Розробка стандартизованого оцінювання

Зміст

Вступ ... 3

Розділ 1. Термінологія ... 4

Тестове завдання як вимірювання конструкту ... 4

Критерії якісних завдань ... 5

Припущення щодо тестових завдань ... 5

Поняття тестового завдання ... 6

Типи завдань ... 6

Розділ 2. Структура тесту ... 7

Відповідність конструкту ... 7

Документи, що визначають структуру тесту ... 8

Специфікації тестових завдань ... 10

Розробка банку тестових завдань ... 12

Розділ 3. Загальні рекомендації до написання тестових завдань та експертної оцінки ... 13

Розділ 4. Написання та експертна оцінка стимулів ... 15

Рівень складності при читанні ... 15

Упередженість та мовна доступність ... 15

Загальні рекомендації ... 16

Універсальність структури ... 17

Розділ 5. Написання та експертна оцінка завдань множинного вибору ... 17

Поняття тестового завдання ... 18

Основа тестового завдання ... 18

Принципи написання та експертної оцінки ключа до основи завдань множинного вибору ... 21

Принципи написання та експертної оцінки дистракторів до основи завдань множинного вибору ... 21

Написання та експертна оцінка основи завдання множинного вибору ... 22

Написання та експертна оцінка ключів до основ завдань множинного вибору ... 28

Написання та експертна оцінка дистракторів до основи завдання множинного вибору ... 30

Розділ 6. Написання, експертна оцінка та оцінювання завдань відкритого типу ... 35

Складові завдань відкритого типу ... 35

Написання та експертна оцінка інструкцій до завдань відкритого типу ... 35

Написання та експертна оцінка схеми оцінювання завдань відкритого типу 36
Розділ 7. Технічні характеристики тестових завдань 41
Коефіцієнт бісеріальної кореляції 41
Бісериальна кореляція дистракторів 42
Показник Р 42
Квантиль-аналіз 42
Диференційне функціонування завдання (DIF) 43
Теорія відповіді на завдання (IRT) 44
Розділ 8. Укладання та експертна оцінка тестових варіантів 45
Укладання тестових варіантів 45
Технічна якість тесту 46
Контрольні примірники 48
Діазотипні копії тесту 49
Глосарій 52
Додаток 53
Вправа 1 53

ВСТУП

У даному посібнику надаються загальні рекомендації щодо процесу та процедур розробки та проведення експертної оцінки тестових завдань і тестових варіантів, результатом яких є розробка тестів, що мають високу якість і не містять помилок. Надані в даному посібнику рекомендації відносяться до тестів, призначених для усіх років навчання та з усіх предметних галузей.

У посібнику надаються загальні нормативні вимоги стосовно побудови тесту, написання стимулів та питань, а також стосовно розробки тестових варіантів та їх експертної оцінки. Крім цього, у посібнику розглядаються найпоширеніші недоліки, які необхідно навчитися знаходити та виправляти розробникам тестів та експертам, які надають висновки.

Посібник складається з восьми розділів:

- Термінологія
- Структура тесту
- Загальні рекомендації до написання тестових завдань та експертної оцінки
- Написання та експертна оцінка стимулів
- Написання та експертна оцінка завдань множинного вибору
- Написання, експертна оцінка та оцінювання завдань відкритого типу
- Технічні характеристики тестових завдань
- Укладання та експертна оцінка тестів

Кожний розділ починається із загального огляду, іноді містить ілюстративний приклад або обговорення.

РОЗДІЛ 1. ТЕРМІНОЛОГІЯ

Як і в інших технічних або дослідницьких галузях, теорія оцінювання має власний спеціалізований словник. Перш ніж розглядати завдання, необхідно знати та зрозуміти термінологію. Даний розділ починається з пояснення терміну «конструкт», що є предметом дослідження під час виконання тестового завдання. Потім розглядається те, що дозволяє зробити тестове завдання якісним. Наприкінці надаються визначення ключових термінів, якими користуються при розробці тестових завдань.

