вания создается библиотека простейших конструкций транспортных средств различного назначения. Здесь применен метод каркасного моделирования. Путем замены основных узлов конструкции, у пользователя развивается нестандартное мышление. Предложены элементарные звенья и реально существующие узлы, при помощи которых пользователь может создавать различного рода конструкции.

На втором этапе также используется библиотека промежуточных звеньев, которую создает сам пользователь,. Использование различного рода промежуточных звеньев позволяет пользователю определить приемлемые массогабаритные показатели, создать требуемый эстетический вид конструкции.

Библиотека, создаваемая пользователем для выходного звена, наиболее обширна. Все определяется требованиями к разрабатываемой конструкции.

Таким образом, разработанная компьютерная модель исполнительного механизма позволяет определить перспективы и возможности применения волнового механизма в системах буровых установок. Предложенная модель технологична, имеет широкий диапазон регулирования развиваемого момента, содержит функции муфты предельного момента, что положительно сказывается на условиях эксплуатации бурильного оборудования.

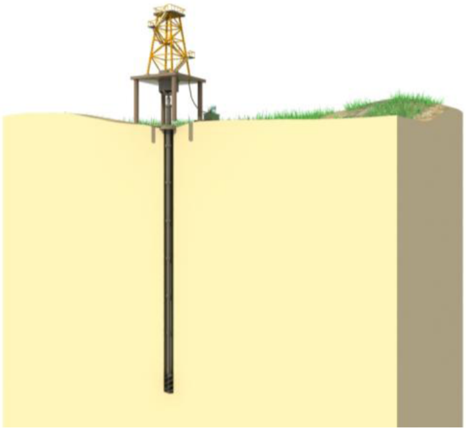

Рис. 3

Литература:

1. А. Н. Геращенко, С. Л. Самсонович Пневматические, гидравлические и электрические приводы летательных аппаратов на основе волновых исполнительных механизмов [Текст] / Геращенко А. Н. , Самсонович С. Л. М. : Машиностроение 2006 – 392с.

О СВЯЗИ МЕЖДУ СТОХАСТИЧЕСКОЙ ТЕОРИЕЙ ТЕСТОВ (ITEM RESPONSE THEORY, IRT) И ТЕОРИЕЙ ВЫБОРА ПО СХОДСТВУ (SIMILARITY CHOICE MODEL, SCM) В АСПЕКТЕ КОНСТРУИРОВАНИЯ СИСТЕМ МОНИТОРИНГА ЗНАНИЙ

Г.В. Иванов
Российский государственный педагогический университет им. А. И. Герцена.
Кафедра клинической психологии (г. Санкт-Петербург)

Применительно к педагогическим измерениям знаний, умений и навыков в системе профессионального образования каждый обучающийся может быть охарактеризован величинами уровней знаний по разным дисциплинам, оцениваемым в соответствии с теорией и практикой тестирования.

Вместе с тем, задача измерения может быть преобразована в оценку понимания ключевых концептов изучаемой дисциплины. В таком случае понятия-концепты описываются точками в семантическом пространстве субъекта [1]. Оси этого пространства – признаки, которые человек приписывает реальности, а расстояния между точками в этом пространстве отражает переживаемое человеком субъективное различие понятий – концептов.

I. Рассмотрим, сформулированную в рамках такого подхода теорию выбора по сходству [2, 3]. Каждое понятие предмета характеризуется координатами в субъективном семантическом пространстве и некой условной характеристикой, отражающей прочность фиксации предмета в памяти-W. Типовой эксперимент в рамках теории выбора – опознание букв. Испытуемому кратковременно многократно в сериях предъявляют букву, которую он прочитывает (верно, или ошибочно). Результаты реакций испытуемого фиксируются в матрице вероятностей его ответов, где строки матрицы – предъявленные буквы, столбцы матрицы – ответы испытуемого. Например в первой строке во втором столбце стоит вероятность того, что испытуемый при предъявлении буквы А отвечает что видел букву Б. Диагональ матрицы заполнена вероятностями правильных ответов (Ответ А при предъявлении А, ответ Б при предъявлении Б и т. д.).

Формула для вероятности ответа:

$$P(S_{\text{предъявл.}} \text{->} S_{\text{названный}}) = \frac{W_{\text{названный}} * e^{-c * D_{\text{между предъявл. и названным}}}}{\sum_{i=1}^{\text{число стимулов}} W_i * e^{-c * D_{\text{между предъявленным и i-тым}}}}$$

D – расстояние между образами в субъективном пространстве, с – коэффициент, характеризующий скорость убывания градиента генерализации для данного испытуемого. Он различен для разных людей, хотя на практике этим часто пренебрегают.

