DOI: 10.31862/2500-297X-2020-3-128-146

Ю.А. Тюменева¹, А.С. Обухов¹, О.Н. Финогенова²

¹ Национальный исследовательский университет
«Высшая школа экономики»,
101000 г. Москва, Российская Федерация

² Сибирский федеральный университет,
660041 г. Красноярск, Российская Федерация

Разработка и валидизация шкалы поведенческих индикаторов для оценки интереса к математике

Авторы статьи преследуют две цели. Первая – показать, что распространенный подход к выявлению предметных интересов и установок через утверждения общедекларативного характера («Мне нравится математика») имеет серьезные дефекты, особенно когда практикуется на детях в массовых опросах. В работе это демонстрируется на примере шкал интереса к математике в крупномасштабном международном мониторинге качества образования Trends in International Mathematics and Science Study (TIMSS). Вторая и главная цель – продемонстрировать альтернативный подход к оценке интересов, а именно, через специфические поведенческие индикаторы, разработанные в рамках существующих теорий интереса и его развития. Разработанная поведенческая шкала интереса к математике (ПГИИМ) имеет не только очень хорошие психометрические характеристики, но и различает детей, обучающихся по углубленной программе, и детей с обычной программой обучения, чего шкала TIMSS не делает. Обсуждается важность теории в разработке личностных шкал.

Ключевые слова: личностный интерес к математике, оценка интереса к математике, поведенческие индикаторы

© Тюменева Ю.А., Обухов А.С., Финогенова О.Н., 2020

Контент доступен по лицензии Creative Commons Attribution 4.0 International License
The content is licensed under a Creative Commons Attribution 4.0 International License

ССЫЛКА НА СТАТЬЮ: Тюменева Ю.А., Обухов А.С., Финогонова О.Н. Разработка и валидизация шкалы поведенческих индикаторов для оценки интереса к математике // Педагогика и психология образования. 2020. № 3. С. 128–146. DOI: 10.31862/2500-297X-2020-3-128-146

DOI: 10.31862/2500-297X-2020-3-128-146

Y. Tyumeneva¹, A. Obukhov¹, O. Finogenova²

¹ National Research University “Higher School of Economics”,
Moscow, 101000, Russian Federation

² Siberian Federal University,
Krasnoyarsk, 660041, Russian Federation

The development and validation of the math interest behavioral indicators scale

The goal of the paper was twofold. First, we intended to show that a prevalent approach to assessing domain-specific interests and attitudes through generalized items (*I like math*) has serious flaws, especially when applying to children in large-scale surveys. We demonstrated these flaws with the Scale of attitudes to math used in Trends in International Mathematics and Science Study (TIMSS). Second, we showed an alternative approach to assessing interests through behaviorally anchored items developed in the framework of the interest development theory. The new Math Interest Behavioral Indicators Scale (MIBI) proved not only excellent psychometric characteristics, but it also discriminated children who were from advanced math programs and who were not, what the TIMSS scale did not do. The importance of theory for scale development is discussed.

Key words: individual math interest, math interest assessment, behaviorally anchored rating scale

CITATION: Tyumeneva Y.A., Obukhov A.S., Finogenova O.N. The development and validation of the math interest behavioral indicators scale. *Pedagogy and Psychology of Education*. 2020. No. 3. Pp. 128–146. (In Russ.). DOI: 10.31862/2500-297X-2020-3-128-146

Дефициты подхода, реализованного в TIMSS-шкале интереса к математике

Опросники установок, интересов и мотивации широко используются в исследованиях образования, когда стоит задача объяснить или предсказать учебные достижения не только когнитивными способностями, но и личностными характеристиками ученика. Такого рода опросники нашли широкое применение как в психологических исследованиях, так и в исследованиях образования. В них используются близкие по содержанию шкалы, которые можно интерпретировать как шкалы установок и интересов по отношению к математике и другим школьным предметам.

К примеру, в анкете TIMSS-2015 для школьников 4-х и 8-х классов предлагаются т.н. мотивационные шкалы, предназначенные для оценки отношения учеников к изучаемым предметам. Одна из них – шкала «Ученикам нравится изучать математику» (“Students Like Learning Mathematics Scale”). Она включает девять утверждений об отношении к изучению математики, например «Я с нетерпением жду уроков математики», «Мне нравится изучать математику», «Уроки математики скучные». Ребенок должен показать степень своего согласия с каждым утверждением, используя одну из четырех категорий «Во многом согласен», «Согласен», «Не согласен», «Во многом не согласен». Эта шкала не уникальная для TIMSS-2015 – пять из девяти заданий используются в TIMSS, начиная с 2007 г.; очень схожие утверждения используются в другом международном мониторинговом исследовании – PISA [18; 22]. Ниже указываются общие риски такого подхода и специфические риски для шкал TIMSS.

