

Корчажкина О.М.

Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, г. Москва, Россия

О ВЕРОЯТНОСТНОМ И НЕЧЁТКОМ СПОСОБАХ ОЦЕНИВАНИЯ МЕТАПРЕДМЕТНЫХ КОМПЕТЕНЦИЙ УЧАЩИХСЯ*

АННОТАЦИЯ

В статье приводится сопоставительный анализ двух способов оценивания латентных образовательных достижений учащихся, основанных на факторной теории оценки (Item Response Theory – IRT) и теории нечётких множеств. Проведённое исследование показало, что нечёткая модель применительно к достижению метапредметных образовательных результатов имеет больший практический потенциал, чем способ IRT, опирающийся на эмпирико-статистические методы, которые используются при вероятностном анализе больших массивов слабоструктурированных данных.

КЛЮЧЕВЫЕ СЛОВА

Латентные достижения, факторная теория оценки, теория нечётких множеств, метапредметные компетенции, образовательные результаты, эмпирико-статистические методы исследования, большие данные.

Olga Korchazhkina

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, Moscow, Russia

ON PROBABILITY AND FUZZY METHODS TO ASSESS STUDENTS' METASUBJECT COMPETENCES

ABSTRACT

The article offers a comparative analysis of two ways of how to assess students' latent achievements based on the Item Response Theory (IRT) and the Theory of Fuzzy Sets. The research shows that the fuzzy model implemented to the assessment of students' metasubject educational results has a greater practical potential than the IRT basis using empirical and statistical methods, which require analysing big data.

KEYWORDS

Latent achievements, Item Response Theory, Fuzzy Set Theory, metasubject competencies, educational results, empirical and statistical research methods, big data.

Вопросы качества педагогических изменений выходят на новый уровень, когда ставится задача оценивания метапредметных компетенций учащихся, или, выражаясь языком Федерального государственного образовательного стандарта (ФГОС) общего образования второго поколения, оценки достижения планируемых метапредметных образовательных результатов при освоении учащимися основной образовательной программы. Очевидно, что метапредметные компетенции представляют собой латентные (не поддающимися непосредственному измерению) характеристики учебно-познавательной деятельности, поскольку они как бы растворены и незримо присутствуют в различных формах этой деятельности, а потому их мониторинг и оценка сопряжены с определёнными трудностями. Решению этих проблем, основными из которых являются определение *содержания метапредметных компетенций*, обсуждение выбора *объектов оценки*, обоснование выбора *формы представления результатов оценки* и *способов оценивания*, посвящена настоящая статья. Причём повышенное внимание предполагается уделить обсуждению выбора **способа оценивания**, поскольку этот вопрос, являясь наиболее дискуссионным в области методологии педагогических измерений, напрямую влияет на «определение сущности, цели, задач, обоснование методов педагогических измерений, шкалирования и правил интерпретации

* Труды XI Международной научно-практической конференции «Современные информационные технологии и ИТ-образование» (SITITO'2016), Москва, Россия, 25-26 ноября, 2016

результатов» [1].

Содержание метапредметных компетенций. Под метапредметными компетенциями ФГОС понимает способность учащихся планировать и использовать в учебной, познавательной и социальной практике межпредметные понятия и универсальные учебные действия (УУД) [11, с. 8].

Межпредметные понятия, или метапредметные знания, есть знания о знаниях и способах их получения, например: система, факт, закономерность, феномен, анализ, синтез и пр. Они не являются целью изучения большинства учебных предметов, где усваиваются понятия, подчинённые межпредметным и соподчинённые между собой [5, с. 216], а помогают формировать УУД, которые представляют собой минимальные содержательные элементы метапредметных компетенций и поэтому выступают в качестве основных деятельностных механизмов метапредметного содержания образования. Без УУД невозможно осуществление учебно-познавательной деятельности [9, с. 22], поскольку они определяются как обобщенные действия, которые открывают возможность широкой ориентации учащихся в различных предметных областях и в строении самой учебной деятельности, включая осознание её целевой направленности, ценностно-смысловых и операциональных (деятельностных) характеристик [11, с. 8].

Выбор объектов оценки. Перевод скрытых качеств в непосредственно наблюдаемые напрямую связан с проблемой формализации. Объекты оценки метапредметных компетенций могут быть выделены только в том случае, если удастся нивелировать их латентные свойства, то есть выразить в явной форме посредством содержательной декомпозиции в виде более мелких составляющих и описать их сущность с помощью естественного языка или представить в числовом выражении. Такими структурными единицами являются, как было отмечено выше, межпредметные понятия, или метапредметные знания, и УУД.

Межпредметные понятия, которые являются знаниевыми компонентами метапредметных компетенций, не имеет смысла рассматривать в качестве отдельных объектов оценки, поскольку они напрямую связаны с УУД и проявляются в их реализации, подтверждая тем самым общеизвестный тезис: кто способен правильно понимать, тот способен правильно действовать.

