

Что может предложить новая методология оценки мышления школьников современному образованию

И. Л. Угланова, И. Н. Погожина

Статья поступила
в редакцию
в июне 2021 г.

Угланова Ирина Львовна — младший научный сотрудник Центра психометрики и измерений в образовании Института образования, Национальный исследовательский университет «Высшая школа экономики». Адрес: 101000, Москва, Потаповский пер., 16, стр. 10. E-mail: iuglanova@hse.ru (контактное лицо для переписки)

Погожина Ирина Николаевна — доктор психологических наук, доцент кафедры психологии образования и педагогики факультета психологии МГУ им. М. В. Ломоносова. Адрес: 125009, Москва, ул. Моховая, 11, стр. 9. E-mail: pogozhina@mail.ru

Аннотация

Для реализации педагогического принципа доступности и посильности обучения и воспитания необходимо учитывать индивидуальные особенности развития логического мышления учеников. Для его объективной и быстрой оценки необходимы инструменты диагностики. Сегодня наиболее полную оценку содержания логического мышления по Пиаже можно получить, используя метод клинической беседы. Однако данный способ диагностики крайне ресурсозатратен и не подходит для масштабного тестирования. Обзор литературы показал, что существующие стандартизированные методы диагностики требуют привлечения большого числа квалифицированных экспертов для проверки результатов и подготовки обратной связи для учителей, методистов, психологов-практиков и исследователей.

В статье описаны методологические принципы разработки инструмента, предназначенного для оценки уровней развития логического мышления. Такой инструмент позволит автоматически обрабатывать получаемые результаты без потери специфики их содержания, что в итоге упростит и ускорит процедуру диагностики. Реализация этих принципов рассматривается на примере компьютеризированных заданий сценарного типа с фокусом на решении и на процессе работы (*computerized performance-based assessment*) в формате скрытого оценивания (*stealth assessment*) для учеников 5-го и 7-го класса.

Ключевые слова

логическое мышление, критическое мышление, диагностика, Пиаже, автоматическая обработка результатов, психометрика, задания сценарного типа, скрытое оценивание.

Для цитирования

Угланова И. Л., Погожина И. Н. (2021) Что может предложить новая методология оценки мышления школьников современному образованию // Вопросы образования / Educational Studies Moscow. № 4. С. 8–34. <https://doi.org/10.17323/1814-9545-2021-4-8-34>

What the New Measure of Thinking in School Students Has to Offer to Contemporary Education

I. L. Uglanova, I. N. Pogozhina

Irina L. Uglanova, Junior Research Fellow, Center for Psychometrics and Measurements in Education, Institute of Education, National Research University Higher School of Economics. Address: Bld. 10, 16 Potapovsky Ln, 101000 Moscow, Russian Federation. E-mail: iuglanova@hse.ru (corresponding author)

Irina N. Pogozhina, Doctor of Sciences in Psychology, Associate Professor, Department of Psychology of Education and Pedagogics, Faculty of Psychology, Lomonosov Moscow State University. Address: Bld. 9, 11 Mokhovaya Str., 125009 Moscow, Russian Federation. E-mail: pogozhina@mail.ru

Abstract For the pedagogical principle of assigning comprehensible and adequate tasks to be implemented, allowance should be made for students' individual levels of logical reasoning, which requires diagnostic measures for objective and quick assessment. Today, the "clinical method" allows the most comprehensive assessment of logical thinking within the Piagetian framework. However, this diagnostic measure is extremely resource-consuming, hence unsuitable for large-scale testing. An overview of literature shows that the existing standardized diagnostic measures require a great number of highly-qualified experts to review the scores and prepare feedback for teachers, instructional designers, practicing psychologists and researchers.

The article describes design methodology of an instrument to evaluate levels of logical reasoning that will allow automated scoring without sacrificing score meaning, eventually facilitating and accelerating the diagnostic measurement procedure. Implementation of these principles is analyzed using the example of computerized performance-based assessment of scenario-based problem solving in the form of stealth assessment of fifth- and seventh-grade pupils.

Keywords automated scoring, critical thinking, diagnostic measurement, formal operations, logical reasoning, logical thinking, performance-based assessment, Piaget, psychometrics, scenario-based problems, stealth assessment.

For citing Uglanova I. L., Pogozhina I. N. (2021) Chto mozhet predlozhit' novaya metodologiya otsenki myshleniya shkol'nikov sovremennomu obrazovaniyu [What the New Measure of Thinking in School Students Has to Offer to Contemporary Education]. *Voprosy obrazovaniya / Educational Studies Moscow*, no 4, pp. 8–34. <https://doi.org/10.17323/1814-9545-2021-4-8-34>

Один из основополагающих дидактических педагогических принципов — принцип доступности и посильности обучения и воспитания. «Все, подлежащее изучению, должно быть распределено сообразно ступеням возраста так, чтобы предлагалось для изучения только то, что доступно восприятию в каждом возрасте» [Коменский, 1939. С. 151]. В современной школе программы обучения в целом сформированы с учетом стадийности психического развития школьников, и «средний» ученик способен овладеть ими. Вместе с тем проблема академической неуспеваемости остается достаточно острой. Одной из ее

причин выступают особенности темпов развития познавательной системы конкретного ребенка, его неготовность к усвоению школьных знаний на определенной ступени обучения. Эмпирически доказана связь между академической успешностью в овладении различными учебными дисциплинами и уровнем развития интеллекта школьников [Malhotra, 2020; Watkins, Lei, Canivez, 2007]. Показано, что понимание математических и физических понятий, социальных норм и правил требует определенного уровня развития логического мышления [Inhelder, Piaget, 1958; Пиаже, Инхельдер, 2003].

Опыт наиболее эффективных систем образования свидетельствует о том, что проблема школьной неуспеваемости может быть решена путем индивидуализации процесса обучения — признания того, что «все дети особые» [Vainikainen et al., 2015; Hautamäki, Thuneberg, 2019; Hienonen, 2020]. Индивидуализация требует учета, в частности, уровня развития логического мышления школьника, показателями которого выступают результаты решения задач Пиаже [DeVries, 1974; Goldschmid, 1967; Lawson, Renner, 1975; Lovell, Shields, 1967; Lawson, Blake, Nordland, 1974]. А для этого современной школе необходимы пригодные для массового мониторинга, объективные, т. е. независимые от субъективной интерпретации результатов, инструменты его диагностики, так как результаты выполнения заданий традиционных тестов интеллекта не содержат достаточной информации о составе и структуре логических операций, необходимой для адаптации учебных программ к уровням развития учащихся [Avila de, Pulos, 1979; Hathaway, Hathaway-Theunissen, 1975; Kaufman, 1972]. Использование нового диагностического инструментария позволит понять когнитивные причины неуспеваемости конкретного учащегося и выстраивать в дальнейшем траектории обучения с учетом индивидуально-психологических характеристик детей и подростков.

Цель представленной работы состоит в выделении и описании этапов построения инструмента, предназначенного для измерения уровней развития логического мышления школьников с автоматической обработкой результатов диагностики (без привлечения экспертов).

**1. Содержание
оцениваемого
конструкта —
состав
и структура
логического
мышления**

Прежде чем перейти к описанию методологии построения инструментов измерения, уточним содержание конструкта, которое эта методология призвана оценивать.

Логическое мышление не дано человеку изначально, оно развивается с момента рождения, проходя ряд стадий и приобретая характеристики все большей объективности: 1) сенсомоторный интеллект; 2) дооперациональное мышление; 3) стадия

конкретных операций; 4) формально-операциональное мышление. От стадии к стадии отображенные мышлением закономерные связи между характеристиками окружающей действительности организуются во все более обобщенные структуры (логические операции). При этом на уровне поведения мы наблюдаем изменение представлений человека об окружающих его объектах, их свойствах, о пространстве, времени, движении, причинности и др. [Пиаже, 1994а; 1994б; 1994в].

Для успешности обучения в средней и старшей школе важны новообразования, появляющиеся на стадии формальных операций:

- 1) логика высказываний — способность мыслить гипотезами, т.е. выделять (отображать) закономерные связи и отношения не только между реальными объектами и их образами, но и между знаками языковых систем (высказываниями). Строятся системы пропозициональных операций: импликация (если..., то...), дизъюнкция (...или... или... или оба), конъюнкция (...и...), отрицание (... не ...), эквивалентность (... , если и только если...) и др.;
- 2) комбинаторные операции — операции систематического сочетания объектов друг с другом во всех возможных вариантах. Например, последовательное комбинирование шести (семи, десяти, ... n) любых объектов по два, по три ... и т.д. всеми возможными способами систематически (без повторов), сохраняя строгий контроль над результатом;
- 3) объединение двух форм обратимости в единую когнитивную структуру четырех трансформаций INRC (I — прямая операция, внесение изменения в ситуацию; N — обратная операция, отрицание изменения, возврат к исходной ситуации; R — операция реципрокности, учет влияния на ситуацию взаимно согласованных факторов, взаимная трансформация; C — отрицание реципрокности, возврат к начальной ситуации). Подросток становится способен анализировать проблемные ситуации, используя все четыре когнитивные трансформации одновременно.

