



# **Освітні вимірювання як інструмент об'єктивного оцінювання якості освіти**

КОВТУНЕЦЬ В.В.

## 2. Як Ви думаєте, навіщо люди здобувають вищу освіту? (2011, Демініціативи ім. Ілька Кучеріва)

Щоб знайти добре оплачувану роботу	61.9
Щоб зробити кар'єру	50.9
Щоб стати хорошим спеціалістом	47.4
щоб реалізувати себе, свій творчий потенціал	21.0
щоб мати можливість жити у культурному, освіченому середовищі	20.8
щоб не йти до армії	8.9
вища освіта - це шлях у науку	8.9
без вищої освіти людину можуть вважати невдахою	8.7
.....	

# ЯКІСТЬ ВИЩОЇ ОСВІТИ. ОСНОВНІ ПОКАЗНИКИ.

- Знання
- Вміння
- Компетентності



Кваліфікація!

- Якість освітнього процесу

# WIKIPEDIA

**Educational measurement** refers to the use of educational assessments and the analysis of data such as scores obtained from educational assessments to infer the abilities and proficiencies of students. The approaches overlap with those in [psychometrics](#). Educational measurement is the assigning of numerals to traits such as achievement, interest, attitudes, aptitudes, [intelligence](#) and performance.

# ХАРАКТЕРИСТИКИ ОСВІТНІХ ВИМІРЮВАНЬ

- Об'єктивність
- Незалежність
- Надійність
- Валідність

# ОЦІНЮВАННЯ ЯК КЛЮЧОВИЙ ЕТАП ВИМІРЮВАНЬ. МЕТОДИ

- Тестування
- Комісійне оцінювання
- Структуроване оцінювання
- Експертне оцінювання

# ВИДИ ОЦІНЮВАНЬ

- Рейтингове
- Критеріальне
- Поточне (формативне)

# ТЕСТУВАННЯ. ВИДИ ТЕСТОВИХ ЗАВДАНЬ

- Однократного вибору
- Багатократного вибору
- Співставлення
- Вписування правильного фрагменту
- Вибір правильного фрагменту
- Коротка відповідь
- Розгорнута відповідь (есе)



# ПРИКЛАД

Що оцінює таке тестове завдання?

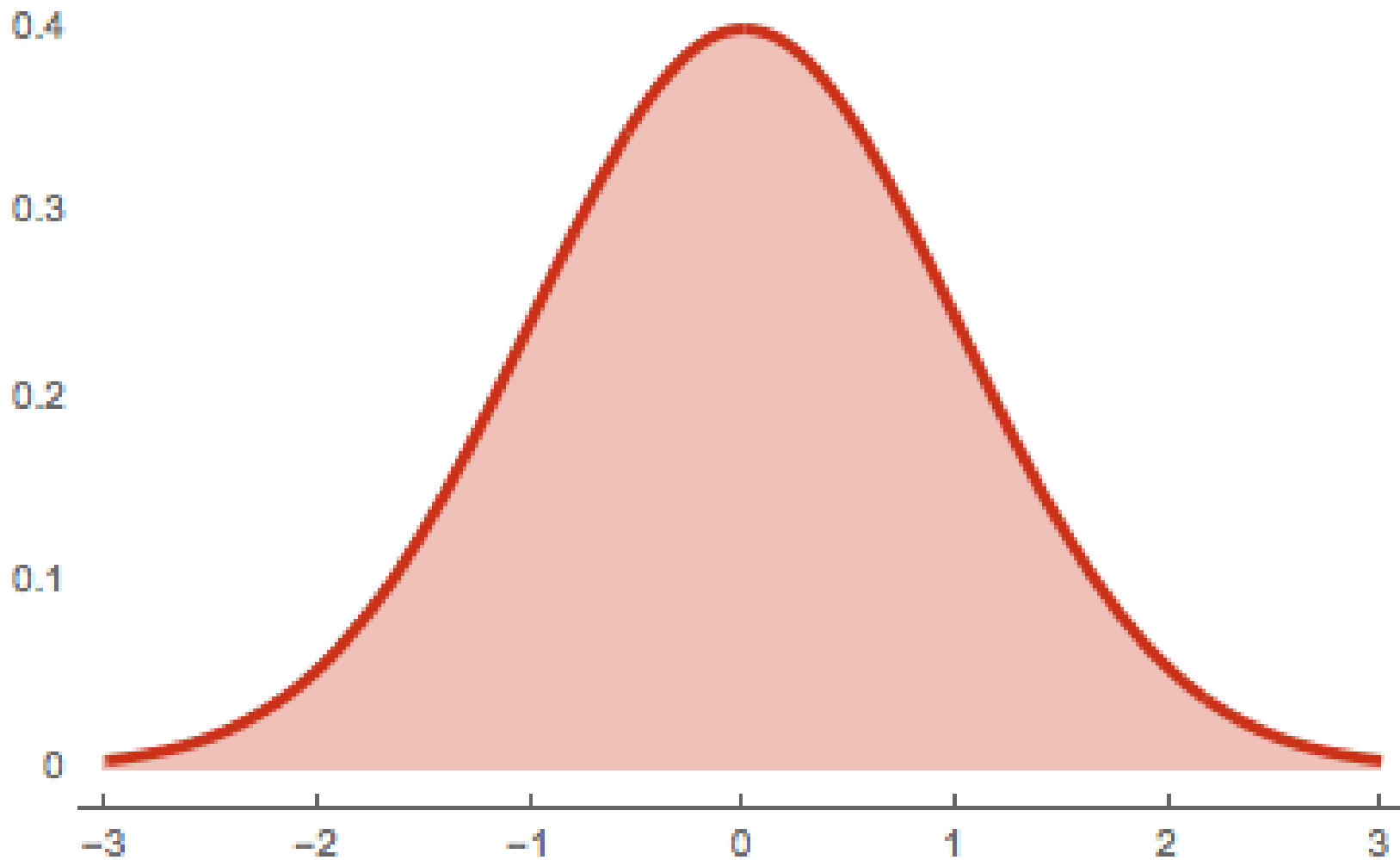
*За який час Земля обертається навколо Сонця:*