Поэтому единственными объектами оценки метапредметных компетенций целесообразно выбрать УУД, которые могут быть вербализованы с помощью глаголов мыслительных операций, совершаемых учащимися при решении учебно-познавательных задач. Мыслительные операции проявляются в рамках познавательных, коммуникативных и регулятивных УУД, которые могут быть чётко расписаны и привязаны к различным видам учебно-познавательных заданий. Это означает, что для оценки метапредметных образовательных результатов необходимо определить, какие УУД привлекает учащийся при выполнении конкретного учебно-познавательного задания – какого качества и в каком объёме. Чтобы поставить в соответствие учебной задаче набор УУД, необходимых для её выполнения, можно воспользоваться классификаций учебно-познавательных компетенций учащихся [6], в котором УУД представлены в виде глаголов мыслительных операций, классифицированных по трём группам – познавательных, коммуникативных и регулятивных УУД.

Если учитель умеет соотносить учебно-познавательные задания с УУД, то это даёт ему возможность сделать подобный подход практико-ориентированным и позволяет осуществлять оценку метапредметных компетенций непосредственно в процессе предметной учебно-познавательной деятельности на каждом уроке, не отводя этому дополнительное время.

Выбор формы представления результатов оценки. Поскольку оценка метапредметных образовательных результатов не может выражаться в виде «процентов» и «показателей», так как является видом экспертной оценки, естественно представлять эти результаты не в точном количественном выражении, а в терминах лексической неопределённости, значениями которой являются не числа, а имена, то есть в терминах *высокий, средний, низкий* уровень с использованием уточняющих наречий *весьма, довольно, определённо, скорее* и пр. [8, с. 204, 208]. То есть оценка сформированности УУД в большей мере субъективна и должна находиться в диапазоне с неопределёнными границами.

Внутри диапазона с неопределёнными границами учитывается как сложность задания, так и более детальные критерии, зависящие от вида заданий, уровня сформированности компетенций учащихся, субъективных требований эксперта и прочих латентных факторов.

Строго говоря, в лингвистических формулировках оценки метапредметных компетенций достаточно учесть два вида неопределённости: 1) неопределённость оценки эксперта, принимающая во внимание субъективные условия осуществления оценочных действий, которая выражает его предпочтения и показывает степень его уверенности в вынесенном решении, то есть достоверность его оценки, и 2) неопределённость, обусловленная неточными формулировками требований к различным видам метапредметных компетенций в соответствии с Кодификатором,

подразделение их на различные уровни по степени сложности и представленность этих показателей в предметных заданиях.

Эти допущения позволяют конкретизировать поставленную задачу: она состоит в том, чтобы **предложить способ оценки метапредметных образовательных результатов на основе учёта неопределённости экспертной оценки и неопределённости, возникающей в результате причисления выполняемого учебно-познавательного задания к тому или иному уровню сложности**. Задачу формализации двух видов неопределённости можно решить, установив не прямое соответствие между балльной системой оценок предметных образовательных результатов и формализованной шкалой оценки метапредметных образовательных результатов. Это позволит установить форму представления результатов оценки определённого набора УУД, которым сопровождается каждое задание. Вопрос состоит в том, **по каким законам должно устанавливаться это не прямое соответствие, или какой способ формализации неопределённости выбрать?**

Как отмечается в [13, с. 74-76], неопределённость может выражаться в виде вероятности и в виде нечёткости: стохастическая (вероятностная) неопределённость «означает неопределённость появления события, которое является само по себе точно описанным», а лексическая неопределённость, или нечёткость, «означает неопределённость в описании события». В [4, с. 9] отмечается, что на основе реально наблюдаемых данных ставится задача воссоздания непосредственно ненаблюдаемого качества, измеряемого с помощью модели. Какую выбрать модель – вероятностную или нечёткую – для того, чтобы перевести скрытые качества в непосредственно наблюдаемые?

Рассмотрим оба способа формализации неопределённости – вероятностный и нечёткий.

Вероятностная модель оценивания на основе факторной теории оценки IRT

Вероятностный способ формализации неопределённости базируется на факторной теории оценки (Item Response Theory – IRT) [12], которую называют также вероятностной или математической теорией измерений [1-4], тогда как дословный перевод IRT можно представить как «теория оценки по результату выполнения задания». Эта теория базируется на вероятностных характеристиках, получаемых в результате обработки обширных массивов данных при тестировании достаточно большой выборки испытуемых. Отметим, что задания предлагаются испытуемым только в тестовом формате. Кроме того, оценки, которые получают испытуемые, демонстрируют их знания в определённой предметной области, то есть являются показателями предметных образовательных результатов.

Исторически факторная теория оценки явилась переложением на область психолого-педагогических измерений теории латентно-структурного анализа (LSA – Latent Structure Analysis), согласно которому на большой выборке испытуемых выявляются скрытые поведенческие факторы и осуществляется прогноз их проявления с помощью вероятностно-статистических моделей. Сама факторная теория оценки раскрывает стохастические (вероятностные) закономерности точно описанного события в виде ответов учащихся на определённые задания. В каждом задании вероятность правильного ответа испытуемых должна расти по мере повышения уровня подготовленности испытуемых. Эти закономерности описываются с помощью функции математического ожидания, полученной по результатам эмпирической апробации и изображённой на рис. 1 (цит. по [4]), где θ – уровень подготовленности испытуемых, или уровень их знаний, а $P(\theta)$ – вероятность успешно выполнить задание.

