
ОБЩИЕ ВОПРОСЫ ОБРАЗОВАНИЯ

УДК 37.013

DOI: 10.17853/1994-5639-2021-6-43-83

ОЦЕНИВАНИЕ В ОБРАЗОВАНИИ: СОВРЕМЕННЫЕ ТЕНДЕНЦИИ, ПРОБЛЕМЫ И ПРОТИВОРЕЧИЯ (ОБЗОР НАУЧНЫХ ПУБЛИКАЦИЙ)

И. Б. Шмигирилова¹, А. С. Рванова²

*Северо-Казахстанский государственный университет им. М. Козыбаева,
Петропавловск, Республика Казахстан.
E-mail: ¹irinankzu@mail.ru; ²alla_rv@mail.ru*

О. В. Григоренко

*Сибирский государственный университет геосистем и технологий,
Новосибирск, Россия.
E-mail: ogrigorenko2311@mail.ru*

Аннотация. *Введение.* Современные взгляды на цели и содержание образования во всем мире меняют точку зрения и на оценивание в образовании: значительно расширяются его цели и функции, внедряются новые формы и инструменты оценивания, пересматриваются роли обучающихся и педагогов в оценочной практике. Мировые тенденции в области образовательной оценки находят свое отражение в национальных системах образования, многие из которых осуществили переход на критериальное оценивание. Этот процесс сопровождается стремительным ростом объема научно-методических публикаций по данной проблеме, уменьшая возможности ученых и практикующих педагогов осмыслить совокупные свидетельства в данной области исследований, что представляет важным на этапе реформирования. Решению данной проблемы может способствовать аналитический обзор источников, представляющий обобщение международного опыта развития оценивания в образовании.

Цель статьи – на основе анализа научных публикаций осмыслить и обобщить современные тенденции, сложившиеся в мировой теории и практике оценивания образовательных результатов обучающихся, выявить проблемы и противоречия как потенциальную основу возникновения рисков в оценочных практиках, определить актуальные направления исследований по совершенствованию оценивания в образовании.

Материалы и методы. Материалами для исследования послужили научные публикации последних 30 лет, раскрывающие вопросы образовательного оценивания. Отбор источников осуществлялся методами дескриптивного и библиометрического анализа. В работе с содержанием материалов применялись сравнительный и аспектный анализ.

Результаты. Широкий диапазон целей оценивания подтверждает значимость оценки как компонента образовательной системы, но при этом влечет за собой возникновение противоречий между отдельными целями, которые проявляются в оценочных дихотомиях: формирующее и суммирующее оценивание, оценка по норме и оценка по критерию. Несмотря на противопоставление формирующей и суммирующей оценок, проведенный анализ свидетельствует о необходимости их интеграции. Многообразие теоретических подходов к толкованию форм оценивания, неоднозначное определение и использование терминологии, множественность моделей критериального оценивания, качественно различных подходов к разработке критериев и созданию на их основе оценочных рубрик являются источниками рисков, проявляющихся в практике образовательного оценивания. Кроме того, препятствием на пути обеспечения эффективного оценивания учебных достижений обучающихся является недостаточный уровень компетентности педагогов как оценщиков, неготовность самих учащихся к принятию на себя ответственности за результаты своего обучения.

Научная новизна. С критических позиций представлены идеи и обобщены результаты дискуссий по проблемам оценки образовательных результатов обучающихся в научных источниках. Проведенный анализ может выступить основой для дальнейших исследований с целью построения упорядоченной теоретической базы оценивания в образовании.

Практическая значимость. Анализ плюсов и минусов международного опыта в области оценивания в образовании может послужить обретению ученым и практикующими педагогами более объективного взгляда на новые подходы к оцениванию образовательных результатов обучающихся, что будет способствовать минимизации проявления возможных рисков при разработке инструментов и процедур оценивания и внедрении их в образовательную практику.

Ключевые слова: оценивание в образовании, реформы оценивания, формативная оценка, оценка для обучения, суммирующая оценка, критериальное оценивание, нормированное оценивание, оценочные рубрики, оценочная грамотность педагога.

Благодарности. Исследование проведено при финансовой поддержке Комитета науки Министерства образования и науки Республики Казахстан (грант № AP08956027 «Инновационные подходы к оцениванию учебных достижений обучающихся в условиях обновления содержания математического образования»). Авторы также выражают искреннюю благодарность рецензентам, замечания которых были полезны в работе над статьей.

Для цитирования: Шмигирилова И. Б., Рванова А. С., Григоренко О. В. Оценка в образовании: современные тенденции, проблемы и противоречия (обзор научных публикаций) // Образование и наука. 2021. Т. 23, № 6. С. 43–83. DOI: 10.17853/1994-5639-2021-6-43-83

ASSESSMENT IN EDUCATION: CURRENT TRENDS, PROBLEMS AND CONTRADICTIONS (REVIEW OF SCIENTIFIC PUBLICATIONS)

I. B. Shmigirilova¹, A. S. Rvanova²

*M. Kozybaev North Kazakhstan State University, Petropavlovsk, Republic of Kazakhstan.
E-mail: ¹irinankzu@mail.ru; ²alla_rv@mail.ru*

O. V. Grigorenko

*Siberian State University of Geosystems and Technologies, Novosibirsk, Russia.
E-mail: ogrigorenko2311@mail.ru*

Abstract. *Introduction.* Modern views on the goals and content of education around the world are changing the viewpoints on educational assessment: the goals and functions of assessment are expanded, new assessment forms and tools are being introduced, the roles of students and teachers in assessment practice are revised. Global trends in educational assessment are reflected in national education systems, many of which have made the transition to criteria-based assessment. This process is accompanied by a rapid increase in the volume of scientific and methodological publications on this issue, reducing the ability of scientists and practising educators to comprehend the cumulative evidence in this area of research, which is important at the stage of reform. The solution to this problem can be facilitated by an analytical review of sources, representing a generalisation of international experience in the development of educational assessment.

Aim. Based on the analysis of scientific publications, the current research is *aimed* to comprehend and generalise the current trends prevailing in the world theory and practice of assessment of student learning outcomes, to identify problems and contradictions as a potential basis for the occurrence of risks in assessment practices, to determine the current research directions towards the improvement of assessment in education.

Methodology and research methods. The scientific papers of the last thirty years, revealing the issues of the education process assessment, have been used in the research. Sources were selected through the methods of descriptive and bibliometric analysis. The comparative and aspect-based analyses were employed in the research.

Results. The wide range of assessment goals confirms the significance of the assessment as the education component; however, it results in some contradictions between the goals, which appear in estimated dichotomies: formative and summative assessment, norm-based and criteria-based assessment. Despite the contrapositive between the formative and summative assessments, the analysis conducted shows the necessity of their integration. The variety of theoretical approaches to the interpretation of assessment forms, the ambiguous definition and use of terminology, the multiplicity of criteria-based assessment models, qualitatively different approaches to the development of criteria and the creation of assessment rubrics on their basis are sources of risks in the practice of educational assessment. In addition, an obstacle to ensuring effective assessment of students' educational achievements is the insufficient level of competence of teachers as assessors, the unpreparedness of students themselves to take responsibility for the results of their learning.

Scientific novelty. From a critical standpoint, the ideas are presented and the results of discussions on the problems of assessing the educational results of students in scientific sources are summarised. The analysis performed can serve as a basis for further research to build an ordered theoretical framework for assessment in education.

Practical significance. An analysis of the pros and cons of international experience in the field of educational assessment can help scientists and practising teachers gain a more objective view of new approaches to the assessment of student learning outcomes, thereby contributing to the minimisation of the manifestation of possible risks in the development of assessment tools and procedures and their implementation in educational practice.

Keywords: educational assessment, assessment reforms, formative assessment, assessment for learning, summative assessment, criteria-based assessment, norm-based assessment, assessment rubrics, assessment literacy of teachers.

Acknowledgements. The present research was performed under the financial support of the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (grant No. A08956027 "Innovative Approaches to the Assessment of Educational Achievements of Students in Terms of Updating the Content of Mathematical Education"). The authors also express their sincere gratitude to the reviewers, whose comments were useful in the work on the article.

For citation: Shmigirilova I. B., Rvanova A. S. Grigorenko O. V. Assessment in education: Current trends, problems and contradictions (review of scientific publications). *The Education and Science Journal*. 2021; 23 (6): 43–83. DOI: 10.17853/1994-5639-2021-6-43-83

Введение

Современные экономические и социокультурные мировые тенденции нашли свое отражение в национальных системах образования: корректируются образовательные цели, выражающие представление об идеальных результатах обучения на всех уровнях; пересматривается содержание образования; образовательная практика ориентируется на инновационные технологии и т. д. Изменения, коснувшиеся различных компонентов образовательной системы, не могли не отразиться на подходах к оценке образовательных достижений обучающихся. Нельзя не согласиться с М. Тагас, когда она говорит, что оценка «пожалуй, центральный, а также самый обременительный и затратный аспект работы многих педагогов» [1, с. 365].

Изменение взглядов на преподавание и обучение в сочетании с новыми требованиями общества к образовательным результатам постепенно меняют культуру обучения. В этих условиях необходимо и оценку увидеть новыми глазами. Неслучайно двадцать прошедших лет этого тысячелетия отмечены реформами в области оценки во всем мире.

Однако обзоры и материалы по проблемам оценивания, представленные Организацией экономического сотрудничества и развития (ОЭСР) и

ЮНЕСКО, прямо указывают на то, что только немногие страны считают исследование проблемы оценивания в образовании не отдельным (параллельным реформе) процессом, а ее важнейшим компонентом. При этом сами реформы не всегда в полной мере учитывают потребности и возможности образовательной практики, поэтому политика и практика оценивания являются предметом многочисленных споров. Административные, социальные и образовательные дискуссии, отраженные в подходах к оценке, сделали ее неоднозначным инструментом поддержки образования: нерешенность многих теоретических вопросов влечет и методологические, административные, профессиональные проблемы. Как отмечает член Британской академии S. Ball, «политика оспаривается, интерпретируется и принимается в различных сферах практики, а риторика, тексты и смысл политиков не всегда прямо и явно переводятся в институциональные практики... Они (практики) склонны неявно сопротивляться и неверно понимать, а в некоторых случаях просто оказываются неработоспособными» [2, с. 10]. Поиск адекватного выхода из сложившейся ситуации актуализирует необходимость всестороннего системного анализа тенденций, сложившихся на сегодняшний день в мировой теории и практике оценивания образовательных результатов.

В силу большого интереса к вопросам оценивания в образовании стремительно растет объем научно-методических публикаций по данной проблеме, что уменьшает возможности ученых и практикующих педагогов оценить совокупные свидетельства в данной области исследований. Поэтому аналитический обзор публикаций является важным и даже необходимым этапом научного процесса. R. W. Palmatier с соавторами отмечают, что такие обзоры «обеспечивают платформу для новых концепций, выявляют несоответствия в существующих исследованиях, синтезируют различные результаты и, как правило, дают другим ученым „современный“ снимок предметной области» [3, с. 1].

Ранее выполненные обзоры публикаций по вопросам образовательного оценивания обращены к рассмотрению значимости оценивания для образовательной практики (J. Baird, T. N. Hopfenbeck, P. E. Newton, A. T. Steen-Utheim [4]); проблемных аспектов формативной оценки и ее сравнения с суммативным оцениванием (R. E. Bennett [5], A. M. Lau [6]); использования информационных технологий в формативном оценивании (J. W. Gikandi, D. Mogowa, N. E. Davisa [7], M. A. Бодоньи [8]); оценки компетенций и навыков XXI века (D. Pepper [9], H. Siarova, D. Sternadel, R. Mašidlauskaitė [10]).

Вопросы образовательного оценивания в различных мировых практиках также в большей или меньшей степени затрагиваются в сравнительных обзорах, исследующих категорию «качество образования» (Т. В. Потемки-

на, Т. И. Пуденко [11]), подходы к оценке эффективности работы учителей (С. DeLuca, D. LaPointe-McEwan, U. Luhanga [12], О. В. Темняткина, Д. В. Токменинова [13]).

Однако направленность этих работ не охватывает многих сторон оценивания как компонента образовательного процесса, максимально объективное понимание которых необходимо для системного изучения данного проблемного поля.

Цель статьи – на основе анализа научных публикаций осмыслить и обобщить современные тенденции, сложившиеся в мировой теории и практике оценивания образовательных результатов обучающихся, выявить проблемы и противоречия как потенциальную основу возникновения рисков в оценочных практиках, определить актуальные направления исследований по совершенствованию оценивания в образовании.

В. Д. Шадриков и И. А. Шадрикова, описывая процесс педагогического оценивания, отмечают, что «по отношению к ученику данный вид оценивания выступает в форме внешнего оценивания» [14, с. 13], однако современное понимание оценивания в образовании выходит за рамки педагогической оценки, интегрируя в себе в том числе и оценочные практики само- и взаимооценивания, поэтому в рамках данной статьи этот процесс будет обозначаться наиболее часто встречающимся в источниках Scopus и Web of Science понятием *educational assessment*, которое в зависимости от контекста интерпретируется как «образовательное оценивание / образовательная оценка». Кроме того, поскольку в различных научных источниках для обозначения процесса оценивания наряду с термином «оценивание» как его синоним используется понятие «оценка», в контексте данного исследования будем придерживаться того же подхода.

Обращаясь к мировому опыту исследования оценки в образовании, в котором все большая роль отводится критериальному оцениванию и сопутствующим ему оценочным мероприятиям и процедурам, авторы данной статьи прежде всего ориентировались на аспекты, непосредственно имеющие отношение к использованию такого оценивания в практике обучения.

Материалы и методы

Основными материалами для данной работы послужили статьи, предоставляемые базами данных Scopus и Web of Science, за период с 1990 по 2020 год.

На первом этапе использовался дескриптивный метод. Отбор источников осуществлялся по ключевым словам: *educational assessment*, *classroom assessment*, *assessment purposes*, *summative assessment*, *assessment for*

learning, formative assessment, assessment to support learning, criteria-based assessment, norm-based assessment, assessment rubrics.

