

# Проблема оптимального объема оценочной шкалы как инструмента деятельности педагога

**М. И. Яновский\***,  
кандидат психологических наук, доцент  
кафедры психологии Донецкого националь-  
ного университета

Исследование было направлено на выявление степени погрешности, даваемой известными в системе образования оценочными шкалами: 5, 12 и 100 баллов. Предполагалось, что такая погрешность возрастает, если объем шкалы превышает границы объема внимания оценивающего («пропускную способность наблюдателя», по Дж.А. Миллеру):  $7 \pm 2$ . Для проверки гипотезы был проведен эксперимент, в котором испытуемым предлагалось оценить в 5-, 12- и 100-балльной системах выполнение письменной школьной работы. Определялась степень согласованности оценок у разных испытуемых, возникающая при применении каждой из этих трех шкал. В исследовании участвовали 68 студентов-филологов третьего курса Донецкого национального университета. Установлено, что высокая согласованность в оценках наблюдается при использовании 5- и 12-балльной шкал, низкая – при использовании 100-балльной шкалы. Таким образом, качество оценочной шкалы как инструмента оценивания значительно падает, если ее объем существенно превышает объем внимания. При незначительном превышении (12-балльная шкала) такого падения качества не наблюдается, по-видимому, в силу легкости компенсаторных когнитивных процессов по отношению к такой шкале.

**Ключевые слова:** оценочная шкала, отметка, объем внимания, погрешность, естественная оценочная шкала.

Проблема отметок в системе обучения – предмет давних споров. Показательно, например, что отметки иногда воспринимаются как инструмент репрессивной государственной системы, так что это даже может приводить к «революционной» их отмене [4].

Однако споры вокруг данной проблемы небезосновательны. По результатам некоторых современных исследований, лишь чуть

больше половины школьников считают свои отметки справедливыми [8]. Остальные считают их либо заниженными, либо завышенными, либо просто не могут оценить степень их адекватности и справедливости [там же]. Ясно, что такая ситуация лишает отметки мотивирующей и ориентировочной функции, а значит, обесмысливает их для учащихся. Тем самым культивируется почва для воз-

\* yanovs@rambler.ru

можной «демонизации» образа школы (или вуза) в сознании части общества.

При этом несмотря на нашу склонность к личностной атрибуции любых проблемных сфер общества, не всегда в наличии трудностей, связанных с отметками в системе обучения, было бы правильно искать виноватого – предвзятого или некомпетентного педагога или «лукавого» в оценке себя учащегося. Оценивание (деятельность, продуктом которой является педагогическая оценка и, в частности, отметка как ее количественное выражение [1]), – процесс психологически сложный, и его реализация сопровождается рядом проблем.

Например, оценивание может рассматриваться как разновидность социального восприятия и поэтому оно может быть подвержено влиянию соответствующих психологических факторов. Исследователи выделяют целый ряд типичных ошибок, возникающих на такой почве:

- ошибка снисходительности – склонность завышать средние оценки, что может происходить, например, при давлении на оценивающего системы отношений (социальных, межличностных и т. д.);

- ошибка центральной тенденции – склонность избегать высоких и низких оценок, которая может возникать по той же причине;

- эффект ореола – склонность оценщика генерализовать частные успехи или неудачи объекта оценивания на любые его проявления [2]. Нередко встречается также ошибка контраста, когда оценщик дает более высокую оценку в силу того, что предыдущая оценка была низка [9].

Эти факторы часто не осознаются, что может быть причиной неадекватности оценок.

С другой стороны, действие оценивания – сложный когнитивный процесс. Для своей успешной реализации он требует формирования и удержания во внутреннем плане определенной шкалы, которая выступает фактически как рабочий инструмент оценивающего действия преподавателя, как виртуальный измерительный прибор. Так же как в любом приборе, ценность его определяется степенью погрешности в измерениях (достоверности его показаний), так и уровень погрешности в оценочной деятельности преподавателя – переменная, за-

висимая не только от личности преподавателя и его позиции в системе отношений, но и от самой используемой оценочной шкалы.