Во-первых, существуют серьезные межиндивидуальные различия в самой интерпретации респондентами слишком обобщенных утверждений [28]. В TIMSS утверждения «Мне нравится изучать математику», «Я люблю математику», «Математика – один из моих любимых предметов», «Математика скучная» сформированы как раз в такой общей манере, не позволяющей однозначно интерпретировать смысл высказывания. Можно любить математику больше русского языка, потому что это не так скучно, например; или потому, что у ребенка сложились дружественные отношения с учителем. Но различия в причинах, приводящих к согласию или несогласию с утверждением, – критически важны с точки зрения валидности инструмента. В данном примере за формальным согласием может стоять либо отношение к другому предмету, а не к математике, либо отношения с конкретным учителем.

Более того, столь общие формулировки могут провоцировать ответы, хотя и отражающие отношение к математике, но действующие в противоположном направлении, чем рассчитывают получить авторы. Например, из-за невысоких требований учителя ребенку, не увлеченному математикой, может «нравиться изучать математику», а ребенок, увлеченный математикой, не согласится с этим утверждением, как раз потому, что на этих уроках такому ребенку скучно.

Во-вторых, такой подход никак не защищен от искажающего влияния социальной желательности и индивидуального стиля ответов. Это может иметь разрушительные эффекты, прежде всего, для валидности этих инструментов и для последующей оценки корреляций между измеренными чертами [26; 27; 37]. Риски этих искажений еще больше усиливаются в применении к детской выборке, или к опросу, который является недобровольным и/или имеющим последствия для респондентов [2]. Именно такой является ситуация мониторинговых исследований качества образования, когда респонденты лишены возможности отказаться от участия в исследовании, и когда последствия его результатов воспринимаются (особенно младшими детьми) совершенно неконтролируемым образом. Так, было показано влияние социальной желательности на ответы учителей в другом мониторинговом исследовании образования TALIS, где использовался схожий декларативно-обобщенный подход к формулировке утверждений, и где само исследование формально не имело никаких последствий для респондентов [1].

В-третьих, такого рода декларативно общие утверждения часто разрабатываются без серьезных теоретических оснований, связывающих утверждения с природой исследуемого конструкта. Как отмечает Г. Марш, педагогический психолог и один из самых продуктивных и влиятельных исследователей Я-концепции, самооценки и мотивации: «К сожалению, высокие стандарты для тестов достижений, разработанные TIMSS, не отражаются в измерении мотивации ученика, поскольку мало подтверждений для сильной теоретической базы, стоящей за отбором и конструированием утверждений. В каждом последовательном цикле [TIMSS] природа этих утверждений изменялась, и психометрические доказательства в поддержку метода, используемого для начисления баллов, были неполные, неадекватные или просто не предоставлялись» [3, с. 111].

В самом деле, в качестве теоретической основы для этой шкалы в TIMSS указывается теория самодетерминации Э. Диси и Р. Райана [6; 7]. Однако ни в самой шкале, ни в технической документации TIMSS не просматривается никаких отсылок к составляющим этой теории. В частности, формулировки не учитывают субъективно

воспринимаемую причинность поведения, внешнего или внутреннего локуса контроля, произвольности выбора, эмоциональную связь со значимыми другими и иных компонентов, критически важных для теории самодетерминации.

В-четвертых, у подростков требует большой осмотрительности применение любых самооценочных шкал, и тем более самооценочных шкал относительно установок и интересов. Это связано с тем, что оценка собственного отношения к изучению какой-то предметной области интегрирована с представлением человека о себе, о своих склонностях, способностях и интересах, т.е. с общей самооценкой и Я-концепцией. Для возможности оценить эти личностные характеристики нужны одновременно две вещи. С одной стороны, требуется рефлексия, умение обобщать и умение различать ситуативные, временные состояния от стабильных личностных характеристик. Такого рода интеграция требует развитых когнитивных способностей, поэтому такая Я-система не может появиться раньше, чем когда произойдут основные изменения в когнитивных структурах детей, прежде всего их способности к абстрагированию и обобщению [2]. С другой стороны, сами оцениваемые характеристики, в данном случае, интерес и отношение к предметным дисциплинам, должны стать устойчивыми и дифференцированными.

Однако есть основания сомневаться, что в 10–11 лет (когда проводится TIMSS – 4 класс) эти личностные новообразования, интересы и Я-концепция, сформированы [5; 30; 34]. В эмпирических работах показано, что появление способности к абстрагированию и связанные с ней наиболее сильные трансформации в Я-системе подростка появляются после 15 лет [1]. С другой стороны, флуктуации в уровнях и стабильности самооценки достигают пика в период раннего подросткового возраста [4]. В соответствии с С. Хартер [13], ранний и средний подростковый возраст характеризуется нереалистичной положительной самооценкой и неспособностью отличить образ себя реального от себя идеального [9; 11; 15]. В период между средним и поздним подростничеством исследователи отмечают появление экстремально-негативных оценочных суждений [14]. Структура самооценки в этом возрасте еще изменяется от простого, глобального конструкта ко все более дифференцированной и многомерной [3; 29].

Таким образом, теоретические и эмпирические работы не дают, к сожалению, оснований считать, что ребенок 10–12 лет в состоянии согласованно оценивать свои личностные интересы и отношения, и что сами эти личностные образования сформировались к этому возрасту (см. напр., [6]).

