

КВАНТИТАТИВНА ПЕДАГОГИЧЕСКА ОЦЕНКА

Йожеф Надь

В Венгрии развитие экспериментальной педагогики по различным причинам было заторможено в начале первой мировой войны и не могло далее развиваться. Только в середине 60-х годов созрели общественные предпосылки для того, чтобы измерения, методы и средства количественной педагогической оценки завоевали гражданское право. Это положение потребовало наверстание большой отсталости; в международной специальной литературе накопились результаты полустолетней исследовательской работы, которые нам нужно было изучить, и в то же время нужно было адаптировать их к социалистическим условиям.

Преимущество означало тот факт, что на этот период пришлось начало усовершенствования количественной педагогической оценки, ориентировался критерий, формативная оценка, разработка понятий полного усвоения и развертывание многих новых течений, возникновение недовольства стандартными традиционными тестами. Была предоставлена возможность для того, чтобы не повторились первоначальная косность и экстремизм количественной педагогической оценки. Одновременно в этот период наблюдались успешные усилия Ласло Кальмара в организации сегедской Кибернетической лаборатории (вычислительного центра). Предоставилась возможность для того, чтобы мы работали с большими образцами, чтобы мы применяли такие методы оценки, которые, хотя и давно уже были в нашем распоряжении, но на большой модели ручными методами были практически неприменимы.

Благоприятным обстоятельством мы считаем тот факт, что в педагогическом общественном мнении проявился очень большой интерес: сотни, тысячи педагогов принимали участие в этой работе без какого-либо возмездия. Одновременно и в других научных педагогических учреждениях наступил период быстрого изучения.

Для этого положения характерно то, что и в математический статистический справочник, написанный для психологов, увидел свет лишь в 1968 году.

Опираясь на первоначальные опыты, мы написали первое руководство на венгерском языке для педагогов под названием

"Измерительные методы в педагогике" (Агостон - Надь - Орос. 1971).

На основе первоначальных результатов мы получили сочувствие и поддержку от сотрудников министерства в развертывании исследований.

Наши исследования можно разделить на два периода: работа, проведенная с середины 60-х годов до середины 70-х годов, и работа, проведенная с середины 70-х годов по настоящее время. Для упрощения мы продемонстрируем наши задачи и результаты по этим двум периодам.

I.

Приняв во внимание международный опыт созревших проблем, Мы исходили из того, что возникла необходимость создания такой количественной педагогической оценки, которая способствовала бы как развитию практики педагогики и ее теории, так и развитию системы общественного воспитания. Поэтому мы стремились разработать такие тесты и провести такие репрезентативные измерения, которые вместе (со всеми основными) функциями обеспечивают работу. В наши дни подобные исследования называются мультифункциональными (а средства называются мультифункциональными тестами).

Поскольку мы располагали небольшим опытом, в первую очередь мы работали в мире знаний, навыков. И в рамках этого мы отметили две области. Картографирование развития навыков (в отношении определенных навыков) и разработка тестов на заключительную тему, которые оценивают результаты разработки тем по учебным дисциплинам.

Согласно вышесказанному, в этих областях мы желали достичь следующих целей:

- создание типов тестов, которые пригодны для обоих видов оценки, разработка методов составления тестов;
- дать в руки педагогов тесты, помогающие им в процессе их педагогической деятельности при разработке учебных тем;

- тщательное и надежное раскрытие уровня усвоения на основе точных всеобщих измерений с помощью тестов;

- кроме этого, с помощью тестов, измеряющих навыки, - исследование процесса развития в популяции (практически это означает, что, начиная с первого года обучения и кончая 12-м годом обучения, в каждом классе определенным образом мы измеряли уровень развития данного навыка одной и той же серией тестов).

При составлении тестов за исходную рабочую гипотезу мы принимали так называемый структурный тотальный принцип, который мы постепенно развивали на основе опытов.