Действительно, оценочная шкала для субъекта поделена на определенные участки, каждый из которых соответствует определенному значению (баллу) оцениваемого параметра. Чтобы правильно оценить объект, нужно правильно определить место для объекта на шкале, т. е. выбрать из континуума как бы предстоящих перед мысленным взором участков шкалы участок, максимально соответствующий качеству объекта. Как дробно должна быть поделена на участки шкала, чтобы она вся, во всех своих частях, могла охватываться и удерживаться сознанием (т. е. чтобы она вся была для сознания «активирована» и чтобы в ней не было «мертвых», неактивных зон)? Общеизвестно, что в среднем объеме внимания взрослого человека составляет 5–6 единиц информации [6]. По Дж. А. Миллеру, «пропускная способность» наблюдателя составляет  $7 \pm 2$  [3]. При этом «пропускная способность» понимается как «верхняя граница области, в пределах которой наблюдатель может согласовывать свои реакции с предъявляемыми ему стимулами» [3, с. 204]. Следовательно, естественно полагать, что и шкала, которую мы могли бы всю, без «слепых пятен» использовать, не должна состоять из большего числа градаций.

Несоответствие шкалы характеристикам сознания будет определять неэргономичность шкалы как рабочего инструмента и, в силу этого, приводить к росту количества случайных флуктуаций оценок, т. е. фактически к возрастанию погрешности шкалы, и, следовательно, снижению ее качества.

Справедливость этого предположения подтверждается экспериментальными данными о границах дробности оценочных шкал, формируемых испытуемыми для оценки стимулов разного рода, например, временных отрезков [5]. Предел количества градаций, при котором испытуемые еще способны адекватно использовать шкалу, обычно близок к «числу Миллера».

Нами был разработан и проведен эксперимент, позволяющий оценить качество, т. е. погрешность, 5-, 12- и 100-балльной шкал. 5-балльную мы взяли в силу того, что она ис-

пользовалась и используется в российском школьном образовании, при этом периодически подвергается жесткой критике; 12-балльную – поскольку она введена и используется в системе среднего образования Украины, а 100-балльную – как применяемую в современных вузах в результате «болонской» реформы.

По нашей гипотезе, в деятельности оценивания наиболее эргономична и обеспечивает наибольшую точность и надежность шкала, количество делений в которой не превышает границы объема внимания оценивающего субъекта 12-балльная (отчасти) и особенно 100-балльная шкалы дают высокую степень погрешности при их использовании для оценивания.

Процедура эксперимента состояла в следующем: испытуемым-экспертам (в роли которых выступили 68 студентов-филологов третьего курса Донецкого национального университета) предоставлялась одинаковая копия письменной работы ученика средней школы, выполненная на среднем уровне (работа оценена школьным учителем на «6» по 12-балльной системе). Задачей экспертов было оценить – грамматически и орфографически – данную работу в трех разных шкалах оценивания (5-, 12- и 100-балльной). Выборка условно делилась на три группы, и в каждой группе очередность использования шкал была разной. Таким образом, каждый участник исследования оценивал одну и ту же работу три раза по трем разным шкалам. С испытуемыми заранее оговаривалась возможность использования в 5-балльной шкале добавочных знаков «+» и «-».

Отметим, что испытуемые были хорошо знакомы со 100-балльной (2,5 года учебы в университете) и с 12-балльной системами оценивания (11 лет учебы в средней школе). 5-балльная система в их обучении не использовалась (по крайней мере, официально).

Погрешность шкал устанавливалась посредством определения согласованности даваемых экспертами отметок при использовании каждой из трех шкал. Для этого все множество отметок были поделено на три условных подмножества – по используемой шкале. В каждом подмножестве фиксировались все задействованные конкретные отметки и часто-

та их использования экспертами. Далее определялся коэффициент вариации – мера, позволяющая сводить к единой шкале и адекватно сопоставлять относительный размах вариаций, колебаний *разномасштабных* величин [7]. Коэффициент рассчитывается по формуле

$$v = \frac{\sigma}{x_{cp}} \times 100\%$$

где  $\sigma$  – стандартное отклонение;  $x_{cp}$  – среднее от  $x$ .

Если  $v$  больше 100 %, это свидетельствует о чрезмерном разбросе данных, если  $v$  меньше 30 %, однородность результатов достоверна.

Мы, таким образом, определяли, каков разброс оценок (в процентах) в каждой из использованных шкал.