Оценка интересов с помощью поведенческих индикаторов

С учетом вышеуказанных ограничений в использовании общих самооценочных утверждений в применении к мониторинговым опросам и несомненной важности их изучения встает вопрос об альтернативных способах оценки интересов. В этой работе мы предлагаем такой альтернативный подход, основанный на идее использования поведенческих характеристик для оценки личностных черт [12; 32]. Этот подход заключается в том, что выделяются поведенческие признаки, которые, как правило, проявляются, если выражена интересующая нас черта (в нашем случае это – интерес к математике). Источниками сведений о сопутствующих поведенческих признаках становятся люди, близко и длительно наблюдающие поведение человека, а сам факт выраженности у него черты обычно так или иначе зафиксирован документально.

Наиболее удобен этот подход в клинической психологии, где он и появился. Там наблюдателями симптомов поведенческих или психических особенностей являлись обычно родственники и медицинский персонал, а фактом выраженности черты – установленный медицинский диагноз. Однако, как показывают исследования, такой подход может быть с успехом применен и в оценке, так скажем, нормальной личности [10; 20; 24; 33; 35]. Основная сложность в реализации этого подхода заключается в трудоемком процессе интервьюирования «наблюдателей» – людей, близко знакомых с «носителями» интересующей нас черты, и определения наблюдаемых поведенческих характеристик, которые были бы общими всех случаев выраженной черты. В случае «интереса к математике» у детей «наблюдателями» являются, конечно, их родители. Выраженность черты может быть определена по факту длительного и желанного обучения на программах углубленной математики, причем не только в общеобразовательной школе, но и на дополнительных обучающих программах.

Чем подход через поведенческие признаки более полезен для оценки интереса к математике подростков, чем подход через согласие с утверждениями общедекларативного плана? Поведенческие факты – это гораздо более однозначные признаки, и поэтому для подростка легче определить их наличие или отсутствие (например, «Я ищу случая поучаствовать в олимпиадах по математике»), в отличие от довольно туманных определений, вроде «Мне нравится математика». Во-вторых, оценка отдельных поведенческих проявлений более уместна в случае с подростками, чьи рефлексивные суждения, как было отмечено

выше, отличаются неустойчивостью, подверженности текущим оценкам со стороны окружающих, экстремальностью, и в целом могут быть еще недостаточно интегрированы.

Теоретические основания шкалы поведенческих индикаторов (ПИИМ)

Принципиально важно, что отбираемые поведенческие индикаторы интереса к математике должны быть обоснованы с точки зрения теории. Это должна быть теория, которая причинным образом объясняет, как связан личностный интерес и определенное поведение.

В качестве теоретических оснований поведенческих индикаторов интереса к математике послужили представления о четырехэтапном пути развития предметного интереса, от ситуационного к личностному [5; 9]. С точки зрения теории М. Митчела, изначальный предметный интерес всегда ситуационный – он вызывается внешним по отношению к ребенку действием, которое на какое-то время захватывает внимание ребенка и приглашает к ответному действию. В случае математики это могут быть, например, действия учителя, которые привлекают и удерживают внимание ребенка во время урока. Ситуационный интерес может возникать в разных контекстах, задаваться разными способами и задевать разные потребности ребенка. Однако для того, чтобы на основе ситуационного мог возникнуть личностный интерес, ситуационный интерес, какими бы способами он не достигался, должен возникать стабильно в связи с определенной деятельностью [Там же].

В рамках теории развития личностного интереса С. Хиди и А. Реннингер [16], о личностном интересе можно говорить, когда человек сам начинает искать (а не просто ждать) возможности включиться в интересующую его деятельность и пребывать в ней как можно более часто и долго. То есть первое принципиальное отличие личностного интереса от ситуативного – в его активной природе. Другой отличительной чертой личностного интереса является стремление к все более разнообразным и все более трудным задачам. Кроме того, при личностном интересе человек становится чувствительным к любой информации, связанной с его интересом, и стремится получать эту информацию, т.е. теоретически должен вовлекаться в социальное взаимодействие по поводу интересующей его деятельности.

Таким образом, ключевыми характеристиками личностного интереса к математике, в соответствии с теориями интереса [16; 25], являются активный поиск возможности заняться математикой, поиск новых сложных задач и новых знаний в математике и готовность к общению по поводу математики.

Разработка утверждений

На первом этапе разработки утверждений был проведен опрос родителей старших подростков (от 14 до 17 лет), которые были постоянными участниками олимпиад, конкурсов, летних школ по математике и чьи успехи в школьной математике были высокими. Кроме того, эти подростки, по мнению родителей, имели сильный и длительный интерес к математике. Сочетание этих признаков, на наш взгляд, являлось достаточным, чтобы утверждать, что мы имели дело с детьми с выраженным интересом к математике. Родителям (матерям) таких подростков были заданы вопросы: «Что в поведении вашего ребенка отличает его от сверстников, которые не интересуются математикой? Какие признаки в поведении вашего ребенка указывают на его интерес к математике? Как вы заметили, что ваш ребенок проявляет интерес к математике?». Всего было опрошено шесть матерей, и опрос был прекращен, т.к. два последних случая не добавили ничего к уже названным признакам. Итоговый список состоял из 16 признаков, соответствующих теории личностного интереса, но только 8 из них повторялись во всех случаях и при этом могли быть однозначно оценены самим ребенком. Эти поведенческие признаки были сформулированы как утверждения и прошли апробацию на 8 подростках (от 11 до 17 лет) на предмет понятности, однозначности и легкости идентификации этого поведения.