Полученные в психометрических экспериментах матрицы вероятностей ответов хорошо согласуются с данной теорией [3, 4].

Рассмотрим важное подмножество частных случаев, когда характеристики W одинаковы для любого образа. Тогда W в числителе и знаменателе сократится:

$$P(S_{\text{предъявл.}} \text{->} S_{\text{названный}}) = \frac{e^{-c * D_{\text{между предъявл. и названным}}}}{\sum_{i=1}^{\text{число стимулов}} e^{-c * D_{\text{между предъявленным и i-тым}}}}$$

II. Рассмотрим далее стохастическую логистическую модель Георга Раша, предназначенную для измерений уровня знаний, и являющуюся частью Стохастической Теории Тестов [5, 6, 7, 8].

Обычный прием для измерения уровня знаний – задавать учащемуся вопросы, на которые он отвечает. Задав, к примеру, 20 вопросов, подсчитывают, на сколько из них человек ответил правильно, и это число (*первичный балл экзаменуемого*) становится мерой уровня знаний. Если проэкзаменовать выборку студентов, можно для каждого задания вычислить, сколько человек на него ответили правильно. Это – первичный балл задания, мера трудности задания.

Однако, первичные баллы как средство измерения имеют несколько недостатков:

1)Вообразим 2 выборки студентов: 20 чел. хорошо успевающих студентов и и 20 чел. неуспевающих. Очевидно, что на один и тот же вопрос в первой выборке правильно ответит куда больше испытуемых, чем во второй. Получится, что задание для первой выборки имеет меньшую трудность, чем для второй.

2)Аналогичная ситуация с оценками величин уровня знаний студентов. Пусть есть 20 трудных (по оценке преподавателя) вопросов и 20 легких. Очевидно один и тот же студент в первом тесте наберет меньше первичных баллов, чем во втором. То есть по результатам первого теста студент знающий, а по результатам второго – незнающий.

Решение этих трудностей предполагается в рамках стохастической теории тестов. В рамках модели Раша для конкретного человека существует вероятность ответить на конкретный вопрос.

Вводится две величины: уровень знаний данного человека θ, и трудность данного задания δ. Вероятность того, что данный человек правильно ответит на данное задание, дается функцией:

$$P(\text{правильного ответа}) = \frac{e^{\theta}}{e^{\theta} + e^{\delta}}$$

Или эквивалентно:

$$P(\text{правильного ответа}) = \frac{1}{1 + e^{(\delta - \theta)}}$$

Зависимость вероятности решения конкретного задания от того, какой у человека уровень знаний называется *Характеристической кривой пункта*.

Сложность пункта δ соответствует такому уровню знаний у человека, при котором этот человек отвечает на пункт правильно с вероятностью 50%.

Другая характеристика кривой – крутизна или *дискриминативность* пункта. Она показывает, насколько быстро изменяется вероятность правильного ответа, по мере роста уровня знаний (например, когда человек учится, или когда мы заменяем менее знающего студента более знающим). Если крутизна кривой вопроса мала (пологая кривая), вероятности правильного ответа на вопрос у троечника Иванова и у отличника Сидорова не так уж и различаются. Следовательно, если студент правильно ответил на вопрос, это дает не много информации о том, троечник он или отличник.

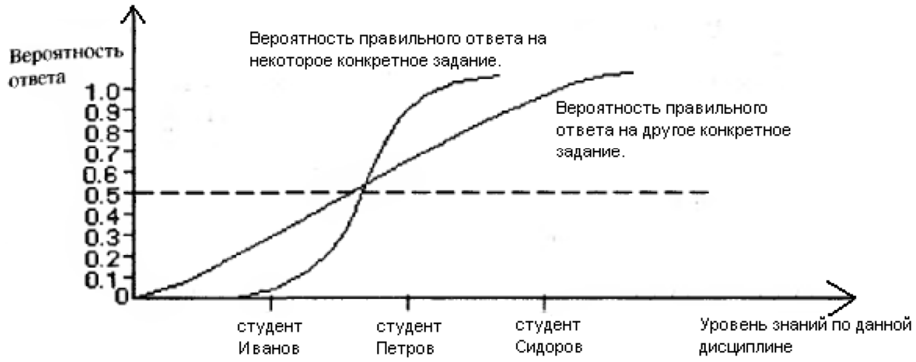

Рис. 1. Характеристические кривые двух пунктов с одинаковой сложностью и разной дискриминативностью (воспроизведено из [5], с изменениями)

Наоборот, для вопроса с высокой дискриминативностью вероятности правильного ответа на вопрос у троечника и отличника сильно различаются. Значит, если студент ответил на вопрос, можно почти наверняка утверждать что он имеет достаточно высокий уровень знаний. Если не ответил – почти наверняка его уровень знаний низок. Понятно, что вопросы с высокой дискриминативностью желательны.