- *За 365 діб*
- *Три роки за 365 діб, на четвертий - за 366*
- *За  $365\frac{1}{4}$  доби*
- *Приблизно за  $365\frac{1}{4}$  доби*

# ЕТАТИ ПІДГОТОВКИ ТЕСТУ

- Постановка мети і завдання
- Підготовка специфікації тесту
- Апробація тестових завдань
- Уточнення тестових завдань
- Проведення тестування
- Аналіз результатів
- Аналіз якості тесту

# Шкалювання (нормалізація) результатів



# СФЕРИ ЗАСТОСУВАННЯ ТЕСТІВ

- Масові стандартизовані іспити
- Експрес-оцінювання у студентському колективі
- Найпростіший етап оцінювання професійних кваліфікацій

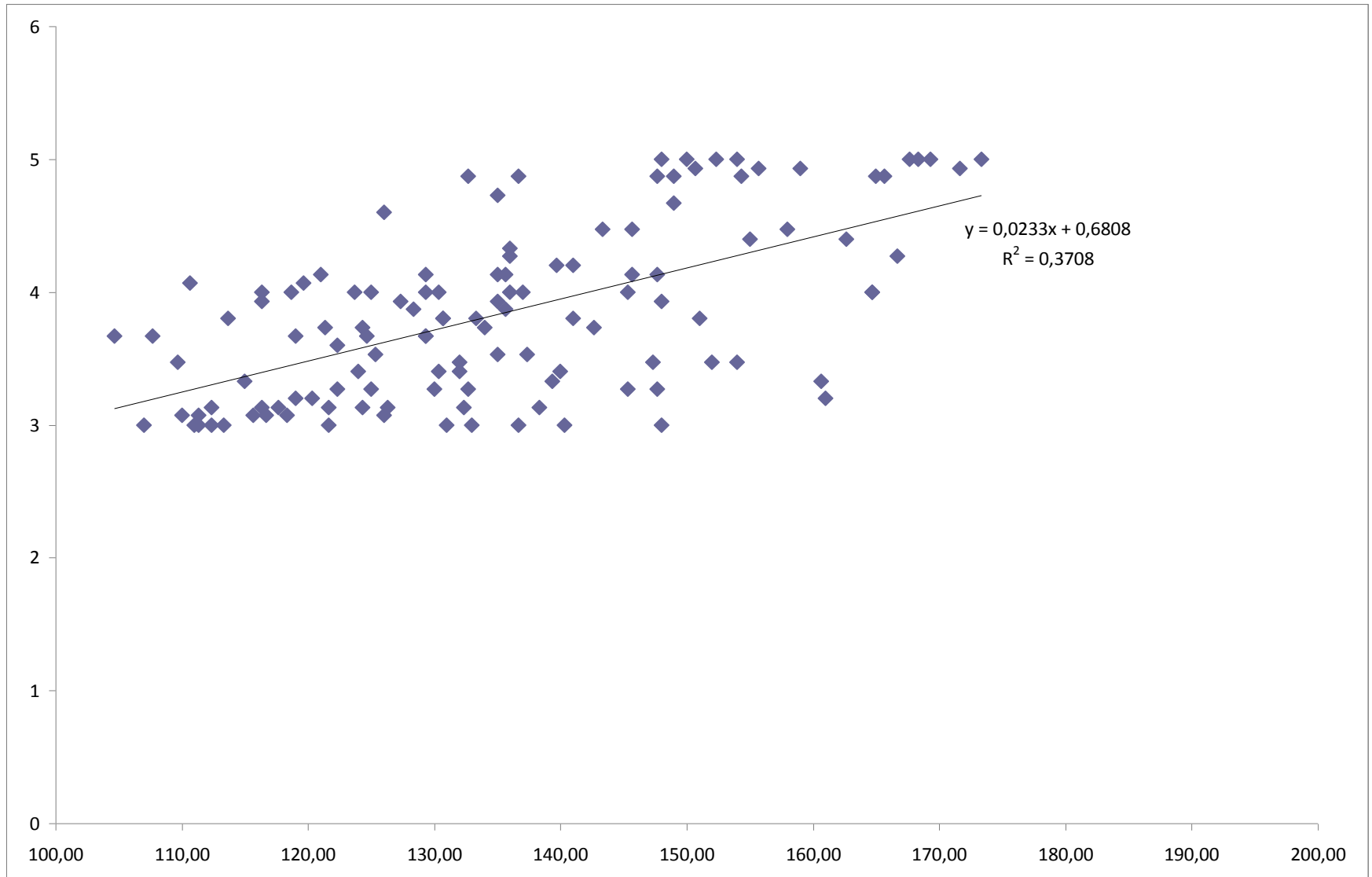
# ПРЕДМЕТНА ВАЛІДНІСТЬ ТЕСТУ

- Відповідність предмету оцінювання
- Мінімізація побічних ефектів

# ПРОГНОСТИЧНА ВАЛІДНІСТЬ ТЕСТУ НА ВІДБІР

- Коефіцієнт кореляції між результатами тесту і подальшою успішністю

# ПРИКЛАД. ЗНО і навчання на магістерських програмах з права



ОЦІНЮВАННЯ РЕЗУЛЬТАТІВ  
НАВЧАННЯ БАКАЛАВРІВ.  
КЛАСИЧНІ І СУЧАСНІ ТЕХНОЛОГІЇ.



# Приклади ЗНО бакалаврів

- **GRE (Graduate Record Examinations)**
- **CLA (The Collegiate Learning Assessment)**
- **Проект ANELO OECD (Assessment of Higher Education Learning Outcomes)**
- **ЗНО бакалаврів медичних ВНЗ в Україні**

# GRE ETS

- Загальний тест
  - Аналітичне письмо (analytical writing)
  - Вербальна аналітика (міркування) (verbal reasoning)
  - Кількісна аналітика (міркування) (quantitative reasoning)