Кроме того, мы опросили еще трех матерей, чьи дети (подростки 13 и 14 лет), по признанию этих матерей, «очень не любят математику». Два поведенческих признака, выделенных этими родителями как свидетельство «нелюбви к математике», были добавлены в итоговую шкалу как обратные утверждения.

Мы добавили одно утверждение, доступное только для самонаблюдения («Я решаю математические задачи просто из интереса»), но оно касалось внутренней мотивации и поэтому было важным для оценки личностного интереса. По результату этой же апробации было принято решение оставить шестибальльную ответную шкалу без средней категории, с крайними категориями «Это полностью похоже на меня» и «Это совершенно непохоже на меня». Промежуточные четыре категории ответной шкалы не были поименованы.

Итоговые утверждения шкалы поведенческих признаков интереса к математике (ПИИМ) представлены ниже:

- 1) «Меня трудно оторвать от занятий математикой»;
- 2) «Я могу биться над трудной задачей по математике час и больше»;
- 3) «Я ищу возможности заниматься углубленно математикой»;
- 4) «По доброй воле я не открываю учебник или книгу по математике»;

- 5) «Я ищу случая поучаствовать в математических конкурсах или олимпиадах»;
- 6) «Я смотрю видео или читаю что-то про математику, когда никто не требует этого от меня»;
- 7) «Я решаю математические задачи просто из интереса»;
- 8) «Если есть возможность поговорить про математику, я делаю это»;
- 9) «Я избегаю любой деятельности, связанной с математикой»;
- 10) «Любой разговор, как-то связанный с математикой, привлекает мое внимание»;
- 11) «Я выискиваю в интернете или книгах интересные математические задачи».

Обоснование критерия

Поскольку нашей задачей было не только разработать новую шкалу интереса к математике на принципах иных, чем сделана шкала TIMSS, но и продемонстрировать, что она обладает лучшими различительными свойствами, нам нужно было сопоставить показатели по двум шкалам с некоторым критериальным поведением. Им выступил факт обучения детей по углубленной математической программе.

Для того чтобы снизить шансы случайного, недобровольного или необоснованного нахождения ребенка на углубленной программе обучения математике и чтобы создать более контрастные критериальные группы, мы использовали двойной контроль. Во-первых, опрашивались, с одной стороны, дети из классов, где реализуется «профильное» обучение математике, из классов специализированных математических школ; дети – участники летних математических школ или программ. С другой стороны, опрашивались дети из школ и классов «гуманитарной» направленности.

Во-вторых, чтобы быть уверенным в том, что ребенок отнесен к группе «углубленная математика» на достаточных основаниях, особенно при оценке детей из летних математических школ, мы включили три вопроса-фильтра в анкету, предваряющую опрос: «Учишься ли ты в математической школе?»; «Учишься ли ты в профильном математическом классе?»; «Учишься ли ты по углубленной математической программе?». При любых двух положительных ответах ребенок приписывался к группе «углубленной математики».

Метод

Выборка. В исследовании приняли участие 525 подростков – учащихся математических школ и классов, а также школ и классов гуманитарного профиля. Однако перед включением в анализ были удалены случаи с ответами, несоответствующие критериям качества:

1) выбор одной и той же ответной опции для всех утверждений в опроснике TIMSS или ПИИМ;

2) согласие или несогласие одновременно с двумя противоположно направленными утверждениями, например, «Я избегаю любой деятельности, связанной с математикой» и «Меня трудно оторвать от занятий математикой»;

3) более двух пропущенных ответов в одном из опросников.

После удаления «некачественных» профилей в анализ было включено 461 случай. Возраст детей варьировался от 11 до 18 лет ($M = 14,5$, $SD = 1,8$); девочек 50%.

Инструмент. Весь опрос состоял из трех частей. Первая часть, анкета, включала вопросы про пол и возраст, а также про факт обучения в математической школе, в профильном математическом классе и по углубленной математической программе. Вторая часть – Шкала установок к математике TIMSS-2015. Она стояла из девяти утверждений и ответной шкалы с четырьмя категориями (табл. 1). Третья часть – разработанная шкала поведенческих индикаторов интереса к математике (ПИИМ) из 11 утверждений и ответной шкалой с шестью категориями.

Таблица 1

Шкала TIMSS «Ученикам нравится изучать математику»

1	Мне нравится изучать математику
2	Хотелось бы, чтобы мне не надо было изучать математику
3	Математика – это скучно
4	Изучая математику, я узнаю много интересного
5	Я люблю математику
6	Я особенно люблю задания, где есть числа
7	Я люблю решать математические задачи
8	Я с нетерпением жду уроков математики
9	Математика – один из моих любимых предметов

Процедура. Исследование было добровольным и анонимным. Проводилось преимущественно через письменный опрос (80% случаев) во время занятий на уроке или в летних лагерях и летних школах; или онлайн (20% случаев) через Google form, где ссылку на опрос дети получали через своих учителей или через родителей. Время опроса не ограничивалось, но оно редко превышало 15 минут.

