

КАЧЕСТВО
ОБРАЗОВАНИЯ
В ЕВРАЗИИ



МНОГОМЕРНЫЕ ИССЛЕДОВАНИЯ В ОБЛАСТИ ОЦЕНКИ КАЧЕСТВА ОБРАЗОВАНИЯ

БАГДАСАРЯН АРСЕН ГЕВОРКОВИЧ

Заместитель директора Центра оценки и тестирования при Правительстве Республики Армения, д-р физ.-мат. наук, проф.

E-mail: angen@arminco.com

Ереван, Республика Армения

АННОТАЦИЯ. В статье рассматриваются некоторые особенности масштабных исследований (многомерных исследований) в области оценки качества образования. Выделяются три основных типа масштабных исследований (V-исследования, S-исследования и D-исследования), и описываются некоторые их свойства и связи. Обсуждается и обосновывается необходимость реализации исследовательских проектов в рамках содружеств государств (например, СНГ).

КЛЮЧЕВЫЕ СЛОВА: исследования в области оценки качества образования, тест, анкета, ТИМСС.

1. ВВЕДЕНИЕ

С середины прошлого века во многих странах проводятся масштабные педагогические исследования в области оценки качества образования. Важность таких исследований достаточно очевидна: реализовывая собственные образовательные программы и оценивая их эффективность, страна неизбежно сталкивается с необходимостью сравнения своих достижений с изученным мировым опытом и достижениями других стран. Отсутствие международных исследований привело бы к ситуации, напоминающей корабль, плывущий в открытом океане в туманную погоду, без навигационной карты, без ориентиров и плывущих мимо кораблей. В такой ситуации невозможно быть уверенным, что «корабль» системы образования движется в правильном направлении и с нужной скоростью.

Как и всякое масштабное мероприятие с большим количеством участников, международные исследования в области оценки качества образования должны проводиться методами, приемлемыми для всех участников исследования. Должен быть хорошо продуманный инструмент исследования, который в целом годился бы для каждой страны в отдельности. Учитывая, что количество стран-участниц (с разными программами обучения и традициями преподавания) с каждым годом растет, становится понятно, что каждое международное исследование, в той или иной степени, основано на некоем компромиссе, который, с одной стороны, обеспечивает приемлемость инструментов и методов для каждой страны-участницы, с другой стороны, дает некоторую погрешность при описании ситуации в конкретной стране. К такой погрешности может привести, например, неполное соответствие программ обучения: допустим, в международных исследованиях по математике большой процент составляют задания по разделам, которым до недавнего времени в наших программах практически не уделялось внимания (элементы теории вероятностей и математической статистики, комбинаторика, представление и интерпретация данных и др.). С другой стороны, традиционно находящиеся в поле зрения наших программ разделы (квадратные уравнения, системы уравнений, тригонометрия и др.) не в достаточной мере отражены в заданиях международных исследований.

Другой причиной, которая способна привести к погрешности результатов исследования, могут стать различия в выборе, формулировках, дизайне инструмента оценивания, сформированного на многолетних традициях преподавания и оценивания в разных странах. Такие различия, особенно в формулировках заданий, не могут не отразиться на общей картине успеваемости, даже не учитывая, что метод тестирования сам по себе уже является довольно новым в наших странах.

Так, не секрет, что ученики из бывшего СССР больше привыкли к сухим и четким формулировкам заданий («решить уравнение», «упростить выра-

жение»), в то время как большинство заданий международных исследований формулируется в реальных ситуациях, приближенных к повседневной жизни. Зачастую ученик непосредственно перед выполнением задания должен прочитать некий сценарий и затем, вникнув в суть, выполнить задание. Во многих случаях наш ученик, имеющий необходимые знания и умения для того, чтобы справиться с таким заданием, оставляет его недочитанным из-за недостаточной мотивации, непривычности ситуации или недостаточных навыков чтения и понимания прочитанного.

Чаще подобная ситуация наблюдается в тестах по предметам естествознания. Нередко для выполнения заданий с открытым ответом требуется изложить свое мнение или описать некий рецепт решения предложенной проблемы. Обычно их требования формулируются словосочетаниями типа: «на твой взгляд», «как тебе кажется», «что бы ты предложил», «что бы ты посоветовал». Наши же ученики привыкли производить многошаговые расчеты с использованием нескольких формул, находить значение неизвестной переменной и записывать его в качестве ответа в отведенном месте. Зачастую вследствие этого наш ученик, усвоивший теоретический материал и умеющий с помощью формул производить необходимые расчеты при решении задач, не отвечает на вопросы типа «что бы ты сделал, чтобы не заразиться от одноклассника, больного гриппом?».