На втором этапе с целью уменьшения массива источников осуществлялись просмотр аннотаций, сопоставление ключевых слов. Так, на данном этапе обнаружилось, что в качестве ключевых слов, кроме тех, в соответствии с которыми осуществлялся первичный отбор источников, достаточно часто используются термины *assessment methods*, *conceptions of assessment*, *learning outcomes*, *student achievement*, *educational policy*, *assessment capability*, *achievement standards*, *academic standards*, *norm reference testing*, *criterion reference testing*, *design of rubrics*, *holistic rubrics*, *analytic rubrics*.

На третьем этапе осуществлялся просмотр содержания статей для установления необходимости дальнейшего их использования. На этом этапе был осуществлен поверхностный библиометрический анализ, в результате которого возникла необходимость интернет-поиска дополнительных источников, не включенных в БД Scopus и Web of Science, но цитируемых многими авторами: статей, публикации которых относятся к более раннему периоду; работ, представленных в материалах конференций; справочников и пособий по практике оценки; документов, освещающих международные оценочные тенденции, и т. п. Результатом данного этапа стал отбор 273 источников, которые наиболее часто цитировались другими авторами, представляя более распространенную точку зрения, либо содержали факты или идеи, которые не встречались у других авторов.

Четвертый этап был посвящен изучению, сравнительному и аспектно-му анализу отобранных материалов. Использование элементов голографического метода на данном этапе позволило не только представить основные структурные компоненты проблемы образовательного оценивания, но и отметить ее взаимосвязи с системой обучения в целом.

Результаты исследования и их обсуждение

Оценивание вне зависимости от его вида понимается в широком смысле как процесс формирования суждения о качестве и степени успеваемости или производительности учащегося. В этой связи термин *evaluation* соотносится с суждениями об эффективности школ, школьных систем, а категория *assessment* используется, главным образом, в отношении описания результативности учебных мероприятий. Этот термин охватывает оценивание в классе, а также масштабные стандартизированные внешние тесты и экзамены.

Оба приведенных термина описывают процесс генерирования и интерпретации доказательств результативности обучения для какой-либо

цели. Как отмечает W. Harlen [15], теоретически различаются цели оценивания как причины для проведения самой процедуры оценки и цели использования результатов оценивания. Сам термин «цель оценки» интерпретируется различным образом: на уровне оценочного суждения, выраженного в некотором диапазоне, что характерно для официальных документов; на уровне решения, принимаемого на основе оценочного суждения; на уровне воздействия, связанного с разработкой и использованием системы оценивания непосредственно в учебном процессе. Подчеркивая, что достижение различных целей оценки требует зачастую процедур, существенно отличающихся друг от друга, Р. Е. Newton [16] выделил и охарактеризовал 18 целей образовательного оценивания. Е. Archer [17], проанализировав ряд статей, разбивает множество целей на три группы: оценивание для обеспечения подотчетности; оценивание с целью выявления прогресса в обучении или сертификации выпускника; оценивание для поддержки обучения. При этом общую цель оценки автор усматривает в помощи «всем заинтересованным сторонам увидеть более широкую образовательную картину и определить срочность и важность распределения ресурсов в соответствии с текущими потребностями образовательных систем» [17, с. 2].

Резюмируя идеи и выводы Е. Archer, Р. Е. Newton, W. Harlen в отношении целевых ориентиров оценивания, можно отметить:

- 1) различие целей оценки особенно значимо в аспектах разработки ее принципов, технологий и инструментов;
- 2) стремление к системе оценивания, поддерживающей ряд различных целей (при условии, что они не являются логически несовместимыми), потребует четкого определения их приоритетности;
- 3) основная цель и (насколько это возможно) другие высокоприоритетные цели будут определять оптимальные проектные характеристики системы оценивания, которые позволят пользователям делать достаточно достоверные выводы из результатов, но при этом безусловный приоритет одной из целей может негативно сказаться на качестве образования.

Таким образом, разнообразие целей оценивания, с одной стороны, свидетельствует о значимости оценки как компонента образовательной системы, с другой стороны, акцентирует внимание на возможных противоречиях между отдельными целями, разрешение которых является важной задачей образования.

Выделение различных видов оценки также в значительной мере связано с целями ее использования. В этой связи одним из самых обсуждаемых вопросов в литературе, посвященной оцениванию, является вопрос о дихотомии формирующей и суммирующей оценок. Несмотря на отдельные нюансы в определениях этих оценочных практик, Р. Black, D. Wiliam [18–20],

D. D. Dixon, F. C. Worrell [21], М. А. Пинская [22] интерпретируют формирующее оценивание как процесс, неотделимый от обучения, основным предназначением которого является определение текущих успехов учащихся и содействие достижению ими намеченных образовательных целей. При этом в определении суммирующего оценивания акцентируется факт фиксации (констатации) результатов обучающегося по итогам освоения конкретного содержания или за определенный временной промежуток.

Истоки различий между формирующими и суммирующими практиками оценивания принято связывать с работами М. Scriven¹ и авторского коллектива под руководством В. Bloom². При этом М. Scriven формирующую оценку соотносит с процессом улучшения и разработки учебных программ, а В. Bloom этот вид оценивания рассматривает как составляющую обучения, необходимую для движения учащихся к умению, мастерству. Несмотря на то что трудно избежать ссылки на формирующее или суммирующее оценивание, как если бы они были различными формами или видами оценки, Р. Е. Newton [16], Р. Black, D. Wiliam [18], W. Harlen [23] считают, что дихотомия формирующей и суммирующей оценок относится не к ее видам, а к целям оценивания. Отсюда наряду с категориями «формирующая оценка» и «суммирующая оценка» появляются термины, которые часто используются как синонимы этих понятий: «оценка для обучения» (AfL) и «оценка обучения» (AoL).

При сравнении формирующей и суммирующей оценочных практик исследователи подчеркивают:

- «оценка для обучения» действительно является целью, «оценка обучения» является не целью, а констатацией предмета оценки, которая может иметь ряд целей и применений (Р. Е. Newton [16]);
- в случае формирующей оценки есть только одно основное использование информации – это помощь в обучении, и если оценка не используется для этой цели, то и не может считаться формирующей. И, напротив, результаты итогового оценивания учебных достижений (или «оценки обучения») могут быть использованы несколькими способами (Р. Е. Newton [24]);
- основное различие между ними относится к цели и следствиям (т. е. последствиям, которые оценка может иметь для ученика) (D. R. Sadler [25]);
- рассмотрение оценки не только как продукта (результата) оценочных мероприятий, но и как процесса оценивания (цикла, связывающего три

¹Scriven M. The methodology of Evaluation // Perspectives of Curriculum Evaluation (AERA Monograph Series on Curriculum Evaluation, No. 1) / R. Tyler, R. Gagné, M. Scriven (eds.). Chicago: Rand McNally; 1967. P. 39–83.

²Bloom B. S., Madaus G. F., Hastings J. T. Handbook on Formative and Summative Evaluation of Student Learning. New York: McGraw-Hill; 1971. 925 p.

фазы: сбор доказательств, интерпретацию доказательств, то есть вынесение суждения, и предпринимаемые действия) указывает на то, что «оценка однозначно является суммирующей, если останавливается на суждении. Тем не менее оценка не может быть исключительно формирующей, без предшествующего суммирующего суждения» (M. Taras [26, с. 468]);

– формирующее оценивание ориентируется на качественное, развернутое суждение, а суммирующая оценка – на количественное выражение степени достижения цели (D. D. Dixon, F. C. Worrell [21]);

– приоритетом суммирующей оценки определяется надежность, а для формирующей оценки – достоверность и полезность (W. Harlen [23]).

Суммирующие (итоговые) оценки часто представляют собой оценки «с высокими ставками», поскольку результаты таких оценочных мероприятий используются для суждения о квалификации или качестве образования отдельного обучающегося или их групп, эффективности работы образовательного учреждения; принятия решения о возможности (или невозможности) обучения учащегося в конкретном учебном заведении, по конкретному направлению и т. д.; решения о сертификации; формирования национальных и международных рейтингов; принятия решения об изменении образовательной политики¹. Однако, как указывает А. М. Lau [6], преимущественная ориентация на суммирующее оценивание может привести к поверхностному подходу в обучении, поскольку сопряжена с риском того, что учителя и учащиеся сосредоточат свое внимание только на тех разделах и видах заданий, которые будут подвергнуты итоговой оценке. Критикуя суммирующее оценивание, автор также указывает на его малопригодность для улучшения образовательной практики. Заметим, что такую критику нельзя считать корректной: во-первых, она не учитывает основную функцию суммирующей оценки; во-вторых, анализ результатов суммирующего оценивания и принятые в связи с этим решения могут служить и образовательным целям повышения качества обучения.

Суммирующую оценку зачастую соотносят с практикой стандартизованного тестирования, результаты которого оказывают влияние на образовательную политику. Преимущества стандартизированного тестирования D. Перрег связывает с его надежностью, легкой интерпретацией результатов, независимостью от конкретного эксперта или оценщика [9], а А. Sjögren также с его ролью в обеспечении внешней мотивации обучения [27]. В то же время стандартизованные тесты как инструменты отчетности и реформирования образовательной политики и практики подвергаются критике

¹OECD Synergies for Better Learning. An International Perspective on Evaluation and Assessment. Paris: OECD Publishing. 2013. 18 p. Available from: http://www.oecd.org/education/school/Synergies%20for%20Better%20Learning_Summary.pdf (date of access: 25.07.2020).

за то, что даже самые известные тесты с высокими ставками не соответствуют поставленной цели точной оценки уровня успеваемости учащихся (G. Butt [28]); такие тесты обычно концентрируются на фактической информации и не дают возможности проявить более высокий уровень мышления (E. Silfver, G. Sjoberg, A. Bagger [29]); стандартизованное тестирование не принимает во внимание особенностей образовательной практики и личных обстоятельств обучающихся, влияющих на итоги тестирования (E. Silfver, G. Sjoberg, A. Bagger [28], S. Courtney [30]).

Кроме того, общей отрицательной тенденцией многих мировых образовательных практик является неправомерное применение результатов стандартизированных тестов, которые, как верно замечает в этой связи В. А. Болотов, становятся основанием для «оценки эффективности деятельности учителя, школы, муниципалитетов, регионов и даже губернаторов» [31, с. 291].

Программы международной оценки учебных достижений, которые тоже, как правило, осуществляются с использованием стандартизированных тестов, могут обеспечить сравнительные данные, отличные от данных национальных оценок. Однако при принятии решений на основе результатов такого оценивания необходимо помнить, что содержание тестов, используемых в международной оценке, может не соответствовать учебному плану каждой отдельной страны. А в силу того, что тесты международной оценки, как и другие стандартизированные тесты, имеют ряд недостатков, нельзя быть уверенным в том, что подходы, принятые за основу в международных исследованиях, могут быть одинаково эффективны в разных странах. Непринятие во внимание этого аспекта может привести к неправильным решениям в образовательной политике и, следовательно, нанести вред образовательной практике.

Формирующая оценка, как это было отмечено ранее, характеризуется с позиции совершенствования практики преподавания и прямой поддержки учащихся в улучшении их обучения. В работах О. Н. Шаповаловой, Н. Ф. Ефремовой [32], Т. В. Четвертных [33] особо подчеркивается существенный потенциал формирующего оценивания в достижении как предметных, так и метапредметных образовательных результатов.

В то же время Р. Black, D. Wiliam [18, 19] особо указывают, что основным фактором, влияющим на успешность формирующего оценивания, является осознание учеником разрыва между его текущими образовательными достижениями и их желаемым уровнем и, что важно, наличие реакции обучающегося, выраженной в предпринятых им действиях для устранения этого разрыва. Роль учителя в этой последовательности заключается в обеспечении участия обучающихся в оценочных мероприятиях через интерпре-

тацию образовательных целей, целей оценивания и его результатов в формате, понятном учащимся, налаживание образовательного взаимодействия, обратной связи, предоставление обучающимся возможностей для само- и взаимооценок. D. Cisterna, A. W. Gotwals [34], J. Hill, H. West [35], описывая приемы, методы и стратегии формирующей оценки, в качестве ключевого фактора ее эффективности выделяют качество обратной связи. Более того, М. Taras [26], указывая на связь формирующей оценки и обратной связи, предлагает формулу:

суммирующая оценка + обратная связь = формирующая оценка.

Ряд исследователей различают «оценку для обучения» и «оценку как обучение» (AaL) как два вида оценки, организуемой в образовательных целях. Идея «оценки как обучения», исходит из того, что понимание учащимися целей обучения, вовлеченность их в оценку собственных достижений и достижений других учащихся, принятие решений об улучшении собственного обучения само по себе является познавательным актом.

Сравнение методов формирующего и итогового оценивания и критика последнего за отстраненность от учебного процесса и связанные с ними страхи и тревоги учащихся породили в образовательной литературе начала XXI века тенденцию, отмеченную М. Taras [26]: формирующая оценка все чаще изображалась как «хорошая» оценка, к которой должны стремиться учителя, тогда как суммирующая оценка – это «плохая» оценка, на которой преподаватели должны минимизировать свое внимание, вместо этого переходя к формирующему оцениванию.

Однако такое положение нельзя признать единственно верным. Примечательной в этой связи является работа R. P. Phelps [36], в которой он анализирует материалы, собранные в рамках проекта REAFISO (Review on Evaluation and Assessment Frameworks for Improving School Outcomes). Автор, критически оценивая представленные доказательства того, что оценки с высокими ставками плохие, потому что они «препятствуют инновациям», подчеркивает, что такой вывод ориентируется только на одну сторону истории, поскольку полностью игнорирует значительную массу источников в несколько сотен исследований.