В рамках модели Раша дискриминативность предполагается одинаковой у всех вопросов.

Основным фактором, снижающим дискриминативность вопросов по [7] является политематичность вопросов. Пример, приводимый данными авторами:

Есть 2 вопроса на знание математики:

1) Сколько аршинов в 3 саженях ?
2) Каков квадратный корень из 10 000 ?

Для ответа на 2й вопрос необходимы познания в математике. Для ответа на первый вопрос необходимо некоторое знание математики и знание того, сколько аршин содержится в одной сажени.

Джон может не знать, как умножать на три (низкий уровень знаний) и не знать, сколько аршин в одной сажени. Джон, скорее всего не ответит на вопрос правильно. Сьюзи может знать, как умножать на три (высокий уровень математических знаний), но не знать, сколько аршин в сажени. Сьюзи тоже, скорее всего не ответит на вопрос. Итак, величины уровня математических знаний сильно различаются, а вероятности правильного ответа на вопрос почти не отличаются. То есть ответ на политематичный вопрос не позволяет говорить о том, что уровень знаний по какой то из входящих в вопрос тем высок или низок.

IRT как частный случай SCM.

Пусть экзаменуемому задают вопрос с двумя вариантами ответа.

По Рашу:

$$P(\text{правильного ответа}) = \frac{e^{\theta}}{e^{\theta}+e^{\delta}}$$

Величины θ и δ измерены на интервальной шкале. Значит:

$\theta=\theta_{\text{истинное}}+\text{const}$

$\delta=\delta_{\text{истинное}}+\text{const}$

const может быть и таким, что $\theta_{\text{истинное}}$ и $\delta_{\text{истинное}}$ –отрицательные величины.

$$P(\text{правильного ответа}) = \frac{e^{\theta\text{ист}+\text{const}}}{e^{(\theta\text{ист}+\text{const})}+e^{(\delta\text{ист}+\text{const})}}$$

$$P(\text{правильного ответа}) = \frac{e^{\text{const}} * e^{-|\theta\text{ист}|}}{e^{\text{const}} * (e^{-|\theta\text{ист}|}+e^{-|\delta\text{ист}|})}$$

$$P(\text{правильного ответа}) = \frac{e^{-|\theta\text{ист}|}}{e^{-|\theta\text{ист}|}+e^{-|\delta\text{ист}|}}$$

Рассмотрим вопрос и варианты ответа на него в семантическом пространстве экзаменуемого (рис. 2).

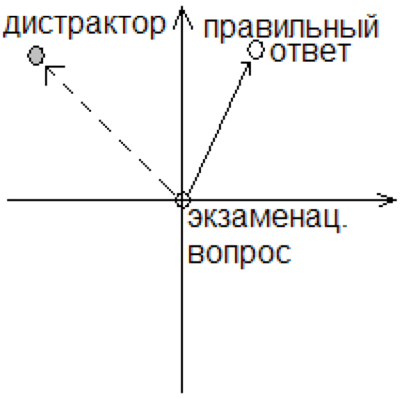

Рис. 2. Семантическое пространство. Начало координат совмещено с образом экзаменационного вопроса. Представлены также образ правильного ответа и образ дистрактора (неправильного варианта ответа)

Положив:

$$|\theta_{ист}| = c * D_{\text{между вопросом и правильным ответом}}$$

$$|\delta_{ист}| = c * D_{\text{между вопросом и дистрактором}}$$

Получим из уравнения IRT уравнение SCM для двух вариантов ответа с равными весами W правильного и неправильного вариантов.

Большее $|\theta_{ист}|$ (большее расстояние между вопросом и правильным ответом), означает меньшее $\theta = -|\theta_{истинное}| + const$ то есть меньший уровень знаний в терминах Раша.

Большее $|\delta_{ист}|$ (большее расстояние между вопросом и дистрактором), означает меньшее $\delta = |\delta_{истинное}| + const$ то есть меньшую трудность задания в терминах Раша.

Для корректного представления IRT как SCM необходимы еще 2 допущения:

1) Параметр θ а значит и $|\theta_{ист}|$ одинаков у одного испытуемого для всех вопросов. То есть субъективное расстояние от вопроса до правильного ответа одинаково у одного человека для всех вопросов.