В содержании его мышления появляются новые операциональные схемы, необходимые для успешного усвоения школьных дисциплин:

- пропорции (соразмерность, равенство двух и более отношений);
- механическое и гомеостатическое равновесие (постоянство состояния системы при воздействии разнонаправленных переменных);

- относительное движение (движение относительно перемещающейся системы отсчета);
- вероятность (возможность наступления события при определенных условиях);
- способность выходить за границы реально наблюдаемого, выдвигать гипотезы (что будет, если...?), мысленно создавая систему вероятных закономерностей и др. [Пиаже, 1994а; 1994б; 1994в; Пиаже, Инхельдер, 2003; Пиаже, 2008].

Таким образом, содержание логического мышления представляет собой сложно организованную систему связанных структур.

2. Оценивание логического мышления методом клинической беседы

Традиционно для изучения логического мышления применяется методология клинической беседы — структурированное или полуструктурированное интервью, в котором интервьюер (эксперт со специальной подготовкой) задает вопросы респонденту в разговоре один на один или в небольших группах. Цель клинической беседы — не только зафиксировать результат решения задачи, как в стандартизированных методах диагностики, но и выяснить, каким образом человек достигает решения: какие мыслительные действия он выполняет или не выполняет и в какой последовательности. Особое значение при интерпретации результатов приобретают допускаемые при решении ошибки как индикаторы несформированности у человека элементов логической операции. На основе полученных данных эксперт делает вывод об уровне сформированности тех или иных логических операций и уровне развития логического мышления в целом [Брангье, 2000; Пиаже, 1994а; 1994б; 1994в; Пиаже, Инхельдер, 2003].

Недостаток данного метода — его высокая ресурсозатратность, особенно при проведении исследований на больших выборках или крупномасштабных мониторинговых срезах. Во-первых, необходимо подготовить достаточное число квалифицированных экспертов, которые смогут провести интервью с минимальным вкладом субъективности в результаты диагностики. Во-вторых, индивидуальные беседы существенно увеличивают время проведения исследования. В большинстве образовательных ситуаций просто не хватает ни специалистов, ни времени для оценки логического мышления методом клинической беседы [Avila de, Pulos, 1979; Meyer, 1972].

Поэтому перед исследователями стоит задача создать методологию для изучения логического мышления, которая бы стала достойной альтернативой клинической беседе. Новая методология должна быть удобной для учителей, методистов, психологов-практиков и исследователей при проведении крупномасштаб-

ного сбора данных. Она должна исключать потерю специфики изучаемого конструкта, т.е. давать возможность учитывать, каким образом мыслительные действия, входящие в состав логической операции, включаются в процесс решения задачи.

Достижение масштабируемости измерений возможно за счет внедрения стандартизированного тестирования, при котором все наблюдаемое в ходе тестирования поведение человека интерпретируется согласно единым, заранее установленным объективным (независимым от эксперта) критериям.

3. Оценивание логического мышления стандартизированными методами

Попытки перевода индивидуального формата тестирования (клинической беседы) в групповой предпринимались последователями Пиаже начиная с 1960-х годов с целью проанализировать соответствие содержания образовательных программ возможностям мышления школьников, например при преподавании предметов естественнонаучного цикла [Lovell, 1961; Shayer, 1978; Shayer, Küchemann, Wylam, 1976].

Исследователи ставили вопрос: позволяет ли групповой формат проведения теста получить такие же результаты, как и клиническая беседа? Показано, что применение клинической беседы и фронтального тестирования (например, стимульный материал демонстрируется на экране, а респонденты записывают свои ответы) дает сопоставимые результаты [Faust, 1983; Renner et al., 1978; Rowell, Hoffmann, 1975; Shayer, 1979]. Но такое фронтальное тестирование не является эквивалентом стандартизированному тесту, так как допускает вариативность в демонстрации стимульного материала и интерпретации результатов.

Мы проанализировали работы, в которых описываются стандартизированные инструменты, предназначенные для измерения уровня развития логических операций на этапе перехода от конкретно-операциональной стадии к формально-операциональной по Пиаже при проведении масштабных количественных исследований и одновременном администрировании теста для групп учеников. Основная задача их авторов заключалась в том, чтобы обеспечить такую же полноту интерпретации результатов, какую дает метод клинической беседы, уменьшая при этом время и затраты на проведение тестирования. Рассмотренные стандартизированные инструменты различаются по ряду важных характеристик¹:

¹ В материалах по ссылке doi:10.17632/vxt3237yvt.1 представлена таблица с систематизацией тестов, психометрические характеристики которых опубликованы в рецензируемых источниках и находятся в открытом доступе (Uglanova I., Pogozhina I. (2021) How can a new methodology for the assessment of logical thinking contribute to modern education? // Mendeley Data, V1).

- 1) в содержании теста различаются:
 - оцениваемый конструкт (логические операции),
 - полнота оценки — результат решения задачи и/или процесс;
- 2) формат ответа может предусматривать:
 - выбор правильного ответа из предложенных (закрытые вопросы, *multiple-choice*),
 - развернутый или краткий самостоятельно формулируемый ответ (открытые вопросы);
- 3) способ обработки результатов может быть:
 - автоматический,
 - с привлечением экспертов;
- 4) стимульный материал может представлять собой: физические объекты, рисунок, текст, видеозапись.

Содержание тестов различается как по набору логических операций, так и по полноте оценки. Несмотря на некоторые различия в представленных в них логических операциях, все тесты диагностируют как конкретно-операциональное, так и формально-операциональное мышление. Что касается полноты оценки, то в тестах может оцениваться только итоговый ответ тестируемого (результат) [Tisher, 1971; Raven, 1973; Milakofsky, Patterson, 1979; Avila de, Pulos, 1979; Roberge, Flexer, 1982; Bergling, 1998; Bakken et al., 2001]. Однако некоторые авторы разработали систему оценивания, в которой учитывается и результат, и процесс решения [Longeot, 1962; 1965; Staver, Gabel, 1979; Lawson, 1978; Tobin, Capie, 1981; Roadranga, 1991]. При этом процесс решения оценивается через задания открытого типа — а значит, требует привлечения экспертов.

С одной стороны, задания закрытой формы имеют очевидные преимущества: сокращают время тестирования, снижают нагрузку на ученика, поскольку не требуют привлечения навыков письма. Кроме того, закрытая форма обеспечивает объективность полученных оценок, тогда как оценки открытых ответов, данные разными экспертами, могут быть несогласованными, что снижает их надежность [Roadranga, 1991]. С другой стороны, закрытый формат не позволяет получить от тестируемого обоснование, почему он выбрал тот или иной ответ, а это сокращает диагностический потенциал инструмента и оставляет место случайному ответу.

Стимульный материал в тестах предъявляется не только в текстовом формате: используются сочетание текста и рисунков [Longeot, 1962; 1965; Tobin, Capie, 1981; Bergling, 1998], только рисунки [Milakofsky, Patterson, 1979; Avila de, Pulos, 1979; Bitner-Corvin, 1988], сочетание текста и видеозаписи [Staver, Gabel, 1979; Tobin, Capie, 1981], сочетание текстового описания и физических объектов [Roberge, Flexer, 1982], только физические объекты [Tisher, 1971; Lawson, 1978].

Полностью текстовый стимульный материал валиден для оценки уровня развития формально-операционального мышления, но увеличивает нагрузку на навыки чтения учеников и поэтому может содержательно исказить результаты оценки конкретно-операциональных структур. Кроме того, отсутствие физических объектов как стимульного материала может приводить к потере чувства значимости задачи [Lawson, 1978]. Использование текстовых стимульных материалов и открытых заданий наиболее проблематично для детей младшего возраста [Lawson, 1978; Roadranga, 1991].

В целом стандартизированные тесты для оценки уровня развития логических структур показывают удовлетворительную надежность и валидность. Как один из способов доказательства их валидности проводится сравнение результатов, полученных в ходе стандартизированного тестирования, с данными клинической беседы. Корреляции итоговых показателей варьируют от незначимых по некоторым шкалам [Staver, Gabel, 1979] до значимых с силой связи 0,88 [Avila de, Pulos, 1979; Bitner-Corvin, 1988].

Таким образом, использование стандартизированных методов позволяет сократить время тестирования, не жертвуя при этом полнотой содержания логического мышления, необходимого для диагностики. Однако проблема оценки результатов выполнения заданий без привлечения экспертов до сих пор не решена. Кроме того, в литературе не представлены инструменты измерения логических структур в ситуациях, включенных в реалистичный и насыщенный, а не лабораторный контекст.

Перспективным направлением диагностики логического мышления (процесса и результата) без привлечения экспертов в насыщенной среде выступают компьютеризированные задания в формате *performance-based*. Этот тип заданий позволяет использовать автоматизированную систему начисления баллов, оценивать процесс и результат решения и создавать ситуацию тестирования, приближенную к реальным жизненным задачам [Wang, Shute, Moore, 2015].