  - Експериментальна або дослідницька частина

# GMAT

- Graduate Management Admission Test -  
математика і англійська;
  - комп'ютерно-адаптивний тест для вступу на програми MBA

# ПРЕДМЕТНІ GRE

- Біохімія
- Біологія
- Клітинна і молекулярна біологія
- Математика
- Фізика
- Хімія
- Мова та література (англійська)
- Психологія

# ПРЕДМЕТНІ GRE (В МИНУЛІ РОКИ)

- Комп'ютерні науки
- Економіка
- Інженерія
- Історія
- Політологія
- Геологія
- Музика
- Соціологія

CLA (The **Collegiate Learning Assessment**) – Тест для оцінювання навчання у коледжах США. З 2000 року

- Завдання з повною відповіддю
- Оцінювання “доданої педагогічної вартості”
- Розділи:
  - Критичне мислення
  - Аналітика
  - Вміння письмової комунікації.

# ПРОЕКТ АНЕЛО

- 17 країн (*Абу-Дабі, Австралія, Бельгія, Колумбія, Єгипет, Фінляндія, Італія, Японія, Корея, Кувейт, Мексика, Нідерланди, Норвегія, Росія, Словаччина, Швеція та Сполучені Штати*)
- До 10 ВНЗ з країни
- Спеціальності:
  - Економіка
  - Інженерія

# АНЕЛО – ЕКОНОМІКА. ОСНОВНІ ВИДИ РЕЗУЛЬТАТІВ НАВЧАЛЬНОЇ ДІЯЛЬНОСТІ

- I. Знання та розуміння предмету
- II. Знання предмету та застосування знань до проблем реального життя
- III. Здатність ефективно використовувати відповідні дані та кількісні методи
- IV. Здатність спілкуватися зі спеціалістами та неспеціалістами в даній галузі
- V. Здатність до самостійного навчання, постановки і вирішення проблем



