

Центр психометрики и измерений в образовании,
Институт образования

Оценивание открытых заданий в тесте смыслочного чтения

Соболева Ольга

Антипкина Инна

Красноярск, 2020



ЗАЧЕМ ЕГО ОЦЕНИВАТЬ?

«Первая причина, по которой нам необходимо анализировать международные тесты и результаты их применения, относится к **статусу психолого-педагогических наук** в современном отечественном образовании. Этот статус низкий не только из-за низкого финансирования. Если наука претендует на изучение всеобщих закономерностей, она может быть только мировой, международной. Если она к этому не стремится, то страдает провинциализмом.

Вторая причина прагматическая, она относится к умению **использовать научные достижения** в практике образования».

Г. А. Цукерман, Г. С. Ковалева, В. Ю. Баранова

ЭКРАН ТЕКСТА

История про Севу

История про Севу

Когда произошла эта история, Сева жил у бабушки в деревне. Лето уже близилось к концу. Сева радовался, что скоро увидит школьных друзей, расскажет им про жизнь в деревне и покажет подарок бабушки – настоящий бинокль. А пока он был один, сидел на скамейке у дома и разглядывал в бинокль окрестности. В поле зрения попадало всё то, что давно было знакомо: то забор, то большая липа, то соседский дом, то сам сосед с сумкой.

Вдруг Сева заметил птицу, но вот только летела она почему-то не вверх, а вниз! «Неужели птица упала на землю?» – поразился Сева и бросился в ту сторону. И правда, за забором на куче сухой травы Сева нашёл маленькую птичку. Она бессильно раскинула крылья и слегка шевелила головой.

Далее

Пример задания с выделением ответа (начало)

История про Севу

Где Сева нашёл птичку? Выдели ответ в тексте.

Открыть текст

Далее

Пример задания с выделением ответа (окончание)

История про Севу

Когда произошла эта история, Сева жил у бабушки в деревне. Лето уже близилось к концу. Сева радовался, что скоро увидит школьных друзей, расскажет им про жизнь в деревне и покажет подарок бабушки – настоящий бинокль. А пока он был один, сидел на скамейке у дома и разглядывал в бинокль окрестности. В поле зрения попадало всё то, что давно было знакомо: то забор, то большая липа, то соседский дом, то сам сосед с сумкой.

Вдруг Сева заметил птицу, но вот только летела она почему-то не вверх, а вниз! «Неужели птица упала на землю?» – поразился Сева и бросился в ту сторону. И правда, **за забором на куче сухой травы** Сева нашёл маленькую птичку. Она бессильно раскинула крылья и слегка шевелила головой.