В других странах может наблюдаться противоположная картина. Так (см. [1]), норвежские ученые на основании данных, полученных с помощью исследований TIMSS и TIMSSAdvanced, пришли к выводу, что в Норвегии, как в начальной школе, так и в старшей, не уделяется достаточного внимания преподаванию чисто теоретических разделов математики (арифметике, алгебре).

Изложенные доводы не являются критикой международных исследований (на основании выводов, сделанных в процессе международных исследований, в конечном итоге корректируются программы обучения, методы преподавания, совершенствуются учебные материалы и т.д.), однако убеждают в том, что **наряду с международными исследованиями в области оценки качества образования есть необходимость проведения исследований на более узкой целевой группе, связанной достаточной общностью культурных и исторических особенностей, где нет значительных различий в программах обучения, традициях и приоритетах отношения к предмету и способу исследования.** В качестве таких исследований в настоящей статье рассматриваются педагогические исследования в области оценки качества образования, проводимые в рамках некоторого содружества государств (например, СНГ). Полученные результаты помогут понять, интерпретировать и дополнить данные международных исследований, откорректировать соответствующие выводы и рекомендации и получить полноценную картину состояния системы образования как в мировом контексте, так и с учетом региональных и национальных, культурных и исторических особенностей.

2. ТИПЫ ИССЛЕДОВАНИЙ

В области оценки качества образования (в зависимости от основных целей, контекста, тематики, целевой группы и других характеристик) могут существовать, а во многих случаях уже существуют, различные типы исследований. Последние могут проводиться как в рамках отдельно взятого учебного заведения (школы, колледжа, вуза), так и в масштабах всего мира (международные исследования). Если рассматривать эти два случая (I_{\min} и I_{\max}) как концевые позиции некоторого множества $\{I\}$ исследований, упорядоченного по правилу увеличения целевой группы, то возникает задача описания и изучения промежуточных позиций этого множества:

$$I_{\min} < \dots < I_{\max}$$

где I_{\min} — исследование в рамках отдельно взятого учебного заведения, а I_{\max} — международное. Речь идет об $\{I\}$, направленных на некоторую целевую группу, объединенную по некоторому общему признаку, которая шире рамок отдельного учебного заведения ($I_{\min} \neq I$), но не охватывает всю мировую общность ($I \neq I_{\max}$):

$$I_{\min} < I < I_{\max} \quad (1)$$

Примером этого типа могут служить национальные исследования, проводимые по какой-либо тематике в рамках конкретного государства с конкретной системой образования.

Таким образом, по описанному принципу можно выделить, например, следующие типы исследований:

Международные исследования;

Исследования в рамках содружеств государств;

Национальные исследования;

Региональные исследования;

Исследования в рамках отдельного учебного заведения.

Как отмечалось ранее, исследования из множества $\{I\}$ различаются по принципу выбора целевой группы, однако надо понимать, что это не формальный количественный принцип различения типов исследований. Выбор целевой группы наверняка будет диктовать и другие характеристики исследования. Так, при международных исследованиях I_{\max} одной из основных целей является сравнение систем образования стран-участниц исследования и выявление трендов развития системы образования в сравнении со среднемировыми достижениями в исследуемой тематике, а показатели успеваемости отдельно взятого ученика не столь важны. С другой стороны, при I_{\min} -исследованиях в рамках конкретного учебного заведения на первый план выходят другие цели (эффективность методов преподавания, условия препо-

давания, показатели освоения материала, оснащенность учебными материалами и т. д.), важные для успешного функционирования данного учебного заведения. В этом случае показатели успеваемости каждого ученика будут важны.

3. ИССЛЕДОВАНИЯ В РАМКАХ СОДРУЖЕСТВ ГОСУДАРСТВ

В настоящее время среди приведенных в предыдущем разделе типов наиболее часто применяются международные и национальные исследования. Необходимость международных исследований достаточно очевидна и доказана на протяжении десятилетий успешной реализацией известных международных проектов, таких как TIMSS, PIRLS, PISA. При изучении достижений и тенденций развития страны в любой отрасли важно знать ее место на мировой шкале достижений и иметь возможность сравнить ее результаты с имеющимся международным опытом.