# АНЕЛО – ІНЖЕНЕРІЯ. ОСНОВНІ ВИДИ РЕЗУЛЬТАТІВ НАВЧАЛЬНОЇ ДІЯЛЬНОСТІ

- I. Загальні інженерні вміння
- II. Фундаментальні та інженерні науки
- III. Інженерний аналіз
- IV. Інженерний дизайн
- V. Інженерна практика

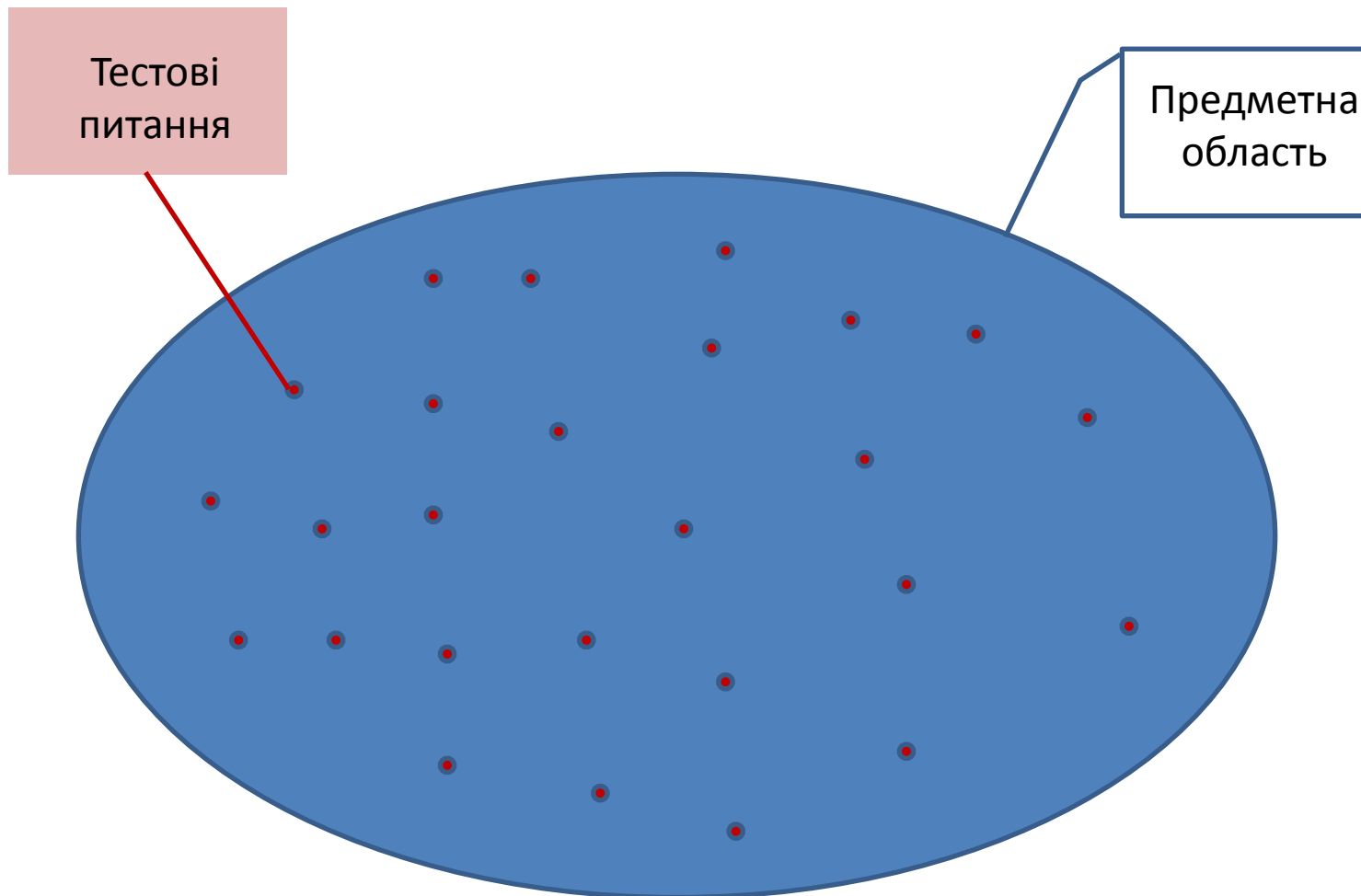
# ОСНОВНІ МЕТОДОЛОГІЧНІ ПРОБЛЕМИ

- Валідність оцінювання
  - конструктна (предметна)
  - конвергентна
  - дискримінантна
  - прогностична
- Надійність оцінювання

