времени ученик научился делать что-то — такая оценка, а ещё что-то — другая, но она может оказаться для каждого ученика своей. Будучи более гуманной, чем абсолютная, относительная шкала тем не менее очень конфликтогенна в условиях совместного обучения группы детей. «Нередко весь класс утверждает, что оценка, объявленная учителем, не соответствует качеству и уровню знаний, которые продемонстрировал отвечающий. Что в других случаях другим учащимся тот же учитель ставит оценку выше. Конфликт в связи с этим во многом зависит от реакции учителя и от мастерства, надёжности, открытости критериев, которыми он руководствуется, оценивая знания и умения...»<sup>4</sup>.

<sup>4</sup> Журавлёв В.И. С. 79–80.

Рассмотрим теперь информационную функцию системы оценивания результатов обучения. Сложившаяся в России четырёхбалльная система была удобна для единого содержания образования, единой и единообразной школы и вообще для тоталитарной системы управления развитием общества. В новых условиях необходима множественность как типов образовательных учреждений, так и вариативность программ. Очевидно, что теперь, при разном содержании в разных классах одного и того же предмета, одна и та же оценка о разных учениках может нести совершенно различную информацию: оценка «удовлетворительно» при общекультурном содержании курса далеко не эквивалентна такой же оценке при содержании профессиональном или прикладном. Следовательно, в оценке надо как-то отразить и профиль обучения — это совсем новая задача.

В рамках уже каким-то образом отобранного содержания базой шкалы оценок может быть уровневая градация по степени достижения того или иного уровня планируемых результатов обучения или овладения содержанием. Если такой подход приемлем, то следует пойти дальше и спроецировать эту шкалу уровней на три профиля содержания образования — получим нечто подобное матрице (см. таблицу).

### Оценка результатов обучения

Уровень планируемых результатов	Профиль содержания образования		
	Общекультурный	Прикладной	Профессиональный
Минимальный	1	2	3
Общий	4	5	6
Продвинутый	7	8	9

Числа в клетках матрицы можно рассматривать как десятибалльную абсолютную шкалу оценок при многоуровневом и многопрофильном обучении (всё, что левее и выше матрицы, оценим баллом «нуль»). По оценкам, проставленным в документе об образовании, сразу можно будет судить, какое содержание и как усвоено выпускником средней школы. Однако имеется масса вопросов, решению которых непосильно не только отдельным школам, но и целым крупным регионам, обладающим достаточной автономией в области образования. Статья 32 российского Закона «Об образовании» практически ничем не может здесь помочь, а скорее мешает.

Очевидно, что переход на такую шкалу оценок возможен только в масштабах всего го-

сударства или по меньшей мере республики в его составе. Кроме того, эту систему должны ещё признать вузы!

Далее мы непосредственно выходим на проблему образовательных стандартов — государственного выражения планируемых результатов обучения своих граждан по всем профилям и уровням. Совершенствование федеральных и создание региональных стандартов образования повлечёт за собой выделение общепризнанных образовательных инвариантов, а следовательно, и приведёт к созданию национальных и региональных тестов, подобных существующим во многих других странах. Это очень перспективное направление развития образовательных систем в России, но, к сожалению, довольно далёкого будущего и поэтому особенно важно не повторить чужих ошибок. Отдельные образовательные системы и даже учреждения могут пойти по пути создания собственных стандартов и разработки внутренних тестов. Их использование для промежуточного контроля как раз и разрешено статьёй 32 Закона «Об образовании». Как образец может быть использован американский тест на готовность к получению высшего образования SAT (Scholastic Aptitude Test) — один из лучших в мире примеров абсолютной количественной шкалы.

Дискретность количественных шкал и небольшая периодичность контрольных процедур вызывают многочисленные нарекания, ставшие уже нормой: «Серьёзными недостатками существующей системы контроля являются: 1) нерегулярность (эпизодичность) обратной связи; 2) неполный охват проверкой всего содержания обучения (выборочность проверки); 3) отсутствие проверки самого процесса работы ученика (проверяется в основном конечный результат); 4) недостаточная вовлечённость самих учащихся в самоконтроль (слабость внутренней обратной связи). Эти недостатки нарушают нормальное протекание взаимодействия между учителем и учащимися, снижают эффективность обучения»<sup>5</sup>.

---

<sup>5</sup> Дидактика средней школы: Некоторые проблемы современной дидактики. М.: Просвещение, 1982.

Использование абсолютных шкал целесообразно для контроля итоговых результатов некоторого учебного периода. Для контроля внутри процесса больше подходит относительная шкала.

