

УДК 373

Разработка инструмента оценки эффективности экспериментальной деятельности подростков на примере задач с физическим содержанием

Александр И. Адамский¹, Никита И. Колачев²

¹Московский городской педагогический университет, Москва, Россия;
Институт проблем образовательной политики «Эврика», Москва, Россия
E-mail: aadam93@gmail.com

ORCID: <https://orcid.org/0000-0002-5395-9464>

²Высшая школа экономики, Москва, Россия

E-mail: nkolachev@hse.ru

ORCID: <https://orcid.org/0000-0002-3214-6675>

DOI: 10.26907/esd.19.4.11

EDN: NTLDTB

Дата поступления: 28 мая 2024; Дата принятия в печать: 18 ноября 2024

Аннотация

В фокусе внимания статьи находится проблема повышения эффективности естественно-научного школьного образования (в частности, физического) на основе принципа заочности и построения занятий на экспериментальной деятельности. Базовой единицей построения такой модели является **действие** учащегося, мотивированное, осознанное, направленное на пробу решения задачи, поиск способа решения и возможность рефлексии и обобщения целенаправленных действий. Однако для оценки эффективности действий нужны качественные исследовательские инструменты, позволяющие собирать достоверные данные. Для решения этой задачи мы разработали инструмент и эмпирически проверили требования, предъявляемые к нему. Мы предположили, что такими требованиями являются: 1) соответствие структуре деятельности; 2) многофункциональность и устойчивость к замене оценивающего субъекта. Требования к инструменту были сформулированы авторами этой статьи на основе теории учебной деятельности В. В. Давыдова и теории психолого-педагогических измерений. Сбор данных осуществлялся с помощью невключенного наблюдения, оценка деятельности учащихся проводилась независимыми экспертами – ассессорами – на основе аудио- и видеосъемки экспериментальных занятий в трех школах г. Южно-Сахалинска. Проверка требований осуществлялась на основе анализа экспертных суждений, а также с применением метрик согласованности – доли совпадений результатов и непараметрического коэффициента конкордации Кендалла. Анализ полученных данных показал: а) разработанный инструмент позволяет зафиксировать действия по решению задачи в соответствии с теоретической моделью; б) результаты наблюдений независимыми ассессорами дают близкие оценки; это отразилось в приемлемых значениях показателей согласованности оценок.

Ключевые слова: естественно-научная грамотность, физика, физическая картина мира, задача, экспериментальная деятельность, инструмент оценки, система оценивания, наблюдение.

Development of an Instrument for Assessing the Effectiveness of Adolescents' Experimental Activity Based on Solving Physics-Oriented Tasks

Alexander Adamsky¹, Nikita Kolachev²

¹ *Moscow City University, Moscow, Russia;*

Institute for Problems of Educational Policy "Evrika", Moscow, Russia

E-mail: aadam93@gmail.com

ORCID: <https://orcid.org/0000-0002-5395-9464>

² *HSE University, Moscow, Russia*

E-mail: nkolachev@hse.ru

ORCID: <https://orcid.org/0000-0002-3214-6675>

DOI: 10.26907/esd.19.4.11

EDN: NTLDTB

Submitted: 28 May 2024; Accepted: 18 November 2024

Abstract

This article focuses on addressing the problem of enhancing the effectiveness of natural science education in schools, particularly in physics, by applying the principle of problem-based learning and designing lessons centered on experimental activity. The foundational unit of this model is a student's action that is motivated, deliberate, and aimed at attempting to solve a task, exploring solutions, and engaging in reflection and generalization of goal-directed actions. However, assessing the effectiveness of such actions requires high-quality research instruments capable of collecting reliable data. To address this challenge, we developed an assessment instrument and empirically tested some psychometric requirements. We hypothesized that the key requirements for such an instrument are: (1) alignment with the structure of activity; (2) multifunctionality and robustness against changes in the evaluator. These requirements were formulated by the authors based on V.V. Davydov's theory of learning activity and the theory of psycho-pedagogical measurement. Data collection was conducted through non-participant observation, with student activity evaluated by independent experts—assessors—using audio and video recordings of experimental lessons in three schools in Yuzhno-Sakhalinsk. The validation of the instrument's requirements was carried out through the analysis of expert judgments and the application of consistency metrics, such as the proportion of matching results and Kendall's nonparametric concordance coefficient. The analysis of the collected data demonstrated that: (a) the developed instrument effectively captures actions related to task-solving in accordance with the theoretical model; (b) the assessments by independent observers yielded similar results, as reflected in acceptable consistency metrics for their evaluations.

Keywords: science literacy, physics, physical concept of the world, experimental activity, assessment tool, evaluation system, observation.

Введение

Естественно-научная, в частности физическая, картина мира сегодня перестает быть исключительно фактором продолжения естественно-научного, инженерного, физического и т. п. профессионального образования, но становится обязательным условием социализации каждого человека (Coombs, 1970). Это связано прежде всего с тем, что условия социализации выпускников школ – и даже подростков, если уже не дошкольников – связаны со стремительными технологическими изменениями, развитием орудий как инструментов ежедневного общения человека с другими людьми, социальной и другими инфраструктурами. Личность становится расширенной (Semenov & Ziskin, 2021) благодаря сложным орудиям, которые уже стали практически частью человеческой личности. Образование в принципе

возникло как компенсация потери человеческим видом наследственного механизма общения (взаимодействия, отношений) с миром, природой, нарождающейся культурой и себе подобными (Mikhailov, 1976). Каждое поколение *homo sapiens* вынуждено заново формировать большую часть программы этого взаимодействия, осваивая навыки, приемы, как бы мы сегодня сказали, компетенции, которые выработали предыдущие поколения. И пока сложность орудий менялась медленно, можно было строить образование «по программе минимум». Т. е. давать основной массе школьников примитивные сведения по всем наукам, но реализовать сложную программу по физике или математике, отбирать по-настоящему способных к этим наукам для их дальнейшей специализации. И 2-3% выпускников могли стать учеными, а 10-15% – инженерами хорошей квалификации. Все остальные вообще не понимали, зачем они «проходят» точные науки: согласно еще советской шутке, «прошел еще один день жизни – интеграл так и не пригодился». К тому же научная основа, например, физики формировалась столетиями, а в школьную программу открытия попадали медленно, пройдя десятилетний цикл методической цензуры.