# КЛАСИЧНА ТЕОРІЯ ТЕСТІВ

- Надійність і зв'язаність тесту ( $\alpha$ -Кронбаха)
- Роздільна здатність
- Психометричні характеристики кожного завдання

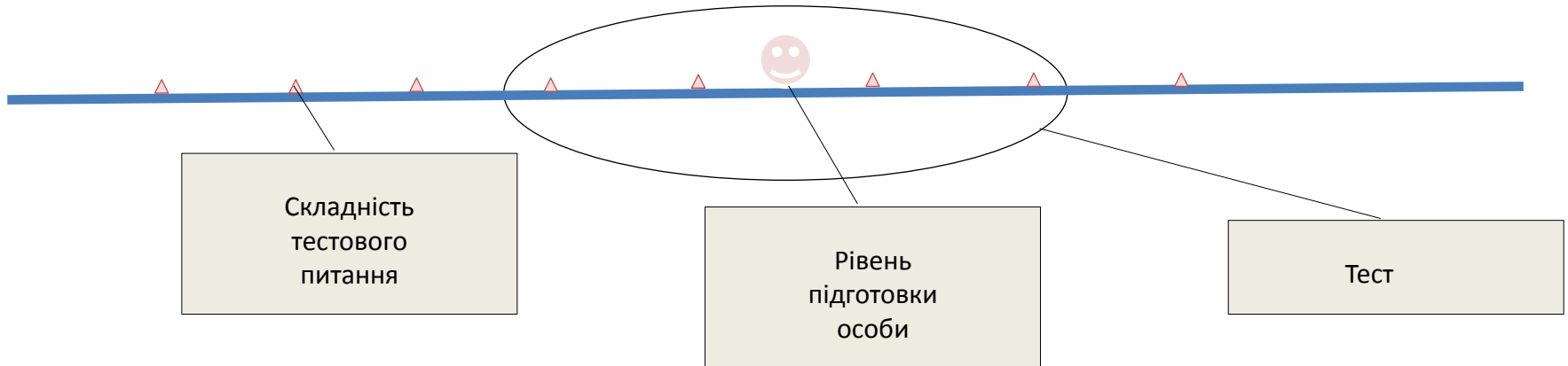
# КЛАСИЧНА ТЕОРІЯ. ПРОБЛЕМИ ВАЛІДНОСТІ І НАДІЙНОСТІ



# ТЕОРІЯ РАША

## (Rush, IRT – Item Response Theory)

- Одновимірний модель тесту – одна латентна змінна
- Завдання статистично незалежні
- Відповідність складності рівню підготовки



# ПЛАН ПІДГОТОВКИ ТЕСТУ

- Визначення предметної області та її розділів
- Визначення основних латентних змінних
- Вибір моделі оцінювання завдань та оцінювання всього тесту
- Побудова специфікації тесту (розділи, кількість завдань, складність завдань)
- Розробка завдань
- Попередня апробація тесту
- Внесення поправок до тесту

# ОЦІНЮВАННЯ

- Проходження тесту
- Оцінювання тестових завдань
- Оцінка за тест (шкалювання)
- Оцінка валідності і надійності тесту

# ПРИКЛАД ПРОСТОГО ТЕСТОВОГО ЗАВДАННЯ З ОСНОВ ІНФОРМАТИКИ

Програма Microsoft Windows є:

- 1) Системою управління базами даних
- 2) Операційною системою
- 3) Текстовим редактором.



# ЩО МОЖЕ ОЗНАЧАТИ ПРАВИЛЬНА ВІДПОВІДЬ:

- Особа володіє основами інформатики
- Особа розуміє відмінність системного і прикладного ПЗ
- Особа перед тестом прочитала словосполучення “операційна система Windows”
- Особа вгадала відповідь

# ЩО МОЖЕ БУТИ ПРИЧИНОЮ НЕПРАВИЛЬНОЇ ВІДПОВІДІ:

- Особа не володіє основами інформатики
- Особа не розуміє відмінність системного і прикладного ПЗ
- Особа не користується комп'ютером з операційною системою Windows
- Особа помилково позначила відповідь

# СУЧАСНІ ТЕНДЕНЦІЇ РОЗВИТКУ ЗНО

- Оцінювання з допомогою комп'ютера
- Адаптивне оцінювання
- Автоматичне формування тесту під час оцінювання
- Автоматичне генерування тестових завдань під час проведення оцінювання
- Оптимізація процесу формування тестової оцінки