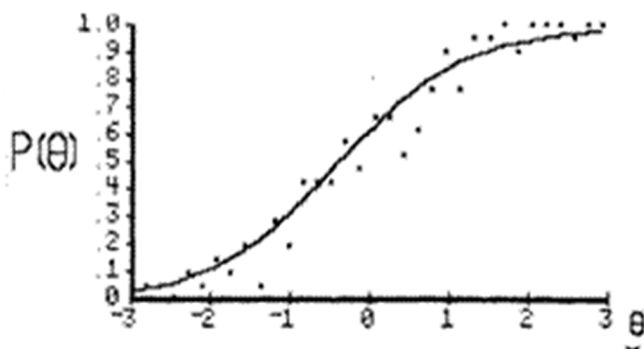


Рис. 1. Подбор графического образа задания по результатам его эмпирической апробации

Очевидно, что вероятность успешного выполнения одного и того же задания для испытуемых с более высоким уровнем знаний выше, чем для испытуемых с низким уровнем

подготовки. Функция $P(\theta)$, построенная путём аппроксимации набора эмпирических данных, представляет собой логистическую функцию экспоненциального типа

$$y(x) = \frac{e^x}{1+e^x}, \quad (1)$$

дошедшую до нас из глубины XIX века, когда её начали использовать для математического прогнозирования биологических изменений в природе. Эта функция представляет собой S-образную кривую, моделирующую рост вероятности наступления некоего события по мере изменения управляющих параметров – увеличения факторов риска, нарастания противоречий, напряжённости, скорости, техногенных изменений и пр.

Теоретические исследования и практическая апробация показали (см. ссылки на литературу в [1-4]), что логистическая функция вида (1) наилучшим образом (то есть с минимальными отклонениями) отражает законы, которым подчиняется не только зависимость вероятности успешно выполнить задание $P(\theta)$ от уровня подготовки испытуемых θ (см. рис. 1), но и зависимость вероятности успешно выполнить задание $P(\beta)$ от сложности самого задания β . На рис. 2 представлены графики зависимости вероятности $P(\theta, \beta)$ выполнения заданий разной степени сложности β_1 и β_2 от уровня подготовки испытуемых θ .



Рис. 2. Зависимость вероятности выполнения двух заданий разной степени сложности β_1 и β_2 от уровня подготовки испытуемых θ

Очевидно, что чем сложнее задание, тем более пологим является график функции $P(\theta, \beta)$.

Эти эмпирические кривые послужили основой математического аппарата теории вероятностей, на которой основан подход IRT.

Но является ли вероятностный подход наиболее адекватно отражающим характеристики латентных процессов, которые сопровождают оценку метапредметных образовательных результатов? Или существуют другие, более эффективные методы оценивания? Для ответа на этот вопрос рассмотрим возможности математического аппарата нечёткой логики [13-15], которая также имеет дело с латентными параметрами и оперирует понятием неопределённости.

Модель оценивания на основе математического аппарата нечёткой логики

Метапредметные компетенции, как уже говорилось, являются латентными характеристиками способности учащихся осуществлять учебно-познавательную (предметную и метапредметную) деятельность, а их оценка сопряжена с двумя формами неопределённости: неопределённости объектов оценки и неопределённости экспертной оценки этих объектов. В нечёткой логике эта неопределённость выражается коэффициентами достоверности, которые определяются через функцию принадлежности $y = \mu(x)$, где x – аргумент, выражающий некоторое числовое множество.

Неопределённость объектов оценки выражается в нечётких лингвистических формулировках УУД по Кодификатору, которые должны быть сгруппированы по уровню их сложности при выполнении учебно-познавательных задач. Для каждого уровня сложности заданий выбирается диапазон чётких оценок (баллов) как область определения соответствующей функции принадлежности, которая входит в зону нечёткости. Соответствующая функция принадлежности показывает, насколько достоверной является причисление некоторой группы УУД, необходимых для выполнения того или иного задания, к выбранному интервалу сложности, что демонстрирует степень уверенности учителя/эксперта в правильности принятой градации балльных оценок на соответствующих интервалах сложности заданий [8, с. 207-208].

Функция принадлежности объектов оценки (УУД) является мерой достоверности, показывающей, насколько валидными являются выбранные задания для оценки того или иного набора УУД. Если исходить из классического определения валидности, то это «мера соответствия

методик и результатов исследования поставленным задачам, адекватность интерпретации результатов выполнения задания по отношению к цели его выполнения, показатель меры соответствия задания цели его выполнения, показатель задания, характеризующий точность измерения исследуемого свойства, оценка адекватности задания исследуемой проблеме» (согласно <http://btimes.ru/dictionary/validnost>). В нашем случае под валидностью понимается мера соответствия, или правомерность использования заданий, которые выбраны в зависимости от конкретных образовательных задач и условий обучения, для измерения метапредметных компетенций, представленных в виде набора соответствующих УУД.