Результаты

Структура и психометрические показатели шкалы ПИИМ. Структура шкалы была оценена с помощью конфирматорного факторного анализа (CFA) методом диагонально взвешенных наименьших квадратов (WLSMV) в программе Mplus, версия 7.3. Метод WLSMV был специально разработан для порядковых данных, когда не делаются допущения относительно распределения, но предполагается, что латентная переменная распределена по нормальному закону [21]. Утверждения оценивались как категориальные переменные, т.к. у нас нет оснований считать показатели по шкале ПИИМ метрическими.

Для оценки соответствия данных модели использовались общепринятые критерии: значения TLI (Tucker–Lewis index) и CFI (comparative fit index) больше, чем 0,90 и 0,95, соответственно отражая приемлемый и превосходный уровень соответствия; значения RMSEA (root mean square error of approximation) 0,05 и 0,08, соответственно отражая близкое соответствие данных модели и минимально приемлемое; значения WRMR (weighted root mean residual) меньше 0,9 [31].

Теоретически мы ожидали, что новая шкала будет иметь однофакторную структуру, т.к. каждый поведенческий индикатор должен был отражать единый латентный конструкт (общий фактор) – интерес к математике. С другой стороны, два утверждения в шкале ПИИМ были сформулированы в негативном ключе, так что можно было ожидать возникновения второго фактора, связанного с эффектом метода [8; 19; 36].

Поэтому мы тестировали две модели: Модель 1, с одним общим фактором интереса к математике (ИМ), и Модель 2, где первый фактор был связан с основным конструктом – интересом к математике, а второй с «эффектом метода» – негативно сформулированными утверждениями (ИМ + НУ). Эффект негативных утверждений контролировался наличием дополнительной размерности – коррелированной уникальной дисперсии двух негативно сформулированных утверждений.

Дополнительно была проверена возможность, что два утверждения, касающихся обсуждения математики («Если есть возможность поговорить про математику, я делаю это» и «Любой разговор, как-то связанный с математикой, привлекает мое внимание») создают дополнительную размерность. Эти утверждения отражают не столько поиск возможности заняться математикой и удовольствие от решения математических проблем, сколько готовность и расположенность к социальному взаимодействию по поводу математики. Мы контролировали присутствие дополнительной «социальной» размерности через коррелированную уникальную дисперсию двух этих утверждений (Модель 3, ИМ + НУ + С).

Статистики согласия Моделей 1, 2 и 3 показаны в табл. 2. Модели 1 и 2 показали плохие статистики согласия. Модель 3, предполагающая наличие общего фактора интереса к математике с дополнительными размерностями для негативных утверждений и для утверждений про социальное взаимодействие (ИМ + НУ + С), имела очень хорошие статистики согласия. Поскольку эта структура находилась также в строгом соответствии с теорией личностного интереса, эта модель была оставлена как итоговая. Нагрузки утверждений в Модели 3 представлены на рис. 1.

Таблица 2

Статистики согласия трех структурных моделей

Модель	CFI	TLI	RMSEA	WRMR
ИМ	0,971	0,964	0,112	1,083
ИМ + НУ	0,974	0,967	0,107	1,028
ИМ + НУ + С	0,992	0,989	0,061	0,641

Исследование различительной способности шкал ПИИМ и TIMSS

Мы сравнили способность шкалы ПИИМ и способность шкалы интереса к математике TIMSS различать детей, обучающихся углубленной математике, и детей, обучающихся по программам общего или гуманитарного профиля. Для того чтобы снизить шансы обучения по тому или иному профилю по случайным причинам, по ошибке или потому, что ребенок осуществлял выбор программы самостоятельно, первоначальная выборка была сокращена до детей 14 лет и старше.

Оставшаяся выборка была разделена на две группы. В группу «Обучающиеся по углубленной программе» попали дети из классов, реализующих программу профильной или углубленной математики, либо дети из летних математических школ, отметивших в анкете два варианта из трех:

- 1) «Учусь в математической школе»;
- 2) «Учусь в профильном математическом классе»;
- 3) «Учусь по углубленной математической программе».

Обучающиеся, выбравшие другие варианты ответов, а также дети из гуманитарных классов и классов общего профиля, были отнесены в группу «Обычная программа».

Таким образом, в исследовании различительной способности шкал приняли участие 327 человек, 54% девочек. Средний возраст 15 лет и 3 месяца ($SD = 1,3$). Выборка была разделена на две подгруппы. В подгруппе «углубленная математика» было 106 детей (33,5%).