Этот принцип прежде всего гласит о том, что измеряемое содержание необходимо подвергнуть тщательному структурному анализу, чтобы получить полную, детальную и логически законченную систему данной темы, знаний, понятия, навыка.

Мы подготовили задания ко всем "логическим вершинам, точкам" таким образом раскрытой структуры.

Итак, с одной стороны, к одной средней теме (около 10-20 учебных часов) готовилось 100-200 задач (200-400 итем), с другой стороны, к одному средне-сложному навыку - несколько десятков, к более сложным навыкам - несколько сотен задач, требующих времени.

Если мы пригодным методом желаем детально познать уровень усвоения данной темы, т.е. диагноз, служащий для управления развитием учебного плана, педагогического процесса, то этого мы не достигнем исходными тезисами традиционной теории тестов. Поэтому из возможного содержания теста, из популяции мы берем модель (если возможно, то методом случайного выбора) и допустим, что таким образом взятые элементы представляют всё измеряемое содержание. Этот метод может быть пригодным для того, чтобы с его помощью можно было определить разницу между учащимися, чтобы охарактеризовать работоспособность учащегося. Однако нам ничего неизвестно об элементах уровня усвоения, не фигурирующих в тесте; поэтому они мало пригодны для диагнозов таких тестов, для способствования управлению учебным процессом, для

развития учебного плана, для управления системой общественно-го воспитания.

Следовательно, структурный тотальный принцип превосходит традиционную теорию теста и служит тезису, способствующему правильной оценке тестов.

Естественно, невозможно, что задачу всех измеряемых элементов решает один ученик. Да для этого нет и необходимости. Задачи они могут выбрать (разыграть) в параллельные тесты. Если эти варианты тестов мы измерим по независимым моделям, то получим необходимую оценку для уровня усвоения любой задачи (любой логической вершины данной структуры).

Если результаты, полученные в определенных параллельных тестах выразим на какой-либо стандартной шкале, то результаты, полученные во всех вариантах тестов, станут сопоставивыми.

Понятие банка задач (итем-банк) при введении принципа структурной тотальности был еще неизвестным. Позже выяснилось, что этот принцип привел нас к формированию понятия своеобразного банка задач.

Это распознавание вы в дальнейшем сознательно использовали при раскрытии всех возможных элементарных текстовых типов задач (384 типа задач). Из них мы составили 48 параллельных тестов, и с каждым тестом по независимой модели (каждый учащийся в классе получал отдельный тест с наперед указанным номером) мы провели всеобщее репрезентативное измерение, начиная с 4-ого класса и кончая 12-ым классом. В результате этого мы смогли получить процесс развития каждого типа задач. Думается, что нет необходимости объяснять диагностическую силу, регулирующую роль таких тестов и измерений: (Надь - Чаки: Текстовый банк задач в младших классах. 1976 г.).

Из банка задач педагог по своему усмотрению может составить оценочное средство, с помощью простых вычислений может получить всевенгерский уровень, с ним может соотнести свои результаты.

Кроме указанной в качестве примера темы в следующих областях мы раскрыли процессы развития навыков и умений.

В качестве прецедента текстового решения задачи являются элементарные навыки счета (Надь, 1971) и навыки основных вычислительных действий, понимая под этим и навыки цифрового написания и навыки меры.

В области родного языка предмет наших исследований представляли умения в технике сочинения (Орос, 1972) и навыки в правописании (Орос, 1974). (В наших планах фигурировала также картография развития навыков письма, но сотрудник, который работал над этой темой, по своей занятости не мог закончить свои исследования).

На этих предметах исследования сформировались своеобразные особенности тестов, измеряющих навыки, методы их составления, своеобразные способы разработки и применение моделей, вытекающие из структурной тотальности.

Тесты (точнее, варианты тестов, составленные определенными методами из банка задач,) показывают степень полного усвоения. Эту степень, с одной стороны, намечает содержание, определенное структурной тотальностью, а с другой стороны - из 100 % результатов, полученных в тесте, вычтем количество допустимых ошибок (в баллах), (приняв во внимание показатель реабилитации и анализ распределений). Например, полное усвоение изменения меры показывается, по крайней мере, на 90 %, усвоение деления - на 85 %.