Проблематика кризиса в образовании: время изменений меньше цикла получения базовых знаний

Всеобщая естественно-научная грамотность мало кого беспокоила, хотя тревожные сигналы были еще в XIX веке. Например, рассказ А. П. Чехова «Злоумышленник» был написан в 1880-х годах и посвящен как раз тому, как один малограмотный крестьянин может пустить под откос (в прямом смысле слова) гениальное творение физической и инженерной мысли. О естественно-научной грамотности речи идти не могло, поскольку к началу XX века большая часть населения (в основном сельские жители) России была более чем на 90% неграмотна (Strannolyubsky, 1893).

Но с середины XX века началась гонка технологий. И сегодня мы переживаем критическую для образования ситуацию: период системных технологических изменений становится меньше, чем образовательный цикл (Chulkova, 2022). Иными словами, если технологии меняются каждые 7 лет, а образовательный цикл «детский сад – школа» – 16 лет, то задача образования уже не в том, чтобы освоить действующие орудия, а в том, чтобы подготовить ребёнка к жизни с неведомыми, но уже возникшими технологиями. И это не военные технологии или освоение космоса – это технологии повседневной жизни: общения, передвижения, учения и лечения, творчества, всех видов производств, финансовой и культурной жизни.

Это первая причина того, что без ясной, внутренне непротиворечивой физической картины мира, понимания сущности физических процессов и закономерностей человеку будет практически невозможно стать успешным или самореализоваться максимально. Более того, человек рискует стать злоумышленником по отношению не только к чужой, но прежде всего к своей жизни. Сегодня это уже можно видеть на примере «антиваксеров»¹, или сверхэнергичных борцов с озоновой дырой, или противников атомной энергии, сотовой связи и т. д.

Массовая естественно-научная, прежде всего физическая, компетентность – условие выживания общества, а не только фактор продолжения образования. Именно поэтому эксперты подчеркивают, что повышение эффективности естественно-научного образования является одной из приоритетных задач российской школы (Lvovsky et al., 2021; Pentin et al., 2019).

¹ «Антиваксер» – сторонник общественного движения, которое оспаривает эффективность, безопасность и правомерность вакцинации населения.

Однако в образовании всегда есть разрыв между значимостью компетенций, необходимыми навыками социализации – и инструментами, которыми пользуется система образования, для решения современных задач. Ф. Кумбс именно эту ситуацию назвал кризисом в образовании (Coombs, 1970), признавая при этом, что этот кризис носит универсальный характер. Система обречена на то, чтобы быть в постоянном поиске инструментов подготовки выпускников к жизни.

Одной из таких проблем является низкий уровень мотивации школьников, особенно подростков, к изучению сложного материала. В российской системе образования это связано, на наш взгляд, прежде всего с архаичной системой преподавания, основанной на репродуктивной методике и методологии, когда основным является запоминание большого массива сведений и отработка навыка действия по образцу, алгоритму (Basyuk & Kovaleva, 2019). Это мнение складывается на основе результатов международных сравнительных исследований качества образования. Специалисты отмечают противоречие между высокими результатами российских учащихся в международном сравнительном исследовании TIMSS² и невысокими результатами в исследовании PISA³: «...российский подход к изучению естественнонаучных предметов больше ориентирован на воспроизведение знаний, а не на их применение или освоение способов действий, присущих естественным наукам: исследования и научной аргументации» (Pentin et al., 2018, p. 102). То есть эксперты фиксируют серьезные пробелы у обучающихся в навыках исследовательской деятельности, а также ориентацию отечественной системы образования на репродуктивный способ освоения содержания.

Принцип задачности – стратегия преодоления кризиса

Одним из направлений совершенствования школьного естественно-научного образования (в частности, физического) мы считаем внедрение принципа **задачности**, то есть включение в учебный процесс экспериментальных, открытых, исследовательских, поисковых задач и построение на их основе содержания образования (Adamskiy & Lvovsky, 2023). В этой логике *результатом решения* задачи является прежде всего *способ* ее решения, наряду с предметным результатом (ответом). Согласно этой концепции, экспериментальная задача выступает акселератором (генератором) разнообразных (нелинейных) действий обучающихся: от пробы, манипуляций с орудиями, смыслами – через выдвижение гипотез и моделирование решений, взаимодействие со взрослым – до постановки и достижения совместных целей деятельности, рефлексии своих действий, формирования первичных представлений и обобщения результатов до уровня понятий. Кроме того, принципиальным условием постановки задачи является невозможность ее решения вне индивидуальной и совместной деятельности (Rubtsov & Ulanovskaya, 2021), в противоположность возможности решения задачи по инструкции учителя, для нахождения «правильного», заранее известного ответа. Таким образом, экспериментальная задача призвана помочь инициировать индивидуальную и совместную деятельность для достижения необходимых результатов, а задачный принцип в построении курсов может способствовать улучшению ситуации с исследовательской деятельностью (в частности, экспериментальной).

Если говорить о структуре экспериментальной деятельности, то можно выделить определенные компоненты (этапы развертывания). Теоретической основой

² Trends in Mathematics and Science Study (Международное исследование качества математического и естественно-научного образования).

³ Programme for International Student Assessment (Международная программа по оценке образовательных достижений учащихся).