### **Порядковые шкалы. «Безоценочное» обучение**

В слабо структурированных областях — таких, как искусство или, скажем, личностное развитие, — чаще используются порядковые шкалы. Их особенность состоит в том, что объект сравнивается с подобными объектами. Если множество величин является само данное множество, то речь идёт о его упорядочивании, иерархизации. Экспертная оценка всегда связана с порядковой шкалой и состоит в упорядочивании измеряемых объектов по каким-то признакам (критериям). Этот способ распространён в западных системах образования и когда-то широко применялся в России. Все знают, что Владимир Ульянов закончил гимназию первым учеником, хотя и не имел высших баллов по всем предметам: он был первым и лучшим в данной гимназии, но это не значит, что в других гимназиях не было учеников более умных и талантливых, чем он. Такой вариант оценочной шкалы называется *ранговым*: каждому объекту присваивается номер, ранг в иерархии. Типичные примеры ранговых шкал — система воинских званий или система должностей на университетской кафедре.

Частный случай ранговой шкалы — рейтинговая система. «Рейтинговая система имеет давние традиции в культуре российской школы. При использовании этой системы четвертные оценки выставляются не традиционным методом, когда возникают устойчивые группы «троечников», «хорошистов», «отличников», а путём составления общего списка класса, в котором в начале списка идут наилучшие, в конце — неудачники. При этом у ученика возникает вполне оправданное ощущение, что ему нужно совсем не так уж много усилий, чтобы передвинуться с 16-го места на 15-е, и успех вполне возможен. Усилия же, необходимые для того, чтобы из «троечника» стать «хорошистом», могут не без основания показаться ему неоправданно значительными. Несмотря на то что данная система имеет свои «плюсы» и «минусы», она представляется вполне приемлемой для решения поставленной задачи»<sup>6</sup>. Таким

образом, рейтинговую и ранговую системы иногда и вовсе не различают.

---

<sup>6</sup> Ушаков К.М. Управление школьной организацией: организационные и человеческие ресурсы. М.: Сентябрь, 1995. С. 91.

Рейтинговая система имеет большое сходство с количественной шкалой, но, что очень важно, таковой не является. Рейтинг — это действительно число. Но получается оно либо путём опроса субъективных мнений экспертов, как рейтинги политических деятелей, либо путем набора очков (пунктов, баллов). Ответил ученик правильно на уроке — получил очко, написал реферат — ещё пару сотен очков, сдал зачёт — ещё сотню, помог товарищу — ещё несколько и так далее. В конце учебного периода все очки, набранные учеником, суммируются и получается его рейтинг. После этого множество учеников можно ранжировать по возрастанию рейтингов. В некоторых школах и вузах такая рейтинговая система промежуточных оценок уже применяется, но, видимо, из-за вполне оправданной осторожности она всё-таки привязана к традиционной балльной. Недостаток рейтинговой системы очевиден: количество очков за то или иное учебное достижение назначается экспертным способом и может сильно варьироваться в разных школах, отражая в своей произвольности вкусы и пристрастия этих педагогических коллективов.

Хорошо известен и другой тип порядковых шкал — дескриптивные, то есть описательные. Суть их в том, что объект, с которым сравнивается данный, хоть и принадлежит тому же множеству, что и измеряемый, но замаскирован, как бы скрыт. Знакомая дескриптивная шкала — это описание знаковой системой. Известная знаковая система — язык, менее очевидная — образная, в том числе графические модели, например, полосовой код на зарубежных товарах или нашивки за выслугу лет на форменной одежде военнослужащих. Оценки с помощью дескриптивных шкал давно уже дискутируются, у них есть горячие сторонники, а кое-где такие системы уже применяются.

Вам знакомы, конечно, многочисленные призывы к безоценочному обучению. Но с точки зрения науки управления никакой кибернетический процесс не может быть регулируемым без обратной связи и оценки. Все сторонники «безоценочного обучения» немного лукавят. На самом деле они ратуют за переход от количественных шкал к дескриптивным — и только!

«Ученик А имеет словарный запас около сорока тысяч слов. Может заменять слова и выражения их синонимами. Тонко различает оттенки в близких по смыслу предложениях. Знает типовые грамматические конструкции и грамотно применяет их в устной речи. Умеет заменять предложения сложной структуры несколькими предложениями простой структуры и наоборот. В устной речи уверенно использует распространённые идиомы...» Примерно такие оценки сообщают родителям своих учеников американские учителя. Заметьте, речь идёт только о достоинствах и достижениях: читать между строк об ошибках и недостатках — «А-а, писать-то не умеет!» — там никому и в голову не приходит. Это один из постулатов воспитания свободной личности. И работу потом ему будут подбирать, исходя из умений и достижений. Мало ли чего мы не умеем!

Как видим, типичная языковая дескриптивная оценка — это некоторая характеристика. В.М. Кларин замечает: «...оценочным может, по сути дела, являться любое суждение, выражающее отношение учителя, — например, похвала, одобрение, сомнение или же прямое неодобрение. Оценочный характер могут носить и невербальные реакции учителя — его одобрительные или неодобрительные, скептические или насмешливые интонации, взгляды, жесты».