В свою очередь, несмотря на то что эмпирические данные показали положительное влияние формирующих методов оценки на результаты обучения, в литературе появляются предупреждения о том, что заявления об эффективности формирующей оценки следует рассматривать с осторожностью из-за ограничений методологии проведенных исследований. Так, например, Н. Torrance [37] отмечает, что воздействие процессов формирующей оценки на самом деле может быть и «деформирующим». R. E. Bennett [5] в статье, являющейся одной из наиболее цитируемых работ по данной тематике по-

следнего десятилетия и представляющих собой критический обзор формирующего оценивания, утверждает, что существующие определения этой оценки допускают такое широкое разнообразие реализаций, что ее эффекты будут широко варьироваться от одной образовательной практики и группы учащихся к другим. Автор, отмечая отсутствие всеобъемлющей теории формирующего оценивания, считает, что, разрабатывая подходы к его реализации в образовательной практике, необходимо учитывать особенности конкретных предметных областей и другие характеристики учебного процесса. Его мнение поддерживается и другими исследователями: D. Cisterna, A. W. Gotwals [34], J. Harrison с соавторами [38]. M. Henderson, T. Ryan, M. C. Phillips [39].

R. Dann [40], рассматривая «оценку как обучение» с социокультурных позиций, отмечает расширение роли формирующей оценки, поскольку участие обучающихся в оценочных мероприятиях является «неотъемлемой частью обучения», но предостерегает от заблуждений о ее неограниченных возможностях в направлении повышения качества обучения. Он подчеркивает, что оценочного акта учителя, состоящего в выявлении и отражении в виде обратной связи сильных и слабых сторон обучения учащегося, недостаточно для развития значимых образовательных результатов, и аргументированно доказывает необходимость специальных исследований того, как обучающиеся воспринимают и интерпретируют обратную связь. Исследователи M. Henderson, T. Ryan, M. C. Phillips [39], R. Dann [40] важнейшим фактором эффективности формирующей оценки также считают готовность учащихся к обучающему и оценочному взаимодействию.

Анализируя практику формирующего оценивания, нельзя обойти вниманием еще одну проблему, которая обозначена у ряда авторов (Л. В. Вилкова [41], Т. В. Четвертных [42], И. Б. Шмигирилова, А. С. Рванова, А. А. Таджигитов [43], M. T. Florez, P. Sammons [44], J. L. Herman, E. Osmundson, D. Silver [45]): приверженность учителей парадигме итоговой оценки и традиционным оценочным практикам и, как следствие, затруднения педагогов в интерпретации результатов оценивания в формате обратной связи, определении дальнейших шагов в обучении, направленных на ликвидацию разрыва между учебными результатами и образовательными целями, а также склонность к формальному использованию методов формирующего оценивания вследствие отсутствия у педагогов специальной подготовки и опыта их применения. Справедливости ради заметим, что в традициях советской школы формирующее оценивание без указания данного термина всегда имело место, по крайней мере, в практике компетентных и любящих профессию педагогов.

Широта дискуссий о взаимосвязях суммирующей и формирующей оценок, их значимости в различных системах отсчета не приводит к ре-

шению вопроса о приоритетности между ними. В этой связи, по мнению исследователей, выход видится в синергии рассматриваемых оценочных практик, что поможет сохранить их преимущества и компенсировать недостатки. Более того, Н. Siarova, D. Sternadel, R. Mašidlauskaitė [10] считают, что интеграция различных подходов к оценке позволит найти сбалансированную и последовательную практику оценки ключевых компетенций, а также помочь в более полном учете физиологических и социальных возможностей и потребностей учащихся, обеспечивая реализацию их потенциального прогресса. При этом, по мнению L. C. Velasco, J.-C. Tojar [46], оценка компетенций требует системного подхода к разработке модели оценивания, комплексного использования методов оценки, разнообразия инновационных инструментов, обеспечивающих получение доказательств в отношении приобретения этих компетенций. Пример такой модели, в которой используются паттерн-дизайны оценочных средств, описан в работе Н. В. Ефремовой [47].

Еще одна дихотомия оценочных практик – нормированное оценивание и оценка на основе критериев – связана с тем, что суждения о качестве работы обучающихся, об их учебных достижениях, а также интерпретация таких суждений в форме числовых или буквенных маркеров всегда осуществляются в некоторой системе координат, точка отсчета и единицы измерения в которой могут быть представлены по-разному.

В общепринятом понимании нормативно-ориентированное оценивание – это оценивание, состоящее в выявлении степени соответствия индивидуальных результатов обучающихся заранее установленной норме. Однако толкование нормативной оценки образовательных результатов в анализируемых источниках имеет различные нюансы. В одном случае в качестве основного признака такого оценивания упоминается только ранжирование обучающихся на основе сравнения их результатов с результатами сверстников для принятия определенных типов решений, например, для того чтобы выявить лучшего кандидата по сравнению с другими кандидатами. При этом распределение оценок в группе учащихся может быть очень далеко от нормального. Этот вид оценивания подвергался существенной критике: интерпретация такой оценки ничего не говорит о реальных уровнях достижений отдельных обучающихся, а только об их относительных достижениях в группе. Однако, как справедливо заметил D. Wiliam [48], большая часть критики, как правило, завышает количество ссылок на нормы, фактически используемых на практике, поскольку вид оценки, когда результаты одного учащегося сравниваются с результатами других членов данного учебного сообщества, следует считать не нормированной, а групповой (когортной) оценкой.

В другом случае нормой считается среднее значение по выборке обучающихся, а пропорции каждого вида оценок жестко контролируются точно в соответствии с кривой нормального распределения. D. Sadler отмечает, что в университетах, где оценка в соответствии с кривой нормального распределения является официальной политикой, ее обычно защищают, указывая на намерения «поддерживать стандарты», уменьшая тенденцию к инфляции оценок [49, с. 187]. Одна из опасностей в этом случае состоит в том, что строго нормированная оценка является самокорректирующейся и не может продемонстрировать, насколько хорошо разработан курс, организован процесс обучения или насколько в целом успешны учащиеся. Как замечает R. Dann [50], например, при использовании этого подхода обучающийся может быть оценен на высший балл, в то время как оценка того же уровня образовательных достижений в предыдущем году была гораздо ниже, поскольку имелись учащиеся, показавшие лучшие результаты. Для того чтобы преодолеть эту тенденцию для нормативного оценивания, С. Wikström [51] предлагает использовать задания, которые различаются по сложности.

Критикуя эту оценочную модель, W. Hornby [52], кроме ряда ее статистических допущений, отмечает ее «ложную точность» и несправедливость, выраженную в том, что «пропорции учащихся, получающих различные оценки, определяются заранее как политика, которая, как отмечают M. Henderson, T. Ryan, M. C. Phillips, затем применяется к классу учащихся» [39, с. 186] независимо от абсолютных результатов учащегося. Кроме того, последние эмпирические исследования L. Ling, B. Yuen, W. Loo, C. Prinsloo [53] доказывают, что обучающиеся склонны считать отрицательным влияние такого оценивания на обучение.

Модель оценки, которую также считают нормированной, соединяющая два этих подхода, состоит в том, что обучающиеся оцениваются в сравнении друг с другом, но распределение оценок, как правило, не является жестко установленным: пропорции каждого вида оценок определяются, по существу, произвольно с отклонением от нормального распределения в ту или иную сторону. Считается, что именно такая модель оценивания с некоторым сдвигом в положительную сторону была в традициях советской школы, поэтому именно эту модель в отечественных источниках принято называть традиционной. В качестве нормы при этом выступали обязательные результаты обучения, которые устанавливались в нормативных документах.

Термин «оценка по критерию» был введен R. Glaser, D. J. Klaus¹ более 50 лет назад. Наиболее развернутое определение критериального оценива-

¹Glaser R., Klaus D. J. Proficiency measurement: Assessing human performance // Psychological principles in system development / R. Gagne (eds.). New York: Holt, Rinehart and Winston, 1962. P. 419–474.

ния представлено в работах А. А. Красноборовой [54]. Оценивание на основе критериев «заключается в сравнении достижений учащегося с четко определенными, коллективно выработанными, заранее известными всем участникам процесса критериями» [54, с. 92]. Оценивание, основанное на критериях, было и остается широко обсуждаемым вопросом. В анализируемых работах обнаруживается более 20 различных определений категории «критерий», многие из которых отличаются друг от друга лишь незначительными деталями и в целом соответствуют этимологии данного понятия: от греческого κριτήριον – «средство для оценки». Авторы определяют критерий как отличительное свойство или характеристику любого объекта, по которой можно судить или оценивать его качество или может быть принято решение или классификация. Аргументы, которые приводятся в пользу использования оценки на основе критериев, приведенные различными авторами, сводятся к двум основным положениям, указывающим на образовательную и этическую основу этого вида оценивания:

1) оценка на основе критериев каждого обучающегося не зависит от того, как ту же работу выполнил другой обучающийся;

2) оценка на основе критериев имеет перспективную цель, заключающуюся в том, чтобы дать возможность обучающимся использовать знание критериев оценки его работы для разумного и правильного ее выполнения.

Указанные преимущества оценивания, основанного на критериях, проявляются в ряде факторов, которые, по мнению А. А. Красноборовой [54], М. А. Пинской, А. В. Иванова [55], а также Ж. А. Абековой с соавторами [56], определяют повышение эффективности процесса обучения.

Примечательным является то, что нормативное оценивание, как правило, соотносится с суммирующей (итоговой, констатирующей) оценкой, в то время как критериально-ориентированное оценивание исследуется в аспекте формирующей оценки либо с позиций и суммирующей, и формирующей оценок.

Внедрение критериального оценивания породило и целый ряд исследовательских проблем, связанных с надежностью и достоверностью использования различных моделей оценки, основанной на критериях. D. Sadler [49] на основе анализа программных документов различных университетов Австралии, Канады, Новой Зеландии, Южной Африки, Швеции, Великобритании, США и работ по проблемам оценивания, тестирования и аттестации в высшем образовании выявил четыре основные модели оценивания на основе критериев (таблица 1), которые, в свою очередь, представляют широкий спектр вариантов и гибридов, отличающихся друг от друга незначительными деталями. При этом автор указывает, что многие образовательные организации используют идентичные или аналогичные модели, не

указывая на то, что они относятся к критериальному оцениванию.

И. Г. Липатникова [56], Е. В. Соколова [57], предлагая различные подходы к разработке критериев, указывают ориентацию на образовательные цели (модель 1), однако при внимательном рассмотрении описания самих критериев и особенностей их использования в процессе оценивания в авторских подходах обнаруживаются элементы и других моделей оценки.

Анализ моделей критериального оценивания, представленных в таблице, поддерживает утверждение D. Sadler [48] о том, что, с одной стороны, интерпретация критериев в этих моделях попадает под общее определение, а с другой стороны, наглядно демонстрирует, что на практике они представляют собой совершенно разные вещи. Так, например, вполне можно утверждать, что критериальное оценивание было присуще и практике советской школы, просто критерии определялись не каждым педагогом в отдельности, а рекомендовались в нормативных документах. В этом легко убедиться, сравнив примеры описания критериев оценки письменных работ, приведенные D. Sadler, и рекомендации по оценке письменных контрольных работ по математике из сборника нормативных документов «Математика в школе» за 1988 год:

– «отметка А ставится, если «работа очень высокого качества во всем; имеются явные доказательства владения предметом; материал организован логически; переход между различными частями является плавным, части взаимосвязаны и вся работа представлена почти безошибочно» (D. Sadler [48, с. 186]);

– «отметка 5 ставится, если работа выполнена полностью; в логических рассуждениях и обосновании решения нет пробелов и ошибок; в решении нет математических ошибок (возможна одна неточность, не являющаяся следствием незнания или непонимания учебного материала)»¹.

Если учесть поправку на контекст оцениваемой области, то совершенно очевидно просматривается сходство подходов в определении оснований для оценивания.

Характеризуя описанные выше модели, D. Sadler высказывает мысль, примечательную в аспекте разрешения противоречия между оценками, ориентированными на критерии или нормы, о том, что и в нормированной оценке можно разглядеть критерий, в качестве которого можно считать основу ранжирования достижений обучающихся по сравнению с другими учениками или соответствие кривой нормального распределения [49, с. 187]. С этим нельзя не согласиться: такое понимание критерия не противоречит общей сути термина.

¹ Математика в школе: сборник нормативных документов / Сост. М. Р. Леонтьева [и др.]. Москва: Просвещение, 1988. С. 42.

Таблица 1

Модели критериального оценивания D. Sadler [49] (обобщено авторами)

Table 1

Criteria-based assessment models by D. Sadler [49]
(summarized by the authors)

№ п/п	Название модели	Суть модели	Обоснование отнесения модели к критериальному оцениванию	Плюсы модели	Минусы модели
1	Достижение целей курса	Оценка указывает на то, насколько обучающиеся достигли целей изучения курса. Цели служат основой для критериев, но сами критерии, по сути, остаются неопределенными	Устанавливает связь между достижением целей курса и присужденными оценками без ссылки на образовательные достижения других учащихся	Может быть выражена в различных формах: общее изложение принципа; широкие устные заявления; двумерная таблица соответствия уровней А, В, С, D, Е и показателей «в основном цель достигнута», «в основном цель не достигнута» и др.	Обучающиеся затрудняются в установлении тесной связи между целями курса, преимущественно ориентированными на педагога, и конкретными предметами оценки. Требуется четкое определение того, достигнута цель или нет, при этом «многие, а возможно, большинство образовательных результатов не могут быть оценены как дихотомические состояния» [49, с. 181]
2	Общее достижение, измеряемое по сумме баллов	Оценки за выполнение различных работ складываются и проецируются на 100-балльную шкалу, которая делится на сегменты, чаще разного размера, в зависимости от количества желаемых оценок	Все обучающиеся, достигшие соответствующих пороговых значений, получают оценки, которые соответствуют им, и баллы распределяются независимо от успеваемости других обучающихся	Соблюдение единообразных оценок по всему учебному заведению обеспечивает высокую степень сопоставимости.	Использование числовых диапазонов создает впечатление определенности и точности, но оценочные баллы обычно напрямую не связаны с овладением конкретными знаниями или навыками.