для выделения структуры экспериментальной деятельности является структура учебной деятельности, описанная В. В. Давыдовым (Davydov, 1996) и схематично представляемая как «понимание учебной задачи – осуществление учебных действий – контроль и оценка действий». Мы считаем, что экспериментальная деятельность включает следующие этапы: принятие задачи, выдвижение гипотез, осуществление проб, проверка (оценка) проб, моделирование физического принципа, создание «продукта», моделирующего физический принцип действия, представление «продукта» другим, а также рефлексия и обобщение (выход на первичные представления физических понятий, закономерности). Под принятием задачи понимается интерес учащихся, вовлечение в действие по решению задачи и понимание предъявляемых к ученикам требований в рамках экспериментальной задачи. Выдвижение гипотез подразумевает любые предположения, связанные с решением задачи, которые могут опираться как на бытовые представления, так и на научные знания. Под осуществлением проб понимаются попытки решить экспериментальную задачу. Проверка (оценка) проб означает сопоставление действий и полученных результатов с поставленной задачей. Моделирование включает выделение соотношения/зависимости между исследуемыми объектами; моделирование может проходить как в предметной, так и в обобщенной форме. Создание «продукта» означает фиксацию результата экспериментирования в предметной форме (в виде рисунка, видеоролика, фотографии и пр.). Представление «продукта» означает публичное выступление и презентацию полученных результатов. Культурно-историческое обобщение и рефлексия предполагают соотнесение полученных результатов с известным всеобщим законом или закономерностью.

Онтология результативности экспериментальной деятельности, задачного подхода

Необходимо добавить: описанные элементы не только являются составляющими структуры экспериментальной деятельности, но и отражают образовательные результаты. Так, в части личностных результатов, отраженных в Примерной рабочей программе по физике Основного общего образования (Institute for Strategy of Education Development, 2021), можно выделить ценность научного познания, то есть на основе решения экспериментальных задач возможно развивать интерес к научному исследованию, способствовать осознанию ценности физики как мощного инструмента познания мира. В части метапредметных результатов можно отметить, что почти все универсальные учебные действия вовлечены в предлагаемую экспериментальную деятельность: базовые логические действия (особенно компонент моделирования в структуре деятельности), базовые исследовательские действия, работа с информацией, совместная деятельность, самоконтроль (рефлексия), эмоциональный интеллект (за счет совместности) и принятие себя и других (тоже за счет совместности). Предметные результаты определяются в основном теми понятиями, которые формируются в ходе решения конкретной экспериментальной задачи.

В то же время необходимо отметить, что успешность решения экспериментальных задач не полностью зависит только от способностей и уровня подготовленности учащихся. Немаловажную роль играют действия учителя, создающие в системном виде возможности для удовлетворения естественной потребности подростков в исследовании и способствующие продуктивному протеканию процесса экспериментирования. Как отмечают А. И. Адамский и В. А. Львовский (Adamskiy & Lvovsky, 2023, p. 104), «необходимо построить квазиисследовательскую (в том числе экспериментальную) деятельность детей таким образом, чтобы ключевые поня-

тия не были даны в виде готовых знаний». За организацию и осуществление такой деятельности во многом ответственен учитель.

При этом существенной частью реализации описанного принципа задачи является оценка экспериментальной деятельности с помощью тщательно проработанного и психометрически проверенного инструмента. Благодаря такому инструменту можно не только оценивать текущее состояние экспериментальной деятельности, но и отслеживать динамику развертывания экспериментальной деятельности с помощью как «поперечных»⁴, так и «продольных»⁵ срезов. Кроме того, по-настоящему учебная задача несет в себе потенциал развития (Davydov, 1996), а исследование развития с помощью качественного инструмента может помочь учителю осуществлять отраженные в профессиональном стандарте трудовые функции (Ministry of Labor and Social Protection..., 2013). В связи с этим разрабатываемый инструмент должен соответствовать ряду требований. Во-первых, инструмент должен быть *моделью структуры экспериментальной деятельности*, то есть позволять исследовать экспериментальную деятельность по указанной структуре. Во-вторых, в силу того что ключевой характеристикой деятельности является ее совместность (ученика и учителя, ученика со сверстниками), инструмент оценки деятельности должен быть *многофункциональным*, то есть должен позволять оценивать деятельность учеников и учителя одновременно в их взаимосвязи. В-третьих, поскольку оценка деятельности нередко предполагает использование метода невключенного наблюдения, важным свойством отчуждаемых инструментов исследования в области педагогики и психологии является *устойчивость к изменению (замене) оценивающего субъекта*. Это требование означает, что у разных ассессоров (оценщиков, экспертов, учителей) при наблюдении за одними и теми же учащимися должен получиться сходный (в идеале одинаковый) результат.

Таким образом, *целью* данного исследования является проверка требований, предъявляемых к инструменту оценки деятельности учащихся (об оценке деятельности учителя в этой работе речь не пойдет). Содержательные гипотезы относительно требований, предъявляемых к инструменту (по-другому, карте или матрице наблюдений), следующие:

Гипотеза №1 – разработанный инструмент позволяет моделировать структуру экспериментальной деятельности;

Гипотез №2 – разработанный инструмент позволяет оценивать деятельность учеников и учителя одновременно в их взаимосвязи;

Гипотеза №3 – разработанный инструмент устойчив к изменению (замене) оценивающего субъекта.

⁴ Под «поперечными» срезами в науке понимается метод исследования, включающий изучение интересующих показателей либо на сходных по характеристикам, но разных по составу, группах людей, либо на людях разного возраста. Например, это может быть ежегодный мониторинг достижений учащихся четвертого класса или сравнение результатов у четвероклассников и шестиклассников и т. д.

⁵ Под «продольными» срезами в науке понимается лонгитюдный метод исследования, то есть изучение интересующих показателей на одних и тех же людях в течение некоторого промежутка времени.

Методология и методы исследования

В Таблице 1 представлено краткое содержание инструмента оценки в виде заложенной в него системы оценивания. Можно заметить, что для оценки критериальных действий часто используется политомическая шкала. Это сделано для того, чтобы, во-первых, максимально дифференцировать действия учащихся, и, во-вторых, отразить потенциальные образовательные результаты.