4. Требования к новой методологии

Новая методология разработки инструментов для диагностики логического мышления должна, во-первых, позволять диагностировать процесс решения задачи и, во-вторых, предоставлять обратную связь — а значит, анализировать результаты, — в том числе учителям, родителям учеников, методистам, психологам-практикам, без привлечения экспертов. При этом новые инструменты измерения должны отвечать требованиям стандартов тестирования [American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 2014; Messick, 1992].

4.1. Реализация диагностики процесса решения задачи

Именно «природа исследуемого конструкта должна руководить подбором и созданием релевантных тестовых заданий, а также установкой целесообразных критериев и рубрик оценивания» [Messick, 1992. Р. 17]. Традиционные задания закрытой формы (*multiple choice*) не подходят для оценки процесса решения, так как они фиксируют только результат (верно/неверно), не учитывая специфику способа достижения результата [Griffin, McGaw, Care, 2012; Messick, 1994; Razzouk, 2011].

Альтернативным форматом выступают задания типа *performance-based* [Messick, 1994], которые фокусируются как на самом решении, так и на процессе. Сегодня задания *performance-based* успешно реализованы для измерения сложных конструктов в компьютерной форме в формате игр, симуляций и заданий сценарного типа [Klerk de, Eggen, Veldkamp, 2016; Graesser, Kuo, Liao, 2017; Sun et al., 2020]. Однако инструменты данного типа для измерения логического мышления в рамках теории Пиаже в литературе пока не представлены.

4.2. Реализация обратной связи без привлечения экспертов

Переход от методологии клинической беседы к стандартизированному оцениванию в формате *performance-based* обеспечивает уникальную комбинацию достоинств заданий закрытой и открытой формы: мы преодолеваем ограничение анализа только определением соответствия продукта деятельности тестируемого правильному ответу и оцениваем процесс решения без обращения к экспертам.

Еще одно преимущество заданий *performance-based* состоит в том, что они могут быть реализованы в формате скрытого оценивания (*stealth assessment*), при котором ситуация тестирования встраивается в насыщенную учебную или игровую среду, и респондент обычно не знает, что его оценивают [Wang, Shute, Moore, 2015]. Такая организация тестирования позволяет, с одной стороны, снизить тревожность тестируемых и повысить их мотивацию, а с другой — подобрать такой контекст заданий, который позволит приблизить интерпретацию результатов к повседневным учебным и бытовым задачам.

Наконец, диагностическая рамка должна нести в себе возможности использования результатов тестирования для создания в будущем программ психологической помощи тем ученикам, у которых уровень развития логического мышления не в полной мере соответствует критериям овладения учебной дисциплиной. Предлагаемый нами подход к диагностике формально-операциональных структур направлен на определение их состава (деятельностную диагностику в терминологии ряда авторов [Ильясов, 1986; Талызина, 2018]). Он открывает возможность оценить сформированность или несформированность не только логической структуры как таковой, но и отдельных

элементов — познавательных действий, входящих в ее состав. Как и в методе клинической беседы, здесь очень важны анализ и интерпретация ошибок, которые допускает ученик в ходе решения. Под ошибкой мы понимаем невыполнение или неверное выполнение того или иного познавательного действия, которое входит в состав оцениваемой логической структуры и необходимо для успешного продвижения в сценарном контексте. На основе собранных данных определяется стадия развития логической структуры.

Полученные результаты диагностики в дальнейшем могут стать основой для построения ориентировочной основы действия для формирования недостающих элементов «проблемной» логической операции, разработки и реализации системы педагогических условий ее интериоризации в рамках деятельностного подхода к учению и овладению социальным опытом. Предлагаемая методология диагностики состава и структуры логических операций позволит точно оценить «проблемные зоны» у максимально большого числа учеников, чтобы затем предложить каждому соответствующую формирующую программу.

**5. Предлагаемая
новая методо-
логия оценки
логического
мышления
школьников**

Предлагаемая нами новая методология оценки логического мышления основана:

- на выделении содержания логических операций в периодизации Пиаже [Пиаже, 1994а; 1994б; 1994в], их состава и структуры как диагностической рамки деятельностной диагностики и основы построения будущих формирующих программ [Ильясов, 1986; Талызина, 2018];
- анализе достижений и недостатков применяемых в настоящее время методов диагностики логического мышления;
- современных психометрических методах и цифровых технологиях.

Для оценки логического мышления школьников мы предлагаем использовать задания сценарного типа, погружающие учеников в насыщенную среду, приближенную к учебным и повседневным задачам. В заданиях сценарного типа тестируемые последовательно выполняют связанные между собой задачи в рамках подготовленной разработчиками контекстно-насыщенной истории (сюжета). Процесс и результат решения этих задач выступают поведенческими индикаторами, на основе которых делается вывод о сформированности того или иного познавательного действия, входящего в состав логической операции. Формат заданий сценарного типа приближен к компьютерным играм, однако, в отличие от игр, обладает меньшей вариативностью, что

делает оценивание более стандартизированным. Такой подход позволяет реализовать принципы скрытого оценивания для диагностики логического мышления.

Итоговые диагностические профили содержат уровни каждой логической структуры (комбинаторики, INRC), заложенной в систему оценивания. Учитель, психолог или родитель получает информацию о сформированности каждой логической операции по отдельности для выстраивания полноценного индивидуального портрета логического мышления ученика. В настоящее время разработаны задания сценарного типа для учеников 5-х и 7-х классов².

Таким образом, новая методология реализует как требования к заданиям для измерения логического мышления, так и требования к оценке результатов диагностики без привлечения экспертов.

6. Реализация требований к заданиям для измерения логического мышления

6.1. Особые условия предъявления задания

Для оценки конкретно-операционального мышления задания предъявляются с опорой на реальные объекты или их изображения (видео, рисунки), так как мышление на этой стадии развития преимущественно осуществляется в образах. Появление в мышлении подростка формальных операций позволяет ему мысленно, т. е. без обращения к реальным объектам или их изображениям, устанавливать сложные логические отношения. Поэтому для диагностики уровня развития формальных операций текстовое предъявление задания, на которое необходимо получить ответ, более релевантно по сравнению с образным [Avila de, Pulos, 1979; Пиаже, 1994а; 1994б; 1994в].

Разработанные нами компьютеризированные задания сценарного типа позволяют комбинировать текстовое и рисуночное предъявление стимульного материала, а также имитировать работу с физическими объектами — а значит, дают возможность оценить умение ученика мысленно оперировать материалами разного уровня (физическими, образными, знаковыми) и точнее определить уровень развития его логического мышления (стадия конкретных или формальных операций).

6.2. Использование в диагностике задания на комбинаторику

Появление комбинаторных операций — одно из новообразований стадии формально-операционального мышления. На рис. 1

² Инструмент для оценки навыков XXI века «4К» (критическое мышление, креативность, коммуникация и кооперация), разработанный сотрудниками Центра психометрики и измерений в образовании (лаборатория измерения новых конструктов и дизайна тестов) Института образования НИУ ВШЭ в рамках договора о научно-исследовательской работе с фондом «Вклад в будущее».

представлен пример задания на комбинаторику для учеников 5-го класса. Сюжет задания фантастический: ученику нужно создать топливо для космического корабля. Часть инструкции дается в текстовой форме, стимульный материал имитирует работу с реальными объектами, позволяя систематически сочетать объекты друг с другом во всех возможных вариантах, сознательно не допуская повторов и пропусков в сочетаниях.

В процессе решения учеником задачи последовательно оценивается сформированность познавательных действий, входящих в состав комбинаторной операции: 1) действие выделения в описанной проблемной ситуации переменных, которые необходимо комбинировать для решения задачи; 2) действие комбинирования переменных, представленных в проблемной ситуации, с учетом требований задания (по два, по три и т.д.); 3) действие контроля за комбинированием переменных (последовательно составлять все возможные комбинации переменных в соответствии с требованиями задания, не допускать повторных комбинаций).

На первом шаге тестируемый исследует планету и собирает ингредиенты для топлива (рис. 1а) — производит выделение переменных. После этого ученик получает иллюстрированную инструкцию, согласно которой нужно провести исследование (систематический перебор собранных ингредиентов) для создания топлива. Особое внимание уделяется тому, чтобы он точно понял инструкцию (рис. 1б), т.е. требования к комбинированию переменных как элементу действия. Далее ученик решает поставленную задачу (рис. 1в) — комбинирует переменные с учетом требований задания и осуществляет контроль за комбинированием переменных.