Это может быть следствием стратегии подготовки к экзамену. Студент, готовящийся к экзамену, заучивая очередной вопрос рефлексирует свое знание этого вопроса (то есть фактически расстояние между вопросом и правильным ответом). Студент устанавливает величину ощущения знания вопроса, которой он должен достигнуть. Это ощущение знания может соответствовать, к примеру субъективному расстоянию в 2,6 ед. между вопросом и ответом. До подготовки это расстояние может быть, к примеру 10 ед. По мере заучивания оно сокращается, и как только становится равным 2,6 ед. и начинает вызывать заранее определенную величину ощущения знания, студент переходит к следующему вопросу, который тоже заучивает до такой степени, чтобы расстояние между вопросом и правильным ответом стало 2,6 ед. и так далее.

Такая стратегия обеспечивает одинаковость параметра $|\theta_{ист}|$ для любого вопроса, и позволяет модели Раша удовлетворительно аппроксимировать эмпирические результаты экзаменов.

2) Параметр трудности вопроса одинаков для всех испытуемых. То есть расстояние от вопроса до дистрактора одинаково для всех испытуемых.

Разумно предположить, что каждый из экзаменуемых, по крайней мере, понимает экзаменационные вопросы. “Понимает вопрос” с точки зрения экзаменатора означает, что вопрос отобразился в определенную точку СП испытуемого. Если все студенты “понимают вопрос” с точки зрения экзаменатора, значит у любого из них вопрос отображается в одну и ту же точку семантического пространства

Идеально соответствующий IRT вопрос мог бы звучать так:

Психология это:

1)Наука о душе

2)Наука о бсзакм.

Здесь первый вариант – правильный ответ. Второй вариант – наукообразный термин, например сгенерированный компьютером.

Поскольку любого студента на лекциях не учили тому, что такое “бсзакм”, любой студент отображает “бсзакм” в свое семантическое пространство на основе его несемантических свойств. Поскольку несемантические (фонологические и т. д.) свойства дистрактора зависят от его физических свойств, дистрактор должен примерно одинаково отображаться в семантическое пространство любого студента.

Из этих допущений (и вопрос и дистрактор всеми воспринимаются одинаково) следует, что расстояние от вопроса до дистрактора должно быть одинаковым у любого студента.

Интересным выводом из представления IRT как частного случая SCM является то, что трудности вопросов, измеренные по Рашу зависят только от правдоподобности дистракторов, а не от самого вопроса и правильного ответа на него.

Больше 2х вариантов ответа.

Возникает вопрос: как должна быть изменена IRT, что бы корректно описывать ответы на вопросы с множеством альтернатив, являясь в то же время частным случаем SCM?

Постулируем, что расстояние от вопроса до любого из дистракторов одинаково. Это должно приводить к равновероятности выбора любого из дистракторов если испытуемый дал неправильный ответ. Такой равновероятности добиваются при формулировании основанного на IRT теста [6].

Тогда основная формула IRT как SCM преобразуется в:

$$\text{P(правильного ответа)} = \frac{e^{-|\theta_{ист}|}}{e^{-|\theta_{ист}|} + [e^{-|\delta_{ист}|}]*(k-1)}$$

Где k – число вариантов ответа на данный вопрос.

$$\frac{\text{P(правильного ответа)}}{1-\text{P(правильного ответа)}} = \frac{e^{-|\theta_{ист}|}}{[e^{-|\delta_{ист}|}]*(k-1)}$$

$$\ln\frac{\text{P(правильного ответа)}}{1-\text{P(правильного ответа)}} = [-|\theta_{ист}|]-[-|\delta_{ист}|] - \ln(k-1)$$

$$\ln\frac{\text{P(правильного ответа)}}{1-\text{P(правильного ответа)}} + \ln(k-1) = [-|\theta_{ист}|]-[-|\delta_{ист}|]$$

$$\ln\frac{\text{P(правильного ответа)}}{1-\text{P(правильного ответа)}} + \ln(k-1) = [-|\theta_{ист}|+const]-[-|\delta_{ист}|+const]$$

$$\ln\frac{\text{P(правильного ответа)}}{1-\text{P(правильного ответа)}} + \ln(k-1) = \theta-\delta$$

Сравним это с формулой, используемой при вычислении по исходной модели Раша.

$$\ln\frac{\text{P(правильного ответа)}}{1-\text{P(правильного ответа)}} = \theta-\delta$$

Можно видеть, что предложенное усовершенствование позволяет учитывать еще одну количественную характеристику вопросов – количество вариантов ответа, чего не может исходная модель.

Учет случайного угадывания.