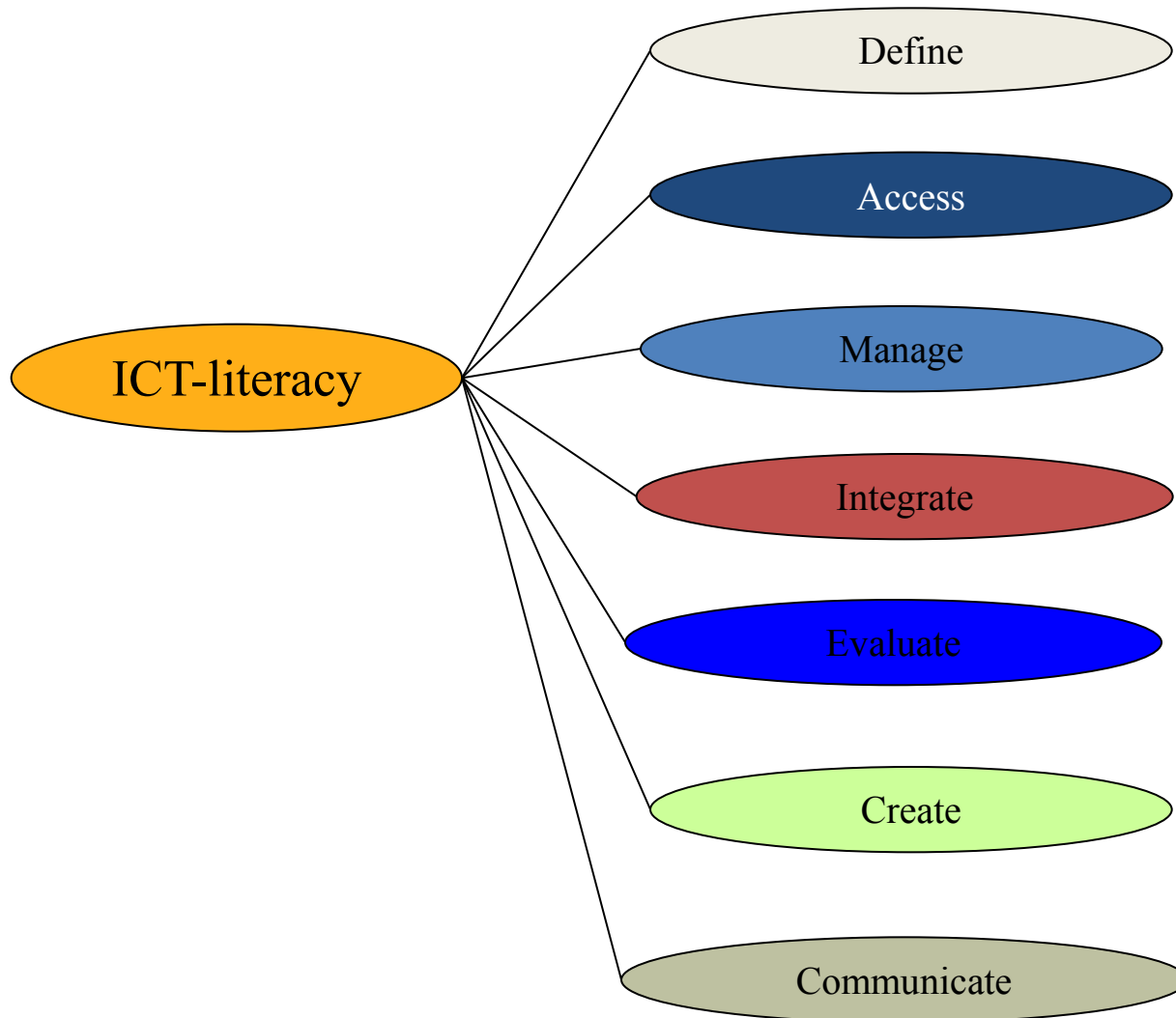
# КОГНІТИВНЕ ОЦІНЮВАННЯ (М.ЗЕЛМАН)

- Комп'ютерне оцінювання
- Адаптивне оцінювання
- Оптимізація оцінки
- Автоматичне генерування завдань

# КОГНІТИВНЕ ОЦІНЮВАННЯ. ТЕОРЕТИЧНІ ОСНОВИ

- Метод ієрархій прийняття рішень (Т.Сааті)
- Теорія ланцюгів Маркова
- Байєсові мережі

# СТРУКТУРА ІНФОРМАЙНО-КОМУНІКАТИВНОЇ ГРАМОТНОСТІ (М.ЗЕЛМАН). ЛАТЕНТНІ ЗМІННІ



Тестовое задание «Работа с почтовым ящиком»:  
автоматическая обработка результатов тестирования

## Измеряемая компетенция: УПРАВЛЕНИЕ

Наблюдаемые признаки/Измеряемые переменные, по которым можно судить о наличии данной компетенции (Observables):

1. Точность в управлении информацией
2. Сохранение информации для того, чтобы ее не потерять
3. Сохранение информации для того, чтобы в дальнейшем ее легко и быстро можно было найти.