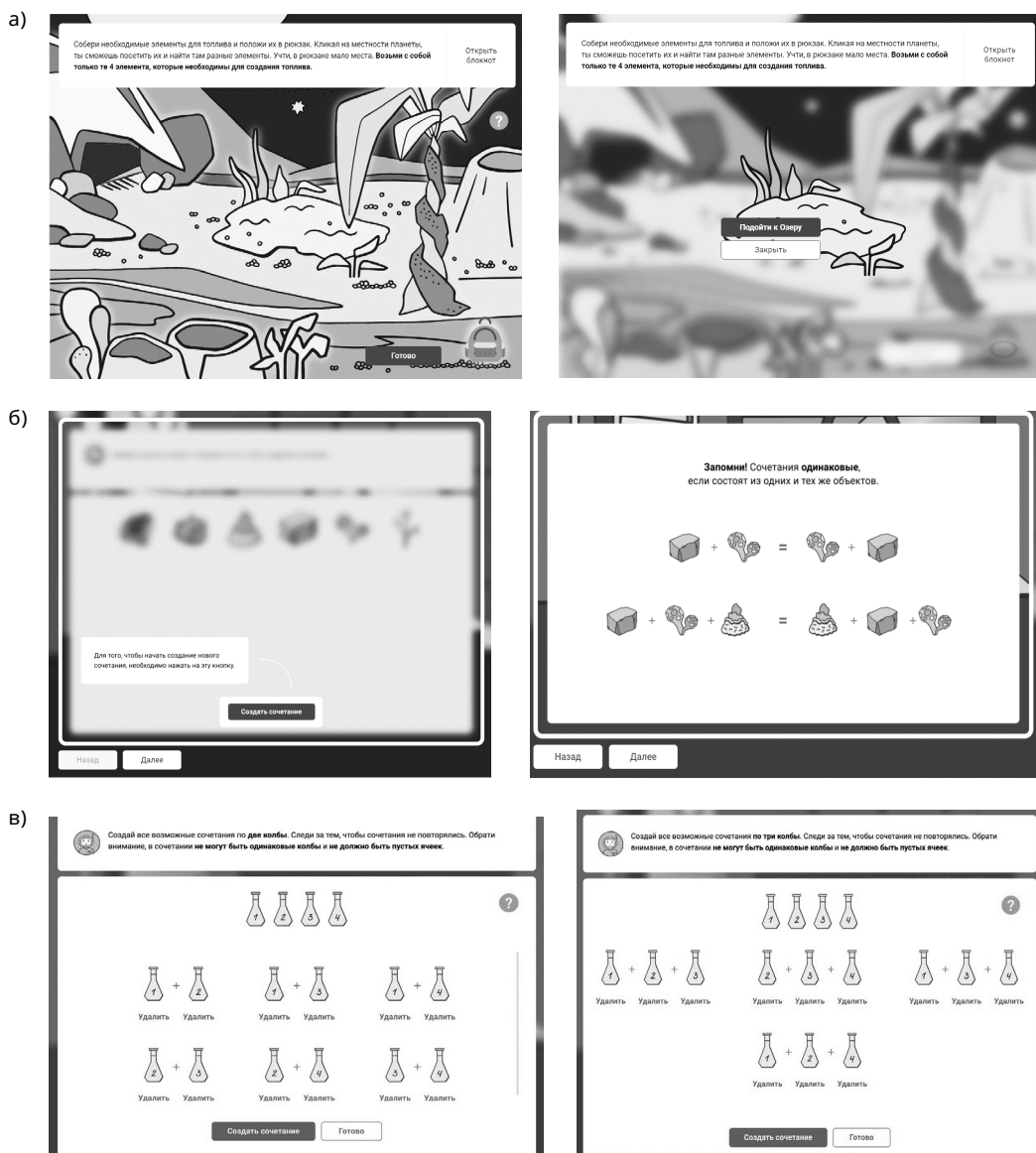
Предоставление учащемуся возможности последовательного перебора стимульного материала в соответствии с заданными условиями сценарной задачи и учета допущенных ошибок создает условия для оценки не только результата решения, но и его процесса.

Итоговая оценка уровня развития операции комбинаторики выставляется на основе результата, полученного при комбинировании переменных (количество верных комбинаций) и его процесса (отсутствие повторяющихся комбинаций). В диагностическом профиле выделяют три стадии развития:

1-я стадия — операция комбинаторики не сформирована: ученик может выделить в описании проблемной ситуации переменные, которые необходимо комбинировать (1), но не комбинирует их с учетом требований задания (2) и не может контролировать данный процесс, совершает повторные комбинации (3);

2-я стадия — операция комбинаторики сформирована частично: ученик может выделить в описании проблемной си-

Рис. 1. Пример элемента задания сценарного типа для диагностики комбинаторных операций



туации переменные, которые необходимо комбинировать (1), но допускает ошибки при комбинировании или комбинирует не все переменные (2), не может контролировать комбинаторный процесс, допускает повторные комбинации (3);

3-я стадия — все познавательные действия, входящие в состав операции комбинаторики, сформированы в полном объе-

ме: ученик последовательно составляет все верные комбинации представленных в проблемной ситуации переменных в соответствии с требованиями задачи, исключает повторные комбинации.

6.3. Использование в диагностике задач на INRC

Появление у учащегося когнитивной структуры четырех трансформаций INRC проявляется в способности анализировать проблемную ситуацию, последовательно учитывая влияние на нее всех заданных условий. Проверая, влияет ли тот или иной фактор на ситуацию, подросток мысленно выполняет одновременно два разнонаправленных действия: варьирует характеристики проверяемого фактора (изменяет переменные) и оставляет неизменными характеристики всех остальных факторов (уравнивает переменные).

Например, в традиционной задаче Пиаже про стержни [Пиаже, Инхельдер, 2003] для того, чтобы определить, влияют ли на гибкость стержней их длина, толщина, форма сечения и материал, из которого они сделаны, подросток последовательно проверяет каждый фактор. Исследуя влияние длины на гибкость, он берет два стержня, отличающиеся по длине — длинный и короткий, — но одинаковые по всем остальным признакам: оба стержня толстые, круглые, сделаны из стали. А затем смотрит на результат воздействия. Если стержни гнутся одинаково, делается вывод о том, что длина не влияет на гибкость. Если гнутся неодинаково — значит, длина влияет. Схожим образом последовательно проверяется влияние всех остальных признаков, заданных в ситуации.

По такой же схеме разработано задание для 7-го класса (рис. 2). Сюжет задания фантастический: ученику нужно выяснить, кто из обитателей старого особняка имеет шанс превратиться в привидение.

Так же как в заданиях на комбинаторику, в этом задании есть возможность оценить не только конечный результат (правильное/неправильное решение), но и процесс его получения, т. е. сделать вывод об уровне сформированности действий, входящих в состав диагностируемой логической операции уравнивания переменных как индикатора структуры INRC [Ильясов, 1986; Балдина, 1987; Погожина, 2006]:

- 1) действия выделения переменных и их значений, представленных в проблемной ситуации;
- 2) действия уравнивания переменных: для определения влияния переменных на результат необходимо сравнить ситуации, в которых все переменные, кроме одной, имеют одинаковые значения;
- 3) действия логического вывода: если результаты действия проверяемой переменной в сравниваемых ситуациях оди-

Рис. 2. Пример элемента задания сценарного типа для диагностики INRC



наковые, то она не оказывает влияния на ситуацию; если результаты различаются, влияние имеет место.

В сценарии ученик сначала собирает информацию обо всех жителях дома (о тех, кто стал и не стал привидением, рис. 2а, 2б) — совершает действие выделения переменных и их значений (1). Затем он уравнивает различные характеристики (рис. 2в) — выполняет правило уравнивания (2), чтобы принять решение, какой именно показатель приводит к тому, что житель дома становится привидением, — действие логического вывода (3).

По результатам последовательного решения сценарных заданий определяются диагностические профили — уровни развития логической операции уравнивания переменных как индикатора структуры INRC:

1-я стадия — операция не сформирована: ученик может выделить группы переменных, описанных в проблемной ситуации (1), но не может уравнивать разные факторы (2) и делать вывод об их влиянии (3);

2-я стадия — операция сформирована частично: верно выделив переменные и их значения (1), ученик затем последовательно уравнивает только часть из них, допускает ошибки уравнивания (2) и логического вывода (3);

3-я стадия — все познавательные действия, входящие в состав операции уравнивания переменных, сформированы в полном объеме: ученик выделяет в проблемной ситуации все представленные переменные и их значения, последовательно уравнивает все имеющиеся переменные, кроме одной, оценивает результат действия «неуравненной» переменной и делает верный вывод о ее влиянии.

**7. Реализация
требований
к оценке
результатов
диагностики без
привлечения
экспертов**

Главное преимущество стандартизированного оценивания — легкость и удобство проверки результатов тестирования. Подготовка системы оценивания, которая может работать без привлечения экспертов, позволяет решить эту задачу.

В предлагаемых нами компьютеризированных заданиях сценарного типа у тестируемого нет возможности ввести текст ответа на задание, но его действия могут быть проинтерпретированы по результатам выборов, сделанных из предложенных вариантов. Эти выборы позволяют понять, какие умственные действия лежат в основе решения задачи.