Нельзя не согласиться с Д.А. Ивановым, утверждающим, что «действие всегда может быть оценено относительно какой-либо нормы, в то время как чувство в принципе не нормируемо. Каждый чувствует по-своему». А мы между тем всё ещё нередко пытаемся измерить «тройками — пятёрками» эстетический вкус или понимание музыки.

В последнее время всё чаще за рубежом обсуждается другой способ, который многим представляется наиболее адекватным для определения личностного развития. Это идея так называемой папки достижений (portfolio assessment)<sup>7</sup>. Она родилась в США как компромисс

между принятой в этой стране жёсткой системой тестового контроля и значительно более мягкой европейской системой. Внешне эта техника оценки схожа с теми папками детских рисунков и стихов, которые собирают любящие родители. В «папку» (собственно папку, коробку, шкафчик и т.д.) собираются лучшие, наиболее значимые работы учеников по предмету или группе предметов. Сравнение однотипных работ, выполненных в разное время или при различных обстоятельствах, и является оценкой. Например, сравнивая два рисунка девочки, говорят: «Так она рисовала год назад, а так рисует сегодня». Или, рассматривая два автографа мальчика, заключают: «Вот такой у него почерк, когда пишет не спеша, а такой — когда надо что-то записать быстро». В формировании такой «папки» может участвовать и сам оцениваемый ученик, помогая учителю выбрать лучшие, наиболее адекватные достижения. В отечественной педагогике аналогичная оценочная шкала (так называемый пакет) используется в опыте работы начальной школы Ш.А. Амонашвили. Рассмотрим эту технику оценки на конкретном примере обучения иностранным языкам.

---

<sup>7</sup> Точно перевести на русский язык это словосочетание довольно трудно. Йост Кок утверждает со ссылкой на латинский словарь, что *assessment* означает «оценка человека, выполняющего определённые действия» (Кок Й. Тот, кто присутствует и судит // Директор школы. 1996. № 1. С. 28).

Винченцо Монда<sup>8</sup> предлагает для обобщения и систематизации знаний по каждой изученной теме давать ученикам задание изготовить по этой теме рекламный плакат, буклет, афишу с текстами, рисунками, открытками, фото. Очевидно, что степень усвоения материала также хорошо проявится в такой работе. Если изучена тема «Италия», то ученикам придётся тогда для коротких текстов рекламного буклета отобрать всё самое существенное и красиво описать. Даже если тема небольшая — «Моя семья», например, то сделать рекламный плакатик «Приглашаем в гости» не очень просто. Собирая все эти буклеты, афиши и плакаты в папку в течение всего курса, мы и получим в первом приближении «папку достижений». Подобную систему отслеживания успехов детей использует педагог Ш.А. Амонашвили: *Каждый ребёнок начнёт составлять также «том» произведений собственного творчества, в который войдут листки с записью первых слов и предложений, первое «сочинение», решения математических задач и примеров, чертежи, рисунки, аппликации.* Конечно, в отличие от «родительских» папок «папка достижений» как оценочная шкала в образовательной технологии имеет свои правила и принципы.

---

<sup>8</sup> *Monda V. Good idea! // EFL Gazette, 1992. June. P. 23.*

Преимущество количественных шкал — их простота и определённость, но их применение связано с заметной потерей внутренней информативности. Порядковые шкалы, особенно дескриптивные, очень информативны и содержательны, но при их использовании надо иметь сложные и дорогие измерительные инструменты — компьютерные экспертные системы или толстые инструкции, необходимы также профессиональные эксперты. Развитие некоторых областей математики (математика нечётких множеств, теория некорректно поставленных задач, математическая статистика и другие), возможно, в какой-то мере решит эти проблемы.

Оценка — это акт, являющийся частью процесса построения модели обучаемого. В основе оценочной деятельности для социальных систем практически всегда лежат статистические критерии. Например, кого мы считаем сильным учеником? Вот один условный пример: «Ученик занимается на продвинутом уровне, если он с вероятностью, не меньшей 80%, справляется с заданиями, доступными не более чем 10% учащихся в классе». Чаще всего модели обучаемых являются многопараметрическими. Самый простой образец — обычный аттестат о среднем образовании. Следовательно, приходится вести речь о многокритериальной задаче оценки. Это достаточно сложное явление, поскольку оно связано ещё с тем, что разные параметры в различных ситуациях могут значительно отличаться по важности, значимости. Сравните требования к математической подготовке будущих абитуриентов математического и филологического факультетов. Как определить весовые характеристики разных параметров модели? Один из тривиальных случаев нами уже упоминался — назначение числа

очков за те или иные виды деятельности при построении рейтинговой шкалы.