При проведении исследования также были учтены некоторые особенности реального использования оценочных шкал в учебных заведениях. Во-первых, мы учли, что 5-балльная шкала на практике имеет более сложную структуру. В ней к цифровым значениям могут добавляться знаки «+» или «-». Мы приравнивали знаки «+» или «-» к 0.5 балла (например, «4+» = «4,5»), что совпало с мнением опрошенных нами педагогов. В итоге это означает, что 5-балльная шкала может рассматриваться как имеющая 7 градаций, т. е. как 7-балльная (рис. 1). Во-вторых, в некоторых шкалах есть недействующие зоны – отрезок шкалы, который не применяется на практике. Такой зоной является отметка «1» в 5-балльной шкале и целый отрезок 100-балльной шкалы: от 1–49, который является «незачетом». Для этой функции на практике, т. е. реально, используется одна отметка «49», ее мы и оставили в рамках эксперимента. Следовательно, для чистоты исследования были вычтены (сдвинуты) эти «недействующие зоны»: в 5-балльной каждая оценка сдвинулась на единицу, а в 100-балльной – на 49 соответственно. 12-балльную шкалу не меняли, так как опрошенные нами учителя утверждали, что в этой шкале применяются на практике все баллы.

Результаты проведенного исследования на выявление степени согласованности приведены в диаграммах (рис. 1–3).

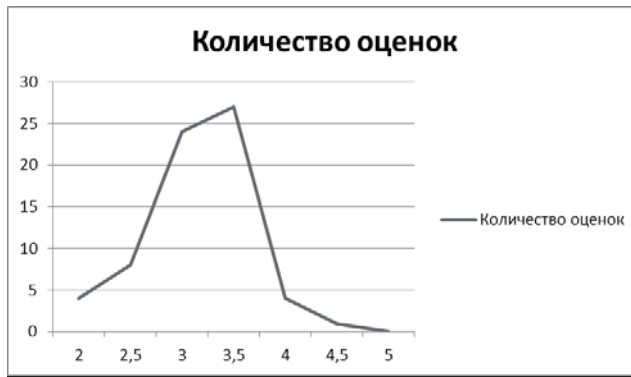


Рис. 1. Распределение количества отметок по баллам в 5-балльной шкале

Коэффициент вариации по 5-балльной шкале равен 23,5%. Это достаточно высокий уровень согласованности оценок внутри группы. Значит, данная оценочная система обладает относительно неплохой надежностью и застрахованностью от вмешательства фактора случайных ошибок оценивающего.

Что касается 12-балльной шкалы оценивания, здесь ситуация практически та же самая: показатель коэффициента вариации равен 21,3%. Таким образом, данная оценочная система также дает довольно высокую степень точности. Это в некоторой степени расходится с нашим предположением, что если размер шкалы превышает объем внимания ( $> 7 \pm 2$  единиц), то шкала дает повышенную погрешность. Отметим, однако, что, во-первых, 12-балльная шкала превышает объем внимания ненамного, и, во-вторых, она, воз-

можно, несколько сжата, сокращена испытуемыми. Так, хотя по 5-балльной системе 1 испытуемый выставил отметку «4+», в 12-балльной эквивалентные ей «10» баллов не встречаются. Также по 5-балльной системе 4 испытуемых выставили отметку «2», но в 12-балльной «1» как вариант эквивалента ей не встречается (встречаются «2» и «3»). В целом можно предположить, что в данной шкале при ее использовании достаточно легко реализуется компенсаторный процесс: для оценивания создается рабочая зона, близкая к границам объема внимания. Эта зона может при необходимости немного сдвигаться в верхнюю или нижнюю часть шкалы. Однако поскольку отличие общего размера 12-балльной шкалы от объема внимания невелико, такое усложнение действия оценивания не отражается на эффективности шкалы.