Подготовка обратной связи для тестируемых базируется на количественном анализе данных с использованием передовых психометрических методов [Almond et al., 2015; Jeon et al., 2020]. В существующих тестах логического мышления применяются разные варианты интерпретации результатов: в качестве

свидетельства продвижения в овладении логической операцией может рассматриваться увеличение количества суммарных тестовых баллов [Staver, Gabel, 1979] или может быть установлен пороговый балл, преодоление которого означает переход на уровень формальных операций [Lawson, 1978; Roadrangka, 1991]. В предлагаемой нами системе диагностики на основании результатов выполнения заданий строится профиль развития логического мышления тестируемого, который отражает уровень владения умственными действиями, входящими в состав изучаемой логической операции.

Профиль включает диагностированный уровень развития логической операции комбинаторики (1-я, 2-я или 3-я стадия) и уровень развития логической операции уравнивания переменных как индикатора структуры INRC (1-я, 2-я или 3-я стадия). Каждая стадия характеризует овладение умственными действиями, входящими в состав формально-операциональной структуры.

В отличие от традиционных подходов к определению уровней — установления порогового балла по сумме набранных баллов за тест в целом или по определенной шкале — мы предлагаем обратиться к современным психометрическим методам, т.е. рассматривать изучаемый конструкт как дискретную переменную и определять для тестируемого вероятность оказаться на определенной стадии [Almond et al., 2015]. Такое дискретное построение и представление результатов позволяет давать удобную для использования учителем, родителем или психологом автоматизированную обратную связь. Однако валидизировать выделение стадий еще предстоит в ходе дальнейшего исследования качества инструмента.

8. Выводы

Доступность и посильность школьных образовательных программ может быть обеспечена путем индивидуализации обучения, для чего необходимо массовое объективное оценивание уровня развития логического мышления учеников и учет его результатов в процессе обучения и разработки программ оказания психологической помощи.

Предложенный Ж.Пиаже метод клинической беседы позволяет наиболее полно и точно оценить логическое мышление на определенной стадии его развития. Вместе с тем метод имеет ряд ограничений при массовой диагностике: необходимость большого числа квалифицированных экспертов, субъективность экспертной оценки, продолжительное время, затрачиваемое на одну диагностическую процедуру. Еще одно ограничение — невозможность скрыть от ученика ситуацию тестирования.

Преодолеть ограничения метода клинической беседы позволяют стандартизированные инструменты. Существующие сегодня групповые тесты логического мышления показывают удовлетворительную надежность и валидность. Однако в них сохраняется необходимость привлечения экспертов для оценки результатов диагностики. Правильность выполнения заданий закрытой формы может оцениваться и без их участия, но эксперты неизбежно привлекаются для проверки обоснования выбора того или иного ответа, т.е. для проверки процесса решения. Кроме того, пока в литературе не представлены инструменты измерения логических операций в ситуациях, включенных в содержательно насыщенный контекст.

Для преодоления существующих недостатков в диагностике стадий логического мышления необходима новая методология, удобная при проведении крупномасштабного сбора данных и позволяющая учитывать, каким образом логические операции «включаются» в процесс решения задачи.

Предлагаемая в статье методология учитывает как специфику логического мышления, так и особенности массового тестирования с автоматической обработкой результатов без привлечения экспертов, а также открывает возможность построения на основе полученных диагностических данных индивидуальных программ психологической помощи ученикам в рамках деятельностного подхода к учению и овладению социальным опытом: в данной методологии реализованы требования к диагностическим заданиям, обусловленные сложностью оцениваемого содержания, а также требования к оценке результатов диагностики, обусловленные необходимостью проведения массового стандартизированного оценивания без привлечения экспертов.

Требования к диагностическим заданиям, обусловленные сложностью оцениваемого содержания — состава и структуры логического мышления, реализованы за счет:

- особых условий предъявления заданий, а именно компьютеризированных заданий сценарного типа. Такое предъявление позволяет подобрать стимульный материал такого формата (текст, изображения или симуляция объектов), который помогает валидно оценить уровень развития логического мышления;
- включения в сценарий заданий, требующих для их успешного решения применения комбинаторных операций;
- включения в сценарий таких задач, для решения которых требуется опора на когнитивную структуру четырех трансформаций INRC. Например, задания должны создавать условия, в которых необходимо произвести уравнивание пере-

менных для выбора оптимального решения в заданном контексте;

- построения для предоставления обратной связи индивидуальных профилей, в которых отражается детализированная информация об уровне сформированности каждой логической операции.

В предлагаемой методологии реализованы требования к оценке результатов диагностики, обусловленные необходимостью проведения массового стандартизированного оценивания без привлечения экспертов. Обратная связь по результатам тестирования с опорой на количественный анализ данных позволяет избежать как затрат, связанных с привлечением большого числа экспертов высокого уровня квалификации, так и искажений данных, вызванных субъективностью экспертов, а также значительно уменьшить общее время диагностики на больших выборках учеников.

9. Рекомендации по использованию результатов диагностики в практике образования

Современная школа в целом учитывает стадии когнитивного развития ученика, и программы сформированы таким образом, что «средний» ученик способен овладеть ими. Тем не менее исследования показывают: распространенные сегодня представления о возрасте, в котором у учащихся формируются формальные операции, приходится признать крайне упрощенными [Shayer, Küchemann, Wylam, 1976]. Действительно, в научной литературе есть данные о том, что формальные логические операции появляются к раннему подростковому возрасту, однако даже не у всех студентов высших учебных заведений они сформированы в полном объеме [Lawson, 1978; Tobin, Capie, 1981; Tisher, 1971]. Индивидуальная диагностика логического мышления поможет определить уровень готовности каждого подростка к освоению учебного материала.

Результаты диагностики логического мышления позволяют установить не только наличие тех или иных логических структур, но и стадию их развития, т. е. покажут, какие именно познавательные действия пока не сформированы у данного конкретного ученика. Использование представленного инструментария даст возможность в дальнейшем выстроить для школьников индивидуальные программы психологической помощи, направленные на формирование недостающих структур мышления.

10. Ограничения и дальнейшие направления исследований

Важная характеристика предложенной нами методологии — погружение ученика в определенную проблемную ситуацию, которая требует проявления оцениваемых логических операций.

С одной стороны, такой подход позволяет интерпретировать результаты тестирования не изолированно, как в «лабораторных» условиях традиционных тестов Пиаже, а включенными в жизненный контекст. С другой стороны, контекст диагностического сценария накладывает ограничения на интерпретацию уровня обобщенности полученных результатов. Мы не можем гарантировать, что ученик будет применять те же логические операции в широком спектре ситуаций.

Эмпирически установлено, что, решая задачи, предъявленные в разных контекстах, один и тот же ученик проявляет навыки логического мышления разного уровня [Bart, 1972; 1978; Cohen, 1980; Twidle, 2006]. Такая особенность формирования логического мышления требует дополнительных исследований как самого феномена, так и специфики его измерения.

Перспективным направлением дальнейших исследований и практической работы по индивидуализации обучения является установление более специфических закономерностей связи между уровнями развития логического мышления и усвоением отдельных школьных дисциплин. Интерпретация причин академической неуспешности по школьным предметам может проводиться более детально и связывать результаты обучения не только с работой учителя или сложностью школьной программы, но и с темпами и индивидуальными особенностями развития логических структур ученика.

Статья подготовлена в рамках гранта, предоставленного Министерством науки и высшего образования Российской Федерации (соглашение о предоставлении гранта № 075-15-2020-928).

Авторы выражают благодарность коллективу проекта «4К» лаборатории измерения новых конструктов и дизайна тестов Центра психометрики и измерений в образовании Института образования НИУ ВШЭ, и в особенности А. А. Камаевой, за творческий и профессиональный подход к разработке заданий.

Литература

1. Балдина Н. П. (1987) Усвоение логических приемов при разных типах учения: дис. ... канд. психол. наук. М.: МГУ им. М. В. Ломоносова.
2. Брангье Ж. К. (2000) Беседы с Жаном Пиаже // Психологический журнал. Т. 21. № 2. С. 138–144.
3. Ильясов И. И. (1986) Структура процесса учения. М.: Издательство Московского университета.
4. Коменский Я. А. (1939) Великая дидактика. Т. 1. М.: Государственное учебно-педагогическое издательство Наркомпроса РСФСР.
5. Пиаже Ж. (2008) Речь и мышление ребенка. М.: Римис.
6. Пиаже Ж. (1994а) Генезис числа у ребенка // Пиаже Ж. Избранные психологические труды. М.: Международная педагогическая академия. С. 237–582.
7. Пиаже Ж. (1994б) Логика и психология // Пиаже Ж. Избранные психологические труды. М.: Международная педагогическая академия. С. 583–628.