Из года в год процесс развития, происходящий в популяции учащихся, запишем серией рассеиваний (типичная наглядность которой представлена на рис. I.).

Тесты на заключительную тему были составлены по пяти учебным дисциплинам старших классов школы-восемилетки (венгерская грамматика, математика, физика, химия, зоология). К каждой теме каждого класса. К одной определенной теме мы разработали по крайней мере четыре, а иногда пять или шесть параллельных тестов.

Вопросы составления тестов, их использования мы изложили в книге "Практические вопросы изучения уровня знаний по заключительной теме" (Надь, 1972). Теоретический опыт многолетней работы мы изложили в издании "Реабилитация и законная сила тестов по заключительной теме" (Надь, 1975).

Тесты по пяти учебным дисциплинам и результаты всевенгерских измерений вышли в свет в 17-и томах из-под пера руководителей исследовательских тем: 4 тома Шандора Ороса по венгерской грамматике (5-8 классы), тома Иштвана Гажо (5-8 классы) по математике (последний том вышел под руководством Кунштар Яношне из-за смерти автора), 4 тома Гезы Добо по зоологии, 3 издания по физике Яноша Вейднера (6-8 классы) и две книги по химии Элемера Куншаги - Вида Михайне (7-8 классы).

В качестве анализа возьмем 6 параллельных тестов по химии - Вариант Г и относящиеся к нему результаты измерения (Куншаги - Видане, 1973 г.).

В тесте бросается в глаза, что мы применяли не технику выбора ответов на вопросы. На это имеются законные причины, которые подробно мы здесь рассматривать не можем. Эту проблему разрешает принцип адекватности, функциональной силы. "Функциональная законная сила тогда и только тогда тривиальна, если работа по решению задач адекватна работе, реализующей действие измеряемой психологической структуры". (Надь, 1975, стр. 32, см. здесь же и дальнейшую информацию).

Другая особенность состоит в том, что каждую задачу мы разбили на так называемые альтернативные элементы, которые могут быть охарактеризованы только двумя способами: I. хорошо,

безупречно, 2. ошибочно, плохо, недостаточно. Это дает возможность (с помощью ключа исправления) дать объективную оценку (что, впрочем, необходимо контролировать соответствующими методами в ходе составления тестов).

Измерительный лист
на заключительную тему
Общеобразовательная школа

Вариант Г/
Имя
Класс

Химия. 7-ой класс

НАИБОЛЕЕ ВАЖНЫЕ УГЛЕРОДНЫЕ СОЕДИНЕНИЯ

I. В пространстве, лишенном воздуха, нагреем дерево (или уголь). Как называют этот процесс?

а/

Назови также и полученный продукт!

Название продукта:

Агрегатное состояние

б/

в/

г/

д/

е/

ж/

а	б	в	г	д	е	ж
2	2	2	I	I	I	I

2. Дополни:

Агрегатное состояние
углекислого газа:

а/

Его удельный вес по
отношению к воздуху

б/

Горение:

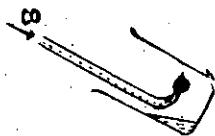
в/

Его воздействие на
живой организм:

г/

а	б	в	г
I	I	I	I

3.



Какое важное свойство угарного газа показывает этот рисунок?

6. Таблица, данная ниже, изображает свойства жиров и масел.
Дополни ее !

	Жиры	Масла
Конструктивные атомы	а/	е/
Агрегатное состояние при комнатной температуре	б/	ж/
удельный вес по отношению к воде	в/	и/
его растворитель	г/	к/
его применение	д/	л/

а	б	в	г	д	е	ж	з	и	к	л
1	1	1	1	1	1	1	1	1	1	1

7. Какой химический процесс:

карбонизация сахаров: а/

их продукты: б/

и какой химический процесс

сжигание сахаров: в/

его продукт:+ =

г/

д/

ж/

а	б	в	г	д	ж
2	2	2	1	1	1

8. Какой живой организмы может образовать только белок?