№ п/п	Название модели	Суть модели	Обоснование отнесения модели к кри- териальному оцениванию	Плюсы модели	Минусы модели
				Если все уч-реждения со-блюдают одно и то же прави-ло распределе-ния оценок, до-стигается на-циональная со-гласованность академических стандартов	Некачественное выполнение од-ного задания, предназначенно-го для оценива-ния, может быть компенсировано отличным выпол-нением другого задания
3	Оценка, от-ражающая модели до-стижений	Модель схожа с предыдущей моделью, но при выставле-нии итоговой оценки могут приниматься во внимание такие аспекты, как последователь-ность в испол-нении задания; соотношение сложности и значимости оцениваемых заданий; уста-новка мини-мального уров-ня в качестве выполнения задания и т. п.	В широком смысле правило комбиниро-вания оценок за отдельные задания в ито-говой оценке и определяет кри-терии оценки	Преподаватель может разра-ботать инди-видуальные правила для комбинирова-ния наборов оценок при выполнении различных заданий для получения ито-говой оценки курса	Значительные трудности в практической ре-ализации модели. Разнообразие оценочных пра-вил (критериев) у различных преподавателей в одном учеб-ном заведении затрудняют сопоставление достижений обу-чающихся
4	Описание качествен-ных крите-риев	Формируются общие крите-рии, предна-значенные для характеристики выпускников. Педагоги указы-вают критерии, которые будут учитываться при вынесении суждения о ка-честве выполне-ния оценочного	Непосредствен-ная разработка и использова-ние критериев	При оценива-нии работ обу-чающихся ис-пользуется ма-трица оценива-ния, имеющая значительную диагностиче-скую ценность для учащего-ся. Критерии могут приме-няться ко всем заданиям	Как правило, используемые критерии «не отображаются каким-либо про-стым способом на цели курса» [49, с. 185].

№ п/п	Название модели	Суть модели	Обоснование отнесения модели к критериальному оцениванию	Плюсы модели	Минусы модели
		задания. Список критериев разворачивается в оценочной шкале (с числовой или словесной разметкой)		определенного типа или каждый конкретный вид заданий может быть снабжен различными наборами критериев	

Углубляя эту мысль, В. Lok, С. McNaught, К. Young [59] указывают, что, вместо того чтобы рассматривать эти формы оценки как дихотомию, необходимо их разумно сочетать через цикл обратной связи. Это, по мнению авторов, поможет избежать инфляции в учебном процессе, а также предоставит полезную информацию педагогам, учащимся и другим заинтересованным в такой информации лицам.

Еще одним источником путаницы понимания сущности критериального оценивания является множественность качественно различных подходов к разработке и представлению критериев оценки образовательных достижений, а также взаимозаменяемое использование терминов «стандарты» и «критерии», «критерии» и «дескрипторы». Например, G. Brown и его соавторы утверждают, что в оценке по критерию «важна только точка прохождения/провала» [60, с. 12], так как ее цель состоит в том, чтобы определить, выполнили ли обучающиеся определенные требования в ходе выполнения задания или нет, и в этом смысле критерий рассматривают как норму того, что можно ожидать от ученика, получившего оценку А, в отличие от тех, кто учится на В.

S. Barrie, A. Brew, M. McCulloch [61] в ходе опроса преподавателей и структурирования его результатов выявили семь различных концепций разработки представления критериев оценивания:

- 1) набор правил, которым необходимо следовать при выполнении работы;
- 2) характеристика образовательного продукта;
- 3) описание аспектов, подлежащих оценке;
- 4) личная реакция оценщика на работу обучающегося;
- 5) абстрактные заявления о качестве работы с использованием слов «хорошо», «эффективно», «интересно», «недостаточно» и т. п.;
- 6) абстрактные формулировки результатов, разъясняемые конкретными примерами (контекстно-ориентированная концепция);

7) сложные и подробные описания количественных измерений, характеризующих конкретный уровень достижения.

D. Wiliam [62, с. 61], вводя понятие «оценка со ссылкой на конструкцию», указывает, что, если заявленные описания совместно определяют уровень требуемой производительности (или, более точно, определяют оправданные выводы об уровне производительности), тогда оценка осуществляется на основе критериев; если такие заявления лишь иллюстрируют те виды логических выводов, которые оправданы, тогда оценка основывается на конструктивной ссылке. Таким образом, пятая и шестая концепции, а возможно, и четвертая, в списке S. Barrie [61] и его соавторов в соответствии с подходом D. Wiliam [62] не могут быть отнесены к критериальному оцениванию. И это вполне справедливо, поскольку одной из основных характеристик критериального оценивания, представленных в подавляющем большинстве источников, является четкость, прозрачность и единое толкование критериев. Указывая на значимость этой характеристики, S. Green [63] приводит высказывание A. Pollitt о том, что измерение с использованием критериев, допускающих различное понимание и истолкование, будет напоминать работу термометров, которые притворяются, что измеряют температуру по шкале Цельсия, на самом деле имеют свое понимание того, что есть точка замерзания, и что есть хороший летний день. Таким образом, встает вопрос: можно ли считать оценивание критериальным, если критерии допускают существенную разницу в их понимании?

W. J. Porham [64], исследуя историю критериальной оценки, отмечает четыре ошибки, искажившие практику ее использования:

- 1) неправильное представление о том, что ссылка на критерий относится к тестам, а не к интерпретациям тестов;
- 2) неправильное представление о том, что критерием является уровень производительности, а не область знаний или навыков;
- 3) неспособность измерить разумное количество доменов;
- 4) тенденция предоставлять слишком много или слишком мало деталей в описаниях доменов.

Указывая, по сути, на те же проблемы, D. Sadler объясняет их иначе: причина многих затруднений, связанных с разработкой критериев, состоит в том, что само это понятие является более широким по охвату и в этой связи более часто используемым, чем термин «стандарт». Это приводит к тому, что критерии упоминаются в соотнесении с оценкой и отметкой, охватывая как «истинные» критерии, так и то, что на самом деле является стандартами [49, с. 188]. Поэтому исследователь считает: вместо того чтобы предусматривать определенную интерпретацию критериев как значимую в контексте оценивания, было бы предпочтительным сделать четкое и последовательное

различие между критериями и стандартами [49, с. 187]. Представляя свое понимание стандарта, автор рассматривает его как «определенный уровень совершенства или достижения, или определенную степень любого качества как предписанный объект усилий или как признанную меру того, что является адекватным для какой-либо цели, установленной авторитетом, обычаем или консенсусом» [49, с. 189]. Анализ приведенного определения позволяет отметить: во-первых, такое толкование стандарт хотя и согласуется с одним из привычных значений этого термина, представленных в словаре С. И. Ожегова и Н. Ю. Шведовой, где он выступает как «образец, которому должно соответствовать, удовлетворять что-нибудь по своим признакам, свойствам, качествам»¹, однако делает упор, прежде всего, на уровень, степень этого соответствия; во-вторых, в указанной интерпретации стандарта просматривается и идея нормы как основы оценки.

Принятие определения стандарта в понимании D. Sadler также приводит к мысли о том, что «критерий» и «стандарт» можно различить по тому, что первый из терминов относится к оцениванию (оценке), а второй – к отметке (в том или ином формате) как форме обозначения результатов оценивания. Так, в примерах, приведенных выше, уровни А и 5 – это и есть стандарты, а их содержание раскрыто через дескрипторы, относящиеся к критериям, описывающим два различных качества работы: знание теоретического материала, структуру и логическую организацию самой работы.

Однако замечание D. Sadler о том, что «существует четыре основных подхода к установлению и опубликованию стандартов в целом: числовые отсечки в надлежащей шкале измерения, неявные знания, устные описания и использование образцов» [49, с. 191–192], как нам кажется, усложняет процесс уяснения различий между критериями и стандартами. Проблемность использования числовых отсечек в аспекте оценки работы обучающегося автор видит в ненадежности существующих шкал измерения и субъективности оценивания. Неявное (молчаливое) знание D. Sadler отождествляет с тем опытом, который существует у оценщика. Его можно назвать личным «стандартом» педагога, который берет свое начало из накопленных знаний в дисциплине, воспоминаний о собственном обучении, предыдущего личного опыта по оценке работы обучающихся. «Молчаливые» стандарты наиболее часто становятся основой для оценочных суждений, когда речь идет об оценивании компетенций, критерии оценки которых требуют многословного выражения. Автор считает, что именно молчаливое знание может быть использовано для создания системы стандартов. В целом соглашаясь с этим мнением, заметим, что в ряде случаев личные стандарты педагога ориенти-

¹ Ожегов С. И., Шведова Н. Ю. Толковый словарь русского языка. 4-е изд., дополненное. Москва: Азбуковник, 1999. С. 762.

руются и на некогнитивные факторы, такие как интерес учащихся к предмету, их настойчивость, примерное поведение и т. п. Говоря об образцах как ключевых примерах продуктов или процессов, выбранных таким образом, чтобы они были типичными для обозначенных уровней качества или компетенции, D. Sadler характеризует их как особенно удобные для обсуждения с коллегами и обучающимися и упоминает, что образцы можно использовать одним из двух способов: для иллюстрации ожидаемого качества или для его оценки. По нашему мнению, здесь обнаруживается опасность: вряд ли возможно создать единый образец хоть сколько-нибудь сложного образовательного продукта или процесса, поэтому образцы, относящиеся к одному и тому же стандарту, могут значительно отличаться друг от друга. Вербальные описания D. Sadler представляет как набор утверждений, которые максимально подробно определяют что-то из обозначенных уровней качества, при этом важно найти экономичный набор словесных описаний.

Решение вопроса четкости и однозначности толкования критериев исследователи видят в их спецификации с использованием дескрипторов (от латинского *descriptor* – «описывающий») как поисковых образов конкретного критерия, использующегося при анализе оцениваемой работы. Более того, как показывает опыт внедрения критериального оценивания в массовую школьную практику (О. В. Григоренко, А. С. Рванова, И. Б. Шмигирилова [65]), при принятии учащимися оценки их учебной деятельности и реализации ее формирующей функции на первый план выходят именно дескрипторы. Однако, по мнению R. L. Linn [66] и D. Sadler [49, 66], ключевые термины, применяемые в дескрипторах, особенно при оценке заданий творческой направленности, часто не имеют универсального значения независимо от тщательности детализации формулировок, поэтому и проблема однозначного понимания критериев остается неразрешенной. Кроме того, поскольку спецификации критериев через дескрипторы ориентированы прежде всего на оценщиков, перед педагогами встает задача их адаптации для использования обучающимися. Целенаправленное обсуждение с учащимися критериев и дескрипторов, в ходе которого они уясняют суть конкретного образовательного результата и того, что составляет определенное качество этого результата, является хорошей эффективной стратегией для формирования у обучающихся навыков самооценки, но требует значительных временных затрат.

D. Sadler для характеристики акта «разработки четких описаний стандартов» вводит термин «кодификация» [67, с. 274] и отмечает, что обычно кодификация стандартов высшего образования состоит в выделении качественных характеристик академических достижений, определяющих критерии, и характеристик количественных – стандартов, то есть уровней,

необходимых для определения квалификации обучающегося. Такое структурирование в наиболее полном виде представляет собой оценочную рубрику.

P. Dawson [68], проанализировав литературу по проблеме разработки и использования рубрик, отметил, что этот термин имеет различные значения. Однако пояснение автора к этому утверждению приводит к мысли о том, что речь идет прежде всего не о значении термина, а об особенностях дизайна самих рубрик, формах их практического воплощения. Автор приводит обобщенное определение рубрики, в качестве которой рассматривает «инструмент, используемый в процессе оценки работы учащегося, который обычно включает в себя три основные составляющие: критерии оценки, определение качества для этих критериев на определенных уровнях и стратегию оценки» [68, с. 351].

В большинстве исследований приводятся два основания для классификации рубрик:

1) аналитическая (оценка приводится на основе суммирования баллов по каждому критерию) и целостная (оценка представляет собой общий балл по заданию или предмету на основе предварительно определенных уровней достижений);

2) общая (применяемые при оценивании различных задач) и специфичная (используемые для оценки конкретной задачи и ряда задач одного типа).

В наиболее полном описании, приводимом P. Dawson [68], насчитывается 14 элементов, относящихся как к характеристикам самих рубрик, так и к особенностям их разработки, которыми рубрики могут отличаться друг от друга: специфика; уровень секретности; образцы; стратегия подсчета баллов; оценочные критерии; уровни качества; сложность суждения; создатели рубрики; процессы качества; конструкция (презентация) рубрики; сопровождающая информация; пояснения.

М. А. Пинская, А. В. Иванов [55], P. Dawson [68], G. Wiggins, J. McTighe [69] указывают на экономию времени оценивания при использовании рубрик, отмечают эффективность рубрик, содержащих четкие критерии, как инструментов объективной оценки учебной производительности, а также подчеркивают их значимость в процессе формирующего оценивания. Как заключают G. Wiggins, J. McTighe, «они позволяют учителям „учить, проверять и надеяться на лучшее“» [69, с. 314].

Однако в ряде источников встречаются и более настороженные отзывы об использовании рубрик. D. Haraway, D. Flinders, не умаляя полезности рубрик для взаимодействия обучающихся с педагогом и учебным модулем, указывают на разнообразие ограничений, накладываемых на их применение, и предостерегают: «многое теряется, когда мы решаем требовать, чтобы все учителя или все учащиеся использовали одни и те же рубрики одинаково» [70, с. 133].

Обосновывая важность постепенного совершенствования рубрик с целью обеспечения их конструктивного соответствия учебной программе, исследователи отмечают потенциальную способность обучающихся играть активную роль в определении того, что характеризует успешную работу. В этой связи создание рубрик обязательно должно предусматривать этапы апробирования, анализа и пересмотра, совершенствования.