Таблица 1. Система оценивания матрицы структуры экспериментальной деятельности

Критерий	Система оценивания
Принятие задачи	<p>0 баллов – школьники не понимают условия полученной задачи, приступают к выполнению задания, не понимая, что от них требуется;</p> <p>1 балл – школьники изучают задание, но не до конца, приступают к выполнению задания, не понимая, что от них требуется;</p> <p>2 балла – школьники внимательно изучают задание</p> <p>ИЛИ</p> <p>школьники внимательно изучают задание и задают вопросы;</p> <p>3 балла – школьники внимательно изучают задание</p> <p>И</p> <p>планируют экспериментальную деятельность (рассуждают, что потребуется для эксперимента, есть ли все необходимое для проведения серии экспериментов и т. д.)</p>
Выдвижение гипотез	<p>0 баллов – школьники не выдвигают гипотезы</p> <p>ИЛИ</p> <p>выдвигают гипотезы, не связанные с поставленной задачей;</p> <p>1 балл – школьниками выдвигаются любые предположения и гипотезы, связанные с поставленной задачей и опирающиеся на бытовое знание или имеющийся опыт;</p> <p>2 балла – школьниками выдвигаются любые предположения и гипотезы, связанные с поставленной задачей и опирающиеся на имеющееся научное знание и/или на поиск аналогии в других областях</p>
Осуществление проб	<p>0 баллов – школьники осуществляют хаотичные, бессистемные пробы (действия);</p> <p>1 балл – школьники совершают действия, связанные с попыткой решить экспериментальную задачу (случайные пробы);</p> <p>2 балла – школьники совершают пробные действия с последующей рефлексией каждого шага</p>
Проверка (оценка) проб	<p>0 баллов – осуществляемые пробы (действия) не сопоставляются с задачей и гипотезами;</p> <p>1 балл – только часть проб (действий) сопоставляется с задачей и гипотезами;</p> <p>2 балла – все пробы (действия) сопоставляются с задачей и гипотезами</p>
Моделирование физического принципа	<p>0 баллов – отсутствует моделирование выделенного отношения;</p> <p>1 балл – осуществляется моделирование выделенного отношения для изучения его свойств в предметной форме;</p> <p>2 балла – осуществляется преобразование модели для изучения свойств отношения в обобщенном виде</p>
Создание «продукта», моделирующего физический принцип действия	<p>0 баллов – школьники не приходят ни к какому умозаключению (результату)</p> <p>ИЛИ</p> <p>сделанное ими умозаключение (результат) не связано с поставленной задачей и совершенными пробами;</p> <p>1 балл – школьники приходят к верному умозаключению (результату) экспериментальным путем.</p>

Критерий	Система оценивания
Представление «продукта» другим учащимся (группам)	<p>0 баллов – школьники не высказывают никакие версии (молчат) ИЛИ произносимые ими тексты не имеют отношения к поставленной задаче;</p> <p>1 балл – школьники выдвигают различные предположения, связанные с поставленной задачей, но не связывают свои экспериментальные пробы с задачей;</p> <p>2 балла – школьники говорят о различных предположениях, инициативно пытаются связать задачу с экспериментальными пробами</p>
Рефлексия и обобщение – выход на первичные представления физических понятий, закономерности	<p>0 баллов – не происходит обобщение и соотнесение с известным общим законом (закономерностью);</p> <p>1 балл – происходит обобщение и соотнесение с известным общим законом (закономерностью)</p>

Технология проведения исследования, сбора данных

Сбор данных для исследования включал два этапа. Первый этап состоял в проведении экспериментальных сессий с учащимися различных школ и классов, в ходе которых деятельность учащихся и учителя фиксировалась с помощью высокотехнологичных устройств – видеокamеры (или смартфона) и диктофона. Второй этап характеризовался кабинетным исследованием в виде фиксации и оценки действий учеников и учителя на основе аудио- и видеоинформации независимыми исследователями (мы их называем ассессорами).

Процедура экспериментальной деятельности предполагает работу в команде. Для этого учащиеся одного класса или нескольких классов распределялись на четыре группы (острова). Рядом с каждой группой находилось видеозаписывающее устройство, а также диктофон для более качественной фиксации речи. Кроме того, использовались панорамные съемки, охватывающие все помещение, по которому перемещался учитель. Учитель снабжался микрофоном-петличкой для лучшего распознавания его речи при дальнейшем кабинетном исследовании.

В кабинетном исследовании были задействованы 19 ассессоров, среди которых были как доктора и кандидаты наук, так и лица без ученой степени. Ассессорами выступали некоторые исполнители проекта «Новая физика», в рамках которого проводилось данное исследование. Все ассессоры прошли обучение по использованию инструмента, провели предварительную сессию оценивания, прежде чем приступили к оценке деятельности учащихся в рамках основного этапа. Предварительная сессия заключалась в коллективном просмотре тестового видеоролика и в оценке деятельности учащихся с помощью разработанного инструмента, завершалась сессия вопросами по его содержанию и процедуре оценивания. Все ассессоры по основной должности были преподавателями или научными сотрудниками ведущих вузов России. Ассессоры работали независимо друг от друга. В ходе просмотра роликов они заполняли карту (матрицу) наблюдений в электронном формате.

Необходимо отметить, что оценка действий на основе просмотра видеороликов и прослушивания аудиозаписей является нетривиальным решением для психолого-педагогических исследований. Опора на аудио- и видеоисточники информации, с одной стороны, позволяет более тщательно провести оценку действий (за счет возможности перемотки, возврата к необходимым местам видеоролика, обращения за помощью к коллегам); с другой – помогает сделать итоги исследования реплицируемыми (воспроизводимыми) и верифицируемыми. Ведь именно невоз-

возможность репликации результатов на сегодняшний день является одной из наиболее острых проблем социальных наук.