Неопределённость экспертной оценки выражается в нечётких вербальных оценках эксперта типа *скорее низкий (чем средний)* уровень компетенций, *довольно высокий* уровень компетенций и т. п., которые коррелируются с чёткими балльными оценками в соответствующих интервалах. А отвечающая им функция принадлежности показывает, насколько достоверной является вынесенная экспертная оценка [7, с. 111].

Одной из самых сложных процедур при построении нечёткой модели оценивания метапредметных образовательных результатов является процесс формализации двух форм лексической неопределённости – субъективности экспертной оценки, выраженной в нечётких суждениях, и неопределённости объекта оценки, которая представлена нечёткими лингвистическими формулировками УУД по Кодификатору:

1. Вербальная формализация метапредметных компетенций в виде словесных формулировок, соответствующих УУД, отражённых в Кодификаторе элементов содержания учебно-познавательных компетенций учащихся [6];

2. Математическая формализация уровня сложности заданий, для выполнения которых требуется владение метапредметными компетенциями, в виде соответствующих функций принадлежности $\mu(k)$, где k – потенциальная оценка, которую может получить учащийся за выполненное задание;

3. Математическая формализация субъективной оценки эксперта в виде соответствующих функций принадлежности $\mu(a)$, где a – реальная оценка, выставленная экспертом;

4. Комбинированная формализация итоговой оценки метапредметных образовательных результатов через мультипликацию функций принадлежности $\mu(k)$ и $\mu(a)$ в единой области их определения. Это позволяет в рамках одного диапазона качественных значений (высокий, средний, низкий уровень), производить более детальное различие в соответствии с субъективными требованиями эксперта и уровнем сложности заданий, для которых требуется владение измеряемыми компетенциями.

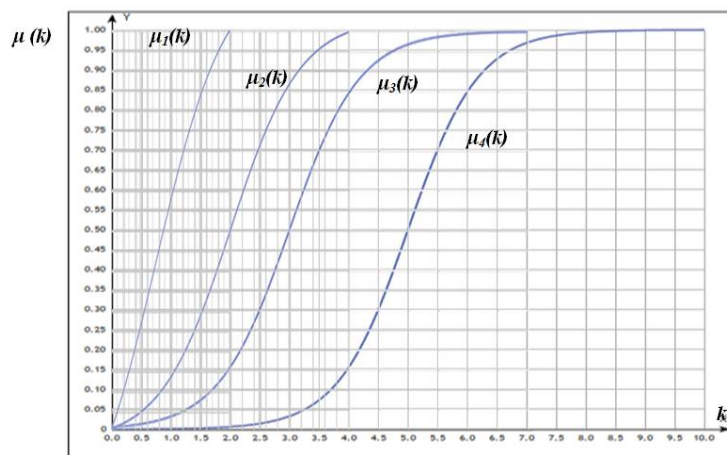


Рис. 3. Зависимость функции принадлежности $\mu(k)$, определяющей степень соответствия наборов УУД заданиям четырёх интервалов сложности от потенциальных оценок в интервале от 0 до 10

Математическое описание нечётких характеристик на примере треугольных и трапециевидальных функций принадлежности $\mu(k)$ и $\mu(a)$ подробно представлено в работах [7; 8]. Однако выбор трапециевидальных и тем более треугольных функций принадлежности не позволяет получить достоверный результат, поскольку вносит весьма значительную погрешность в процедуру оценки. Опора на эмпирические данные, предоставленные IRT, позволили выбрать в качестве функций принадлежности экспоненциальные логистические кривые, основой которых является формула (1), корректируемая весовыми коэффициентами и линейными сдвигами по осям абсцисс и ординат.

Так функции принадлежности $\mu(k)$ на рис. 3 показывают, насколько достоверно выбранный набор УУД соответствует уровню сложности конкретного задания. Здесь k – потенциальная оценка, которую может получить учащийся за выполнение задания: чем сложнее задание, тем выше может быть максимальная оценка и тем более пологой является логистическая функция. Можно также сказать, что эти функции принадлежности демонстрируют степень уверенности учителя/эксперта в правильности принятой градации балльных оценок на соответствующих интервалах сложности заданий.

Каждая из функций принадлежности на рис. 3 описывается экспоненциальной логистической функцией в соответствующем интервале оценок, которые может получить учащийся за выполнение задания той степени сложности, которая укладывается в один из выбранных интервалов – И1, И2, И3 или И4 (по мере роста сложности заданий):

1. Интервал сложности заданий 1 (И1): $\mu_1(k) = \frac{0.72(2e^{k-1}-1)}{1+2e^{k-1}} + 0.5$ (оценки 0-2);
2. Интервал сложности заданий 2 (И2): $\mu_2(k) = \frac{0.53(2e^{k-2}-1)}{1+2e^{k-2}} + 0.5$ (оценки 0-4);
3. Интервал сложности заданий 3 (И3): $\mu_3(k) = \frac{2e^{k-3}-1}{1+2e^{k-3}} + 0.5$ (оценки 0 – 7);
4. Интервал сложности заданий 4 (И4): $\mu_4(k) = \frac{2e^{k-5}-1}{1+2e^{k-5}} + 0.5$ (оценки 0 – 10).