D. Sadler [67], Z. Chan, S. Ho [71], S. M. Humphry, S. Heldsinger [72] критикуют рубрики (прежде всего аналитические), выдвигая в качестве аргументов:

- отсутствие полного совпадения характеристик отдельных образцов, на основании которых выбираются критерии для оценивания, с характеристиками всей массы работ, которые будут оцениваться на основе выбранных критериев;

- невозможность с уверенностью предсказать, а тем более отразить в кодификаторах разнообразие форм и содержания представления результатов выполнения обучающимися задания, которое подлежит оценке;

- трудность определения и, главное, словесного выражения границ между уровнями образовательных достижений;

- эластичность дескрипторов (спецификаций), используемых при кодификации стандартов, приводящую чаще не к абсолютной их интерпретации, а к относительной: их толкование легко расширяется или сжимается, чтобы соответствовать конкретной реальности;

- чрезмерную конкретность структуры кодификации даже при значительной абстрактности стандарта, устанавливающую собственные взаимосвязи между первичным свидетельством качества учебных достижений и оценочным суждением.

На основании приведенных аргументов и убеждения в том, что любое оценивание, в том числе критериальное, базируется на профессиональном мнении и суждениях оценщиков (учителей и преподавателей) о стандарте, D. Sadler делает вывод: кодификация по своей природе является сдерживающей, тем более в сопровождении «образцов», потому что их формулировки не допускают новых типов и конфигураций качеств, составляющих общее качество. А это может привести к потенциально значимым издержкам – жесткости и единообразию [67, с. 282]. Автор предлагает альтернативу кодификации: уровни достижений (то есть стандарты) рекомендует выявлять после обсуждения совокупности работ обучающихся компетентными оценщиками. Когда соглашение между оценщиками достигается, следующий шаг требует точного объяснения того, почему оправдана принятая классификация уровней достижений обучающихся, при этом текст этого обоснования будет функционировать как оценочное описание реального конкретного объекта, а значение его спецификаторов будет полностью контекстуализировано.

Ряд исследователей разделяют мнение D. Sadler, о чем, например, свидетельствуют данные, приведенные P. Chadha [73] в отчете по результатам исследовательского проекта, выполненного в Дублинском технологическом университете.

L. N. Jescovitch с соавторами [74] считают, что рубрики трудно использовать при аргументации уровня достижения критерия (или стандарта), описание которого выстроено в аддитивных или пересекающихся дескрипторах. Для решения этой проблемы авторы разработали системный подход, который определили как деконструкцию: процесс преобразования целостной рубрики в совокупность отдельных концептуальных компонентов, которые могут быть использованы для разработки аналитической рубрики, осуществляемый несколькими экспертами. Эти отдельные компоненты при необходимости могут быть объединены в оценку, которая сохраняет свойства целостной рубрики, что позволяет максимизировать выгоды и минимизировать недостатки каждого типа рубрик.

Особо категорично в отношении структуры аналитических рубрик высказываются S. M. Humphry, S. Heldsinger [72] и, отмечая недостаточное количество эмпирических исследований, подтверждающих валидность оценок образовательных результатов на основе рубрик, в свою очередь, представляют данные о существовании потенциальной угрозы валидности такого оценивания, вытекающей из особенностей конструкции самих рубрик. В двухэтапном методе оценки работ обучающихся S. M. Humphry, S. Heldsinger прослеживаются схожие с предложенными D. Sadler идеи: первый этап состоит в осуществлении парных сравнений работ, а на втором этапе происходит оценка всей массы работ учащихся по калиброванным образцам, выбранным на первом этапе с учетом дескрипторов, определенных ранее. Эксперимент, описанный авторами, подтверждает достаточно высокий уровень надежности этого метода оценки. Однако, по нашему мнению, парное сравнение большого количества разнообразных образцов затруднит практическое применение данного метода.

Таким образом, использование рубрик для оценивания достижений обучающихся хотя и имеет много положительных сторон, однако требует вдумчивого подхода педагогов, понимания ими того, что при необходимости стоит отказаться от изначально заданной структуры рубрики, которая, может, и выглядит упорядоченно и стройно, но не способствует достижению той цели, для которой она была разработана.

Еще одним значимым моментом в аспекте рассмотрения теории и практики образовательного оценивания является проблема компетентности педагогов как оценщиков и их подготовка в этом направлении, поскольку одним из значимых факторов, угрожающих надежности оценки образова-

тельных достижений обучающихся, являются того или иного рода ошибки оценщика. Недостаточный уровень компетентности педагогов в этом аспекте J. Hill, H. West [35], а также B. Randel с соавторами [75] связывают:

- с ограниченностью осознания роли оценивания как основы для принятия различного рода педагогических решений, ориентированных на решение проблемы повышения учебных достижений обучающихся;
- незнанием концептуальных основ использования оценки для обучения;
- отсутствием полного представления о разнообразии эффективных оценочных практик: у многих оценщиков такое представление базируется исключительно на личном оценочном опыте.

О значимости этого аспекта для образовательной практик говорит тот факт, что именно ему отводится место в ряду тех немногих проблем, которые являются предметом многих эмпирических исследований.

J. Hill, H. West [35], И. Н. Емельянова, О. А. Теплякова, Г. З. Ефимова [76] считают, что для изменения такого положения необходимо переосмыслить подготовку педагогических кадров в контексте реализации более интегративного подхода, в котором оценка рассматривается не как отдельный этап деятельности педагога, а в единстве со всеми компонентами образовательного процесса, а также развивать и другие направления: организовывать работу профессиональных сообществ как площадку для обмена опытом эффективных практик оценки; предоставлять педагогам многоаспектную обратную связь на основе анализа результатов оценивания на национальном и региональном уровнях; обеспечивать непрерывное профессиональное развитие.

Именно осуществление комплекса различных инициатив поможет педагогам не только освоить конкретные теории и практики оценивания во всем их многообразии, но и продвигаться к уровню порождения идей и замыслов по дальнейшему совершенствованию образовательной оценки.

Заключение

Оценка, являясь компонентом процесса обучения, оказывает на него значительное влияние, поэтому развитие теории обучения требует развития теории оценивания, изменения методов оценки, пересмотра роли обучающихся и преподавателей в оценочных практиках.

Обобщая результаты проведенного анализа, можно отметить:

1) несмотря на довольно значительную эклектичность исследований в области образовательного оценивания, прослеживается общая тенденция смещения акцентов оценки с констатирующей (суммирующей) на формирующую, с нормативной на критериально-ориентированную;

2) образовательная практика по-прежнему в большей степени привержена суммативному оцениванию, что объясняется предрасположенностью педагогов к традиционной итоговой оценке, недостаточным владением ими инструментами формирующего оценивания, а также ограниченностью исследовательской базы относительно того, как поддержать учителей в реализации практики формирующей оценки;

3) попытки разработать общую теорию формирующего оценивания приводят к столкновению с необходимостью учитывать уникальность учебной дисциплины, проявляющуюся как в ее содержании, так и в деятельностном компоненте, при разработке и реализации процедур оценивания на практике, в том числе интерпретации результатов оценивания в формате обратной связи;

4) существуют различные модели оценивания на основе критериев, которые, с одной стороны, в той или иной мере позволяют учесть особенности оценочной практики различных учебных заведений и целевые ориентиры различных предметных областей, с другой стороны, часто даже самими оценщиками не осознаются как критериально-ориентированные;

5) оценивание на основе критериев призвано обеспечить возможность сверять учебные достижения отдельного обучающегося с целями освоения образовательной программы вне зависимости от результатов других обучающихся, однако процесс такой оценки зависит от множества переменных (однозначность и четкость критериев оценивания, их единое понимание оценщиками; соответствие методов, инструментов и оценочных материалов предметному контексту; компетентность педагогов как оценщиков и их «молчаливые» стандарты оценивания; готовность обучающихся к пониманию критериев и др.), которые необходимо учитывать в оценочной практике;

6) множественность качественно различных подходов к разработке и представлению критериев оценивания, а также отсутствие в научных источниках четкого различения терминов «стандарты» и «критерии», «критерии» и «дескрипторы» вносит путаницу в понимание самой сущности критериального оценивания и отражается на качестве инструментов оценки, среди которых, прежде всего, выделяются рубрики;

7) рубрики как инструменты критериальной оценки, с одной стороны, могут позволить улучшить ее объективность, справедливость, обеспечить возможность привлекать обучающихся к оцениванию, с другой стороны, в случае принятия единообразия конструкции содержат потенциальную угрозу валидности оценки;

8) препятствием на пути обеспечения образовательного оценивания, отвечающего запросам современности, также является неготовность обуча-

ющихся к активному включению в процесс обратной связи, к принятию на себя ответственности за результаты своего обучения;

9) форсирование внедрения реформ в образовательную практику во многих аспектах является достаточно обременительным и для педагогов, и для обучающихся.

Таким образом, несмотря на широкие потенциальные преимущества современных подходов к оценке достижений, существует множество рисков непродуманного и поспешного внедрения оценочных реформ в образовательную практику. Реформирование системы оценивания будет вести к ее совершенствованию, если будет исходить не столько из административных решений, сколько из потребностей образовательной практики и иметь в своей основе надежную теоретически обоснованную и экспериментально проверенную модель оценивания, согласованную с другими компонентами процесса обучения, разумно сочетающую преимущества суммирующего и формирующего оценивания, критериальной и нормированной оценок и минимизирующую их недостатки.

Приходится констатировать, что необходимого внимания не уделяется изучению именно проблемных аспектов различных оценочных практик. Как правило, критическому анализу подвергается только традиционный подход к оценке образовательных результатов; суммирующее и формирующее оценивание на основе критериев рассматриваются в основном как позитивное явление. Представленный в статье анализ не только плюсов, но и минусов, связанных с практиками суммирующей и формирующей оценок, выявленных в научно-методических публикациях, может способствовать обретению исследователями и педагогами более объективного взгляда на новые подходы к образовательному оцениванию. Также следует подчеркнуть, что в русскоязычном сегменте научных статей, освещающих оценку в образовании, очень незначительный процент эмпирических исследований, которые далеко не перекрывают всех аспектов проблемы.

Специального исследования требуют вопросы интегративного использования одних и тех же методов оценки для реализации различных целей оценивания; разработки механизмов обратной связи; обеспечения надежности и валидности различных видов процедур и инструментов оценки; учета контекста дисциплин в процессе оценивания; разработки критериев для оценки компетенций и в этой связи сравнения критериев оценивания в академической и профессиональных областях, формирования и развития оценочной грамотности педагога.

Проведенный анализ не претендует на представление всеобъемлющей картины оценивания в образовании, но может выступить основой для дальнейших исследований с целью дополнения и упорядочивания теории образовательной оценки, разработки на ее основе процедур оценивания и эффективного их использования в образовательной практике.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Taras M. Assessment for Learning: understanding theory to improve practice // *Journal of Further and Higher Education*. 2007. Vol. 31 (4). P. 363–371. DOI: 10.1080/03098770701625746
2. Ball S. Introduction: The Great Education Debate (1976–2016) // In: *The education debate*. 3rd ed. Bristol, UK; Chicago, IL, USA: Bristol University Press, 2017. P. 1–12. Available from: <https://discovery.ucl.ac.uk/id/eprint/10018968/1/250102.pdf> (date of access: 19.08.2020).
3. Palmatier R. W., Houston M. B., Hulland J. Review articles: purpose, process, and structure // *Journal of the Academy of Marketing Science*. 2018. Vol. 46. P. 1–5. DOI: 10.1007/s11747-017-0563-4
4. Baird J., Hopfenbeck T. N., Newton P. E., Steen-Utheim A. T. State of the Field Review: Assessment and Learning, Oxford University Centre for Educational Assessment Report OUCEA/14/2. 2014. Available from: <https://www.researchgate.net/publication/263654863> (date of access: 21.08.2020).
5. Bennett R. E. Formative assessment: a critical review // *Assessment in Education: Principles, Policy and Practice*. 2011. Vol. 18 (1). P. 5–25. DOI: 10.1080/0969594X.2010.513678
6. Lau A. M. Formative good, summative bad? – A review of the dichotomy in assessment literature // *Journal of Further and Higher Education*. 2016. Vol. 40 (4). P. 509–525. DOI: 10.1080/0309877X.2014.984600
7. Gikandi J. W., Morrowa D., Davisa N. E. Online formative assessment in higher education: A review of the literature // *Computers & Education*. 2011. Vol. 57 (4). P. 2333–2351. DOI: 10.1016/j.compedu.2011.06.004
8. Бодоньи М. А. Зарубежный и отечественный опыт использования современных технологий в формирующем оценивании [Электрон. ресурс] // *Отечественная и зарубежная педагогика*. 2020. Т. 1. № 3 (68). С. 78–95. Режим доступа: <https://publications.hse.ru/mirror/pubs/share/direct/382850089.pdf> (дата обращения: 23.07.2020).
9. Pepper D. KeyCoNet. 2013 Literature Review: Assessment for key competences. 2013. 35 p. Available from: http://keyconet.eun.org/c/document_library/get_file?uuid=b1475317-108c-4cf5-a650-dae772a7d943&groupId=11028 (date of access: 12.07.2020).
10. Siarova H., Sternadel D., Mašidlauskaitė R. Assessment practices for 21st century learning: review of evidence, NESET II report. Luxembourg: Publications Office of the European Union. 2017. Available from: https://nesetweb.eu/wp-content/uploads/2019/06/AR1_20172.pdf (date of access: 08.07.2020).
11. Потемкина Т. В., Пуденко Т. И. Независимая оценка качества образования: общественная экспертиза в зарубежной и отечественной практике // *Образование и наука*. 2015. № 10 (129). С. 64–79. DOI: 10.17853/1994-5639-2015-10-64-79
12. DeLuca C., LaPointe-McEwan D., Luhanga U. Teacher assessment literacy: A review of international standards and measures // *Educational Assessment, Evaluation and Accountability*. 2016. Vol. 28 (3). P. 251–272. DOI: 10.1007/s11092-015-9233-6
13. Темняткина О. В., Токменинова Д. В. Современные подходы к оценке эффективности работы учителей. Обзор зарубежных публикаций // *Вопросы образования*. 2018. № 3. С. 180–193. DOI: 10.17323/1814-9545-2018-3-180-195
14. Шадриков В. Д., Шадрикова И. А., Педагогическое оценивание: учебное пособие [Электрон. ресурс] Москва: Университетская книга. 2018. 156 с. Режим доступа: <https://publications.hse.ru/mirror/pubs/share/direct/286820966.pdf> (дата обращения: 07.12.2020).