Копировать

Готово

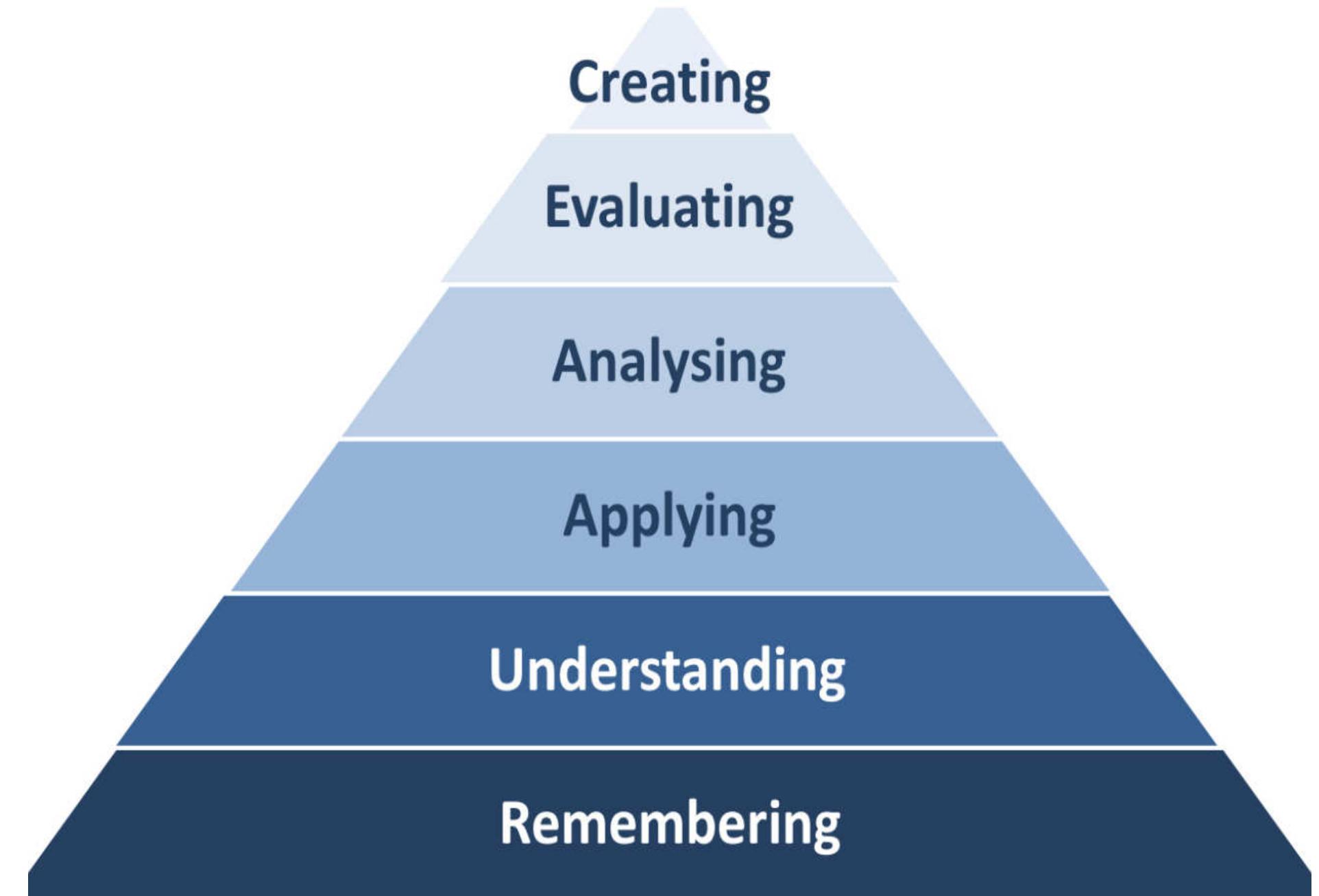
Выбранный текст

Пусто

ТЕОРЕТИЧЕСКИЕ РАМКИ ЧТЕНИЯ

PIRLS, PISA опираются на рамки, основанные на таксономии Блума (Anderson & Krathwohl, 2000)

Уровни таксономии учитываются при экспертной разработке систем оценивания (особенно для политомических заданий)



ГИПОТЕЗЫ

- 1) Существует иерархия паттернов ответов на задания:
 - a) Точный и лаконичный ответ (уровень оценивания);
 - b) Точный, но слишком развернутый ответ (уровень синтеза);
 - c) Неполный ответ: выделена часть правильного ответа (уровень вспоминания);
 - d) Слишком большое «перевыделение» текста (уровень вспоминания);
 - e) Другой нерелевантный ответ.
- 2) Дети используют преимущественно один паттерн ответа на вопросы одного формата.

МЕТОДОЛОГИЯ

ВЫБОРКА

1369 третьеклассников московских школ;
50,3% - мальчики, 49,7% - девочки.

ИНСТРУМЕНТ

Тест смыслового чтения ПРОГРЕСС. Состоит из двух блоков (базовый тестлет и тестлет повышенной сложности) и 35 заданий.

ПРОЦЕДУРА

Ученики проходили тестирования в компьютерных классах своих школ. Перед началом был зачитан текст стандартной инструкции. Ограничение по времени – 1 урок.