Как измерить 1-й наблюдаемый признак/ переменную №1:

- Если участник тестирования сохранил все нужные письма и удалил все остальные (ненужные), то он достигает уровня «ВЫСОКИЙ» по 1-ому наблюдаемому признаку;
- Если задание выполнено на 80-99%, то участник тестирования достигает уровня «СРЕДНИЙ» по данному признаку.
- Если менее 80% писем рассортированы/обработаны правильно, то уровень «НИЗКИЙ».

# ДИЗАЙН ТЕСТУ ОСНОВАНИЙ НА ДОКАЗАХ (EVIDENCE-CENTERED)

- Компетентності
- Докази цих компетентностей
- Побудова байєсової мережі: апріорні оцінки
- Уточнення ймовірностей - уточнення підсумкової оцінки



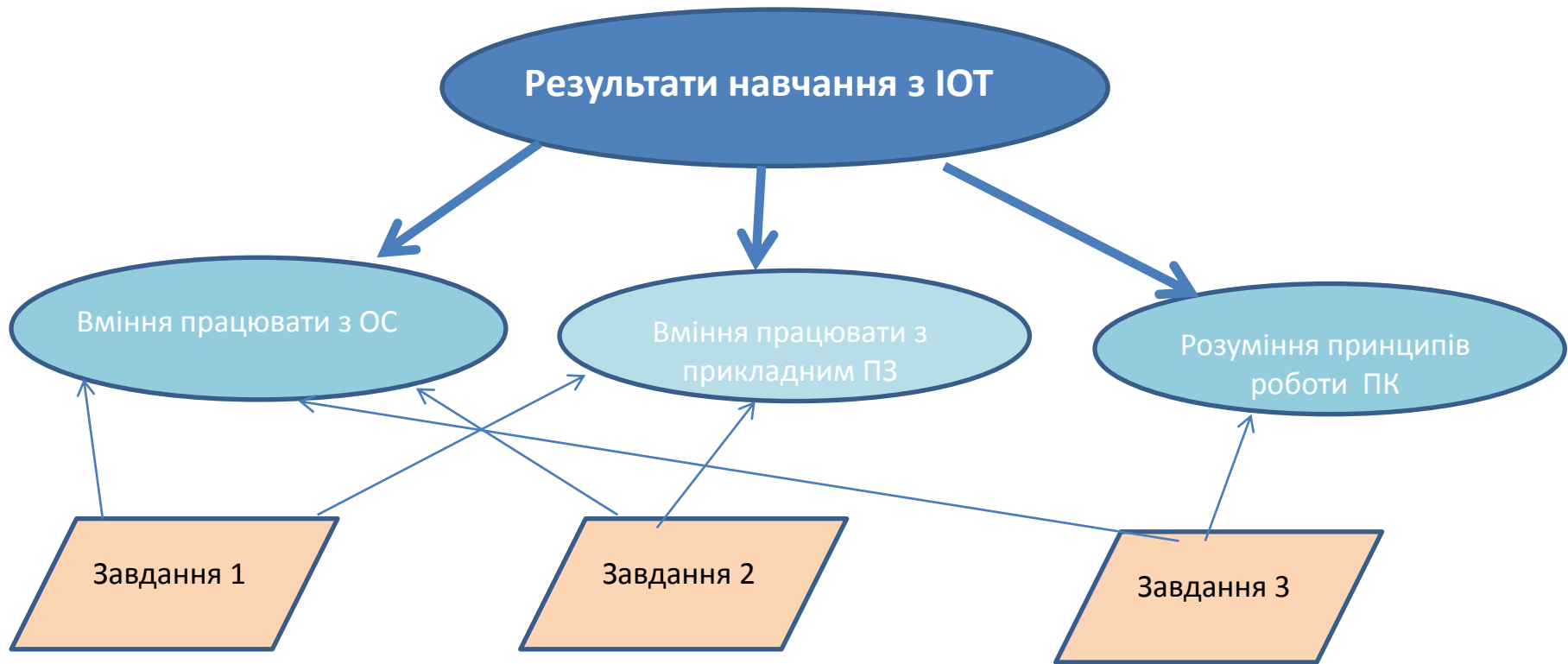
# АПРІОРНЕ ВСТАНОВЛЕННЯ ЗНАЧИМОСТІ ОКРЕМИХ ЛАТЕНТНИХ ЗМІННИХ МЕТОДОМ АНАЛІЗУ ІЄРАРХІЙ Т.СААТІ