а/

Дополни:

б/ белок, характерный для че-
ловеческого организма

в/

а	б	в
2	2	2

9. Назови три вида промышленного сырья, содержащего белок, и промышленность, в которой они применяются!

Название сырья, содержащего белок	В какой промышленности применяются
а/	б/
в/	г/
д/	е/

а	б	в	г	д	е
1	1	1	1	1	1

10. Назови окислы углерода!

- а/ б/
 Напиши окончательную форму этих окислов:
- в/ г/
 В каком химическом процессе они могут возникнуть?
- д/ е/

а	б	в	г	д	е
3	2	4	4	5	6

Результат: баллы в %

РЕКОМЕНДУЕМОЕ ЗАДАНИЕ

11. Здесь мы перечисляем вещества и характерные для них свойства. Напиши рядом с названием материала количество соответствующих ему свойств!

нефть: а/
 белок: б/

углевод: в/

жиры, масла г/

1. Питательные вещества
2. Смесь углеводов
3. Носители жизни
4. Прогорклость: разложение

а	б	в	г
2	2	2	2

Оценка рекомендуемого задания:

..... баллы в %

РЕЗУЛЬТАТ:

НАИБОЛЕЕ ВАЖНЫЕ УГЛЕРОДНЫЕ СОЕДИНЕНИЯ

Вариант Г/

- I. а/ сухая дистилляция /испарение/
 б/ древесный газ
 в/ газообразное состояние
 г/ кокс /древесный уголь/
 д/ твердое состояние
 е/ смола
 ж/ жидкое состояние
2. а/ газообразное состояние
 б/ тяжелее
 в/ не питает
 г/ ядовитый
3. а/ горючий газ
 б/ I
 в/ CO
 г/ I
 д/ O
 е/ I
 ж/ CO₂
 з/ 2
 и/ CO
 к/ I
 л/ O₂
 м/ 2
 н/ CO₂
4. а/ природный газ
 б/ смесь углеводородов
 в/ метан
 г/ CH₄
5. а/ топливо
 б/ очистка ран
 в/ очистительное средство
7. а/ сухое испарение /разложение/
 б/ сахарный уголь /искусственный уголь/
 в/ окисление
 г/ CO₂
 д/ H₂O
 е/ температура
8. а/ растительный
 б/ растительный
 в/ животный
9. а, б, в, г, д, е
 среди буквенных обозначений, считая от начала, оставим столько без зачеркивания, сколько правильных ответов дал ученик.
10. а, б, в, г, д, е
 среди буквенных обозначений, считая от начала, оставим столько без зачеркивания, сколько правильных ответов дал ученик.
- II. Рекомендуемое задание
 а/ 2
 б/ 3
 в/ I
 г/ 4
- Выставление оценок
 отлично: от 84,6 до 100,0
 хорошо: от 61,2 до 84,5

6. а/ С, Н, О	среднее: от 37,9 до
б/ твердый	<u>61,1</u>
в/ более легкий	удовлетв.: от 14,6 до
г/ бензин, алкоголь	37,8
д/ питание	неудовлетв.: от 0,0
е/ жидкий	до 14,5
ж/ более легкий	
з/ бензин	
и/ производство мыла, красок	