С другой стороны, очевидна необходимость всестороннего исследования ситуации в национальном контексте перед предъявлением результатов для международного сравнения. Неестественно участие страны, например, в чемпионате мира по футболу, если она не проводит свой собственный национальный чемпионат.

Если для международных исследований важны проявления отдельной страны в общем мировом пространстве систем образования, то национальные исследования призваны демонстрировать проекции этих проявлений на национальную действительность. В первом случае важны тенденции развития, выявляемые в процессе сравнения с другими участниками исследования (с другими странами), во втором — реальные проявления этих тенденций в рамках страны.

Однако большой спектр различий систем образования стран — участников исследования (различия программ исследуемого материала, уровней образования, глубины и методов преподавания, национальные и исторические различия) в некоторых случаях приводит к уменьшению эффективности международного исследования, основанного на сравнении этих систем. Естественно предположить, что более эффективным будет сравнение систем стран со схожими культурными и историческими традициями. Так, по очевидным причинам неэффективным будет участие, например, Армении в международном исследовании по английскому языку среди следующего списка стран-участниц: США, Великобритания, Австралия, Армения. С другой стороны, интересно было бы провести исследование по русскому языку, например, среди следующего списка стран: Молдавия, Казахстан, Грузия, Армения.

Перечисленные соображения показывают необходимость проведения международных исследований в рамках стран, объединенных по некоторым (культурным, историческим и другим) критериям. В приведенной выше типологии мы условно назвали их «исследования в рамках содружеств госу-

дарств». Они имеют шанс включить в себя преимущества международных исследований, сохраняя эффект сравнения, и преимущества национальных, где нет пропасти культурных и исторических различий. Здесь могут быть успешно использованы методы международных исследований с учетом особенностей национальных.

4. ГЕОМЕТРИЧЕСКАЯ ИНТЕРПРЕТАЦИЯ

В предыдущем параграфе мы постарались обосновать, что, наряду с проявлениями отдельной страны в общем мировом пространстве систем образования (международные исследования) и изучением проекций этих проявлений на национальную действительность (национальные исследования), важное значение имеют промежуточные (в смысле формулы (1)) исследования в рамках содружеств государств, которые будут отражать проекцию системы образования рассматриваемой страны на плоскость достижений в рамках содружества государств.

Напрашивается трехмерная геометрическая интерпретация обсуждаемых понятий: если систему образования конкретной страны представить в виде тела в трехмерном пространстве, то определяющими становятся все три составляющие (оси) полноценного изучения ситуации:

- **международная составляющая;**
- **составляющая содружества государств;**
- **национальная составляющая.**

Таким образом, получаем следующую схему описания системы образования рассматриваемой страны,

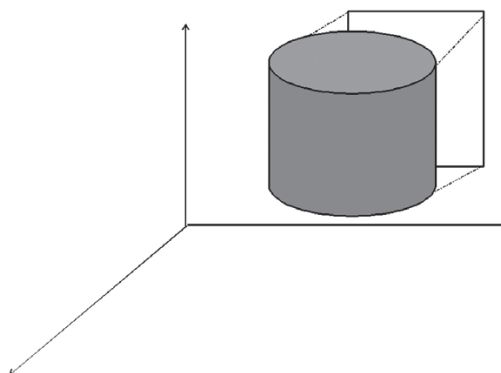


Рисунок 1

где: абсцисса — национальная составляющая, ордината — составляющая содружества, аппликата — международная составляющая.

В контексте рисунка 1, следуя предложенной геометрической аналогии, можно условно поставить задачу вычисления объема исследуемого тела (международные исследования), нахождения площади его плоской проекции (исследования в рамках содружества государств) и задачу вычисления длины его одномерной проекции на ось национальной составляющей (национальные исследования). Исходя из этого, следуя общепринятым в геометрии обозначениям, будем пользоваться следующей терминологией:

V-исследованиями будем называть международные исследования;

S-исследованиями обозначим исследования в рамках содружеств государств;

D-исследованиями будем называть национальные исследования.