Рис. 1. Структурная модель шкалы ПИИМ

Надежность шкалы ПИИМ 0,928 (Альфа Кронбаха)

Результаты

В этом исследовании определялось, как хорошо шкалы ПИИМ и TIMSS могут различить детей из двух групп: углубленная и неуглубленная программа обучения по математике. Описательная статистика распределения показателей шкал показана в табл. 3. Мы использовали медиану распределений и непараметрический критерий сравнения распределений Манна–Уитни, т.к. у нас нет оснований считать показатели по шкалам ПИИМ и TIMSS метрическими.

Статистический тест Манна–Уитни показал, что распределение показателей по шкале ПИИМ в группах углубленной и обычной математики значительно отличаются ($p < 0,0001$) с более высокими значениями для группы «углубленной математики». По шкале TIMSS показатели в двух группах не отличались значимо ($p < 0,105$).

Таблица 3

**Характеристики распределения показателей шкалы ПИИМ
и шкалы интереса к математике TIMSS**

Шкалы	Углубленная программа	Обычная программа
<i>ПИИМ</i>		
Медиана	4	3
Skewness	–0,111 (0,235)	0,187 (0,168)
Kurtosis	–0,711 (0,465)	–1,091 (0,334)
<i>TIMSS</i>		
Медиана	2	2
Skewness	0,404 (0,235)	0,173 (0,168)
Kurtosis	–0,267 (0,465)	–0,887 (0,334)

Обсуждение

Предварительной целью нашей работы было продемонстрировать, что подход к оцениванию интересов и установок через утверждения общедекларативного характера несет в себе риск множественной интерпретации утверждений и искажению ответов из-за социальной желательности. Эти риски особенно усугубляются, когда факт опроса не связан с желанием респондентов и имеет неясные последствия для них. Однако, когда общедекларативный подход применяется к оценке интересов школьников, особенно детей младшего подросткового возраста, то к этим общим рискам добавляется еще сомнение в том, что сам измеряемый конструкт, личностный интерес, и способность его адекватно оценить, достаточно сформированы в этом возрасте. Как было показано в теоретической части, в отношении детей 10–12 лет вряд ли можно говорить об устойчивых и интегрированных личностных предметных интересах, так же, как и о сформированной способности к самооцениванию себя в реалистичной и устойчивой манере. С учетом этих сомнений остается удивительным, как длительно и широко используются в TIMSS общедекларативные шкалы всего его мотивационного блока анкеты, и в частности, исследуемая здесь шкала.

Второй и главной целью работы было продемонстрировать, что подход к оценке интересов может быть реализован через использование поведенческих индикаторов, т.е. набора признаков, которые могут быть

зафиксированы внешне и которые свидетельствуют о выраженности интереса к математике. Сами признаки были собраны через интервью с родителями подростков, чьи дети давно и постоянно демонстрируют сильный интерес к математике, вовлечены в углубленное ее изучение не только в обычной школе, но и на дополнительных курсах или программах, занимают призовые места на математических олимпиадах и имеют высокие академические достижения по математике. Для того чтобы поведенческие индикаторы соответствовали теоретическому пониманию конструкта личностного интереса, из родительских описаний были оставлены только те характеристики поведения детей, которые отражали активный поиск возможностей заняться математикой, желание новых знаний в математике и готовность к общению по поводу математики. Кроме того, в опросник были включены обратные утверждения. Разработанные в таком подходе утверждения, за счет отсылок к однозначным поведенческим, внешне регистрируемым признакам, могли быть оценены самим подростком с гораздо большей надежностью.

Однофакторная структура новой шкалы поведенческих индикаторов интереса к математике, соответствующая теории личностного интереса [16; 25], показала очень хорошее согласие с данными и высокую надежность. Она также обладала достаточной различительной способностью, чтобы дифференцировать детей из двух групп – обучающихся и не обучающихся по углубленной математической программе. Важным результатом является и то, что шкала интереса к математике TIMSS не смогла различить детей из этих групп.

Таким образом, новая шкала поведенческих индикаторов продемонстрировала хорошие структурные и различительные характеристики и может использоваться для исследования личностного интереса к математике школьников старшего подросткового возраста.

Исключительно важно и то, что эта шкала теоретически обоснована. То есть ее валидность основана не только на эмпирических корреляциях и статических индексах, а на том, что мы можем объяснить причинную связь между конструктом «личностный интерес к математике» и ответами респондента. В данном случае речь идет о том, что сформированный личностный интерес вызывает к жизни активный поиск и самостоятельное создание возможностей вновь и вновь заняться интересной деятельностью.

Нужно сказать, что для многих шкал мотивации и установок такая связь неочевидна. К примеру, в утверждениях шкалы TIMSS и в документации по нему нет эксплицитно высказанных предположений не только о том, как именно латентный конструкт связан с ответами по шкале,

но даже и о том, какой конструкт оценивается. Последнее следует из того факта, что ни в технической документации, ни в отчетах TIMSS не упоминается, что же оценивает эта шкала. Называются «мотивация», «интерес», «позитивные установки», а само название шкалы «Students Like Learning Mathematics Scale» («Ученикам нравится изучать математику») свидетельствует скорее об атеоретическом подходе к его разработке.