РЕЗУЛЬТАТЫ ВАРИАНТА Г/III

Проценты

0 10 20 30 40 50 60 70 80 90 100

1. Сухое испарение и его продукты	а	сухая дистилляция /испар./	75,0
	б	древесный газ	58,0
	в	газообразный	43,0
	г	кокс /древесный уголь/	52,3
	д	твердый	58,0
	е	смола	57,0
	ж	жидкообразный	50,0
2. Характеристики CO_2	а	газообразный	49,0
	б	более тяжелый	56,0
	в	не питает	60,3
	г	ядовитый	47,0
3. Составление уравнения, выражающего горение CO	а	горючий газ	85,0
	б	I	47,0
	в	CO	55,4
	г	I	63,0
	д	0	62,4
	е	I	50,4
	ж	CO_2	58,3
	з	2	51,0
	и	CO	22,0
	к	I	33,0
	л	O_2	38,0
	м	2	32,0
	н	CO_2	45,0
4. Углеводороды	а	природный газ	81,0
	б	смесь углеводородов	58,1
	в	метан	67,0
	г	CH_4	70,2
5. Применение бензина	а	топливо	88,0
	б	очистка рая	89,0
	в	очистительное средство	82,3

6. Сопоставление жиров и растительных масел	а	$C_2 H_2 O$	88,0
	б	твердый	70,2
	в	более легкий	65,0
	г	бензин, алкоголь	80,0
	д	питание	67,2
	е	C_2, H_2, O	93,4
	ж	жидкий	74,0
	з	более легкий	72,0
	и	бензин	89,0
	к	производство мыла, красок	88,3
	7. Обугливание и горение сахаров	а	сухое испарение /разложения/
б		сахарный уголь /ис. уголь/	58,0
в		окисление	60,0
г		CO_2	64,0
д		H_2O	67,3
е		температура	69,0
8. Возникновение белков	а	растительный	85,0
	б	растительный	72,0
	в	животный	57,0
9. Промышленное сырье, содержащее белок	а	пример	78,0
	б	пример	77,3
	в	пример	67,0
	г	пример	65,0
	д	пример	68,1
	е	пример	63,0
10. Окислы углерода	а	пример	62,1
	б	пример	88,0
	в	пример	57,0
	г	пример	46,1
	д	пример	38,4
	е	пример	33,3
0 10 20 30 40 50 60 70 80 90 100			

Итоговые данные к варианту
Г/III

Кол-во учеников	158
Среднее \bar{x}	62,9
Конфиденция интервал $\pm A$	$\pm 3,6$
Требования на точность	$\pm 6,9$
Рассеивание Относительное $\pm s$	$\pm 23,3$
рассеивание	44,4 %

Распределение	
Баллы в %	Ученики %
0,1 - 5,0	2,5
5,1-10,0	1,2
10,1-15,0	4,4
15,1-20,0	3,7
20,1-25,0	3,7
25,1-30,0	3,1
30,1-35,0	6,3
35,1-40,0	3,1
40,1-45,0	5,6
45,1-50,0	5,6
50,1-55,0	8,2
55,1-60,0	10,1
60,1-65,0	7,5
65,1-70,0	8,2
70,1-75,0	10,1
75,1-80,0	4,4
80,1-85,0	5,0
85,1-90,0	4,4
90,1-95,0	0,0
95,1-100,0	1,8

РАСПРЕДЕЛЕНИЕ РЕЗУЛЬТАТОВ ВАРИАНТА Г/III

Относит. частотность
в %

Овладев накопленным опытом в первой половине 70-х годов, мы взялись за решение одной комплексной задачи. Мы разработали исследовательскую систему по своеобразной школьной подготовке, которая служит усовершенствованию определения детей в школу (ПРЕФЕР, Надь, 197), с помощью которой мы провели всеобщие репрезентативные измерения среди 5-6 летних детей (их было 10 000). От обычных тестов на аттестат зрелости исследовательская система отличается в том, что она стремится измерять не способности, интеллигенцию, поведение, а желает быть педагогическим исследованием. То есть она оценивает самую различную подготовку в знании и поведении, необходимые в начале школьного обучения; только такие особенности, которые могут быть развиты, обучены педагогической деятельностью.

Всеенгерские измерения служат и той цели, чтобы мы познали уровень развития способностей, навыков, знаний, поведения, стоящих перед школьным порогом двух поколений, процесс их развития в зависимости от всех возможных факторов (готовится публикация результатов в книге "Подготовка к школе детей 5-6-летнего возраста").