Таким образом, приведенной в разделе 2 типологии исследований мы придали некую размерность, вследствие чего будем называть их многомерными. В рассматриваемой терминологии упорядоченность исследований, предложенная в формуле (1), приводит к следующей упорядоченности для исследований из вышеприведенного списка:

$$D < S < V \quad (2)$$

Естественно в этом случае полагать (см. рис. 1), что международные V-исследования — трехмерны, S-исследования в рамках содружеств государств — двумерны, а национальные D-исследования — одномерны.

По геометрической аналогии рисунка 1 можно заметить, что одно и то же трехмерное тело может иметь разные проекции как на плоскость «содружества государств», так и на «национальную» ось:

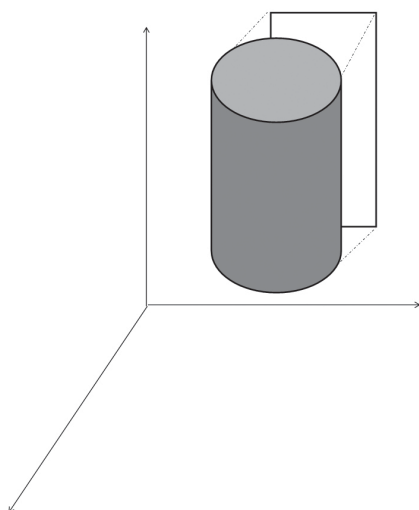


Рисунок 2

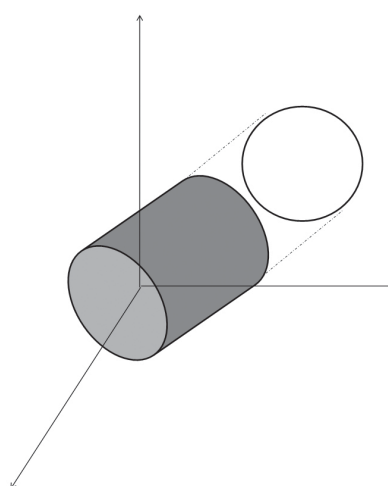


Рисунок 3

На рисунках 2 и 3 показано одно и то же тело, которое имеет совершенно разные проекции на плоскость «содружества государств». По аналогии можно представить, что страны, проявившие себя примерно одинаково в V-исследовании (одинаковые тела, равные объемы) по какому-либо предмету, имеют абсолютно разные показатели при S-исследовании (разные проекции, разные площади проекций) по тому же предмету. Нетрудно привести примеры и обратного явления, когда совершенно разные трехмерные тела имеют одинаковые проекции.

Можно указать множество причин такого явления: различия в программах обучения, в традициях и методах преподавания, в приоритетах прохождения материала, в инструментах и методах исследования и т. д., но основной вывод таков (см. Свойство 3):

Исследования разных размерностей по одному и тому же предмету не могут заменять, а должны лишь дополнять и улучшать друг друга.

5. ОСНОВНЫЕ ТИПЫ ИССЛЕДОВАНИЙ

Основными будем называть V-, S- и D-исследования из приведенной выше терминологии. Перечислим некоторые характеристики и примеры основных исследований.

5.1. V-ИССЛЕДОВАНИЯ

Основную цель V-исследований можно охарактеризовать как сравнение образовательных систем и показателей успеваемости учеников в пространстве международных стандартов на материале общности программ обучения исследуемой тематики с помощью разработанных общеприемлемых инструментов.

V-исследования призваны четко определить место страны по исследуемой тематике среди всего международного разнообразия, сравнить систему образования страны со среднемировым опытом и показателями других участников V-исследования, выявить и указать сильные и слабые стороны системы образования в мировом контексте, продемонстрировать тенденции развития (или, не дай Бог, отставания).

Тематикой V-исследований могут служить международные ценности: математика, предметы естествознания, навыки и умения чтения и т. д.

В качестве примеров приведем три наиболее известных международных исследования:

- TIMSS (Trends in International Mathematics and Science Study) — исследование по математике и предметам естествознания для учеников 4-х и 8-х классов, проводимое ассоциацией IEA (The International Association for the Evaluation of Education Achievement),

- PIRLS (Progress in International Reading Literacy Study) — исследование по чтению и пониманию для учеников начальной школы, проводимое IEA,
- PISA (Program for International Student Assessment) — исследование в трех областях: чтение, математика, предметы естествознания для учеников старшей школы, проводимое организацией OECD (Organization for Economic Cooperation and Development).