ТЕСТОВЕ ЗАВДАННЯ ЯК ВИМІРЮВАННЯ КОНСТРУКТУ

Пізнавальні здібності, якими користуються учні для того, щоб відповісти на завдання, часто називають "психологічним конструктом." До психологічних конструктів відносять такі як вербальні або математичні навички, наукове або історичне розуміння, а до вужчих конструктів – вміння виконати певну математичну дію або написати твір-оповідання. Конструкти можуть існувати тільки гіпотетично, оскільки їх неможливо спостерігати безпосередньо в думках учня. Єдиний спосіб, завдяки якому можна зробити висновок про існування конструкту в думках учня або про його міру – це спробувати виміряти його опосередковано через механізм стимулів та відповідей. Стимулом може бути усне екзаменування учня експертом, який добре знає конструкт або завдання виконати лабораторний експеримент у присутності експерта. Так чи інакше, фахівець оцінює відповіді студента або його поведінку і робить висновок про наявність конструкту, а, можливо, і про ступінь його володіння учнем. Стандартизовані тести не є єдиним способом в технології «стимул – «відповідь», це просто більш ефективний спосіб її застосування.

З іншого боку, значна небезпека стимулу та відповіді полягає в тому, що стимул може стати недосконалим способом отримати бажану відповідь від студента, і що відповідь насправді не буде доказом наявності відповідного конструкту. Ця небезпека є дуже реальною у стандартизованих тестах, особливо, якщо конструкт, яким цікавляться, є комплексним, не може бути адекватно представленим у тестових завданнях чи підбірці завдань, або не дозволяє учаснику тестування надати відповідь, яка б якнайкраще відображала його/її здібності або знання. Навіть якщо можна уникнути цієї небезпеки, виникають подальші проблеми стосовно належної інтерпретації відповіді студента, а відтак і наявності у нього конструкту. Деякі конструкти легше оцінити під час оцінювання процесу виконання, ніж під час оцінювання на папері. Проте деякі конструкти навіть краще оцінювати саме за допомогою олівця та паперу, виконуючи завдання множинного вибору, тоді як інші можна належною мірою оцінити за допомогою завдань відкритого типу.

В даному посібнику не розглядаються психологічні теорії пізнання, проте питання, які вони ставлять перед авторами та експертами тестових завдань, є надзвичайно важливими. Конструкт, який необхідно оцінити під час виконання тестового завдання, повинен якомога ближче відповідати конструкту, який зазначено або який передбачається в специфікації змісту завдання і специфікації певних здібностей, які студент повинен застосувати, щоб виконати завдання. Ця бажана відповідність називається «відповідністю» і докладніше розглядається нижче. Проте навіть маючи гарну відповідність, завдання можуть містити недоліки, що можуть зменшити вірогідність того, що завдання точно оцінить відповідний конструкт. Наприклад, якщо є підказки в завданнях множинного вибору, які дають вказівки на правильну відповідь, або якщо є двозначність в завданні множинного вибору або у завданні відкритого типу, оцінка відповідного конструкту не буде точною. Багато інших аспектів у тестовому завданні можуть приховувати можливості для заплутування. До таких аспектів належить питання правильної постановки запитання, чіткого формулювання питання, а також представлення у завданнях множинного вибору правдоподібних дистракторів, які не дозволяють спрямувати учня до правильної відповіді.

Такі рекомендації стосовно того, що слід робити, і чого потрібно уникати, розглядаються у наступному розділі посібника.

КРИТЕРІЇ ЯКІСНИХ ЗАВДАНЬ

Може здатися дивним, що ми починаємо розділ про що слід робити та чого потрібно уникати із твердження про обмеженість цих понять. Проте, питання полягає саме в тому, що якісні завдання, це дещо більше, ніж наявність того, що слід робити, та відсутність того, чого потрібно уникати при написанні завдання. Якісні завдання вимагають творчості або навіть мистецької майстерності, які не можна недооцінювати. Наприклад, прийнятні дистрактори можуть варіювати від максимально правдоподібного до відверто оманливого. Жодне з зазначених правил стосовно правдоподібності дистракторів не може навчити, як потрібно писати дистрактори або надавати експертну оцінку. Здібність до концептуального мислення, необхідна для написання якісних дистракторів – а це також є психологічним конструктом – є талант, який автор завдання використовує при його написанні. Але цей талант максимально проявляється, якщо дистрактори не порушують жодного з правил.

ПРИПУЩЕННЯ ЩОДО ТЕСТОВИХ ЗАВДАНЬ

Автори завдань та експерти мають усвідомлювати не тільки правила написання якісних завдань, але й певні припущення, які робитимуть стосовно тестових завдань, коли настане час інтерпретувати тестові бали. Жодне з двох припущень, які розглядаються в даному посібнику, не мають бути неочікуваними, оскільки обидва пов'язані з попереднім твердженням про те, що завдання повинні бути якомога чистішим інструментом тестування відповідного психологічного конструкту.