В соответствии с теорией нечётких множеств лингвистические формулировки уровней оценки метапредметных компетенций соотносятся со значениями выбранных функций принадлежности $\mu(a)$, задающих степень уверенности эксперта в принятом решении (рис. 4). Это означает, что явная неопределённость, присутствующая в любом экспертном оценивании и затрудняющая применение точных количественных методов и подходов, а также снижающая уверенность в получении достоверных результатов, приводит к необходимости определения достоверности принимаемых экспертом решений, что и делает функция принадлежности $\mu(a)$. Аргументы a – это множество оценок, присваиваемых экспертом учащемуся за владение УУД.

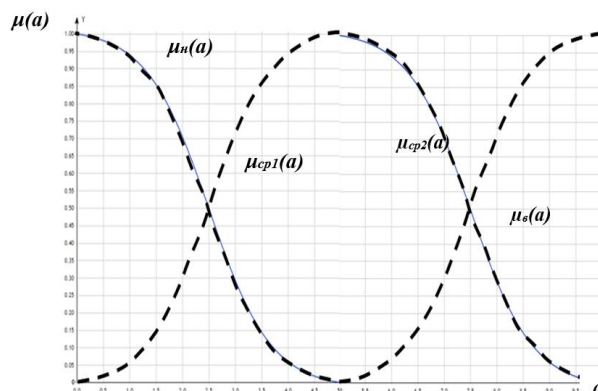


Рис. 4. Функции принадлежности $\mu(a)$ экспертных оценок метапредметных образовательных результатов для низкого, среднего и высокого уровня развития компетенций

Функции принадлежности $\mu(a)$ в пределах четырёх интервалов оценок эксперта a в соответствии с уровнями развития метапредметных компетенций выглядят следующим образом:

1. Низкий уровень развития метапредметных компетенций:

$$\mu_n(a) = \frac{0.515(2e^{-a+2.5}-1)}{1+2e^{-a+2.5}} + 0.5 \text{ (оценки 0-5);}$$

2. Средний уровень 1 развития метапредметных компетенций:

$$\mu_{ср1}(a) = \frac{0.515(2e^{a-2.5}-1)}{1+2e^{a-2.5}} + 0.5 \text{ (оценки 0-5);}$$

3. Средний уровень 2 развития метапредметных компетенций:

$$\mu_{ср2}(a) = \frac{0.515(2e^{-a+7.5}-1)}{1+2e^{-a+7.5}} + 0.5 \text{ (оценки 5-10);}$$

4. Высокий уровень развития метапредметных компетенций:

$$\mu_v(a) = \frac{0.515(2e^{a-7.5}-1)}{1+2e^{a-7.5}} + 0.5 \text{ (оценки 5-10).}$$

В [8, с. 211] было предложено установить градацию экспертных оценок как нечётких суждений с тремя обобщёнными уровнями компетенций – нижним, средним и высоким, каждый из которых имеет деление на три дополнительных подуровня в соответствии с наречиями и

наречными выражениями *определённо, довольно, весьма, более или менее, скорее, крайне* и пр. На основе этой модели в [7] был предложен алгоритм нечёткого автомата оценки метапредметных компетенций учащихся в зависимости от уровня сложности выполненных заданий и оценок, присвоенных экспертами за это задание. Алгоритм реализован в среде *Excel* и основан на наложении (мультипликации) функций принадлежности $\mu(k)$ и $\mu(a)$ (рис. 5). Значения результирующей функции принадлежности в общей области определения (области оценок от 0 до 10) считаются надёжными, если они находятся в интервале уверенной достоверности: $0,625 < (\mu(a) \times \mu_n(k)) < 1,0$.

Результат, получаемый с помощью нечёткого автомата, представляется в двух форматах: либо с помощью нечёткой формулировки *низкий/средний/высокий* уровень компетенций, сопровождаемой коэффициентом достоверности, либо только с помощью нечёткой формулировки *определённо/более или менее/скорее низкий/средний/высокий* уровень компетенций без коэффициента достоверности (рис. 6).

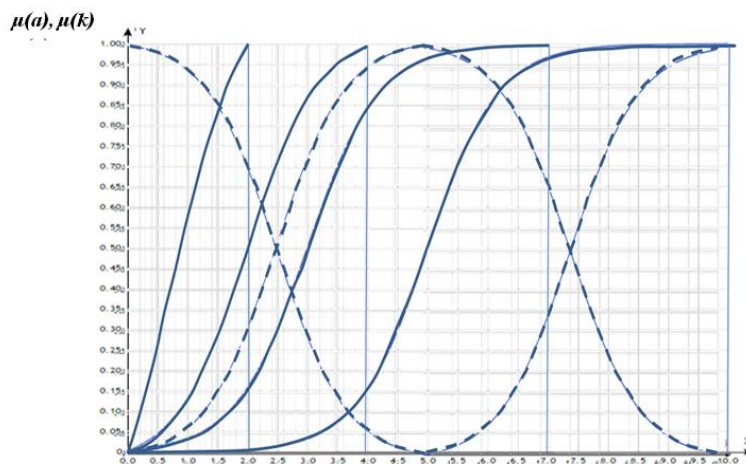


Рис. 5. Наложение функций принадлежности $\mu(k)$ (сплошные линии) и $\mu(a)$ (пунктир) в интервале единой шкалы оценок 0-10 для определения результирующей оценки, учитывающей степень достоверности соответствия наборов УУД заданиям $\mu(k)$ и степень достоверности оценки эксперта $\mu(a)$