Исследование было проведено в трех школах г. Южно-Сахалинска. В школах было организовано 15 групп учащихся 7 класса. В каждой группе было по 4 человека, итого в исследовании приняли участие 60 обучающихся. Экспериментальные сессии независимо друг от друга проводили 4 учителя. Сбор данных проходил в апреле 2023 года.

В данном исследовании соблюдались общепринятые этические принципы на всех этапах его проведения. Участники были полностью проинформированы о целях, процедуре и возможных результатах исследования, а их добровольное участие было обеспечено. Были получены информированные согласия от родителей обучающихся. Конфиденциальность данных строго соблюдалась. Данные хранились в защищенном месте, доступном только для исследовательской группы, и использовались исключительно в рамках задач исследования.

Предметный (физический) материал, задачи, предлагаемые учащимся в ходе исследования

В рамках экспериментальных сессий учащиеся решали две задачи под условным названием «Посади птичку в клетку» и «Волчки». Авторы задач – М. Д. Солдатенкова и К. О. Теплякова, сотрудники Лаборатории квантовых детекторов Московского педагогического государственного университета. Суть первой задачи заключается в том, что учащимся необходимо, вращая игрушку (тауматроп) с изображением птицы и клетки и меняя характеристики видекамеры на телефоне (частоту кадров в секунду), посадить птичку в клетку. Во второй задаче обучающимся нужно, вращая волчки, снять несколько видео, на которых можно увидеть на волчке различное количество точек (одну, две или несколько), тогда как в состоянии покоя видна только одна точка.

В рамках упомянутого проекта «Новая физика» проводились еженедельные семинары, на которых обсуждались направления работы участников, готовность к проведению исследований, результаты исследований и пр. На каждом семинаре велись протоколы. Для проверки гипотезы № 1 использованы данные протоколов семинаров № 4 от 13.04.2023 и № 7 от 04.05.2023. Семинар № 4 проводился в период сбора данных в школах г. Южно-Сахалинска и был посвящен предварительному анализу результатов проб учащихся, а также вопросам работы с инструментом оценивания экспериментальной деятельности. На этом семинаре выступали те участники проекта, которые непосредственно находились в школах г. Южно-Сахалинска и наблюдали за экспериментальной деятельностью учащихся в режиме реального времени. Семинар № 7 проводился после сбора данных и был посвящен уже более глубокому анализу возможности использования инструмента при работе с видео- и аудиозаписями экспериментальных сессий. На этом семинаре выступали ассессоры, которые провели первые пробы оценивания в рамках кабинетного исследования. Проверка гипотезы № 2 в большей степени опиралась на работу всех ассессоров с картой (матрицей) наблюдений, с помощью которой они смогли зафиксировать действия учителей и сопоставить их с действиями учащихся. В данном случае использованы материалы протокола семинара № 11 от 01.06.2023. Гипотеза № 3 проверялась с помощью методов исследования согласованности оценок ассессоров. Для этого каждую группу учащихся оценивали два ассессора независимо друг от друга. В качестве мер согласованности использовались два показателя – доля совпадений оценок и непараметрический коэффициент конкордации Кендалла. Первый показатель рассчитывается на основе оператора условия: *если *совпадение*, то 1; иначе*

0. Затем делится количество единиц на общее число оценок. Коэффициент конкордации Кендалла основан на рангах и высчитывается по следующей формуле:

$$W = \frac{12S}{m^2 \cdot (n^3 - n) - m \cdot \sum_{i=1}^m T_i^2},$$

где S – отклонение рангов от среднего, m – количество экспертов, n – количество оцениваемых факторов (критериев), $\sum_{i=1}^m T_i$ – сумма повторяющихся рангов (связок).

У доли совпадений нет четко установленных пороговых значений, обычно пишут, что чем ближе к 1, тем лучше (Gisev et al., 2013); в то же время для коэффициента конкордации Кендалла они следующие (Moslem et al., 2019):

Диапазон W	Интерпретация
[0.00; 0.09]	Нет согласованности
[0.10; 0.29]	Низкая согласованность
[0.30; 0.59]	Средняя согласованность
[0.60; 1.00]	Высокая согласованность

Долю совпадений можно рассчитать как по каждому критерию матрицы наблюдений, так и в среднем по всем критериям. Коэффициент конкордации рассчитывается только по всем критериям вместе.

Результаты исследования

Проверка гипотезы №1

На семинаре № 4 состоялось выступление участника М. по итогам проведения экспериментальных сессий в школах г. Южно-Сахалинска. Был представлен ответ на вопрос: «Все ли этапы из матрицы наблюдения присутствуют в структуре деятельности? Есть ли что-то, чего нет в карте наблюдения?». Ответ участника М. был положительным: «Предварительно, все основные этапы структуры деятельности присутствуют в карте наблюдения. Точнее можно будет сказать после подробного анализа аудио и видеозаписей исследовательской сессии».

На семинаре № 7 был представлен доклад участника Н. под названием «Первые впечатления от просмотра видеороликов с экспериментальной деятельностью учащихся». По результатам просмотра трех видеоматериалов докладчиком был сделан следующий вывод: структура деятельности воспроизводится не полностью. Не хватало выдвижения гипотез, оценки проб, моделирования. Структура деятельности учеников при решении второй задачи значительно сжата, что усложняет выделение этапов действия в матрице наблюдения. При этом было замечено, что этапы деятельности проявляются нелинейно, это также нужно учитывать при обсуждении структуры деятельности. К примеру, на этапе представления «продукта» можно понять, получилось ли у обучающихся смоделировать выделенное отношение. При этом до этапа представления «продукта» моделирование не фиксировалось наблюдателем.

Таким образом, можно заметить, что структура деятельности потенциально воспроизводима, однако эмпирически (на практике) в естественных условиях полностью проявляется не всегда.