Приведенные международные исследования TIMSS, PIRLS, PISA хорошо известны и изучены, поэтому, не вдаваясь в подробности описания инструментов и методов этих исследований, отметим лишь малую часть публикаций, посвященных им [2–7].

5.2. S-ИССЛЕДОВАНИЯ

Основной целью S-исследований может служить сравнение образовательных систем и показателей успеваемости учеников стран-участниц на плоскости стандартов содружества государств, на материале общности программ обучения исследуемой тематики с помощью разработанных общеприемлемых инструментов.

Тематикой S-исследований могут являться общие ценности стран содружества: история, русский язык и литература (для стран СНГ), английский язык (для стран Британского сообщества) и т. д. Для стран СНГ вопросы могут относиться к общей истории (например, о Великой Отечественной войне) и к общему культурному наследию (например, А. С. Пушкин, Чингиз Айтматов, Арам Хачатурян).

С другой стороны, любая тематика любого V-исследования может служить тематикой и для соответствующего S-исследования (но не наоборот, см. Свойство 1). Так, S-исследования могут проводиться и по математике, и по предметам естествознания, и по исследованию навыков и умений чтения и т. д. К сожалению, нам не известны примеры S-исследований (во всяком случае, среди стран бывшего СССР). Однако есть надежда, что в качестве примеров таких исследований в будущем можно будет назвать, например, следующие проекты:

- ICT (Information Communication Technologies) — исследование в области информационно-коммуникационной грамотности,
- OLAS (аббревиатура на армянском языке) — исследование в области иностранных языков (например, русский и английский — для стран СНГ).

5.3. D-ИССЛЕДОВАНИЯ

D-исследования посвящены изучению особенностей национальной (региональной) системы образования и успеваемости учеников на национальной шкале стандартов, на материале принятой программы обучения с помощью разработанных национальных инструментов.

Тематикой D-исследований могут служить национальные ценности: родной язык и литература, история страны и т. д. Тематикой D-исследования может являться также тематика любого V- или S-исследования, адаптированная к системе образования рассматриваемой страны.

Во всем мире проводится достаточно много национальных D-исследований. В Армении это, в первую очередь, следующие исследования:

- НААС (аббревиатура на армянском языке) — национальное исследование по армянскому языку и литературе, истории Армении,
- ВААС (аббревиатура на армянском языке) — национальное исследование по предметам естествознания.

Отметим также национальные исследования, которые проводятся в настоящее время (см. [8]):

- исследование «Что мы знаем о себе?» по армяноведческим предметам (при финансовой поддержке Государственного комитета по науке Республики Армения),
- исследование по русскому языку (при финансовой поддержке Программы малых грантов центра CISED (The Center for International Cooperation in Education Development)).

6. НЕКОТОРЫЕ СВОЙСТВА МНОГОМЕРНЫХ ИССЛЕДОВАНИЙ

Рассмотрим некоторые особенности, которые свойственны многомерным исследованиям в области оценки качества образования, и некоторые взаимосвязи исследований разных размерностей.

СВОЙСТВО 1. Тематика исследования с большей размерностью может быть применена и для исследования с меньшей размерностью.

Утверждение Свойства 1 достаточно очевидно. На примере исследований из формулы (2) оно означает, что по тематике международного V-исследования (например, по математике и предметам естествознания, как в исследованиях TIMSS) можно проводить соответствующие S- и D-исследования как в рамках содружеств государств, так и в национальном масштабе.

Заметим, что обратное утверждение к Свойству 1, вообще говоря, неверно. Не всякое, например, национальное исследование может быть расширено до размеров международного (например, национальные исследования по родному языку).

СВОЙСТВО 2. Уменьшение размерности исследования не всегда ведет к сужению его тематики.

Свойство 2 означает, в частности, что некоторые разделы рассматриваемого предмета, которые не годятся для оценивания в масштабе

международного V-исследования (не входят в программу значительного числа стран-участниц или не соответствуют приоритетам преподавания рассматриваемого предмета), могут быть включены в тематику S- или D-исследований (вследствие совпадения или схожести программ и традиций). Утверждение, обратное к Свойству 2, тоже верно. Например, в программе исследований TIMSS Advanced по углубленной математике и физике есть разделы, которые составляют часть уже университетского курса (во всяком случае в Армении).