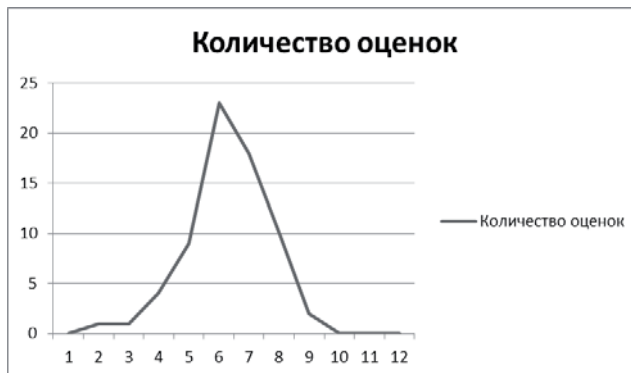


Рис. 2. Распределение количества отметок по баллам в 12-балльной шкале



Рис. 3. Распределение количества отметок по баллам в 100-балльной шкале

И, наконец, самая «точная» (по замыслу авторов данной шкалы) в системе образования – 100-балльная шкала. При расчете коэффициента вариации мы получили 63,9 %. Учитывая, что зона «хорошей» однородности данных приходится на промежуток от 0 % до 30 %, 100-балльная шкала не может рассматриваться как оптимальная (по сравнению с 5- или 12- балльной шкалами) для практического использования. Она дает достаточно большой уровень рассогласованности, и, следовательно, в ней заложена возможность довольно сильного влияния случайных, ситуативных обстоятельств, в том числе субъективизма. По-видимому, чрезмерная дробность оставляет за пределами объема внимания оценщика многие ее участки; это делает ее громоздкой и трудной в использовании, в связи с чем обратим внимание, что некоторые из баллов не были использованы ни одним из 68 испытуемых, даже несмотря на то, что они соседствуют с баллами, по которым был максимум ставившихся отметок. Так, отметку «65» выбрали 13 испытуемых, а соседние отметки «66» и «67» – ни один. Такой результат противоречит закону нормального распределения и означает, что данная оценочная шкала психологически неоднородна: в ней есть «мертвые» зоны, т. е. участки, которые практически не используются (или, по крайней мере, используются с существенно меньшей вероятностью, чем соседние). По-видимому, оценивающий субъект редуцирует чрезмерно дробную шкалу к пси-

хологически более естественной и простой. Действительно, можно обратить внимание, что в распределении отметок наблюдается 8, или, точнее, 9 пиков: «49», «50», «55», «60», «65», «70», «75», «80» и «85» баллов (своего рода округление к кратным пятерке числам). По-видимому, происходит процесс изменения масштаба шкалы: градации укрупняются, чтобы сознанию было возможно их использовать для оценивания как единую шкалу (поскольку объем внимания составляет  $7 \pm 2$ , то 9 пиков – это количество реально используемых делений 100-балльной шкалы, которое совпадает с верхней границей объема внимания).

Итак, если сравнивать используемые в системе образования оценочные шкалы, можно сделать вывод, что 5- и 12-балльная шкалы наиболее точны и надежны в использовании, т. е. обладают наиболее высоким качеством как инструменты оценочной деятельности педагога. Они отличаются малым разбросом отметок при их применении разными людьми, а это значит, что вероятность случайности и субъективности здесь относительно невелика.

100-балльная шкала в силу своей психологической неэргономичности обуславливает большую рассогласованность в оценках, даваемых одному и тому же объекту разными субъектами. Заметим, что неуверенность, которую она порождает у педагога, повышает вероятность неосознанного влияния на отметку и субъективного личного от-

ношения педагога к личности учащегося, а также других феноменов социального восприятия, о которых мы упоминали в начале статьи.

При этом, на наш взгляд, не может иметь принципиального значения аргумент, что 100-балльная шкала чаще используется как накопительная, а не как шкала однократной оценки. При формировании и реализации «разбалловки» в накопительной системе педагог использует 100-балльную шкалу как единый виртуальный измерительный инструмент, и это значит, что психологические особенности данной шкалы, о которых мы говорили, сохраняются.

### Литература

1. Амонашвили Ш.А. Воспитательная и образовательная функция оценки учения школьников: Экспериментально-педагогическое исследование. М., 1984.  
2. Бергер Л. Ошибки оценщика // Психологическая энциклопедия. 2-е изд. / Под ред. Р. Корсини, А. Ауэрбаха. СПб., 2006.  
3. Миллер Дж.А. Магическое число семь плюс или минус два. О некоторых пределах нашей способности перерабатывать информацию // Инженерная психология / Под ред. Д.Ю. Панова и В.П. Зинченко. М., 1964.  
4. Об отмене отметок. Постановление народного комиссара по просвещению, май 1918 года // Народное образование в СССР. Общеобразовательная школа. Сборник документов. 1917–1973 гг. / Сост. А.А. Абакумов,

### Выводы

1. 100-балльная оценочная шкала психологически является наименее адекватной. Поэтому как инструмент деятельности педагога она может считаться неэргономичной и привносящей в оценочную деятельность педагога повышенную степень неточности и, возможно, субъективизма.