8. Пиаже Ж. (1994в) Психология интеллекта // Пиаже Ж. Избранные психологические труды. М.: Международная педагогическая академия. С. 51–236.
9. Пиаже Ж., Инхельдер Б. (2003) Психология ребенка. СПб.: Питер.
10. Погожина И. Н. (2006) Методика диагностики формально-логического мышления: диагностика сформированности структуры INRC // Школьный психолог. № 9. С. 40–43.
11. Талызина Н. Ф. (2018) Деятельностная теория учения. М.: Издательство Московского университета.
12. Almond R. G., Mislevy R. J., Steinberg L. S., Yan D., Williamson D. M. (2015) Bayesian Networks in Educational Assessment. New York: Springer. doi:10.1007/978-1-4939-2125-6
13. American Educational Research Association, American Psychological Association, National Council on Measurement in Education (eds) (2014) Standards for Educational and Psychological Testing. Lanham, MD: American Educational Research Association.
14. Avila de E., Pulos S. (1979) Group Assessment of Cognitive Level by Pictorial Piagetian Tasks // Journal of Educational Measurement. Vol. 16. No 3. P. 167–175. doi:10.1111/j.1745-3984.1979.tb00098.x
15. Bakken L., Thompson J., Clark F. L., Johnson N., Dwyer K. (2001) Making Conservationists and Classifiers of Preoperational Fifth-Grade Children // The Journal of Educational Research. Vol. 95. No 1. P. 56–61. doi:10.1080/00220670109598783
16. Bart W. M. (1978) Issues in Measuring Formal Operations // The Genetic Epistemologist. Vol. 7. P. 3–4.
17. Bart W. M. (1972) Construction and Validation of Formal Reasoning Instruments // Psychological Reports. Vol. 30. No 2. P. 663–670. doi:10.2466/pr0.1972.30.2.663
18. Bergling B. M. (1998) Constructing Items Measuring Logical Operational Thinking: Facet Design-Based Item Construction Using Multiple Categories Scoring // European Journal of Psychological Assessment. Vol. 14. No 2. P. 172–187. doi:10.1027/1015–5759.14.2.172
19. Bitner-Corvin B. L. (1988) Is the GALT a Reliable Instrument for Measuring the Logical Thinking Abilities of Students in Grades Six through Twelve? Paper presented at the 61st Annual Meeting of the National Association for Research in Science Teaching (Lake of the Ozarks, MO, April 10–13, 1988). <https://files.eric.ed.gov/fulltext/ED293716.pdf>
20. Cohen H. G. (1980) Dilemma of the Objective Paper-and-Pencil Assessment within the Piagetian Framework // Science Education. Vol. 64. No 5. P. 741–745. doi:10.1002/sce.3730640521
21. Klerk de S., Eggen T. J. H. M., Veldkamp B. P. (2016) A Methodology for Applying Students' Interactive Task Performance Scores from a Multimedia-Based Performance Assessment in a Bayesian Network // Computers in Human Behavior. Vol. 60. July. P. 264–279. doi:10.1016/j.chb.2016.02.071
22. DeVries R. (1974) Relationships among Piagetian, IQ, and Achievement Assessments // Child Development. Vol. 45. No 3. P. 746–756.
23. Faust D. (1983) A Promising Approach to the Development of a Group Piagetian Measure // Psychological Reports. Vol. 53. No 3. P. 771–774. doi:10.2466/pr0.1983.53.3.771
24. Goldschmid M. L. (1967) Different Types of Conservation and Nonconservation and Their Relation to Age, Sex, IQ, MA, and Vocabulary // Child Development. Vol. 38. No 4. P. 1229–1246. doi:10.1111/j.1467-8624.1967.tb04398.x
25. Graesser A., Kuo B.-C., Liao C.-H. (2017) Complex Problem Solving in Assessments of Collaborative Problem Solving // Journal of Intelligence. Vol. 5. No 2. Art. No 10. doi:10.3390/jintelligence5020010
26. Griffin P., McGaw B., Care E. (2012) Assessment and Teaching of 21st Century Skills. Heidelberg: Springer.

27. Hathaway W. E., Hathaway-Theunissen A. (1975) The Unique Contributions of Piagetian Measurement to Diagnosis, Prognosis, and Research of Children's Mental Development // G. I. Lubin (ed.) *Piagetian Theory and the Helping Professions*. Los Angeles: University of Southern California.
28. Hautamäki J., Thuneberg H. (2019) Koulutuksen Tasa-Arvo-taseet [Equity Balances in Education] // J. Hautamäki, I. Rämä, M.-P. Vainikainen (eds) *Perusopetus, Tasa-Arvo ja Oppimaan Oppiminen: Valtakunnallinen Arviointitutkimus Peruskoulun Päättövaiheesta* [Comprehensive Education, Equality, and Learning to Learn: Nationwide Evaluative Research on the Final Phase of Comprehensive Education]. Helsinki: University of Helsinki. P. 77–96.
29. Hienonen N. (2020) Does Class Placement Matter? Students with Special Educational Needs in Regular and Special Classes // *Helsinki Studies in Education*. No 87. <https://helda.helsinki.fi/bitstream/handle/10138/318683/Does-clas.pdf?sequence=1&isAllowed=y>
30. Inhelder B., Piaget J. (1958) *The Growth of Logical Thinking from Childhood to Adolescence: An Essay on the Construction of Formal Operational Structures*. New York: Basic Books.
31. Jeon M., Draney K., Wilson M., Sun Y. (2020) Investigation of Adolescents' Developmental Stages in Deductive Reasoning: An Application of a Specialized Confirmatory Mixture IRT Approach // *Behavior Research Methods*. Vol. 52. No 1. P. 224–235. doi:10.3758/s13428-019-01221-5
32. Kaufman R. A. (1972) *Educational System Planning*. New York: Prentice-Hall.
33. Lawson A. E. (1978) The Development and Validation of a Classroom Test of Formal Reasoning // *Journal of Research in Science Teaching*. Vol. 15. No 1. P. 11–24. doi:10.1002/tea.3660150103
34. Lawson A. E., Blake A. J. D., Nordland F. H. (1974) Piagetian Tasks Clarified: The Use of Metal Cylinders // *The American Biology Teacher*. Vol. 36. No 4. P. 209–211. doi:10.2307/4444748
35. Lawson A. E., Renner J. W. (1975) Relationships of Science Subject Matter and Developmental Levels of Learners // *Journal of Research in Science Teaching*. Vol. 12. No 4. P. 347–358. doi:10.1002/tea.3660120405
36. Longeot F. (1965) Analyse Statistique de Trois Tests Genetiques Collectifs // *Bulletin de l'institut National D'Etude*. Vol. 20. No 4. P. 219–237.
37. Longeot F. (1962) An Essay of the Application of Genetic Psychology to Differential Psychology // *Bulletin de l'Institute d'Etude Du Travail et d'Orientation Professionnelle*. No 18. P. 153–162.
38. Lovell K. (1961) A Follow-Up Study of Inhelder and Piaget's The Growth of Logical Thinking // *British Journal of Psychology*. Vol. 52. No 2. P. 143–153. doi:10.1111/j.2044-8295.1961.tb00776.x
39. Lovell K., Shields J. B. (1967) Some Aspects of a Study of the Gifted Child // *British Journal of Educational Psychology*. Vol. 37. No 2. P. 201–208. doi:10.1111/j.2044-8279.1967.tb01929.x
40. Malhotra S. (2020) Psychometric Intelligence and Academic Achievement: A Comparative Analysis of Elementary Schools // *EDUTEC: Journal of Education and Technology*. Vol. 3. No 2. P. 83–95. doi:10.29062/edu.v3i2.40
41. Messick S. (1994) *Alternative Modes of Assessment, Uniform Standards of Validity*1. Princeton, NJ: Educational Testing Service. <https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.2333-8504.1994.tb01634.x>
42. Messick S. (1992) The Interplay of Evidence and Consequences in the Validation of Performance Assessments. Paper presented at the Annual Meeting of the National Council on Measurement in Education (San Francisco, CA, April 1992). <https://files.eric.ed.gov/fulltext/ED390891.pdf>
43. Meyer C. E. (1972) Can Piaget's Theory Provide a Better Psychometry? // J. F. Magary, M. Poulsen, G. I. Lubin, G. Coplin (eds) *Piagetian Theory and the Helping Professions*. Los Angeles, CA: Children's Hospital. P. 5–10. <https://files.eric.ed.gov/fulltext/ED085612.pdf>