Першим припущенням є «одновимірність» завдання. Це означає, що тестове завдання перевіряє тільки одну окрему здібність і не більше. Наприклад, за допомогою завдання з математики на основі певного ситуативного стимулу також можна перевірити навички читання, особливо, якщо рівень тексту до завдання є вищим за рівень навичок читання, властивий даному року навчання, разом із математичними здібностями, необхідними для того, щоб правильно відповісти на завдання. Про учня, який надає неправильну відповідь на таке завдання, буде зроблено припущення про недостатність здібностей, якими необхідно скористатися для надання відповіді (хоча такі висновки ніколи не роблять на основі одного завдання), навіть у тому випадку, коли учень і знав би правильну відповідь, якби він зміг належним чином усвідомити прочитаний текст завдання. Можна навести багато інших прикладів стосовно завдань, в яких принцип одномірності не виконується; це такі завдання, коли передбачається, що, надаючи відповідь, учень використовує одні здібності (наприклад, міркування), проте деякі учні у своїй відповіді використовують інші здібності (наприклад, запам'ятовування). Одномірність загалом безпосередньо випливає з характеристик учасників тестування: хто проходить тестування, як учнів навчали і що очікується від учнів при спробі надати відповідь на завдання.

Друге припущення щодо завдань називається принципом «локальної незалежності». Це означає, що відповідь учня на будь-яке завдання не впливає на інші відповіді на завдання тесту і є статистично незалежною від них. Незалежність найчастіше порушують, коли є підказка в одному завданні для відповіді на інше завдання. У такому випадку, дійсні здібності учня на другому завданні не оцінюються. Автори завдання можуть гарантувати, що вимога щодо локальної незалежності задовольняється тільки стосовно окремих завдань, які вони готують. Це виключає завдання, які розробники пишуть як альтернативні одне одному. Саме укладач тесту повинен бути уважним, щоб не допустити порушення принципу локальної незалежності у тесті в процесі підбору завдань із різних джерел.

Звичайно, легко побачити, де одне завдання може давати підказку для іншого. Набагато складніше побачити, де одне завдання може винагородити або покарати студента двічі за володіння або відсутність необхідного знання або навичок. Це так само є порушенням локальної незалежності, що часто трапляється в тих збірках завдань, де два або більше завдання пов'язані зі спільним стимулом. Наприклад, якщо стимулом у тесті із історії мистецтва є репродукція картини, і у першому завданні запитують, виходячи із художнього стилю, хто був автором картини, наступне завдання не повинно містити питання, коли картина була створена; дуже ймовірно, що учень, який правильно назвав художника, назве правильну дату створення картини, тоді як для учня, який назвав художника неправильно, ситуація буде протилежною. Такий подвійний ризик/ подвійна винагорода у написанні завдання є порушенням принципу локальної незалежності. Щоб зробити завдання в цьому прикладі незалежним одне від одного, друге завдання може містити питання про те, яку роль кольору, перспективі, тощо відведено у цьому творі живопису. Але навіть тут, автор завдання повинен зважати, чи учень, який знає художника, отримає перевагу під час аналізу стилістичних елементів живопису.

ПОНЯТТЯ ТЕСТОВОГО ЗАВДАННЯ

Тестове завдання – це кожне окреме питання, а для деяких типів завдань, це також набір потенційних відповідей на тест. Часто в стандартизованих тестах є багато «питань», які потенційно мають кілька відповідей. **Тестове завдання –** це термін, який використовують для позначення як питання, так і відповідей у завданнях множинного вибору. **Тестове завдання** також позначає питання відкритого типу, хоча відповідь учня не вважається частиною завдання.

ТИПИ ЗАВДАНЬ

Існує три типи тестових завдань: завдання множинного вибору, завдання відкритого типу та завдання на виконання.

Завдання множинного вибору зазвичай складають найбільшу частину процесу оцінювання. Ці завдання надають від трьох до п'яти варіантів відповіді, лише один з яких є вірним. Завдання множинного вибору оцінюють головним чином вміння розпізнати інформацію.

Завдання відкритого типу є такими, що спонукають студента створити власну відповідь без наданих варіантів, з яких можна вибрати відповідь. Залежно від галузі знань, відповіддю на таке завдання може бути одне чи кілька слів або число, або письмова відповідь обсягом кілька абзаців чи складне рівняння. Завдання відкритого типу оцінюють головним чином вміння пригадати інформацію.

Завдання на оцінку виконання складається із завдання або серії завдань, які студент має виконати. (В даному посібнику такий тип оцінювання не розглядається.)