Проверка гипотезы № 2

В ходе обсуждения опыта оценивания ассессорами были выявлены проблемы и высказаны следующие замечания:

– при проведении анализа видеоматериалов обнаружена проблема качества звука и распознавания действий ребенка (не видно, что ребенок делает в телефоне во время экспериментальной деятельности);

– в оценке действий учеников и учителей также отмечены следующие сложности: не были представлены участники деятельности (учитель, наблюдающий), неоднозначно деление детей на группы, не обозначена роль планшета на каждом острове (он был нужен для записи звука);

– перед проведением исследовательской сессии была проведена встреча, переданы карточки с заданиями, поставлена задача перед учителями – организовать занятие экспериментальной деятельности. Если предложить учителям подготовить технологическую карту урока, то можно будет оценить наличие или отсутствие действий учителя на этапе планирования и сравнить с этапами экспериментальной деятельности, которые были сформированы в матрице оценивания;

– в некоторых школах дети заполняли карточки, в других – нет. Хорошо бы понять, что дети заполняли и почему не все дети заполняли;

– каждый учитель реализует свой план действий. Поэтому предложено записать интервью с учителем по результатам проведенного исследования о его замысле.

Как можно заметить, все перечисленные замечания и предложения относятся к процедуре исследования и в меньшей – к содержанию разработанного инструмента. Это говорит о том, что содержательно, а не технически, у ассессоров не возникло проблем при установлении соответствия действий учителя и учащихся.

Проверка гипотезы № 3

В Таблице 2 представлены показатели согласованности экспертных оценок. Можно заметить, что доля совпадений по таким критериям, как принятие задачи, моделирование, создание и представление «продукта» находится на уровне случайного угадывания. По другим критериям согласованность существенно выше. При этом среднее значение доли совпадений оценок выше случайного угадывания. В связи с тем что доля совпадений – консервативная оценка согласованности, лучше ориентироваться на показатель конкордации Кендалла. Коэффициент конкордации оценок экспертов оказался равным 0.73, что говорит о достаточно высокой согласованности мнений экспертов в целом по всем критериям.

Таблица 2. Показатели согласованности экспертных оценок при наблюдении за действиями учащихся

Критерий	Доля совпадений оценок	Коэффициент конкордации
Принятие задачи	0.50	–
Выдвижение гипотез	0.75	–
Осуществление проб	1.00	–
Проверка (оценка) проб	1.00	–
Моделирование физического принципа	0.50	–
Создание «продукта», моделирующего физический принцип действия	0.50	–
Представление «продукта» другим учащимся (группам)	0.50	–
Рефлексия и обобщение – выход на первичные представления физических понятий, закономерности	0.75	–
Среднее	0.66	0.73

Обсуждение

По результатам исследования обнаружилось, что специалисты в области физики и педагогики высоко оценивают содержание инструмента. Кроме того, результаты эмпирической оценки экспериментальной деятельности указывают на то, что независимые эксперты (ассесоры) дают близкие (согласованные) оценки при наблюдении за одними и теми же группами учащихся. Все это в совокупности подтверждает содержательную валидность собираемых данных.

Насколько нам известно, наш инструмент пока не имеет аналогов в научной сфере и педагогической практике, поэтому непосредственно сопоставить результаты этого исследования с другими не представляется возможным. Подтвердим эту мысль примерами. Немецкие исследователи (Krüger et al., 2022) сравнивали эффективность классических экспериментов с экспериментами, основанными на интерактивной компьютерной симуляции, на основе авторской анкеты, куда были включены вопросы предметного содержания и вопросы о субъективной когнитивной сложности выполняемых заданий. Коллеги из Индонезии (Jalil et al., 2018) разработали ряд задач и систему оценивания, однако их задачи, на наш взгляд, больше похожи на лабораторную работу, а система оценивания (рубрики оценивания) касается только предметных результатов и не включает метапредметные компоненты.

Несмотря на разнотипность инструментов, косвенные сравнения вполне допустимы. Так, в модели оценки естественнонаучной грамотности исследования PISA присутствует такая компетенция, как понимание особенностей естественнонаучного исследования, которая характеризуется способностью выдвигать гипотезы, оценивать способы научного исследования, предлагать дизайн исследования для проверки выдвинутой гипотезы и др. (OECD, 2023). При этом интерпретация данных исследования, преобразование одной формы представления данных в другую (по сути, моделирование) вынесены в другую компетенцию – интерпретацию данных и использование научных доказательств для получения выводов. Содержательно модель оценки естественнонаучной грамотности по PISA во многом пересекается с содержанием нашего инструмента, так как структура экспериментальной деятельности, по нашему представлению, универсальна. При этом необходимо сказать о различиях между этими инструментами. Во-первых, инструменты собирают разного типа данные (Cattell, 1983): в случае PISA речь идет о Т-данных (objective test data по Р. Кеттелу), то есть о данных объективного тестирования; в нашем случае – о L-данных (life record data), то есть о данных наблюдения. Во-вторых, в исследовании PISA экспериментальная деятельность достаточно ограничена, в силу того что она симулируется в компьютерной среде; наш инструмент предназначен для оценки экспериментальной деятельности в естественной среде (натурный эксперимент). При этом мы не исключаем, что инструмент может быть использован для оценки экспериментальной деятельности, которая протекает в виртуальной или альтернативной реальности.

Необходимо отметить, что предлагаемое исследование содержит ряд ограничений. Во-первых, исследование проводилось только в трех школах г. Южно-Сахалинска, что ограничивает возможность обобщения полученных результатов на другие образовательные учреждения или регионы. Более широкая выборка, охватывающая различные географические регионы и большее количество школ, может дать более репрезентативные результаты. Во-вторых, в нашей модели исследования использование высокотехнологичных орудий для сбора данных (аудио- и видеозаписи) может привести к проблемам, если качество записей будет снижено из-за технических сбоев или некачественного оборудования, что может повлиять

на достоверность данных. Часть данных не была включена в исследование именно из-за низкого качества видеозаписей. Поиск способов снижения «зашумленности» данных при таком способе их сбора является ключевой задачей исследователей.