15. Harlen W. On the relationship between assessment for formative and summative purposes // *Assessment and Learning*. 2nd ed. London: Sage, 2012. P. 87–102. Available from: <http://sk.sagepub.com/books/assessment-and-learning-2e/n6.xml> (date of access: 21.07.2020).
16. Newton P. E. Clarifying the purposes of educational assessment // *Assessment in Education: Principles, Policy & Practice*. 2007. Vol. 14 (2). P. 149–170. DOI: 10.1080/09695940701478321
17. Archer E. The Assessment Purpose Triangle: Balancing the Purposes of Educational Assessment // *Frontiers Education*. 2017. Vol. 4. Article 41. Available from: <https://www.frontiersin.org/articles/10.3389/feduc.2017.00041/full> (date of access: 23.07.2020). DOI: 10.3389/feduc.2017.00041
18. Black P., Wiliam D. Assessment and classroom learning // *Assessment in Education: Principles, Policy & Practice*. 1998. Vol. 5 (1). P. 7–74. DOI: 10.1080/0969595980050102
19. Black P., Wiliam D. Developing the theory of formative assessment // *Educational Assessment Evaluation and Accountability*. 2009. Vol. 21 (1). P. 5–31. DOI: 10.1007/s11092-008-9068-5
20. Black P. Formative assessment – an optimistic but incomplete vision // *Assessment in Education: Principles, Policy & Practice*. 2015. Vol. 22 (1). P. 161–177. DOI: 10.1080/0969594X.2014.999643
21. Dixon D. D., Worrell F. C. Formative and Summative Assessment in the Classroom // *Theory Into Practice*. 2016. Vol. 55 (2). P. 153–159. DOI: 10.1080/00405841.2016.1148989
22. Пинская М. А. Формирующее оценивание: оценивание в классе [Электрон. ресурс]. Москва: Логос, 2010. 264 с. Режим доступа: http://school257-newiz.ucoz.ru/dok/formirujushhee_ocenivanie-avtor_m.a-pinskaja.pdf (дата обращения: 23.07.2020).
23. Harlen W. Teachers' summative practices and assessment for learning-tensions and synergies // *The Curriculum Journal*. 2005. Vol. 16 (2). P. 207–223. DOI: 10.1080/09585170500136093
24. Newton P. E. Validity, purpose and the recycling of results from educational Assessment // *Assessment and Learning*. 2nd ed. London: Sage, 2012. P. 264–276. Available from: <http://sk.sagepub.com/books/assessment-and-learning-2e/n16.xml> (date of access: 21.07.2020).
25. Sadler D. R. Formative assessment and the design of instructional systems // *Instructional Science*. 1989. Vol. 18. P. 119–144. Available from: <http://michiganassessmentconsortium.org/wp-content/uploads/Formative-Assessment-and-Design-of-Instructional-Systems.pdf> (date of access: 25.07.2020).
26. Taras M. Assessment – Summative and Formative – some theoretical reflections // *British Journal of Educational Studies*. 2005. Vol. 53 (3). P. 466–478. DOI: 10.1111/j.1467-8527.2005.00307.x
27. Sjögren A. The long run consequences of being graded in elementary school: paper presented at the Final Conference of the RTN Network “Economic of Education and Education Policy in Europe” (EEEEPE). London: LSE, 2009. Available from: <https://www.researchgate.net/publication/46470357> (date of access: 29.07.2020).
28. Butt G. Making Assessment Matter. London: Bloomsbury Continuum, 2010. 156 p. Available from: <https://scholar.google.com/citations?user=ptgvb4wAAAAJ&hl=en> (date of access: 28.07.2020).
29. Silfver E., Sjöberg G., Bagger A. An ‘appropriate’ test taker: the everyday classroom during the national testing period in school year three in Sweden // *Ethnography and Education*. 2016. Vol. 11 (3). P. 237–252. DOI: 10.1080/17457823.2015.1085323

30. Courtney S. Post-panopticism and school inspection in England // *British Journal of Sociology of Education*. 2016. Vol. 37 (4). P. 623–642. DOI: 10.1080/01425692.2014.965806
31. Болотов В. А. Прошлое, настоящее и возможное будущее российской системы оценки качества образования // *Вопросы образования*. 2018. № 3. С. 287–297. DOI: 10.17323/1814-9545-2018-3-287-297
32. Шаповалова О. Н., Ефремова Н. Ф. Дидактический потенциал формирующего оценивания метапредметных результатов школьников: российский и зарубежный опыт [Электрон. ресурс] // *Мир науки. Педагогика и психология*. 2019. Т. 7. № 6. Режим доступа: <https://mir-nauki.com/PDF/94PDMN619.pdf> (дата обращения: 22.08.2020).
33. Четвертных Т. В. Метапредметное содержание формирующего оценивания // *Вестник Костромского государственного университета*. Серия: Педагогика. Психология. Социокинетика. 2019. № 2. С. 21–25. DOI: 10.34216/2073-1426-2019-25-2-21-25
34. Cisterna D., Gotwals A.W. Enactment of Ongoing Formative Assessment: Challenges and Opportunities for Professional Development and Practice // *Journal of Science Teacher Education*. 2018. Vol. 29 (3). P. 200–222. DOI: 10.1080/1046560X.2018.1432227
35. Hill J., West H. Improving the student learning experience through dialogic feed-forward assessment // *Assessment & Evaluation in Higher Education*. 2020. Vol. 45(1). P. 82–97. DOI: 10.1080/02602938.2019.1608908
36. Phelps R. P. Synergies for better learning: an international perspective on evaluation and assessment // *Assessment in Education: Principles, Policy & Practice*. 2014. Vol. 21 (4). P. 481–493. DOI: 10.1080/0969594X.2014.921091
37. Torrance H. Formative assessment at the crossroads: Conformative, deformative and transformative assessment // *Oxford Review of Education*. 2012. Vol. 38 (3). P. 323–342. DOI: 10.1080/03054985.2012.689693
38. Henderson M., Ryan T., Phillips M. The challenges of feedback in higher education // *Assessment & Evaluation in Higher Education*. 2019. Vol. 44 (8). P. 1237–1252. DOI: 10.1080/02602938.2019.1599815
39. Harrison C. J., Könings K. D., Dannefer E. F., Schuwirth L. W., Wass V., Van der Vleuten C. P. Factors influencing students' receptivity to formative feedback emerging from different assessment cultures // *Perspectives on Medical Education*. 2016. Vol. 5 (5). P. 276–284. DOI: 10.1007/s40037-016-0297-x
40. Dann R. Assessment as learning: blurring the boundaries of assessment and learning for theory, policy and practice // *Assessment in Education: Principles, Policy & Practice*. 2014. Vol. 21 (2). P. 149–166. DOI: 10.1080/0969594X.2014.898128
41. Вилкова А. В. Реализация формирующего оценивания учебных достижений обучающихся: проблемы и пути решения // *Научно-педагогическое обозрение*. 2019. № 4 (26). С. 7–13. DOI: 10.23951/2307-6127-2019-4-7-13
42. Четвертных Т. В. Готовность педагогов к проведению формирующего оценивания в образовательном процессе [Электрон. ресурс] // *Вестник Омского государственного педагогического университета*. Гуманитарные исследования. 2018. № 4 (21). С. 146–150. Режим доступа: <https://www.elibrary.ru/item.asp?id=36961492> (дата обращения: 21.07.2020).
43. Шмигирилова И. Б., Рванова А. С., Таджигитов А. А. Педагогическая мастерская как форма развития оценочной компетентности учителя математики [Электрон. ресурс] // *Современные проблемы науки и образования*. 2020. № 6. С. 87. Режим доступа: <https://www.elibrary.ru/item.asp?id=44595518> (дата обращения: 21.12.2020).
44. Florez M. T., Sammons P. Assessment for learning: effects and impact. Oxford University Department of Education. 2013. 32 p. Available from: <https://www.education->

developmenttrust.com/EducationDevelopmentTrust/files/5a/5a6d6203-ec49-4d33-9c5d-42c188184807.pdf (date of access: 03.08.2020).

45. Herman J. L., Osmundson E., Silver D. Capturing quality in Formative assessment practice: Measurement challenges, CRESST Report 770. Los Angeles: National Centre for Research on Evaluation, Standards and Student Testing. 2010. 26 p. Available from: <https://files.eric.ed.gov/fulltext/ED512648.pdf> (date of access: 16.07.2020).

46. Velasco L. C., Tojar J-C. Competency-Based Evaluation in Higher Education – Design and Use of Competence Rubrics by University Educators // International Education Studies. 2018. Vol. 11 (2). P. 118–132. DOI: 10.5539/ies.v11n2p118

47. Ефремова Н. В. Концептуальная модель оценки компетенций студентов // Современные наукоемкие технологии. 2019. № 7. С. 169–174. DOI: 10.17513/snt.37607

48. Wiliam D. Standards in examinations: a matter of trust? // The Curriculum Journal. 1996. Vol. 7 (3). P. 293–306. DOI: 10.1080/09585179600070303

49. Sadler D. R. Interpretations of criteria-based assessment and grading in higher education // Assessment & Evaluation in Higher Education. 2005. Vol. 30 (2). P. 175–194. DOI: 10.1080/0260293042000264262

50. Dann R. Promoting assessment as learning: Improving the learning process. London: Routledge Falmer, 2002. 176 p. Available from: <https://books.google.ru/books?id=EGVCe-r-bX-wC&hl=ru> (date of access: 18.07.2020).

51. Wikström C. Criterion-Referenced Measurement for Educational Evaluation and Selection. Department of Educational Measurement Umeå University Thesis. 2005. 46 p. Available from: <https://www.diva-portal.org/smash/get/diva2:143588/FULLTEXT01.pdf> (date of access: 21.07.2020).

52. Hornby W. Assessing using grade-related criteria: a single currency for universities? // Assessment & Evaluation in Higher Education. 2003. Vol. 28 (4). P. 435–454. DOI: 10.1080/0260293032000066254

53. Ling L., Yuen B., Loo W., Prinsloo C. Students' Conceptions of Bell Curve Grading Fairness in Relation to Goal Orientation and Motivation // International Journal for the Scholarship of Teaching and Learning 2020. Vol. 14 (1). DOI: 10.20429/ijso.2020.140107.

54. Красноторова А. А. Критериальный подход в оценивании учебных достижений учащихся [Электрон. ресурс] // Педагогическое образование и наука. 2009. № 6. С. 91–94. Режим доступа: <https://www.elibrary.ru/item.asp?id=17062493> (дата обращения: 20.07.2020).

55. Пинская М. А., Иванов А. В. Формирующий подход: критериальное оценивание в действии [Электрон. ресурс] // Народное образование. 2010. № 5. С. 192–201. Режим доступа: <https://www.elibrary.ru/item.asp?id=15516810> (дата обращения: 01.08.2020).

56. Абекова Ж. А., Оралбаев А. Б., Ермаханов М. Н., Джакипов А. С. Совершенствование учебного процесса при критериальном оценивании, его главные преимущества и особенности [Электрон. ресурс] // Успехи современного естествознания. 2015. № 1 (2). С. 295–296. Режим доступа: <https://www.elibrary.ru/item.asp?id=23068809> (дата обращения: 01.08.2020).

57. Липатникова И. Г. Оценивание как диагностическая процедура формирования конечных результатов обучения по математике [Электрон. ресурс] // Педагогическое образование в России. 2016. № 7. С. 177–182. Режим доступа: <https://www.elibrary.ru/item.asp?id=26638542> (дата обращения: 02.08.2020).

58. Соколова Е. В. Конструирование диагностических заданий в условиях критериального оценивания достижений учащихся в изучении школьного курса геометрии [Элек-

трон. ресурс] // Преподаватель XXI век. 2016. № 4–1. С. 277–287. Режим доступа: <https://www.elibrary.ru/item.asp?id=27633963> (дата обращения: 02.08.2020).

59. Lok B., McNaught C., Young K. Criterion-referenced and norm-referenced assessments: compatibility and complementarity // *Assessment & Evaluation in Higher Education*. 2016. Vol. 41 (3). P. 450–465. DOI: 10.1080/02602938.2015.1022136

60. Brown G., Bull J., Pendlebury M. Assessing student learning in higher education. London: Routledge, 1997. 317 p. Available from: https://archive.org/details/assessingstudent1997brow_kl150 (date of access: 03.08.2020).

61. Barrie S., Brew A., McCulloch M. Qualitatively different conceptions of criteria used to assess student learning. Proceedings of the AARE-NZARE Conference. Melbourne. 1999. Available from: <https://www.aare.edu.au/data/publications/1999/bre99209.pdf> (date of access: 03.08.2020).

62. Wiliam D. Assessing authentic tasks: alternatives to mark-schemes [Internet]. // *Nordic Studies in Mathematics Education*. 1994. Vol. 2 (1). P. 48–68. Available from: https://www.dylanwiliam.org/Dylan_Williams_website/Papers_files/Nomad.PDF (date of access: 04.08.2020).

63. Green S. Criterion referenced assessment as a guide to learning – the importance of progression and reliability. Paper presented at the Association for the Study of Evaluation in Education in Southern Africa International Conference. 10–12 July 2002. Johannesburg. Available from: <https://www.cambridgeassessment.org.uk/Images/109693-criterion-referenced-assessment-as-a-guide-to-learning-the-importance-of-progression-and-reliability.pdf> (date of access: 04.06.2020).