РОЗДІЛ 2. СТРУКТУРА ТЕСТУ

ВІДПОВІДНІСТЬ КОНСТРУКТУ

Конструкти можуть бути широкими, як вербальні або кількісні здібності, або конкретними, подібно до компонентів змісту, що містяться в стандартах для певного року навчання. Можна тільки будувати гіпотези щодо конструктів, тому що їх не можна спостерігати безпосередньо. Як було зазначено у Розділі 1, єдиним способом пересвідчитися в існуванні конструкту у свідомості учня є опосередковане вимірювання конструкту шляхом оцінювання.

Проте існує певна небезпека в тому, що оцінювання може бути недосконалим методом виявлення бажаних знань конструкту учня, тобто оцінювання насправді не надає свідчень щодо конструкту, який нас цікавить. Ця небезпека є реальною при стандартизованому тестуванні, особливо, якщо конструкт, що нас цікавить, є складним, не може бути адекватно втіленим у тестовому завданні або низці завдань, не дозволяє учню відповісти якнайкраще або не є чітко та адекватно представленим в структурі тесту.

Крім того, розробники та експерти повинні пам'ятати про мету тестування. Якщо метою тестування є об'єктивне визначення академічних досягнень учнів певного року навчання, то тест має вимірювати тільки матеріал, що відноситься саме до цього року навчання, і тест не повинен містити компоненти змісту, що можуть заплутати учнів. Особливо це стосується рівня читання. Якщо тести є вирішальними, вони повинні відповідати більш суворим вимогам, але це не є предметом обговорення у даному посібнику. Розробники та експерти повинні розуміти, що всі складові тесту перевірятимуться дуже ретельно при виникненні судових позовів.

Розробники тестів також повинні усвідомлювати, що деякі припущення стосовно завдань в тесті відносяться до етапу інтерпретації тестових балів. Одним з таких припущень є те, що питання у тестових варіантах оцінюють тільки той обсяг знань або здібностей, для оцінювання яких призначений даний тест. Наприклад, тест із природничих дисциплін, який включає багато розширених текстів-сценаріїв крім перевірки знань з природничих дисциплін може перевіряти вміння читання. Це особливо стосується випадків, коли рівень складності текстів дорівнює або перевищує рівень, що відповідає року навчання, і це заважає деяким учням надавати правильні відповіді на тестові завдання. Неправильні відповіді учня на такі завдання розцінюються як свідчення недостатніх знань конструкту, на оцінювання якого спрямований тест (природничі дисципліни), навіть якщо учень мав би змогу надати правильну відповідь, якби він чи вона змогли прочитати текст завдання з адекватним розумінням змісту. У такому тесті змішуються знання природничих дисциплін з навичками читання.

Оцінювання на рівні країни або регіону/штату надають корисну інформацію для прийняття валідних рішень та покращення освітніх процесів в тій мірі, в якій оцінювання відповідає вимогам академічних стандартів країни або регіону. Якщо тести не забезпечують адекватного оцінювання знань та вмінь, зазначених в академічних стандартах штату щодо змісту, або якщо вони вимірюють те, що не увійшло до стандарту, буде складно визначити, чи досягли учні бажаного рівня знань та умінь. Внаслідок цього буде складно визначити відповідну політику, тобто рішення стосовно навчальних програм та викладання.

Відповідність має кілька вимірів: щоб показати відповідність системи оцінювання стандартам штату, потрібно більше ніж просто визначити, чи всі завдання в тесті співпадають з одним чи кількома стандартами або чи кожний з визначених стандартів змісту відповідає одному чи кільком завданням в оцінюванні. *Відповідність – це більше ніж двосторонній процес*. Для забезпечення відповідності стандартів та оцінювання

необхідно враховувати наступне:

- ♦ Чи охоплює оцінювання повний обсяг змісту, вказаного у академічних стандартах змісту, що означає, що всі стандарти правомірно представлені в оцінюванні, *та*
- ♦ Чи вимірює оцінювання як зміст (що знають учні), так і процес (що учні можуть зробити), зазначені в академічних стандартах змісту, *та*
- ♦ Чи оцінювання відображає той ступінь та ті акценти, що представлені в академічних стандартах змісту (наприклад, якщо академічні стандарти змісту багато уваги приділяють операціям, так само має і оцінювання), *та*
- ♦ Чи відображає оцінювання повний спектр когнітивної складності та рівні складності і глибина розкриття понять та процесів, зазначених в академічних стандартах змісту, тобто вимоги оцінювання мають бути такими ж, як і вимоги стандартів, *та*
- ♦ Чи надає оцінювання результати, що представляють всі рівні досягнень, описаних в рівнях академічної успішності.