АНАЛИЗ ДАННЫХ

Психометрическое моделирование: Ordered Partition Model (Wilson, 1992)

ORDERED PARTITION MODEL

- Моделирование тестов, в которых способы ответа на задания не являются эквивалентными
- Она применяется, если существуют: 1) предположение о прогрессе между уровнями, отражаемое в заданиях; 2) в рамках одного уровня существуют различные варианты набрать балл
- Может применяться для моделирования тестов способностей, психологических тестов, тестов знаний и т.д.

$$P(X_{ni} = k) = \frac{\exp[\theta_n B_i(k) - \xi_{ik}]}{\sum_{h=1}^{K_i} \exp[\theta_n B_i(h) - \xi_{ih}]}$$

Формальное определение модели

$$\delta_{im} = \ln \left[\frac{\sum_{B_i(t)=m-1} \exp(-\xi_{it})}{\sum_{B_i(t)=m} \exp(-\xi_{it})} \right]$$

Определение «шага»

ПРИМЕР СКОРИНГА

№ задания	Критерий	Стратегия	Балл
16	Нужный фрагмент: «дождевых червей и ягоды бузины»	A	2
	Перевыделение в пределах одного предложения	B	
	Неполный ответ: «дождевых червей и ягоды», «червей и ягоды», «червей и ягоды бузины», «(дождевых) червей», «ягоды (бузины)»	C	1
	Перевыделение на одно предложение	D	0
	Другой нерелевантный ответ	E	

ХАРАКТЕРИСТИКИ ЗАДАНИЙ

ID зад.	Труд.	Ошиб.	Незвеш. стат.	Взвеш. стат	ТБК	
f5_strat	-0.593	0.057	1.22	1.15	0,35	2
f6_strat	1.358	0.249	1.05	1.02	0,22	1
f10_strat	-0.745	0.052	1.12	1.10	0,4	
f11_strat	-0.318	0.058	1.06	1.06	0,36	
f12_strat	-0.593	0.122	1.09	1.03	0,18	
f16_strat	-1.243	0.073	1.24	1.10	0,23	-1
f17_strat	1.617	0.177	1.01	1.00	0,33	
f23_strat	-0.714	0.060	1.32	1.20	0,33	-2