Практически все вышеизложенное приводилось для обоснования следующего основного свойства многомерных исследований (об этом уже говорилось в п. 4).

СВОЙСТВО 3. *Исследования разных размерностей (по фиксированной тематике) не могут заменять, а могут дополнять и улучшать друга.*

Следующее свойство тоже достаточно очевидно и важно.

СВОЙСТВО 4. *Изменение размерности исследования должно сопровождаться адаптацией как контекста, так и формулировки и дизайна инструмента.*

Во многих случаях качество исследования зависит от тщательного соблюдения именно этого свойства. Иначе полученные результаты не будут отражать реальную ситуацию и будут несостоятельны в плане международного (национального, регионального) сравнения.

7. ЗАКЛЮЧЕНИЕ

Удивительно, но факт, что тестология как дисциплина еще 25 лет назад была практически неизвестна на территории бывшего СССР (и достаточно успешно развивалась и применялась во всем цивилизованном мире). Это тем более странно, если иметь в виду, что, несмотря на все недостатки (политические, экономические) обустройства СССР, фундаментальная наука у нас всегда была в почете, а по многим направлениям даже занимала лидирующие позиции в мире. Естественно, ни о каких масштабных исследованиях в области оценки качества образования речи в то время не шло. Видимо, дело было в том, что советская система образования (в отличие от политики и экономики) в целом достаточно эффективно справлялась со своей задачей, и такого рода исследования не были столь актуальны. Сказалась также недостаточная развитость компьютерных технологий.

С распадом СССР распалась и советская система образования, а для систем образования вновь появившихся стран возникла необходимость изучения сложившейся в образовании ситуации для выбора и совершенствования вектора развития. Оказалось, что в мире накоплен достаточный опыт в области оценивания качества образования, который достигается с помощью масштабных международных исследований. Участие в соответствующих международных исследованиях выявило как сильные стороны, так и проблемы систем образования наших стран в контексте общепринятых международных стандартов в области образования. Национальные же особенности в образовании, сложившиеся на основании национальных традиций и приоритетов преподавания, изучаются (должны изучаться) с помощью масштабных национальных исследований в рамках отдельных стран.

Мы считаем, что культурно-исторический и научно-образовательный опыт совместного пребывания наших стран в пределах единого государства содержит богатый потенциал для исследования и при профессиональной качественной его реализации, наряду с международными и национальными исследованиями в области оценки качества образования, может стать мощным средством развития (а при необходимости и реформирования) систем образования стран бывшего Союза. Создание ЕАОКО (Евразийской Ассоциации оценки качества образования) является первым и обнадеживающим шагом в этом направлении. Разработка усилиями ЕАОКО качественных инструментов оценивания и успешное их применение в соответствующих исследовательских проектах могут дать новый толчок в сфере образования стран бывшего СССР.

СПИСОК ЛИТЕРАТУРЫ

- 1) Grønmo L.S., Onstad T. The significance of TIMSS and TIMSS Advanced. 2013. P. 138.
- 2) Mullis I., Martin M., Chrostowski S. TIMSS 2003 Technical Report, TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College, Library of Congress, Catalog Card Number: 2004111600. ISBN: 1-889938-35-1. P. 503.
- 3) Mullis I., Martin M., Kennedy A. PIRLS 2006 Technical Report, TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College, Library of Congress, Catalog Card Number: 2007925071. ISBN: 1-889938-46-7. P. 287.
- 4) Результаты международного исследования TIMSS-2011 / под ред. Г. С. Ковалевой // Сайт Центра оценки качества образования [Электронный ресурс]. — Режим доступа: <http://www.centeroko.ru/public.htm>.
- 5) Результаты международного исследования PIRLS-2011 / под ред. Г. С. Ковалевой // Сайт Центра оценки качества образования [Электронный ресурс]. — Режим доступа: <http://www.centeroko.ru/public.htm>.
- 6) PISA 2003: Technical Report / ed. R. Adams, M. Wu. — OECD, 2004.
- 7) Wu M. Comparing the Similarities and Differences of PISA 2003 and TIMSS // OECD Education Working Papers. 2010. — № 32.
- 8) Тестирование по предметам армяноведения / А. Г. Багдасарян, М. М. Аладжян, С. Г. Меликян, С. М. Мкртчян, В. А. Мкртчян // Педагогика. — 2012. — № 5. — С. 37—54.