Авторы этой работы не имели намерения развернуть целенаправленную критику TIMSS в части этой конкретной шкалы. Важнее было привлечь внимание исследователей к проблеме, неоднократно уже озвученной в международном психометрическом сообществе, но, кажется, не замеченной сообществом российским. Это проблема слабой теоретической обоснованности личностных шкал и подмена теории статистическими корреляциями, которые не могут рассказать ничего о том, что же все-таки измеряется, почему и как это связано с полученными ответами. Этой работой авторы хотели показать, что аккуратная теоретическая проработка каждого утверждения может сделать психометрический анализ более осмысленным, результаты более интерпретируемыми, а весь инструмент более валидным.

Библиографический список / References

1. Капуза А.В., Тюменева Ю.А. Надежность и структура шкалы социальной желательности TALIS: оценка в рамках современной теории тестирования // Мониторинг общественного мнения: Экономические и социальные перемены. 2016. № 6. С. 14–29. [Kapuza A.V., Tyumeneva Yu.A. Reliability and structure of TALIS social desirability scale: Assessment within the framework of modern testing theory. *Monitoring of Public Opinion: Economic and Social Changes Journal (Public Opinion Monitoring)*. 2016. No. 6. Pp. 14–29. (In Russ.)]
2. Пиаже Ж. Эволюция интеллекта в подростковом и юношеском возрасте // Психологическая наука и образование. 1997. № 4. С. 56–64. [Piaget J. Evolution of intelligence in adolescence and adolescence. *Psychological Science and Education*. 1997. No. 4. Pp. 56–64. (In Russ.)]
3. Сайко Э.В., Фельдштейн Д.И. Подросток на дистанции взросления и в пространстве его определения // Мир психологии. 2007. № 4. С. 3–10. [Sayko E.V., Feldstein D.I. Teenager at a distance of growing up and in the space of its definition. *Mir psichologii*. 2007. No. 4. Pp. 3–10. (In Russ.)]
4. Bernstein R.M. The development of the self-system during adolescence. *Journal of Genetic Psychology*. 1980. No. 2 (136). Pp. 231–245.
5. Carlson R. Stability and change in the adolescent's self-image. *Child Development*. 1965. No. 3 (36). P. 659.
6. Deci E.L., Vallerand R.J., Pelletier L.G., Ryan R.M. Motivation and education: The self-determination perspective. *Educational Psychologist*. 1991. No. 3–4 (26). Pp. 325–346.

7. Deci E.L., Ryan R.M. A motivational approach to self: Integration in personality. *Nebraska Symposium on Motivation*. 1990. Vol. 38. Pp. 237–288.
8. DiStefano C., Motl R.W. Further investigating method effects associated with negatively worded items on self-report surveys. *Structural Equation Modeling*. 2006. No. 3 (13). Pp. 440–464.
9. Eccles J. et al. Age and gender differences in children's self- and task perceptions during elementary school. *Child Development*. 1993. No. 3 (64). Pp. 830–847.
10. Einfeld S.L., Tonge B.J. The developmental behavior checklist: The development and validation of an instrument to assess behavioral and emotional disturbance in children and adolescents with mental retardation. *Journal of Autism and Developmental Disorders*. 1995. No. 2 (25). Pp. 81–104.
11. Elkind D. Egocentrism in adolescence. *Child Development*. 1967. No. 4 (38). Pp. 1025.
12. Grote R.C. The complete guide to performance appraisal. New York, 1996.
13. Harter S. Issues in the assessment of the Self-concept of children and adolescents. *Scientific Research*. 1990. C. 292–325.
14. Harter S., Pike R. The pictorial scale of perceived competence and social acceptance for young children. *Child Development*. 1984. No. 6 (55). Pp. 1969.
15. Herzberger S.D. et al. A developmental study of social self-conceptions in adolescence: Impressions and misimpressions. *Merrill-Palmer Quarterly of Behavior and Development*. Vol. 27. Pp. 15–29.
16. Hidi S.A., Renninger K. The four-phase model of interest development. *Educational Psychologist*. 2006. No. 41 (2). Pp. 111–127.
17. Hoff K.A. et al. Normative changes in interests from adolescence to adulthood: A meta-analysis of longitudinal studies. *Psychological Bulletin*. 2018. No. 4 (144). Pp. 426–451.
18. Hooper M., Mullis I.V.S., Martin M.O. TIMSS 2015. Context Questionnaire Framework. *TIMSS 2015 Assessment Framework*. 2015. Pp. 61–83.
19. Kamoen N., Holleman B. Positive, negative and bipolar questions: The effect of question polarity on ratings of text readability. *Survey Research Methods*. 2013. No. 3 (7). Pp. 181–189.
20. Keller M.W., Brummel B.J., Streets V.N. et al. Professional competencies with behaviorally anchored ratings for graduate students. *ASEE Annual Conference and Exposition, Conference Proceedings*. June 2017. DOI: <https://doi.org/10.18260/1-2--28760>
21. Li C.H. Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*. 2016. No. 3 (48). Pp. 936–949.
22. Liem G.A.D., Martin A.J. The motivation and engagement scale: Theoretical framework, psychometric properties, and applied yields. *Australian Psychologist*. 2012. No. 1 (47). Pp. 3–13.
23. Marsh H.W. et al. Factorial, convergent and discriminant validity of TIMSS math and science motivation measures: A comparison of Arab and Anglo-Saxon countries. *Journal of Educational Psychology*. 2013. No. 1 (105). Pp. 108–128.
24. Matosas-López L., Leguey-Galán S., Doncel-Pedrerá L.M. Converting likert scales into behavioral anchored rating scales (BARS) for the evaluation of teaching effectiveness for formative purposes. *Journal of University Teaching and Learning Practice*. 2019. No. 3 (16). Pp. 1–26.