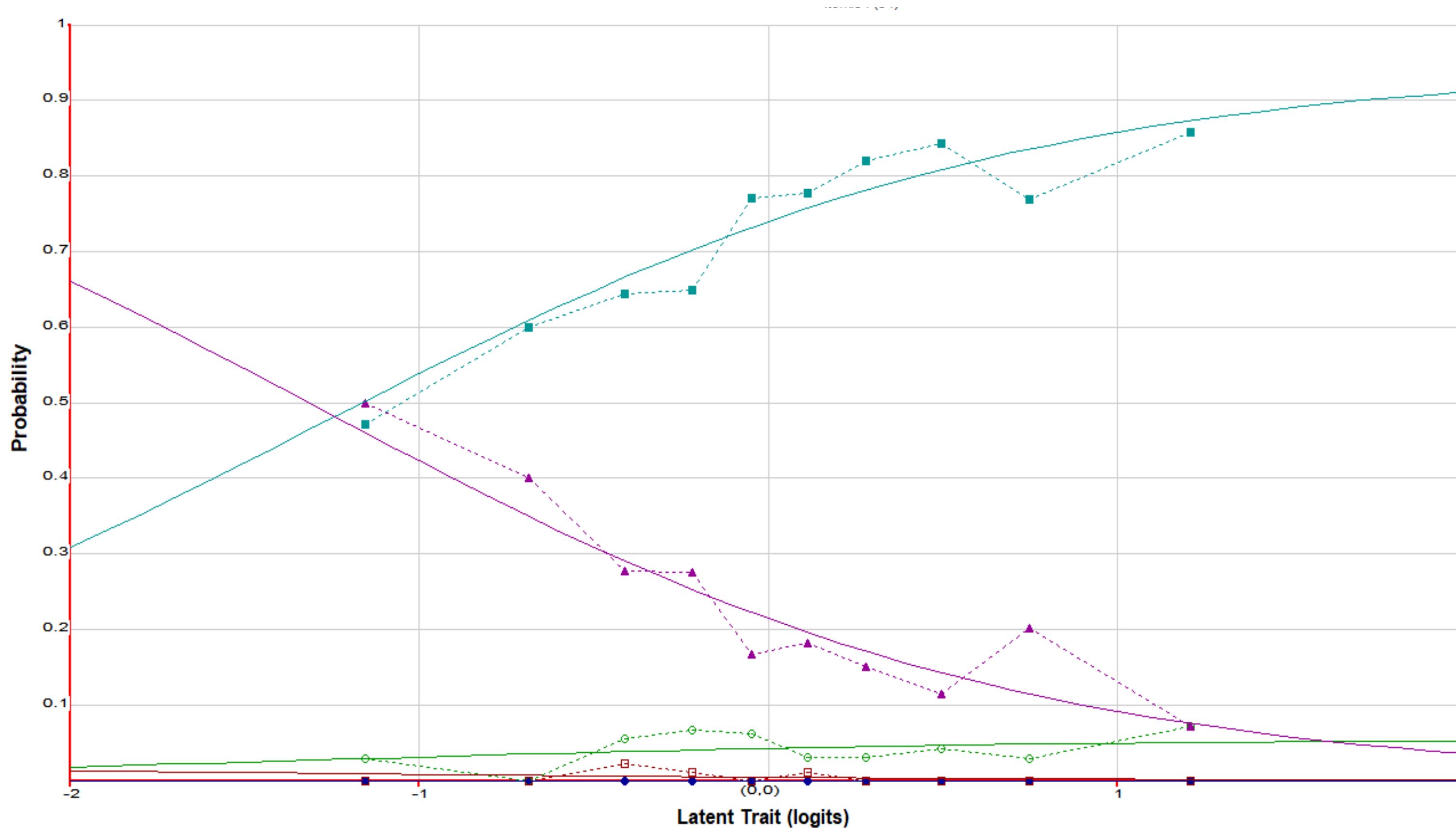
The scatter plot shows the distribution of task characteristics across 34 tasks. The x-axis represents the value of the characteristic, and the y-axis represents the index of the task. Data points are marked with 'x' or 'xx'. Numerical values are labeled next to the points.

- Tasks 1 through 29 are clustered on the right side of the plot, with values ranging from approximately -0.593 to 1.617.
- Tasks 30 through 34 are clustered on the left side of the plot, with values ranging from approximately -1.243 to -0.714.

ХАРАКТЕРИСТИКИ ПАТТЕРНОВ

	Страт.	Порог	Ошибка	ТБК		Страт.	Порог	Ошибка	ТБК
F5	A			0.09		F12		a	
	B	-0,4960	0.077	0.06			A		0.05
	C	-0,5370	0.108	-0.02			B	-1,9490	0.082
	D	0,5560	0.159	-0.05			C	2,7650	0.203
	E	0,3160	0.290	-0.03			D	-2,5940	-0.09
	F	-2,8040		-0.14					
F6	A			-0.06		F16			0.15
	B	-2,4730	0.226	0.16			A		
	C	6,2540	0.307	-0.11			B	-0,9360	0.070
	D	0,2930		-0.09			C	-0,9870	0.139
F10	A			0.17			D	0,2900	0.376
	B	-0,9640	0.072	-0.02			E	-3,3380	-0.14
	C	1,7470	0.253	-0.01		F17			0.02
	D	-3,0180		-0.21			A		
F11	A			0.14			B	-1,2400	0.165
	B	-0,3530	0.092	0.05			C	8,4200	0.531
	C	-1,2300	0.090	-0.02			D	-2,3290	-0.22
	D	0,6280		-0.20		F23			0.12

ПАТТЕРНЫ НЕ УПОРЯДОЧЕНЫ



ОБСУЖДЕНИЕ

СТРАТЕГИИ НЕ ОТРАЖАЮТ КОГНИТИВНУЮ
ИЕРАРХИЮ
С чем связаны способы решения задачий в таком случае?

ОБОСНОВАННЫЙ СКОРИНГ

- В случае обнаружения стратегий за скорингом бы лежала эмпирическая база
- Автоматический скоринг заданий с выделением ответа