Заключение

Предлагаемое исследование посвящено проверке функциональности разработанного на основе теории учебной деятельности инструмента оценки экспериментальной деятельности обучающихся. Если говорить языком психометрики, то проверялась содержательная валидность данных, получаемых с помощью этого инструмента. Необходимость его разработки была обусловлена тем, что в распоряжении исследователей и практиков нет высококачественных методик непосредственной оценки эффективности экспериментальной деятельности учащихся. Это обстоятельство позволяет нам говорить об уникальности и значимости разработанного нами инструмента.

На основе экспертных суждений было подтверждено, что разработанный инструмент наблюдения за экспериментальной деятельностью учащихся позволяет зафиксировать действия по теоретической структуре и сопоставить действия учеников на разных этапах. Более того, результаты наблюдений независимыми ассессорами дают близкие оценки, что отразилось в приемлемых значениях показателей согласованности оценок и говорит об относительной устойчивости инструмента по отношению к замене (смене) субъекта оценивания. Кроме того, необходимо отметить, что представленный инструмент позволяет оценивать деятельность учащихся в рамках педагогического исследования на основе непосредственных и естественных действий, в ходе обычного занятия, а не специальных аттестационных или тестовых процедур. Таким образом, использование разработанного инструмента может стать частью образовательной деятельности, позволяя на основе получаемых данных совершенствовать учебный процесс.

Дальнейшие исследования будут связаны с проверкой системы оценивания – наполняемостью категорий оценивания, уточнением формулировок, а также с общей оценкой экспериментальной деятельности учащихся как на индивидуальном, так и на групповом уровне на достаточных по объему выборках. Кроме того, необходимы дополнительные свидетельства валидности собираемых с помощью предлагаемого инструмента данных – структура, связь с другими близкими переменными и т. д.

Благодарности

Работа выполнена при поддержке проекта «Новая физика»: Научно-методическое обоснование обновления содержания программ по физике основного и среднего общего образования и подготовки педагогов-физиков к его реализации, 122081200114-0.

Список литературы

- Адамский, А. И., Львовский, В. А. Институциональные разрывы деятельностного подхода в физическом образовании // Вестник Московского городского педагогического университета. Серия: Педагогика и психология. – 2023. – Т. 17. – №. 1. – С. 96–113. – DOI:10.25688/2076-9121.2023.17.1.05
- Басюк, В. С., Ковалева, Г. С. Инновационный проект Министерства просвещения «Мониторинг формирования функциональной грамотности»: основные направления и первые результаты // Отечественная и зарубежная педагогика. – 2019. – Т. 1. – №. 4(61). – С. 13–33.
- Давыдов, В. В. Теория развивающего обучения. – М.: ИНТОР, 1996. – 544 с.

- Кумбс, Ф. Г. Кризис образования в современном мире: системный анализ / Пер. с англ. С. Л. Володиной [и др.]. – М.: Прогресс, 1970. – 261 с.
- Львовский, В. А., Янишевская, М. А., Якушина, Е. В. Деятельностный подход к обучению предметам естественно-научного цикла – залог формирования у обучающихся навыков, необходимых в реальной жизни // Медиа. Информация. Коммуникация. – 2021. – Т. 36. – №. 1. – С. 42–46.
- Михайлов, Ф. Т. Загадка человеческого Я. – М.: Политиздат, 1976. – 287 с.
- Пентин, А. Ю., Ковалева, Г. С., Давыдова, Е. И., Смирнова, Е. С. Состояние естественнонаучного образования в российской школе по результатам международных исследований TIMSS и PISA // Вопросы образования. – 2018. – №. 1. – С. 79–109. – DOI:10.17323/1814-9545-2018-1-79-109
- Пентин, А. Ю., Никифоров, Г. Г., Никишова, Е. А. Основные подходы к оценке естественнонаучной грамотности // Отечественная и зарубежная педагогика. – 2019. – Т. 1. – №. 4(61). – С. 80–97.
- Приказ Министерства труда и социальной защиты РФ от 18 октября 2013 г. N 544н "Об утверждении профессионального стандарта "Педагог (педагогическая деятельность в сфере дошкольного, начального общего, основного общего, среднего общего образования) (воспитатель, учитель)". – Электронный ресурс. – Режим доступа: <https://base.garant.ru/70535556/?ysclid=lkc9kpttmg502725260> (дата обращения: 18.07.2023)
- Рабочая программа основного общего образования. Физика. Базовый уровень / Институт стратегии развития образования. – Электронный ресурс. – Режим доступа: https://edsoo.ru/Primernaya_rabochaya_programma_osnovnogo_obschego_obrazovaniya_predmeta_Fizika_proekt_.htm?ysclid=lk7yi2yud61688715464 (дата обращения: 18.07.2023).
- Семенов, А. Л., Зискин, К. Е. Концепция расширенной личности как ориентир цифрового пути образования // Герценовские чтения: психологические исследования в образовании. – 2021. – № 4. – Т. 2289. – С. 530–535.
- Совместная учебная деятельность и развитие детей / Коллективная монография. Под редакцией В. В. Рубцова, И. М. Улановской. – М.: ФГБОУ ВО МГППУ, 2021. – 352 с.
- Страннолюбский, А. Н. Состояние народного образования в селах Европейской России // Русская школа. – 1893. – № 4. – С. 107–143.
- Чулкова, Г. М. Инновации в системе высшего образования // Управление наукой: теория и практика. – 2022. – Т. 4. – № 3. – С. 129–140. – DOI:10.19181/smtp.2022.4.3.9
- Cattel, R. B. Structured Personality-learning Theory: A Wholistic Multivariate Research Approach. – New York, NY: Praeger, 1983. – 466 p.
- Gisev, N., Bell, J. S., Chen, T. F. Interrater agreement and interrater reliability: key concepts, approaches, and applications // Research in Social and Administrative Pharmacy. – 2013. – Vol. 9. – No. 3. – Pp. 330–338. – DOI:10.1016/j.sapharm.2012.04.004
- Jalil, S., Ali, M. S., Haris, A. Development and validation of science process skills instrument in physics // Journal of Physics: Conference Series. – 2018. – Vol. 1028. – No. 1. – URL: <https://iopscience.iop.org/article/10.1088/1742-6596/1028/1/012203/pdf> (accessed: 02.05.2024)
- Krüger J. T., Höffler T. N., Wahl M., Knickmeier K., Parchmann I. Two comparative studies of computer simulations and experiments as learning tools in school and out-of-school education // Instructional Science. – 2022. – Vol. 50. – Pp. 169–197. – DOI:10.1007/s11251-021-09566-1
- Moslem, S., Ghorbanzadeh, O., Blaschke, T., Duleba, S. Analysing stakeholder consensus for a sustainable transport development decision by the fuzzy AHP and interval AHP // Sustainability. – 2019. – Vol. 11. – No. 12. – DOI:10.3390/su11123271
- OECD. PISA 2025 Science Framework (Draft). 2023. – URL: https://pisa-framework.oecd.org/science-2025/assets/docs/PISA_2025_Science_Framework.pdf (accessed: 02.05.2024)