44. Milakofsky L., Patterson H. O. (1979) Chemical Education and Piaget: A New Paper-Pencil Inventory to Assess Cognitive Functioning // *Journal of Chemical Education*. Vol. 56. No 2. P. 87–90. doi:10.1021/ed056p87
45. Raven R.J. (1973) The Development of a Test of Piaget's Logical Operations // *Science Education*. Vol. 57. No 3. P. 377–385. doi:10.1002/sce.3730570316
46. Razzouk R. (2011) Using Evidence-Centered Design for Developing Valid Assessments of 21st Century Skills. Bellevue, WA: Edvation.
47. Renner J.W., Sutherland J., Grant R., Lawson A. E. (1978) Displacement Volume, An Indicator of Early Formal Thought; Developing a Paper-and-Pencil Test // *School Science and Mathematics*. Vol. 78. No 4. P. 297–303. doi:10.1111/j.1949-8594.1978.tb09362.x
48. Roadrangka V. (1991) The Construction of a Group Assessment of Logical Thinking (GALT) // *Kasetsart Journal of Social Sciences*. Vol. 12. No 2. P. 148–154.
49. Roberge J.J., Flexer B.K. (1982) The Formal Operational Reasoning Test // *Journal of General Psychology*. Vol. 106. No 1. P. 61–67. doi:10.1080/00221309.1982.9710973
50. Rowell J. A., Hoffmann P.J. (1975) Group Tests for Distinguishing Formal from Concrete Thinkers // *Journal of Research in Science Teaching*. Vol. 12. No 2. P. 157–164. doi:10.1002/tea.3660120210
51. Shayer M. (1979) Has Piaget's Construct of Formal Operational Thinking Any Utility? // *British Journal of Educational Psychology*. Vol. 49. No 3. P. 265–276. doi:10.1111/j.2044-8279.1979.tb02425.x
52. Shayer M. (1978) The Analysis of Science Curricula for Piagetian Level of Demand // *Studies in Science Education*. Vol. 5. Iss. 1. P. 115–130. doi:10.1080/03057267808559861
53. Shayer M., Küchemann D.E., Wylam H. (1976) The Distribution of Piagetian Stages of Thinking in British Middle and Secondary School Children // *British Journal of Educational Psychology*. Vol. 46. No 2. P. 164–173. doi:10.1111/j.2044-8279.1976.tb02308.x
54. Staver J. R., Gabel D. L. (1979) The Development and Construct Validation of a Group-Administered Test of Formal Thought // *Journal of Research in Science Teaching*. Vol. 16. No 6. P. 535–544. doi:10.1002/tea.3660160607
55. Sun C., Shute V.J., Stewart A., Yonehiro J., Duran N., D'Mello S. (2020) Towards a Generalized Competency Model of Collaborative Problem Solving // *Computers & Education*. Vol. 143. No 1. Art. No 103672. doi:10.1016/j.compedu.2019.103672
56. Tisher R. P. (1971) A Piagetian Questionnaire Applied to Pupils in a Secondary School // *Child Development*. Vol. 42. No 5. P. 1633–1636. doi:10.2307/1127935
57. Tobin K. G., Capie W. (1981) The Development and Validation of a Group Test of Logical Thinking // *Educational and Psychological Measurement*. Vol. 41. No 2. P. 413–423. doi:10.1177/001316448104100220
58. Twidle J. (2006) Is the Concept of Conservation of Volume in Solids Really More Difficult than for Liquids, or Is the Way We Test Giving Us an Unfair Comparison? // *Educational Research*. Vol. 48. No 1. P. 93–109. doi:10.1080/00131880500498511
59. Vainikainen M. P., Hautamäki J., Hotulainen R., Kupiainen S. (2015) General and Specific Thinking Skills and Schooling: Preparing the Mind to New Learning // *Thinking Skills and Creativity*. Vol. 18. April. P. 53–64. doi:10.1016/j.tsc.2015.04.006
60. Wang L., Shute V., Moore G. R. (2015) Lessons Learned and Best Practices of Stealth Assessment // *International Journal of Gaming and Computer-Mediated Simulations*. Vol. 7. No 4. P. 66–87. doi:10.4018/IJGCMS.2015100104
61. Watkins M. W., Lei P.-W., Canivez G. L. (2007) Psychometric Intelligence and Achievement: A Cross-Lagged Panel Analysis // *Intelligence*. Vol. 35. No 1. P. 59–68. doi:10.1016/j.intell.2006.04.005

References

- Almond R. G., Mislevy R. J., Steinberg L. S., Yan D., Williamson D. M. (2015) *Bayesian Networks in Educational Assessment*. New York: Springer. doi:10.1007/978-1-4939-2125-6
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education (eds) (2014) *Standards for Educational and Psychological Testing*. Lanham, MD: American Educational Research Association.
- Avila de E., Pulos S. (1979) Group Assessment of Cognitive Level by Pictorial Piagetian Tasks. *Journal of Educational Measurement*, vol. 16, no 3, pp. 167–175. doi:10.1111/j.1745-3984.1979.tb00098.x
- Bakken L., Thompson J., Clark F. L., Johnson N., Dwyer K. (2001) Making Conservationists and Classifiers of Preoperational Fifth-Grade Children. *The Journal of Educational Research*, vol. 95, no 1, pp. 56–61. doi:10.1080/00220670109598783
- Baldina N. P. (1987) *Usvoenie logicheskikh priyomov pri raznykh tipakh ucheniya* [Mastering Logical Techniques by Different Types of Teaching] (PhD Thesis). Moscow: Lomonosov Moscow State University.
- Bart W. M. (1978) Issues in Measuring Formal Operations. *The Genetic Epistemologist*, vol. 7, pp 3–4.
- Bart W. M. (1972) Construction and Validation of Formal Reasoning Instruments. *Psychological Reports*, vol. 30, no 2, pp. 663–670. doi:10.2466/pr0.1972.30.2.663
- Bergling B. M. (1998) Constructing Items Measuring Logical Operational Thinking: Facet Design-Based Item Construction Using Multiple Categories Scoring. *European Journal of Psychological Assessment*, vol. 14, no 2, pp. 172–187. doi:10.1027/1015-5759.14.2.172
- Bitner-Corvin B. L. (1988) Is the GALT a Reliable Instrument for Measuring the Logical Thinking Abilities of Students in Grades Six through Twelve? Paper presented at the 61st Annual Meeting of the National Association for Research in Science Teaching (Lake of the Ozarks, MO, April 10–13, 1988). Available at: <https://files.eric.ed.gov/fulltext/ED293716.pdf> (accessed 20 October 2021).
- Bringuier J.-C. (2000) Besedy s Zhanom Piazhe [Conversations with Jean Piaget]. *Psikhologicheskyy zhurnal*, vol. 21, no 2, pp. 138–144.
- Cohen H. G. (1980) Dilemma of the Objective Paper-and-Pencil Assessment within the Piagetian Framework. *Science Education*, vol. 64, no 5, pp. 741–745. doi:10.1002/sce.3730640521
- Comenius J. A. (1939) *Velikaya didaktika* [The Great Didactic]. Moscow: State Educational and Pedagogical Publishing House of the People's Commissariat of Education of the RSFSR.
- DeVries R. (1974) Relationships among Piagetian, IQ, and Achievement Assessments. *Child Development*, vol. 45, no 3, pp. 746–756.
- Faust D. (1983) A Promising Approach to the Development of a Group Piagetian Measure. *Psychological Reports*, vol. 53, no 3, pp. 771–774. doi:10.2466/pr0.1983.53.3.771
- Goldschmid M. L. (1967) Different Types of Conservation and Nonconservation and Their Relation to Age, Sex, IQ, MA, and Vocabulary. *Child Development*, vol. 38, no 4, pp. 1229–1246. doi:10.1111/j.1467-8624.1967.tb04398.x
- Graesser A., Kuo B.-C., Liao C.-H. (2017) Complex Problem Solving in Assessments of Collaborative Problem Solving. *Journal of Intelligence*, vol. 5, no 2, art. no 10. doi:10.3390/jintelligence5020010
- Griffin P., McGaw B., Care E. (2012) *Assessment and Teaching of 21st Century Skills*. Heidelberg: Springer.
- Hathaway W. E., Hathaway-Theunissen A. (1975) The Unique Contributions of Piagetian Measurement to Diagnosis, Prognosis, and Research of Children's Mental Development. *Piagetian Theory and the Helping Professions* (ed. G. I. Lubin), Los Angeles: University of Southern California.
- Hautamäki J., Thuneberg H. (2019) Koulutuksen Tasa-Arvoaset [Equity Balances in Education]. *Perusopetus, Tasa-Arvo ja Oppimaan Oppiminen: Valtakunnal-*