	A	B	C	D	E	F
	Фамилия Имя	уровень сложности задания	оценка	результат		
1						
2	ф1	3	6	довольно средний	средний 0.86	
3	ф2	2	4	определенно средний	средний 1.0	
4	ф3	4	9	довольно высокий	высокий 0.9	
5	ф4	4	10	определенно высокий	высокий 1.0	
6	ф5	3	5	скорее средний	средний 0.71	
7	ф6	3	6	довольно средний	средний 0.86	
8	ф7	3	7	скорее средний	средний 0.75	
9	ф8	1	1	возможно низкий	низкий 0.5	
10	ф9	3	4	определенно средний	средний 1.0	

Рис. 6. Интерфейс нечёткого автомата, выполненного с помощью приложения *Excel*, с результатами оценки УУД (метапредметных компетенций) в интервалах сложности заданий И1, И2, И3 и И4

Основные отличия рассматриваемых способов оценки компетенций учащихся – с помощью вероятностной и нечёткой математической модели – сведены в следующую таблицу:

Основные отличия модели на базе IRT от нечёткой модели

	Модель IRT	Нечёткая модель

1. Ведущая идея	обоснование возможности прогнозирования результатов выполнения тестовых заданий различной степени сложности (определение вероятности правильного ответа)	измерение уровня сформированности метапредметных компетенций при выполнении заданий разной степени сложности через привлекаемые УУД
2. Цель	предназначена для вероятностного прогнозирования успешности выполнения тестовых заданий	предназначена для экспертного оценивания выполнения конкретных учебно-познавательных заданий
3. Объект измерения/контроля	предметные знания	метапредметные компетенции
4. Методика измерения и базис математического описания	вероятностное прогнозирование (эмпирическая модель) основана на эмпирических (опытных, усреднённых) данных	нечёткая оценочная модель базируется на априорных данных, основанных на валидности задания (соответствия уровня его сложности уровню сложности привлекаемых для его выполнения УУД, а также предметных и метапредметных знаний)
5. Неопределённость и способы её формализации	визуализация формальных характеристик тестовых заданий	вербальная и математическая визуализация содержательных характеристик учебно-познавательных заданий
6. Число испытуемых	надёжность обеспечивается при большом числе испытуемых	надёжность обеспечивается при любом числе учащихся
7. Типы заданий, используемые при измерении	тестирование (задания в виде тестов)	задания и комплекты заданий в различных форматах
8. Степень и мера сложности заданий	дифференцирующая способность задания (мера крутизны логистической кривой зависимости вероятности успешного выполнения задания от уровня подготовленности испытуемого) определяются по четырём формальным признакам [4, с. 12-13]: 1) доля неправильных ответов испытуемых на каждое задание проектируемого теста; 2) потенциал сложности задания выражен отношением доли неправильных ответов на каждое задание проектируемого теста к вероятности успешного выполнения задания; 3) единая шкала уровня трудности задания и уровня подготовленности испытуемых (как натуральный логарифм от п. 2); 4) параметр сложности задания (окончательная мера сложности задания), корректируемые в процессе шкалирования значения из п.3.	дифференцирующая способность задания (мера крутизны логистической кривой зависимости функции принадлежности в определённом числовом интервале сложности заданий соответствующих метапредметных компетенций по Кодификатору) определяется по степени сложности метапредметных компетенций и предметных знаний, которые необходимы для выполнения задания [7; 8]: 1) доля неправильных ответов испытуемых на каждое задание проектируемого теста не играет существенной роли, поскольку мерой успешности выполнения задания является функция принадлежности результирующей логистической кривой как мера достоверности принятого решения; 2) потенциал сложности задания выражен функцией принадлежности в соответствующем интервале сложности заданий соответствующих метапредметных компетенций по Кодификатору демонстрируют степень уверенности эксперта в правильности принятой градации балльных оценок на соответствующих интервалах сложности заданий; 3) единая балльная шкала, объединяющая баллы сложности задания и баллы как оценки эксперта за

		выполнение этого задания; 4) параметры сложности задания выступают в виде коэффициента достоверности в зависимости от балла в четырёх интервалах сложности и задаются соответствующей функцией принадлежности.
9. Математическое описание	свойства заданий описываются с помощью «характеристических кривых заданий» (Item Characteristic Curves)	свойства заданий описываются с помощью функций принадлежности
10. Процедура шкалирования	шкалирование осуществляется после тестирования и фактически является его результатом	шкалирование осуществляется до процедуры оценки и задаёт общую область определения для двух функций принадлежности

Выводы. Проведя краткий сопоставительный анализ факторной теории оценки IRT и теории нечётких множеств применительно к оцениванию планируемых метапредметных достижений учащихся, можно обозначить следующие ограничения инструментов вероятностного прогнозирования, предоставляемых IRT, по сравнению с возможностями математического аппарата нечёткой логики:

- объектом оценки в IRT являются, как правило, **предметные знания**, проверяемые с помощью **процедуры тестирования**. Тогда как в реальном учебном процессе перед учителем необходимо возникает проблема оценить степень достижения **планируемых метапредметных образовательных результатов** в процессе решения учебно-познавательных **задач любых форматов**, в том числе и с открытым ответом;

- метапредметные образовательные результаты представляют собой **неопределённо описываемые события**, вероятность появления которых не важна или заведомо известна. Преодолеть неопределённость описания события можно путём задания функций принадлежности, определяющих степень **достоверности наступления события**. Метод IRT позволяет преодолеть **неопределённость точно описанного события** через определение **вероятности его наступления** путём задания соответствующей функции распределения;

- нечёткая модель надёжно работает на **произвольном числе учащихся** – от одного до группы учащихся, учеников всего класса или параллели. Тогда как надёжность метода IRT обеспечивается только при **достаточно большой** – до нескольких сотен – **выборке испытуемых**;

- метапредметные образовательные результаты в силу неопределённости описания требуют **высокой степени формализации**, которая может быть достигнута путём их декомпозиции на УУД и дальнейшем выражении через вербальные формулировки в виде глаголов мыслительных операций, что позволяет в максимальной степени учесть нюансы и детали при нивелировании неопределённости. В методе IRT формализация характеристик объектов оценки осуществляется через шкалирование.

Таким образом, нечёткая модель, применяемая для оценки овладения универсальными учебными действиями и основанная на критериях достоверности, гораздо в большей степени отвечает природе неопределённости, заложенной в метапредметных компетенциях, и неопределённости экспертной оценки этих компетенций, чем математическая теория измерений IRT, базирующаяся на вероятностных характеристиках, с помощью которых осуществляется прогнозирование результатов тестирования больших масс испытуемых.

Литература

1. Аванесов В.С. Вопросы методологии педагогических измерений // Педагогические измерения 2005. №1. <http://testolog.narod.ru/EdMeasmt3.html>.
2. Аванесов В.С. Основные понятия и положения математической теории измерений (Item Response Theory). <http://gpo1006.keva.tusur.ru/node/4>.
3. Аванесов В.С. Проблема педагогического измерения латентных качеств <http://viperson.ru/articles/problem-pedagogicheskogo-izmereniya-latentnyh-kachestv>.
4. Аванесов В.С. Item Response Theory: основные понятия и положения. <http://testolog.narod.ru/Theory59.html>.
5. Иванова О.А. Междисциплинарные понятия и формирование универсальных учебных действий на уроках математики // Известия Российского государственного педагогического университета им. А.И. Герцена. 2013. Выпуск № 161. С. 215-219. <http://cyberleninka.ru/article/n/mezhpredmetnye-ponyatiya-i-formirovanie-universalnyh-uchebnyh-deystviy-pri-izuchenii-matematiki>.
6. Кодификатор элементов содержания учебно-познавательных компетенций учащихся. <https://sites.google.com/site/efficiencyolga/home/kodifikator-elementov-soderzania-ucebno-poznavatelnyh-kompetencij-ucasihsa>.
7. Корчажкина О.М. Измерение метапредметных образовательных результатов: постановка задачи моделирования нечёткого автомата / Современные информационные технологии и ИТ-образование. – М.: МГУ. Т. 1 (№ 11), 2015. – 638 с. С. 106-116.

8. Корчажкина О.М. Концепция оценки метапредметных образовательных результатов методами нечёткого моделирования / Современные информационные технологии и ИТ-образование [Электронный ресурс] / Сборник научных трудов VIII Международной научно-практической конференции / под ред. В.А. Сухомлина. – Москва: МГУ, 2013. – Т.2. – 357с. – 1 электрон. опт. диск (CD-ROM). С. 204-217.
9. Корчажкина О.М. Метапредметное содержание образования во ФГОС общего образования // Педагогика. 2016. № 2. С. 16-25.
10. Корчажкина О.М. Типология заданий для электронных учебников: укрупнять или разделять? / Труды Международного конгресса по информатике: информационные системы и технологии (CSIS'2016). г. Минск, Республика Беларусь. 24 – 27 октября 2016 г. (в печати).
11. Федеральный государственный образовательный стандарт среднего (полного) общего образования. / М-во образования и науки Рос. Федерации. – М.: Просвещение, 2013. – 63 с. (Стандарты второго поколения).
12. Mislevy, R.J. How Developments in Psychology and Technology Challenge Assessment. Доклад на заседании учебного комитета Национальных академий естественных, инженерных и медицинских наук США (The National Academies of Sciences, Engineering, and Medicine) <https://vimeo.com/136642747>.
13. Piegat, A. Fuzzy Modeling and Control. – Heidelberg: Physica-Verlag, 2001. – 728 p.
14. Zimmermann, H.-J. Fuzzy set theory // WIREs Comp Stat, John Wiley & Sons, Inc. Vol. 2, May/June. 2010. P. 317-332. http://www.mv.helsinki.fi/home/niskanen/zimmermann_review.pdf
15. Zimmermann, H.-J. Fuzzy set theory and its applications / H.-J. Zimmermann. – 4th ed. – 514 p. [http://kashanu.ac.ir/Files/Content/H_-J_%20Zimmermann%20\(auth_\)%20Fuzzy%20Set%20Theory%E2%80%94and%20Its%20Applications%20%202001.pdf](http://kashanu.ac.ir/Files/Content/H_-J_%20Zimmermann%20(auth_)%20Fuzzy%20Set%20Theory%E2%80%94and%20Its%20Applications%20%202001.pdf).