64. Popham W. J. Criterion-referenced Measurement: Half a Century Wasted? // *Educational Leadership*. 2014. Vol. 71 (6). P. 62–66. Available from: <http://www.educationalleadership-digital.com/educationalleadership/201403?pg=65#pg65> (date of access: 05.08.2020).

65. Григоренко О. В., Рванова А. С., Шмигирилова И. Б. Проблемы внедрения критериального оценивания в школьную практику [Электрон. ресурс] // *Сибирский учитель*. 2020. № 2 (129). С. 13–19. Режим доступа: <https://www.elibrary.ru/item.asp?id=43044448> (дата обращения: 21.07.2020).

66. Linn R. L. Criterion-referenced measurement: a valuable perspective clouded by surplus meaning // *Educational Measurement: Issues and Practice*. 1994. Vol. 13 (4). P. 12–14. DOI: 10.1111/j.1745-3992.1994.tb00564.x

67. Sadler D. R. The futility of attempting to codify academic achievement standards // *Higher Education*. 2014. Vol. 67 (3). P. 273–288. DOI: 10.1007/s10734-013-9649-1

68. Dawson P. Assessment rubrics: Towards clearer and more replicable design, research and practice // *Assessment & Evaluation in Higher Education*. 2017. Vol. 42 (3). P. 347–360. DOI: 10.1080/02602938.2015.1111294

69. Wiggins G., McTighe J. Understanding by Design. Alexandria, VA: Association for Supervision & Curriculum Development, 2005. 370 p. Available from: <https://eric.ed.gov/?id=ED509029> (date of access: 11.08.2020).

70. Haraway D., Flinders D. Rubrics in Context // *Rubric Nation: Critical Inquiries on the Impact of Rubrics in Education*. Charlotte, NC: Information Age, 2015. P. 119–134. Available from: https://nsuworks.nova.edu/fse_facbooks/130 (date of access: 11.08.2020).

71. Chan Z., Ho S. Good and bad practices in rubrics: the perspectives of students and educators // *Assessment & Evaluation in Higher Education*. 2019. Vol. 44 (4). P. 533–545. DOI: 10.1080/02602938.2018.1522528 (date of access: 11.08.2020).

72. Humphry S. M., Heldsinger S. A Two-Stage Method for Classroom Assessments of Essay Writing // *Frontiers in Education*. 2019. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1111/jedm.12223> (дата обращения 11.06.2020).

73. Chadha P., et al. A Rubric Guide for New Academics, Practitioner Research Project Report. TU Dublin. 2019. Available from: <https://arrow.tudublin.ie/ltpcgdprp/23> (date of access: 11.08.2020).

74. Jescovitch L. N., Scott E. E., Cerchiara J. A., Doherty J. H., Wenderoth M. Pat, Merrill J. E., Urban-Lurain M., Haudek K. C. Deconstruction of Holistic Rubrics into Analytic Rubrics for Large-Scale Assessments of Students' Reasoning of Complex Science Concepts // *Practical Assessment, Research, and Evaluation*. 2019. Vol. 24, Article 7. Available from: <https://par.nsf.gov/servlets/purl/10166313> (date of access: 13.06.2020). DOI: 10.7275/9h7f-mp76

75. Randel B., Apthorp H., Beesley A. D., Clark T. F., Wang X. Impacts of professional development in classroom assessment on teacher and student outcomes // *The Journal of Educational Research*. 2016. Vol. 109 (5). P. 491–502. DOI: 10.1080/00220671.2014.992581

76. Емельянова И. Н., Теплякова О. А., Ефимова Г. З. Практика использования современных методов оценки на разных ступенях образования // *Образование и наука*. 2019. № 21 (6). С. 9–28. DOI: 10.17853/1994-5639-2019-6-9-28

References

1. Taras M. Assessment for learning: Understanding theory to improve practice. *Journal of Further and Higher Education*. 2007; 31 (4): 363–371. DOI: 10.1080/03098770701625746

2. Ball S. Introduction: The great education debate (1976–2016). In: *The education debate* [Internet]. 3rd ed. Bristol, UK; Chicago, IL, USA: Bristol University Press; 2017 [cited 2020 Aug 19]. p. 1–12. Available from: <https://discovery.ucl.ac.uk/id/eprint/10018968/1/250102.pdf>

3. Palmatier R. W., Houston M. B., Hulland J. Review articles: Purpose, process, and structure. *Journal of the Academy of Marketing Science* volume. 2018; 46: 1–5. DOI: 10.1007/s11747-017-0563-4

4. Baird J., Hopfenbeck T. N., Newton P. E., Steen-Utheim A. T. State of the field review: Assessment and learning, Oxford University Centre for Educational Assessment Report OUCEA/14/2 [Internet]. 2014 [cited 2020 Aug 21]; Available from: <https://www.researchgate.net/publication/263654863>

5. Bennett R. E. Formative assessment: A critical review. *Assessment in Education: Principles, Policy and Practice*. 2011; 18 (1): 5–25. DOI: 10.1080/0969594X.2010.513678

6. Lau A. M. Formative good, summative bad? – A review of the dichotomy in assessment literature. *Journal of Further and Higher Education*. 2016; 40 (4): 509–525. DOI: 10.1080/0309877X.2014.984600

7. Gikandi J. W., Morrow D., Davisa N. E. Online formative assessment in higher education: A review of the literature. *Computers & Education*. 2011; 57 (4): 2333–2351. DOI: 10.1016/j.compedu.2011.06.004

8. Bodony M. A. Foreign and domestic experience of the use of modern technologies in formative assessment. *Otechestvennaya i zarubezhnaya pedagogika = Domestic and Foreign Pedagogy* [Internet]. 2020 [cited 2020 Jul 23]; 3 (68): 78–95. Available from: <https://publications.hse.ru/mirror/pubs/share/direct/382850089.pdf> (In Russ.)

9. Pepper D. KeyCoNet 2013 literature review: Assessment for key competences [Internet]. 2013 [cited 2020 Jul 12]. 35 p. Available from: http://keyconet.eun.org/c/document_library/get_file?uuid=b1475317-108c-4cf5-a650-dae772a7d943&groupId=11028
10. Siarova H., Sternadel D., Mašidlauskaitė R. Assessment practices for 21st century learning: Review of evidence, NESET II report [Internet]. Luxembourg: Publications Office of the European Union; 2017 [cited 2020 Jul 08]. Available from: https://nesetweb.eu/wp-content/uploads/2019/06/AR1_20172.pdf
11. Potemkina T. V., Pudenko T. I. Independent evaluation of the quality of education: social expertise in foreign and Russian practice. *Obrazovanie i nauka = The Education and Science Journal*. 2015; 10: 64–79. DOI: 10.17853/1994-5639-2015-10-64-79 (In Russ.)
12. DeLuca C., LaPointe-McEwan D., Luhanga U. Teacher assessment literacy: A review of international standards and measures. *Educational Assessment, Evaluation and Accountability*. 2016; 28 (3): 251–272. DOI: 10.1007/s11092-015-9233-6
13. Temnyatkina O. V., Tokmeninova D. V. Modern approaches to teacher performance assessment. An overview of foreign publications. *Voprosy obrazovaniya = Educational Studies Moscow*. 2018; 3: 180–193. DOI: 10.17323/1814-9545-2018-3-180-195 (In Russ.)
14. Shadrikov V. D., Shadrikova I. A. Pedagogicheskoe ocenivanie = Pedagogical assessment [Internet]. Moscow: Publishing House Universitetskaya kniga; 2018 [cited 2020 Dec 07]. 156 p. Available from: <https://publications.hse.ru/mirror/pubs/share/direct/286820966.pdf> (In Russ.)
15. Harlen W. On the relationship between assessment for formative and summative purposes. In: *Assessment and learning* [Internet]. 2nd ed. London: Sage; 2012 [cited 2020 Jul 21]. p. 87–102. Available from: <http://sk.sagepub.com/books/assessment-and-learning-2e/n6.xml>
16. Newton P. E. Clarifying the purposes of educational assessment. *Assessment in Education: Principles, Policy & Practice*. 2007; 14 (2): 149–170. DOI: 10.1080/09695940701478321
17. Archer E. The assessment purpose triangle: Balancing the purposes of educational assessment. *Frontiers Education* [Internet]. 2017 [cited 2020 Jul 23]; 4, Article 41. Available from: <https://www.frontiersin.org/articles/10.3389/feduc.2017.00041/full> DOI: 10.3389/feduc.2017.00041
18. Black P., Wiliam D. Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*. 1998; 5 (1): 7–74. DOI: 10.1080/0969595980050102
19. Black P., Wiliam D. Developing the theory of formative assessment. *Educational Assessment Evaluation and Accountability*. 2009; 21 (1): 5–31. DOI: 10.1007/s11092-008-9068-5
20. Black P. Formative assessment – an optimistic but incomplete vision. *Assessment in Education: Principles, Policy & Practice*. 2015; 22 (1): 161–177. DOI: 10.1080/0969594X.2014.999643
21. Dixson D. D., Worrell F. C. Formative and summative assessment in the classroom. *Theory into Practice*. 2016; 55 (2): 153–159. DOI: 10.1080/00405841.2016.1148989
22. Pinskaya M. A. Formiruyushcheye otsenivaniye: otsenivaniye v klasse = Formative assessment: Classroom assessment [Internet]. Moscow: Publishing House Logos; 2009 [cited 2020 Jul 23]. 264 p. Available from: http://school257-newiz.ucoz.ru/dok/formirujushhee_ocenivanie-avtor_m.a-pinskaja.pdf (In Russ.)
23. Harlen W. Teachers' summative practices and assessment for learning-tensions and synergies. *The Curriculum Journal*. 2005; 16 (2): 207–223. DOI: 10.1080/09585170500136093
24. Newton P. E. Validity, purpose and the recycling of results from educational Assessment. In: *Assessment and learning* [Internet]. 2nd ed. London: Sage; 2012 [cited 2020 Jul 21]. p. 264–276. Available from: <http://sk.sagepub.com/books/assessment-and-learning-2e/n16.xml>

25. Sadler D. R. Formative assessment and the design of instructional systems. *Instructional Science* [Internet]. 1989 [cited 2020 Jul 21]; 18: 119–144. Available from: <http://michiganassessmentconsortium.org/wp-content/uploads/Formative-Assessment-and-Design-of-Instructional-Systems.pdf>
26. Taras M. Assessment – summative and formative – some theoretical reflections. *British Journal of Educational Studies*. 2005; 53 (3): 466–478. DOI: 10.1111/j.1467-8527.2005.00307.x
27. Sjögren A. The long run consequences of being graded in elementary school. In: *Final Conference of the RTN Network “Economic of Education and Education Policy in Europe: (EEEEPE)”* [Internet]. London: LSE; 2009 [cited 2020 Jul 29]. Available from: <https://www.researchgate.net/publication/46470357>
28. Butt G. Making assessment matter [Internet]. London: Bloomsbury Continuum; 2010 [cited 2020 Jul 28]. 156 p. Available from: <https://scholar.google.com/citations?user=ptgyb4wAAAAJ&hl=en>
29. Silfver E., Sjöberg G., Bagger A. An ‘appropriate’ test taker: The everyday classroom during the national testing period in school year three in Sweden. *Ethnography and Education*. 2016; 11 (3): 237–252. DOI: 10.1080/17457823.2015.1085323
30. Courtney S. Post-panopticism and school inspection in England. *British Journal of Sociology of Education*. 2016; 37 (4): 623–642. DOI: 10.1080/01425692.2014.965806
31. Bolotov V. A. The past, present, and possible future of the Russian education assessment system. *Voprosy obrazovaniya = Educational Studies Moscow*. 2018; 3: 287–297. DOI: 10.17323/1814-9545-2018-3-287-297 (In Russ.)
32. Shapovalova O. N., Efremova N. F. The didactic potential of formative assessment of meta-disciplinary results of schoolchildren: Russian and foreign experience. *Mir nauki. Pedagogika i psikhologiya = World of Science. Pedagogy and Psychology* [Internet]. 2019 [cited 2020 Aug 22]; 7 (6). Available from: <https://mir-nauki.com/PDF/94PDMN619.pdf> (In Russ.)
33. Chetvertnykh T. V. Comprehensive content of formative evaluation. *Vestnik Kostromskogo gosudarstvennogo universiteta. Seriya: Pedagogika. Psikhologiya. Sotsiokinetika = Bulletin of the Kostroma State University. Series: Pedagogy. Psychology. Sociokinetics*. 2019; 2: 21–25. DOI: 10.34216/2073-1426-2019-25-2-21-25 (In Russ.)
34. Cisterna D., Gotwals A. W. Enactment of ongoing formative assessment: Challenges and opportunities for professional development and practice. *Journal of Science Teacher Education*. 2018; 29 (3): 200–222. DOI: 10.1080/1046560X.2018.1432227
35. Hill J., West H. Improving the student learning experience through dialogic feed-forward assessment. *Assessment & Evaluation in Higher Education*. 2020; 45 (1): 82–97. DOI: 10.1080/02602938.2019.1608908
36. Phelps R. P. Synergies for better learning: An international perspective on evaluation and assessment. *Assessment in Education: Principles, Policy & Practice*. 2014; 21 (4): 481–493. DOI: 10.1080/0969594X.2014.921091
37. Torrance H. Formative assessment at the crossroads: Conformative, deformative and transformative assessment. *Oxford Review of Education*. 2012; 38 (3): 323–342. DOI: 10.1080/03054985.2012.689693
38. Harrison C. J., Könings K. D., Dannefer E. F., Schuwirth L. W., Wass V., Van der Vleuten C. P. Factors influencing students’ receptivity to formative feedback emerging from different assessment cultures. *Perspectives on Medical Education*. 2016; 5 (5): 276–284. DOI: 10.1007/s40037-016-0297-x