- linen Arviointitutkimus Peruskoulun Päättövaiheesta* [Comprehensive Education, Equality, and Learning to Learn: Nationwide Evaluative Research on the Final Phase of Comprehensive Education] (eds J. Hautamäki, I. Rämä, M.-P. Vainika-inen), Helsinki: University of Helsinki, pp. 77–96.
- Hienonen N. (2020) Does Class Placement Matter? Students with Special Educational Needs in Regular and Special Classes. *Helsinki Studies in Education*, no 87. Available at: <https://helda.helsinki.fi/bitstream/handle/10138/318683/Doesclas.pdf?sequence=1&isAllowed=y> (accessed 20 October 2021).
- Il'yasov I.I. (1986) *Struktura protsessu ucheniya* [Structure of the Teaching Process]. Moscow: Moscow University.
- Inhelder B., Piaget J. (1958) *The Growth of Logical Thinking from Childhood to Adolescence: An Essay on the Construction of Formal Operational Structures*. New York: Basic Books.
- Jeon M., Draney K., Wilson M., Sun Y. (2020) Investigation of Adolescents' Developmental Stages in Deductive Reasoning: An Application of a Specialized Confirmatory Mixture IRT Approach. *Behavior Research Methods*, vol. 52, no 1, pp. 224–235. doi:10.3758/s13428-019-01221-5
- Kaufman R.A. (1972) *Educational System Planning*. New York: Prentice-Hall.
- Klerk de S., Eggen T.J.H.M., Veldkamp B.P. (2016) A Methodology for Applying Students' Interactive Task Performance Scores from a Multimedia-Based Performance Assessment in a Bayesian Network. *Computers in Human Behavior*, vol. 60, July, pp. 264–279. doi:10.1016/j.chb.2016.02.071
- Lawson A.E. (1978) The Development and Validation of a Classroom Test of Formal Reasoning. *Journal of Research in Science Teaching*, vol. 15, no 1, pp. 11–24. doi:10.1002/tea.3660150103
- Lawson A.E., Blake A.J.D., Nordland F.H. (1974) Piagetian Tasks Clarified: The Use of Metal Cylinders. *The American Biology Teacher*, vol. 36, no 4, pp. 209–211. doi:10.2307/4444748
- Lawson A.E., Renner J.W. (1975) Relationships of Science Subject Matter and Developmental Levels of Learners. *Journal of Research in Science Teaching*, vol. 12, no 4, pp. 347–358. doi:10.1002/tea.3660120405
- Longeot F. (1965) Analyse Statistique de Trois Tests Genetiques Collectifs. *Bulletin de l'institut National D'Etude*, vol. 20, no 4, pp. 219–237.
- Longeot F. (1962) An Essay of the Application of Genetic Psychology to Differential Psychology. *Bulletin de l'Institute d'Etude Du Travail et d'Orientation Professionnelle*, no 18, pp. 153–162.
- Lovell K. (1961) A Follow-Up Study of Inhelder and Piaget's The Growth of Logical Thinking. *British Journal of Psychology*, vol. 52, no 2, pp. 143–153. doi:10.1111/j.2044-8295.1961.tb00776.x
- Lovell K., Shields J.B. (1967) Some Aspects of a Study of the Gifted Child. *British Journal of Educational Psychology*, vol. 37, no 2, pp. 201–208. doi:10.1111/j.2044-8279.1967.tb01929.x
- Malhotra S. (2020) Psychometric Intelligence and Academic Achievement: A Comparative Analysis of Elementary Schools. *EDUTECH: Journal of Education and Technology*, vol. 3, no 2, pp. 83–95. doi:10.29062/edu.v3i2.40
- Messick S. (1994) *Alternative Modes of Assessment, Uniform Standards of Validity*. Princeton, NJ: Educational Testing Service. Available at: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.2333-8504.1994.tb01634.x> (accessed 20 October 2021).
- Messick S. (1992) The Interplay of Evidence and Consequences in the Validation of Performance Assessments. Paper presented at the *Annual Meeting of the National Council on Measurement in Education (San Francisco, CA, April 1992)*. Available at: <https://files.eric.ed.gov/fulltext/ED390891.pdf> (accessed 20 October 2021).
- Meyer C.E. (1972) Can Piaget's Theory Provide a Better Psychometry? *Piagetian Theory and the Helping Professions* (eds J.F. Magary, M. Poulsen, G.I. Lubin,

- G. Coplin), Los Angeles, CA: Children's Hospital, pp. 5–10. Available at: <https://files.eric.ed.gov/fulltext/ED085612.pdf> (accessed 20 October 2021).
- Milakofsky L., Patterson H. O. (1979) Chemical Education and Piaget: A New Paper-Pencil Inventory to Assess Cognitive Functioning. *Journal of Chemical Education*, vol. 56, no 2, pp. 87–90. doi:10.1021/ed056p87
- Piaget J. (2008) *Rech i myshlenie rebyonka* [The Language and Thought of the Child]. Moscow: Rimis.
- Piaget J. (1994a) Genesis chisla u rebyonka [The Child's Conception of Number] *Izbrannye psikhologicheskie trudy* [Selected Psychological Works of J. Piaget], Moscow: International Pedagogical Academy, pp. 237–582.
- Piaget J. (1994b) Logika i psikhologiya [Logic and Psychology] *Izbrannye psikhologicheskie Trudy* [Selected Psychological Works of J. Piaget], Moscow: International Pedagogical Academy, pp. 583–628.
- Piaget J. (1994c) Psikhologiya intellekta [The Psychology of Intelligence] *Izbrannye psikhologicheskie Trudy* [Selected Psychological Works of J. Piaget], Moscow: International Pedagogical Academy, pp. 51–236.
- Piaget J., Inhelder B. (2003) *Psikhologiya rebyonka* [The Psychology of the Child]. Saint Petersburg: Piter.
- Pogozhina I. N. (2006) Metodika diagnostiki formal'no-logicheskogo myshleniya: diagnostika sformirovannosti struktury INRC [Methods of Diagnostics of Formal-Logical Thinking: Diagnostics of the Formation of the INRC Structure]. *Shkol'ny psikholog*, no 9, pp. 40–43.
- Raven R. J. (1973) The Development of a Test of Piaget's Logical Operations. *Science Education*, vol. 57, no 3, pp. 377–385. doi:10.1002/sce.3730570316
- Razzouk R. (2011) *Using Evidence-Centered Design for Developing Valid Assessments of 21st Century Skills*. Bellevue, WA: Edvation.
- Renner J. W., Sutherland J., Grant R., Lawson A. E. (1978) Displacement Volume, An Indicator of Early Formal Thought; Developing a Paper-and-Pencil Test. *School Science and Mathematics*, vol. 78, no 4, pp. 297–303. doi:10.1111/j.1949-8594.1978.tb09362.x
- Roadrangka V. (1991) The Construction of a Group Assessment of Logical Thinking (GALT). *Kasetsart Journal of Social Sciences*, vol. 12, no 2, pp. 148–154.
- Roberge J. J., Flexer B. K. (1982) The Formal Operational Reasoning Test. *Journal of General Psychology*, vol. 106, no 1, pp. 61–67. doi:10.1080/00221309.1982.9710973
- Rowell J. A., Hoffmann P. J. (1975) Group Tests for Distinguishing Formal from Concrete Thinkers. *Journal of Research in Science Teaching*, vol. 12, no 2, pp. 157–164. doi:10.1002/tea.3660120210
- Shayer M. (1979) Has Piaget's Construct of Formal Operational Thinking Any Utility? *British Journal of Educational Psychology*, vol. 49, no 3, pp. 265–276. doi:10.1111/j.2044-8279.1979.tb02425.x
- Shayer M. (1978) The Analysis of Science Curricula for Piagetian Level of Demand. *Studies in Science Education*, vol. 5, iss. 1, pp. 115–130. doi:10.1080/03057267808559861
- Shayer M., Küchemann D. E., Wylam H. (1976) The Distribution of Piagetian Stages of Thinking in British Middle and Secondary School Children. *British Journal of Educational Psychology*, vol. 46, no 2, pp. 164–173. doi:10.1111/j.2044-8279.1976.tb02308.x
- Staver J. R., Gabel D. L. (1979) The Development and Construct Validation of a Group-Administered Test of Formal Thought. *Journal of Research in Science Teaching*, vol. 16, no 6, pp. 535–544. doi:10.1002/tea.3660160607
- Sun C., Shute V. J., Stewart A., Yonehiro J., Duran N., D'Mello S. (2020) Towards a Generalized Competency Model of Collaborative Problem Solving. *Computers & Education*, vol. 143, no 1, art. no 103672. doi:10.1016/j.compedu.2019.103672
- Talyzina N. F. (2018) *Deyatel'nostnaya teoriya ucheniya* [Activity Theory of Teaching]. Moscow: Moscow University.

- Tisher R. P. (1971) A Piagetian Questionnaire Applied to Pupils in a Secondary School. *Child Development*, vol. 42, no 5, pp. 1633–1636. doi:10.2307/1127935
- Tobin K. G., Capie W. (1981) The Development and Validation of a Group Test of Logical Thinking. *Educational and Psychological Measurement*, vol. 41, no 2, pp. 413–423. doi:10.1177/001316448104100220
- Twidle J. (2006) Is the Concept of Conservation of Volume in Solids Really More Difficult than for Liquids, or Is the Way We Test Giving Us an Unfair Comparison? *Educational Research*, vol. 48, no 1, pp. 93–109. doi:10.1080/00131880500498511
- Vainikainen M. P., Hautamäki J., Hotulainen R., Kupiainen S. (2015) General and Specific Thinking Skills and Schooling: Preparing the Mind to New Learning. *Thinking Skills and Creativity*, vol. 18, April, pp. 53–64. doi:10.1016/j.tsc.2015.04.006
- Wang L., Shute V., Moore G. R. (2015) Lessons Learned and Best Practices of Stealth Assessment. *International Journal of Gaming and Computer-Mediated Simulations*, vol. 7, no 4, pp. 66–87. doi:10.4018/IJGCMS.2015100104
- Watkins M. W., Lei P.-W., Canivez G. L. (2007) Psychometric Intelligence and Achievement: A Cross-Lagged Panel Analysis. *Intelligence*, vol. 35, no 1, pp. 59–68. doi:10.1016/j.intell.2006.04.005