2. Достаточно высокий уровень согласованности при использовании дают 5- и 12-балльные шкалы.

3. С точки зрения психологической естественности оптимальной шкалой могла бы быть та, количество градаций в которой не превышает объем внимания.

Н.П. Кузин, Ф.И. Пузырев, Л.Ф. Литвинов. М., 1974.  
5. Подвигина Д.Н., Варовин И.А. О структуре субъективных шкал оценки человеком коротких интервалов времени // Экспериментальная психология. 2010. Т. 3, № 3.  
6. Рубинштейн С.Л. Основы общей психологии. СПб., 2012.  
7. Сепетлиев Д.А. Статистические методы в научных медицинских исследованиях / Под ред. А.М. Меркова. М., 1968.  
8. Собкин В.С., Калашникова Е.А. От двойки до пятерки: грани школьной отметки // Национальный психологический журнал. 2011. № 1(5).  
9. Эйкен Л.Р. Оценочные (рейтинговые) шкалы // Психологическая энциклопедия. 2-е изд. / Под ред. Р. Корсини, А. Ауэрбаха. СПб., 2006.

## The problem of optimal scale of assessment as a tool of a teacher's work

M. I. Yanovski,

PhD in Psychology, Assistant Professor, Chair of Psychology,  
Donetsk National University

Our study is aimed to determine the degree of uncertainty given by well-known in the education system rating scales: 5, 12 and 100 points. It was assumed that this error increases when the scale volume exceeds the limits of attention of the evaluating person ("observer's bandwidth" by J.A. Miller):  $7 \pm 2$ . To test the hypothesis, an experiment was conducted in which subjects were asked to estimate a written school work in 5-, 12- and 100-point systems. We determined the degree of consistency of ratings in different subjects, which arises in the application of each of these three scales. The study involved 68 third year Philology students of Donetsk National University. High consistency is found in scores on 5 and 12-point scales, and low when using a 100-point scale. Thus, the quality of the rating scale as a tool for estimating drops significantly as its volume significantly exceeds the capacity of attention. With its slight excess (12-point scale), the quality drop is not observed, presumably, due to ease of compensatory cognitive processes relative to the scale.

**Keywords:** rating scale, school mark, attention capacity, accuracy, natural grading scale.

### References

1. Amonashvili Sh.A. Vospitatel'naja i obrazovatel'naja funkcija ocenki uchenija shkol'nikov: Jeksperimental'no-pedagogicheskoe issledovanie. M., 1984.
2. Berger L. Oshibki ocenshchika // Psihologicheskaja jenciklopedija. 2-e izd. / Pod red. R. Korsini, A. Auferbaha. SPb., 2006.
3. Miller Dzh.A. Magicheskoe chislo sem' pljus ili minus dva. O nekotoryh predelah nashej sposobnosti pererabatyvat' informaciju // Inzhenernaja psihologija / Pod red. D.Ju. Panova i V.P. Zinchenko. M., 1964.
4. Ob otmene otmetok. Postanovlenie narodnogo komissara po prosveshheniju, maj 1918 goda // Narodnoe obrazovanie v SSSR. Obshheobrazovatel'naja shkola. Sbornik dokumentov. 1917–1973 gg. / Sost. A. A. Abakumov, N.P. Kuzin, F.I. Puzyrev, L.F. Litvinov. M., 1974.
5. Podvigina D.N., Varovin I.A. O strukture sub'ektivnyh shkal ocenki chelovekom korotkih intervalov vremeni // Jeksperimental'naja psihologija. 2010. T. 3, № 3.
6. Rubinshtejn S.L. Osnovy obshhej psihologii. SPb., 2012.
7. Sepetliev D.A. Statisticheskie metody v nauchnyh medicinskih issledovanijah / Pod red. A.M. Merkova. M., 1968.
8. Sobkin V.S., Kalashnikova E.A. Ot dvojki do pjaterki: grani shkol'noj otmetki // Nacional'nyj psihologicheskij zhurnal. 2011. № 1(5).
9. Jejken L.R. Ocenochnye (rejtingovyje) shkaly // Psihologicheskaja jenciklopedija. 2-e izd. / Pod red. R. Korsini, A. Auferbaha. SPb., 2006.