39. Henderson M., Ryan T., Phillips M. The challenges of feedback in higher education. *Assessment & Evaluation in Higher Education*. 2019; 44 (8): 1237–1252. DOI: 10.1080/02602938.2019.1599815
40. Dann R. Assessment as learning: Blurring the boundaries of assessment and learning for theory, policy and practice. *Assessment in Education: Principles, Policy & Practice*. 2014; 21 (2): 149–166. DOI: 10.1080/0969594X.2014.898128
41. Vilkova L. V. Implementation of formative assessment of students' educational achievements: problems and solution. *Nauchno-pedagogicheskoye obozreniye = Pedagogical Review*. 2019; 4 (26): 7–13. DOI: 10.23951/2307-6127-2019-4-7-13 (In Russ.)
42. Chetvertnykh T. V. Teachers' readiness for using the forming assessment in the educational process. *Vestnik Omskogo gosudarstvennogo pedagogicheskogo universiteta. Gumanitarnyye issledovaniya = Bulletin of the Omsk State Pedagogical University. Humanities Research* [Internet]. 2018 [cited 2020 Jul 27]; 4 (21): 146–150. Available from: <https://www.elibrary.ru/item.asp?id=36961492> (In Russ.)
43. Shmigirilova I. B., Rvanova A. S., Tadzhigitov A. A. Pedagogical workshop as a form of development of estimated competence of a mathematics teacher. *Sovremennyye problemy nauki i obrazovaniya = Modern Problems of Science and Education* [Internet]. 2020 [cited 2020 Dec 21]; 6: 87. Available from: <https://www.elibrary.ru/item.asp?id=44595518> (In Russ.)
44. Florez M. T., Sammons P. Assessment for learning: Effects and impact [Internet]. Oxford University Department of Education; 2013 [cited 2020 Aug 03]. 32 p. Available from: <https://www.educationdevelopmenttrust.com/EducationDevelopmentTrust/files/5a/5a6d6203-ec49-4d33-9c5d-42c188184807.pdf>
45. Herman J. L., Osmundson E., Silver D. Capturing quality in formative assessment practice: Measurement challenges, CRESST Report 770 [Internet]. Los Angeles: National Centre for Research on Evaluation, Standards and Student Testing; 2010 [cited 2020 Jul 16]. 26 p. Available from: <https://files.eric.ed.gov/fulltext/ED512648.pdf>
46. Velasco L. C., Tojar J-C. Competency-based evaluation in higher education – design and use of competence rubrics by university educators. *International Education Studies*. 2018; 11 (2): 118–132. DOI: 10.5539/ies.v11n2p118
47. Efremova N. Conceptual model of evaluation of competence of students. *Sovremennyye naukoymkiye tekhnologii = Modern High Technologies*. 2019; 7: 169–174. DOI: 10.17513/snt.37607 (In Russ.)
48. Wiliam D. Standards in examinations: A matter of trust? *The Curriculum Journal*. 1996; 7 (3): 293–306. DOI: 10.1080/0958517960070303
49. Sadler D. R. Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education*. 2005; 30 (2): 175–194. DOI: 10.1080/0260293042000264262
50. Dann R. Promoting assessment as learning: Improving the learning process [Internet]. London: Routledge Falmer; 2002 [cited 2020 Jul 18]. 176 p. Available from: <https://books.google.ru/books?id=EGVCe-r-bXwC&hl=ru>
51. Wikström C. Criterion-referenced measurement for educational evaluation and selection [Internet]. Department of Educational Measurement Umeå University Thesis; 2005 [cited 2020 Jul 21]. 46 p. Available from: <https://www.diva-portal.org/smash/get/diva2:143588/FULLTEXT01.pdf>
52. Hornby W. Assessing using grade-related criteria: a single currency for universities? *Assessment & Evaluation in Higher Education*. 2003; 28 (4): 435–454. DOI: 10.1080/0260293032000066254

53. Ling L., Yuen B., Loo W., Prinsloo C. Students' conceptions of bell curve grading fairness in relation to goal orientation and motivation. *International Journal for the Scholarship of Teaching and Learning* 2020; 14 (1). DOI: 10.20429/ijso.2020.140107
54. Krasnoborova A. A. The criteria assessment in the concentration of students. *Pedagogicheskoye obrazovaniye i nauka = Pedagogical Education and Science* [Internet]. 2009 [cited 2020 Jul 20]; 6: 91–94. Available from: <https://www.elibrary.ru/item.asp?id=17062493> (In Russ.)
55. Pinskaya M. A., Ivanov A. V. Formative approach: Criteria-based assessment in action *Narodnoye obrazovaniye = People's Education* [Internet]. 2010 [cited 2020 Jul 20]; 5: 192–201. Available from: <https://www.elibrary.ru/item.asp?id=15516810> (In Russ.)
56. Abekova Zh. A., Oralbaev A. B., Ermahanov M. N., Dzhakipova A. S. Method of forming of educational-cognitive activity student at criterion evaluation knowledge *Uspekhi sovremennogo estestvoznaniya = Advances in Current Natural Sciences* [Internet]. 2015 [cited 2020 Jul 20]; 1 (2): 295–296. Available from: <https://www.elibrary.ru/item.asp?id=23068809> (In Russ.)
57. Lipatnikova I. G. Assessment as a diagnostic procedure of forming the final results of teaching mathematics. *Pedagogicheskoye obrazovaniye v Rossii = Pedagogical Education in Russia* [Internet]. 2016 [cited 2020 Jul 20]; 7: 177–182. Available from: <https://www.elibrary.ru/item.asp?id=26638542> (In Russ.)
58. Sokolova E. V. Construction of diagnostic tasks in the conditions of criteria-based assessment of students' achievements in the study of school geometry course. *Prepodavatel' XXI vek = Teacher XXI Century* [Internet]. 2016 [cited 2020 Jul 20]; 4: 277–287. Available from: <https://www.elibrary.ru/item.asp?id=27633963> (In Russ.)
59. Lok B., McNaught C., Young K. Criterion-referenced and norm-referenced assessments: Compatibility and complementarity. *Assessment & Evaluation in Higher Education*. 2016; 41 (3): 450–465. DOI: 10.1080/02602938.2015.1022136
60. Brown G., Bull J., Pendlebury M. Assessing student learning in higher education [Internet]. London: Routledge; 1997 [cited 2020 Aug 03]. 317 p. Available from: https://archive.org/details/assessingstudent1997brow_kl150
61. Barrie S., Brew A., McCulloch M. Qualitatively different conceptions of criteria used to assess student learning. In: *Proceedings of the AARE-NZARE Conference* [Internet]; 1999; Melbourne. 1999 [cited 2020 Aug 03]. Available from: <https://www.aare.edu.au/data/publications/1999/bre99209.pdf>
62. Wiliam D. Assessing authentic tasks: Alternatives to mark-schemes. *Nordic Studies in Mathematics Education* [Internet]. 1994 [cited 2020 Aug 04]; 2 (1): 48–68. Available from: https://www.dylanwiliam.org/Dylan_Wiliams_website/Papers_files/Nomad.PDF
63. Green S. Criterion referenced assessment as a guide to learning – the importance of progression and reliability. In: *Association for the Study of Evaluation in Education in Southern Africa International Conference* [Internet]; 2002 July 10–12. Johannesburg; 2002 [cited 2020 Jun 04]. Available from: <https://www.cambridgeassessment.org.uk/Images/109693-criterion-referenced-assessment-as-a-guide-to-learning-the-importance-of-progression-and-reliability.pdf>
64. Popham W. J. Criterion-referenced measurement: Half a century wasted? *Educational Leadership* [Internet]. 2014 [cited 2020 Aug 05]; 71 (6): 62–66. Available from: <http://www.educationalleadership-digital.com/educationalleadership/201403?pg=65#pg65>
65. Grigorenko O. V., Rvanova A. S., Shmigirilova I. B. Problems of introduction of criteria-related assessment in school education practice. *Sibirskiy uchitel' = Siberian Teacher* [Internet]. 2020 [cited 2020 Jul 27]; 2 (129): 13–19. Available from: <https://www.elibrary.ru/item.asp?id=43044448> (In Russ.)

66. Linn R. L. Criterion-referenced measurement: A valuable perspective clouded by surplus meaning. *Educational Measurement: Issues and Practice*. 1994; 13 (4): 12–14. DOI: 10.1111/j.1745-3992.1994.tb00564.x
67. Sadler D. R. The futility of attempting to codify academic achievement standards. *Higher Education*. 2014; 67 (3): 273–288. DOI: 10.1007/s10734-013-9649-1
68. Dawson P. Assessment rubrics: Towards clearer and more replicable design, research and practice. *Assessment & Evaluation in Higher Education*. 2017; 42 (3): 347–360. DOI: 10.1080/02602938.2015.1111294
69. Wiggins G., McTighe J. Understanding by design [Internet]. Alexandria, VA: Association for Supervision & Curriculum Development; 2005 [cited 2020 Jun 11]. 370 p. Available from: <https://eric.ed.gov/?id=ED509029>
70. Haraway D., Flinders D. Rubrics in context. In: Rubric nation: Critical inquiries on the impact of rubrics in education [Internet]. Charlotte, NC: Information Age; 2015 [cited 2020 Aug 11]. p. 119–134. Available from: https://nsuworks.nova.edu/fse_facbooks/130/
71. Chan Z., Ho S. Good and bad practices in rubrics: The perspectives of students and educators. *Assessment & Evaluation in Higher Education*. 2019; 44 (4): 533–545. DOI: 10.1080/02602938.2018.1522528
72. Humphry S. M., Heldsinger S. A Two-stage method for classroom assessments of essay writing. *Frontiers in Education*. 2019; DOI: 10.1111/jedm.12223
73. Chadha P., et al. A rubric guide for new academics, practitioner research project report [Internet]. TU Dublin; 2019 [cited 2020 Aug 11]. Available from: <https://arrow.tudublin.ie/ltpgdrprp/23/>
74. Jescovitch L. N., Scott E. E., Cerchiara J. A., Doherty J. H., Wenderoth M. Pat, Merrill J. E., Urban-Lurain M., Haudek K. C. Deconstruction of holistic rubrics into analytic rubrics for large-scale assessments of students' reasoning of complex science concepts. *Practical Assessment, Research, and Evaluation* [Internet]. 2019 [cited 2020 Jun 13]; Vol. 24, Article 7. Available from: <https://par.nsf.gov/servlets/purl/10166313> DOI: 10.7275/9h7f-mp76
75. Randel B., Apthorp H., Beesley A. D., Clark T. F., Wang X. Impacts of professional development in classroom assessment on teacher and student outcomes. *The Journal of Educational Research*. 2016; 109 (5): 491–502. DOI: 10.1080/00220671.2014.992581
76. Emelyanova I. N., Teplyakova O. A., Efimova G. Z. Modern evaluation methods at various levels of education. *Obrazovanie i nauka = The Education and Science Journal*. 2019; 6 (21): 9–28. DOI: 10.17853/1994-5639-2019-6-9-28 (In Russ.)

Информация об авторах:

Шмигирилова Ирина Борисовна – кандидат педагогических наук, доцент кафедры математики и информатики Северо-Казахстанского государственного университета им. М. Козыбаева; ORCID 0000-0003-0235-1640, Scopus ID 57210787308, Researcher ID O-9939-2018; Петропавловск, Республика Казахстан. E-mail: irinankzu@mail.ru

Рванова Алла Сергеевна – кандидат педагогических наук, доцент кафедры математики и информатики Северо-Казахстанского государственного университета им. М. Козыбаева; ORCID 0000-0002-7235-1016; Петропавловск, Республика Казахстан. E-mail: alla_rv@mail.ru

Григоренко Ольга Викторовна – кандидат физико-математических наук, доцент, заведующая кафедрой высшей математики Сибирского государственного университета

геосистем и технологий; ORCID 0000-0001-8957-8746, Researcher ID AAO-9344-2020; Новосибирск, Россия. E-mail: ogrigorenko2311@mail.ru

Вклад соавторов:

И. Б. Шмигирилова – разработка методологии исследования; обзор литературы; анализ данных; формирование первоначальных выводов; оформление результатов исследования.

А. С. Рванова – обзор литературы; обработка данных; критический анализ и доработка текста статьи.

О. В. Григоренко – обзор литературы; обработка данных; критический анализ и доработка текста статьи.

Информация о конфликте интересов. Авторы заявляют об отсутствии конфликта интересов.

Статья поступила в редакцию 10.09.2020; принята в печать 12.05.2021.

Авторы прочитали и одобрили окончательный вариант рукописи.

Information about the authors:

Irina B. Shmigirilova – Cand. Sci. (Education), Associate Professor, Department of Mathematics and Informatics, M. Kozybaev North Kazakhstan State University; ORCID 0000-0003-0235-1640, Scopus ID 57210787308, Researcher ID O-9939-2018; Petropavlovsk, Republic of Kazakhstan. E-mail: irinankzu@mail.ru

Alla S. Rvanova – Cand. Sci. (Education), Associate Professor, Department of Mathematics and Informatics, M. Kozybaev North Kazakhstan State University; ORCID 0000-0002-7235-1016; Petropavlovsk, Republic of Kazakhstan. E-mail: alla_rv@mail.ru

Olga V. Grigorenko – Cand. Sci. (Phys.-Math.), Associate Professor, Head of the Department of Higher Mathematics, Siberian State University of Geosystems and Technologies; ORCID 0000-0001-8957-8746, Researcher ID AAO-9344-2020; Novosibirsk, Russia. E-mail: ogrigorenko2311@mail.ru

Contribution of the authors:

I. B. Shmigirilova – development of the methodological component of the study; literature review; data analysis; formation of initial conclusions; presentation of research results.

A. S. Rvanova – literature review; data processing; critical analysis and revision of the text of the article.

O. V. Grigorenko – literature review; data processing; critical analysis and revision of the text of the article.

Conflict of interest statement. The authors declare that there is no conflict of interest.

Received 10.09.2020; accepted for publication 12.05.2021.

The authors have read and approved the final manuscript.