Попарне порівняння значимості змінних

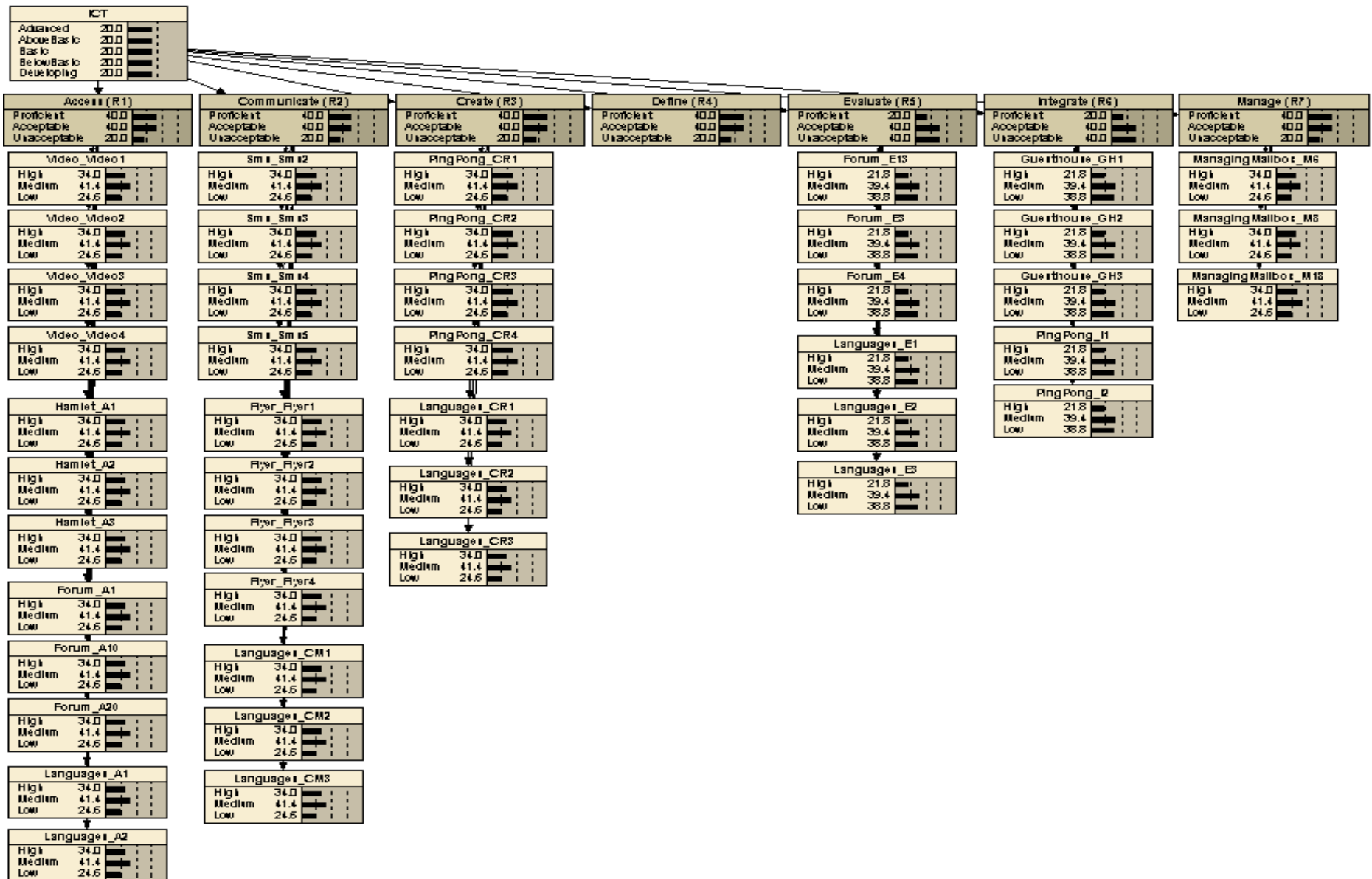
Забезпечення послідовності матриці  
порівнянь

Обчислення значимості змінних

# ПРИКЛАД. ОЦІНЮВАННЯ З ІОТ



# Автоматическая обработка результатов тестирования: сети Байеса



# АВТОМАТИЧНА ГЕНЕРАЦІЯ ТЕСТОВИХ ЗАВДАНЬ

- Банк моделей (замість банку тестових завдань)
- Алгоритми побудови тестових завдань за моделями

# ВИСНОВКИ

- Тестування – лише один з методів оцінювання. Витратний.
- Потрібно визначати доцільність оцінювання у формі тестування.
- Тестування практично незаміниме для ЗНО.
- Старі технології стають вразливими.
- Технології ЗНО мають постійно розвиватися.