25. Mitchell M. Situational interest: its multifaceted structure in the secondary school mathematics classroom. *Journal of Educational Psychology*. 1993. No. 3 (85). Pp. 424–436.
26. Moors G. The effect of response style bias on the measurement of transformational, transactional, and laissez-faire leadership. *European Journal of Work and Organizational Psychology*. 2012. No. 2 (21). Pp. 271–298.
27. Plieninger H. Mountain or molehill? A simulation study on the impact of response styles. *Educational and Psychological Measurement*. 2017. No. 1 (77). Pp. 32–53.
28. Rammstedt B., Farmer R.F. The impact of acquiescence on the evaluation of personality structure. *Psychological Assessment*. 2013. No. 4 (25). Pp. 1137–1145.
29. Rosenberg M. et al. Global self-esteem and specific self-esteem: Different concepts, different outcomes. *American Sociological Review*. 1995. No. 1 (60). Pp. 141.
30. Savin-Williams R.C., Demo D.H. Developmental change and stability in adolescent self-concept. *Developmental Psychology*. 1984. No. 6 (20). Pp. 1100–1110.
31. Schreiber J.B. et al. Reporting structural equation modeling and confirmatory factor analysis results: A review. *Journal of Educational Research*. 2006. Vol. 99. No. 6. Pp. 323–338.
32. Schwab D.P., Heneman H.G., DeCotiis T.A. Behaviorally anchored rating scales: A review of the literature. *Personnel Psychology*. 1975. No. 4 (28). Pp. 549–562.
33. Shapiro E.S. Behavioral assessment in school psychology. USA, 2014.
34. Simmons R.G., Rosenberg F., Rosenberg M. Disturbance in the self-image at adolescence. *American Sociological Review*. 1973. No. 5 (38). Pp. 553–568.
35. Smith P.C., Kendall L.M. Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. *Journal of Applied Psychology*. 1963. No. 2 (47). Pp. 149–155.
36. Weijters B., Baumgartner H. Misresponse to reversed and negated items in surveys: A review. *Journal of Marketing Research*. 2012. Vol. 49. No. 5. Pp. 737–747.
37. Wetzel E., Böhnke J.R., Rose N. A simulation study on methods of correcting for the effects of extreme response style. *Educational and Psychological Measurement*. 2016. No. 2 (76). Pp. 304–324.

Статья поступила в редакцию 12.03.2020, принята к публикации 18.05.2020

The article was received on 12.03.2020, accepted for publication 18.05.2020

Сведения об авторах / About the Authors

Тюменева Юлия Алексеевна – кандидат психологических наук; старший научный сотрудник Международной лаборатории оценки практик и инноваций в образовании Института образования, Национальный исследовательский университет «Высшая школа экономики», г. Москва

Yulia A. Tyumeneva – PhD in Psychology; senior research fellow at the International Laboratory for Evaluation of Practices and Innovations in Education of the Institute of Education, National Research University “Higher School of Economics”, Moscow

E-mail: jtiumeneva@hse.ru, jutu@yandex.ru

Обухов Алексей Сергеевич – кандидат психологических наук, доцент; ведущий эксперт Центра исследований современного детства Института образования, Национальный исследовательский университет «Высшая школа экономики»

Aleksey S. Obukhov – PhD in Psychology; leading expert at the Research Center of Modern Childhood of the Institute of Education, National Research University “Higher School of Economics”, Moscow

E-mail: ao@redu.ru

Финогенова Ольга Николаевна – кандидат психологических наук, доцент; доцент кафедры общей и социальной педагогики Института педагогики, психологии и социологии, Сибирский федеральный университет, г. Красноярск

Olga N. Finogenova – PhD in Psychology; associate professor at the Department of General and Social Pedagogy of the Institute of Pedagogy, Psychology and Sociology, Siberian Federal University, Krasnoyarsk

E-mail: finogenova1971@mail.ru

Заявленный вклад авторов

Тюменева Ю.А. – общее руководство направлением исследования, планирование исследования, анализ данных, подготовка текста статьи.

Обухов А.С. – организация и участие в проведении полевого исследования, участие в подготовке текста статьи.

Финогенова О.Н. – организация и участие в проведении полевого исследования, участие в подготовке текста статьи.

Contribution of the authors

Yu.A. Tyumeneva – general direction of the research, planning of the research, analysis of data, the preparation of the text of the article

A.S. Obukhov – organization and participation in the field gatherings and research, participation in the preparation of the text of the article.

O.N. Finogenova – organization and participation in the field gatherings and research, participation in the preparation of the text of the article.