II

Основным фактором нового исследовательского периода, начавшегося в середине 70- годов, является то, что мы хотели бы еще более непосредственно служить управлению учебным процессом. Однако и в дальнейшем мы остановились на оценке, способствующей тематическому единому обучению, и на оценке, служащей управлению процессом усвоения, продолжающегося долгие годы.

В отношении преподавания темы с использованием понятия "выучивание" / *mastery learning* / для осуществления так называемого комплексного обучения теме необходима выработка оценочных средств, но это уже неотделимо от экспериментирования комплексного обучения теме. Проектную статью и результаты предварительных экспериментов см. в издании 1977 г, "Руково-

дящая оценка учебы с помощью банка задач".

Сущность компенсированного обучения по теме состоит в том, что перед разработкой темы мы проводим так называемую предкомпенсацию для того, чтобы при разработке темы выявить необходимые условия знаний. После проработки темы мы ввели посткомпенсацию; за это время отстающие для достижения желаемого уровня получают дополнительные задания, а с другими — педагог проводит такую работу.

Для того, чтобы система действовала, необходимы предтесты и посттесты. Однако эти традиционные средства не совсем пригодны для управления учебным процессом.

Если мы от педагогов ждем составления тестов, то они могут быть самостоятельными в использовании конкретных условий (в отличие от заранее составленных, стандартных тестов), но такое ожидание утопично, к тому же уровень таких появившихся средств часто сомнителен, нет меры соотношений.

Если задания составляются централизованно к пред- и постизмерению, то каким-то образом нужно выбирать усвояемые элементы, что опасно своей односторонностью, а эта зависимость (ведь заранее составленный стандартный тест нельзя изменить) может втянуть обучение в формализм.

Решению проблемы помогает то понятие банка задач, которое мы разработали в предварительном периоде исследований и которое было выше кратко охарактеризовано. Нам предоставилась возможность составить задачи на все элементы знаний (в предпосылке), на все усвояемые элементы темы к оценке как вступительной, так и заключительной темы; с этими задачами мы проводим необходимые эксперименты, измерения и проверенные задачи на всеевгерском уровне мы предоставили в распоряжение педагогов в форме банка задач. (Формально кажется целесообразным подготовить сборник задач, снабженный порядковыми номерами, из которого намеченные педагогом задачи учащиеся решают в своих тетрадях).

Исследования завершатся к концу 1981 года. Ожидаемые результаты мы хотели бы поместить в банке статей, где было бы

описано конструирование средств, применение методов, действие компенсированного обучения по теме. Кроме того, по некоторым учебным дисциплинам мы сообщим модели о результатах экспериментов.

Располагая многолетним опытом по оценке процессов усвоения в отношении навыков и умений, мы наметили еще более смелые планы. Мы стремимся разработать так называемую систему операционных способностей и раскрыть их процессы развития в школьном возрасте (т.е. на всеобщих репрезентативных моделях). Мы проводим эту работу с надеждой на то, что этим мы можем внести свой вклад в дело сознательного развития этих основных способностей (до тех пор, пока не познаем способности и процесс их развития, об их сознательном развитии не может быть и речи).

В настоящее время на основе предварительной работы мы предполагаем, что в интеллектуальной человеческой мыслительной деятельности действует несколько десятков операционных способностей. Стало также ясным, что эти способности не могут сформироваться как без развития языковой структуры, так и без развития так называемых основополагающих понятий на соответствующем уровне. Поэтому операционные способности можно и даже необходимо изучать вместе с процессом развития языковой структуры и основных понятий.

Это исследование планируется на несколько периодов. Первый период завершится к середине 80-х годов разработкой системы операционных способностей; экспериментированием тестов, банков задач, необходимых к измерению развития этих способностей; показом всеобщего уровня развития в нескольких возрастных разрезах (намечается экспериментирование с детьми от 8-10 лет и от 14-16 лет). Дальнейшие исследования послужат более детальному раскрытию процессов развития и более непосредственно будут способствовать сознательному развитию.