Один из недостатков модели Раша – отсутствие учета случайного угадывания. Алан Бирнбаум сформулировал т. н. трехпараметрическую модель. В ней характеристическая кривая пункта зависит от трех параметров:

Трудности пункта δ, дискриминативности пункта (крутизны кривой) α и вероятности того что субъект при ответе прибегнет к случайному угадыванию r. Отметим сомнительность оценивания столь большого количества параметров по матрице ответов.

Полезно, однако, заметить, что крутизна кривой пункта (Рис. 3) зависит как от величины α, так и от величины r. Большее r сокращает диапазон, в котором может изменяться вероятность правильного ответа (от r до 1), и тем самым уменьшает крутизну кривой, снижая дискриминативность.

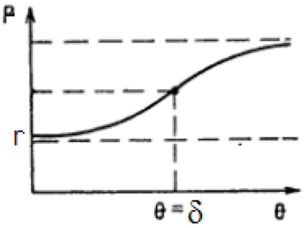

Рис. 3. (воспроизведено из [8])

Оправданно уменьшать r, увеличивая число вариантов ответа на вопрос.

Из случайного угадывания следует, что вероятность правильного ответа не может быть меньше 1/k. Попробуем объяснить этот факт с точки зрения SCM.

Вспомним идеально соответствующий IRT вопрос

Если экзаменуемый не знает, что такое "наука о душе", он формирует свое представление об этом по звучанию, так же как и представление о дистракторе. Следовательно, оба варианта попадают субъективно в разряд "наукообразно звучащих слов" (примерно в одно место семантического пространства). Следовательно $|\theta_{ист}| = |\delta_{ист}|$.

$$P(\text{правильного ответа}) = \frac{e^{-|\theta_{ист}|}}{e^{-|\theta_{ист}|} + e^{-|\theta_{ист}|}} = \frac{1}{2}$$

Последнее ограничение можно записать так:

$|\delta_{ист}| \geqslant |\theta_{ист}|$

IRT представленная как SCM с этим ограничением приводит к тому что вероятность правильного ответа не опускается ниже уровня случайного угадывания.

В рамках классической IRT формулировались задания, получающие оценки трудности, превышающие уровень знаний студентов, в то время как вышеуказанное неравенство запрещает это. По крайней мере, отчасти это можно объяснить множителем (k-1), который в классической модели Раша искажает оценку разности уровня знаний и трудности, увеличивая ее на ln(k-1) т. е. от 0,693 для k=3 до 2,19 логит для k=10

В заключение укажем один из путей проверки модели.

Необходимо задать выборке людей набор вопросов. Другой выборке людей задаются те же вопросы, но у половины вопросов увеличивается число вариантов ответа (скажем с 3 до 7, чтобы не выходить за пределы оперативной памяти испытуемого). В рамках предложенной модели сложности соответствующих вопросов вычисленные по двум выборкам должны совпадать, независимо от того увеличивалось или нет число дистракторов. В классической модели сложности вопросов с большим числом вариантов ответа должны увеличиться. При сравнении целесообразно выбрать среднюю сложность вопросов неизменных от выборки к выборке как начало отсчета как первой так и второй шкалы. Для предлагаемой нами модели необходимо еще сдвинуть начало отсчета на *–ln(исходное число вариантов ответа-1)*. Тогда обе построенные интервальные шкалы будут иметь одинаковое начало отсчета и одинаковую единицу (логит).

Литература:

1. Соколов Е. Н. , Терехина А. Ю. , Ребрик Б. С. Геометрическая модель структуры знания. // Вопросы психологии. 1986. № 6
2. Shepard R. N. Stimulus and response generalization: a stochastic model relating generalization to distance in psychological space. // Psychometrika. 1957. №4
3. Shepard R. N. Toward universal law of generalization in psychological science. // Science. 1987. № 4820
4. Gilmore G. C, Hersh H. , Caramazza A. , Griffin J. Multidimensional letter similarity derived from recognition errors. // Perception and Psychophysics. 1979. №5
5. Клайн П. Справочное руководство по конструированию тестов – Киев. : "ПАН Лтд. ", 1994
6. Нейман Ю. М. , Хлебников В. А. Введение в теорию моделирования и параметризации педагогических тестов. - М. : "Прометей", 2000
7. 7. Фер Р. М. , Бакарак В. Р. Психометрика: Введение. - Челябинск. : "Издательский центр УдГу", 2010.
8. Челышкова М. Б. Разработка педагогических тестов на основе современных математических моделей. - М. : Исследовательский центр проблем качества подготовки специалистов, 1995.