References

1. Avanesov V.S. Voprosyi metodologii pedagogicheskikh izmereniy // Pedagogicheskie izmereniya 2005. #1. <http://testolog.narod.ru/EdMeasmt3.html>.
2. Avanesov V.S. Osnovnyie ponyatiya i polozheniya matematicheskoy teorii izmereniy (Item Response Theory). <http://gpo1006.keva.tusur.ru/node/4>.
3. Avanesov V.S. Problema pedagogicheskogo izmereniya latentnykh kachestv <http://viperson.ru/articles/problema-pedagogicheskogo-izmereniya-latentnyh-kachestv>.
4. Avanesov V.S. Item Response Theory: osnovnyie ponyatiya i polozheniya. <http://testolog.narod.ru/Theory59.html>.
5. Ivanova O.A. Mezhpredmetnyie ponyatiya i formirovanie universalnykh uchebnykh deystviy na urokah matematiki // Izvestiya Rossiyskogo gosudarstvennogo pedagogicheskogo universiteta im. A.I. Gertsena. 2013. Vyipusk # 161. S. 215-219. <http://cyberleninka.ru/article/n/mezhpredmetnye-ponyatiya-i-formirovanie-universalnyh-uchebnyh-deystviy-pri-izuchenii-matematiki>.
6. Kodifikator elementov sodержaniya uchebno-poznavatelnykh kompetentsiy uchastshisya. <https://sites.google.com/site/efficiencyolga/home/kodifikator-elementov-soderzaniya-ucebno-poznavatelnyh-kompetencij-ucasihsa>.
7. Korchazhkina O.M. Izmerenie metapredmetnykh obrazovatelnykh rezultatov: postanovka zadachi modelirovaniya nechYotkogo avtomata / Sovremennyye informatsionnyie tehnologii i IT-obrazovanie. – М.: МГУ. Т. 1 (# 11), 2015. – 638 с. S. 106-116.
8. Korchazhkina O.M. Kontseptsiya otsenki metapredmetnykh obrazovatelnykh rezultatov metodami nechYotkogo modelirovaniya / Sovremennyye informatsionnyie tehnologii i IT-obrazovanie [Elektronnyiy resurs] / Sbornik nauchnykh trudov VIII Mezhdunarodnoy nauchno-prakticheskoy konferentsii / pod red. V.A. Suhomlina. – Moskva: МГУ, 2013. – Т.2. – 357с. – 1 электрон. опт. диск (SD-ROM). S. 204-217.
9. Korchazhkina O.M. Metapredmetnoe sodержanie obrazovaniya vo FGOS obshego obrazovaniya // Pedagogika. 2016. # 2. S. 16-25.
10. Korchazhkina O.M. Tipologiya zadaniy dlya elektronnykh uchebnikov: ukрупnyat ili razdelyat? / Trudy Mezhdunarodnogo kongressa po informatike: informatsionnyie sistemyi i tehnologii (CSIS'2016). g. Minsk, Respublika Belarus. 24 – 27 oktyabrya 2016 g. (v pechati).
11. Federalnyiy gosudarstvennyiy obrazovatelnyiy standart srednego (polnogo) obshego obrazovaniya. / M-vo obrazovaniya i nauki Ros. Federatsii. – М.: Prosveschenie, 2013. – 63 с. (Standartyi vtorogo pokoleniya).
12. Mislevy, R.J. How Developments in Psychology and Technology Challenge Assessment. Доклад на заседании учебного комитета Национальных академий естественных, инженерных и медицинских наук США (The National Academies of Sciences, Engineering, and Medicine) <https://vimeo.com/136642747>.
13. Piegat, A. Fuzzy Modeling and Control. – Heidelberg: Physica-Verlag, 2001. – 728 p.
14. Zimmermann, H.-J. Fuzzy set theory // WIREs Comp Stat, John Wiley & Sons, Inc. Vol. 2, May/June. 2010. P. 317-332. http://www.mv.helsinki.fi/home/niskanen/zimmermann_review.pdf
15. Zimmermann, H.-J. Fuzzy set theory and its applications / H.-J. Zimmermann. – 4th ed. – 514 p. [http://kashanu.ac.ir/Files/Content/H_-J_%20Zimmermann%20\(auth_\)%20Fuzzy%20Set%20Theory%E2%80%94and%20Its%20Applications%20%202001.pdf](http://kashanu.ac.ir/Files/Content/H_-J_%20Zimmermann%20(auth_)%20Fuzzy%20Set%20Theory%E2%80%94and%20Its%20Applications%20%202001.pdf).

Поступила 15.10.2016

Об авторе:

Корчажкина Ольга Максимовна, старший научный сотрудник лаборатории «Проблемы информатизации образования» Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, кандидат технических наук, olgakomax@gmail.com.