References

- Adamskiy, A. I., & Lvovsky, V. A. (2023). Institutional breaks of the activity-based approach in physical education. *Vestnik MGPU. Seriya «Pedagogika i psihologiya» – MCU Journal of Pedagogy and Psychology*, 17(1), 96–113. <https://doi.org/10.25688/2076-9121.2023.17.1.05>

- Basyuk, V. S., & Kovaleva, G. S. (2019). Innovative project of the Ministry of Education “Monitoring of functional literacy development”: The main directions and the first results. *Otechestvennaya i zarubezhnaya pedagogika – Domestic and foreign pedagogy*, 1(4(61)), 13–33.
- Cattel, R. B. (1983). *Structured Personality-learning Theory: A Wholistic Multivariate Research Approach*. Praeger.
- Chulkova, G. M. (2022). Innovations in the System of Higher Education. *Upravlenie naukoj: teoriya i praktika – Science Management: Theory and Practice*, 4(3), 129–140. <https://doi.org/10.19181/sntp.2022.4.3.9>
- Coombs, Ph. H. (1970). *The world educational crisis. A systems analysis*. Progress.
- Davydov, V. V. (1996). Developmental learning theory. INTOR.
- Gisev, N., Bell, J. S., & Chen, T. F. (2013). Interrater agreement and interrater reliability: key concepts, approaches, and applications. *Research in Social and Administrative Pharmacy*, 9(3), 330–338. <https://doi.org/10.1016/j.sapharm.2012.04.004>
- Institute for Strategy of Education Development. (2021). *Working program of basic general education. Physics. Basic level*. https://edsoo.ru/Primernaya_rabochaya_programma_osnovnogo_obshchego_obrazovaniya_predmeta_Fizika_proekt_.htm?ysclid=lk7yi2yd61688715464
- Jalil, S., Herman, Sidin Ali, M., & Haris, A. (2018). Development and validation of science process skills instrument in physics. *Journal of Physics: Conference Series*, 1028(1). <https://doi.org/10.1088/1742-6596/1028/1/012203>
- Krüger, J. T., Höfler, T. N., Wahl, M., Knickmeier, K., & Parchmann, I. (2022). Two comparative studies of computer simulations and experiments as learning tools in school and out-of-school education. *Instructional Science*, 50, 69–197. <https://doi.org/10.1007/s11251-021-09566-1>
- Lvovsky, V., Yanishevskaya M., & Yakushina, E. (2021). An active approach to teaching the subjects of the natural science cycle is the key to the formation of students' skills necessary in real life. *Media. Informaciya. Kommunikaciya – Media. Information. Communication*, 36(1), 42–46.
- Mikhailov, F. T. (1976). *The mystery of the human self*. Politizdat.
- Ministry of Labor and Social Protection of the Russian Federation. (2013). On approval of the professional standard “Pedagogue (pedagogical activity in the sphere of preschool, primary general, basic general, secondary general education) (tutor, teacher)”; Order No. 544н, 18th October. <https://base.garant.ru/70535556/?ysclid=lkc9kptmg502725260>
- Moslem, S., Ghorbanzadeh, O., Blaschke, T., & Duleba, S. (2019). Analysing stakeholder consensus for a sustainable transport development decision by the fuzzy AHP and interval AHP. *Sustainability*, 11(12), 3271. <https://doi.org/10.3390/su11123271>
- OECD. (2023). PISA 2025 Science Framework (Draft). https://pisa-framework.oecd.org/science-2025/assets/docs/PISA_2025_Science_Framework.pdf
- Pentin, A. Yu., Nikiforof, G. G., & Nikishova, E. A. (2019). Main approaches to the assessment of scientific literacy. *Otechestvennaya i zarubezhnaya pedagogika – Domestic and foreign pedagogy*, 1(4(61)), 80–97.
- Pentin, A., Kovaleva, G., Davidova, E., & Smirnova, E. (2018). Science Education in Russia According to the Results of the TIMSS and PISA International Studies. *Voprosy obrazovaniya – Educational Studies Moscow*, 1, 79–109. <https://doi.org/10.17323/1814-9545-2018-1-79-109>
- Rubtsov, V. V., & Ulanovskaya, I. M. (Eds.). (2021). *Cooperative learning activity and children's development*. MGPPU.
- Semenov, A. L. & Ziskin, K. E. (2021). The concept of an expanded personality as a reference point of the digital path of education. *International Research and Practice Conference “The Herzen University Conference on Psychology in Education”*, 4, 530–535.
- Strannolyubsky, A. N. (1893). The State of Public Education in the Villages of European Russia. *Russkaya shkola – Russian school*, 4, 107–143.