Щоб допомогти забезпечити відповідність академічних стандартів змісту та оцінювання, потрібно ретельно продумувати структуру тесту та розробити документи з рекомендаціями для укладання та експертної оцінки кожного варіанту тестів. Документи, що визначають структуру тесту, використовуються для укладання збору варіантів тестів та для створення банку завдань, яким розробники тестів користуються для розробки варіантів тестів. Ці документи забезпечують те, що кожний новий варіант тестів можна порівняти з будь-яким іншим варіантом, який використовувався у тестуванні раніше.

ДОКУМЕНТИ, ЩО ВИЗНАЧАЮТЬ СТРУКТУРУ ТЕСТУ

Щоб зменшити ймовірність того, що тест оцінює не той конструкт, який він призначений оцінювати, необхідно ретельно дотримуватися рекомендацій, зазначених в документах, що визначають структуру тесту. Відповідність таким документам забезпечує те, що варіанти тестів під час проведення різних тестувань є порівнюваними.

Документи, що регламентують структуру тесту, включають інформацію про бали, які будуть повідомлені учасникам тестування, про типи тестових завдань, час, відведений на виконання тесту та іншу інформацію, яка залишається незмінною для всіх тестових сесій. Наприклад, якщо тест призначений оцінити навички читання, документ, що визначає структуру тесту, може включати інформацію про типи уривків для читання у кожному варіанті тесту, кількість уривків у тесті, кількість слів в уривках кожного типу, загальну кількість слів в усіх уривках в тесті, кількість та типи завдань, пов'язаних з кожним уривком, стандарти, що оцінюються за допомогою уривків кожного типу, рівень складності уривків в тесті, а також статистичні характеристики тесту.

ПРИКЛАД

Структура тесту

Мета структури тесту

Опис структури тесту має використовуватися працівниками відділу освіти штату Огайо, а також підрядниками відділу освіти з тестування, які здійснюють контроль за укладанням варіантів тестів для забезпечення паралельності варіантів, коли це є потрібним, та найширшого охоплення змісту в усіх варіантах тесту.

Обсяг тесту

Як зазначалося вище у специфікаціях тесту, кожний варіант тесту досягнень з читання для 8-го класу містить загалом 38 завдань (32 завдання множинного вибору, 4 завдання на коротку відповідь та 2 завдання на розгорнуту відповідь) із загальною кількістю балів 48.

	Завдання множинного вибору	Завдання з короткою відповіддю	Завдання з розгорнутою відповіддю	Загалом
Загальна кількість завдань	32	4	2	38
Бали	32	8	8	48

Формат тесту та послідовність тестових завдань

Уривки для читання розташовані у послідовності, описаній в тестових специфікаціях.

Уривок 1 – довгий літературний чи інформативний (може включати парні уривки)

Уривок 2 – середній інформативний

Уривок 3 – середній літературний

Уривок для апробації та блок завдань до нього

Уривок 4 – короткий інформативний

Уривок 5 – короткий літературний

Кожний уривок для читання супроводжується завданнями. Завдання множинного вибору розташовуються згідно з послідовністю подій чи інформації в тексті. Відкриті завдання не повинні розташовуватися на початку, в кінці, або у кілька завдань разом. Наскільки можливо, з урахувань обмежень щодо послідовності розташування завдань, відкриті завдання мають розташовуватися згідно з послідовністю подій чи інформації в уривку.

Кількість і тип включених для апробації завдань

Як зазначалося вище, кожний варіант тесту включатиме уривок для апробації та блок з шести завдань до нього. Блок завдань міститиме чотири або п’ять завдань множинного вибору, міститиме одне або жодного завдання на коротку відповідь та одне або жодного завдання на розгорнуту відповідь. Блок завдань не міститиме водночас і завдання на коротку відповідь, і завдання на розгорнуту відповідь. Якщо операційний варіант (варіант, по якому зараховуються бали) міститиме три завдання на розгорнуту відповідь, блок завдань для апробації не міститиме додаткового завдання на розгорнуту відповідь.

Охоплення змісту

Завдання використовуватимуться для оцінювання чотирьох стандартів: володіння лексикою; поняття друку, стратегії самоконтролю та розуміння прочитаного; інформативний, технічний, спонукальний тексти; та літературний текст.

Кожний стандарт матиме мінімальну кількість балів в кожному варіанті.

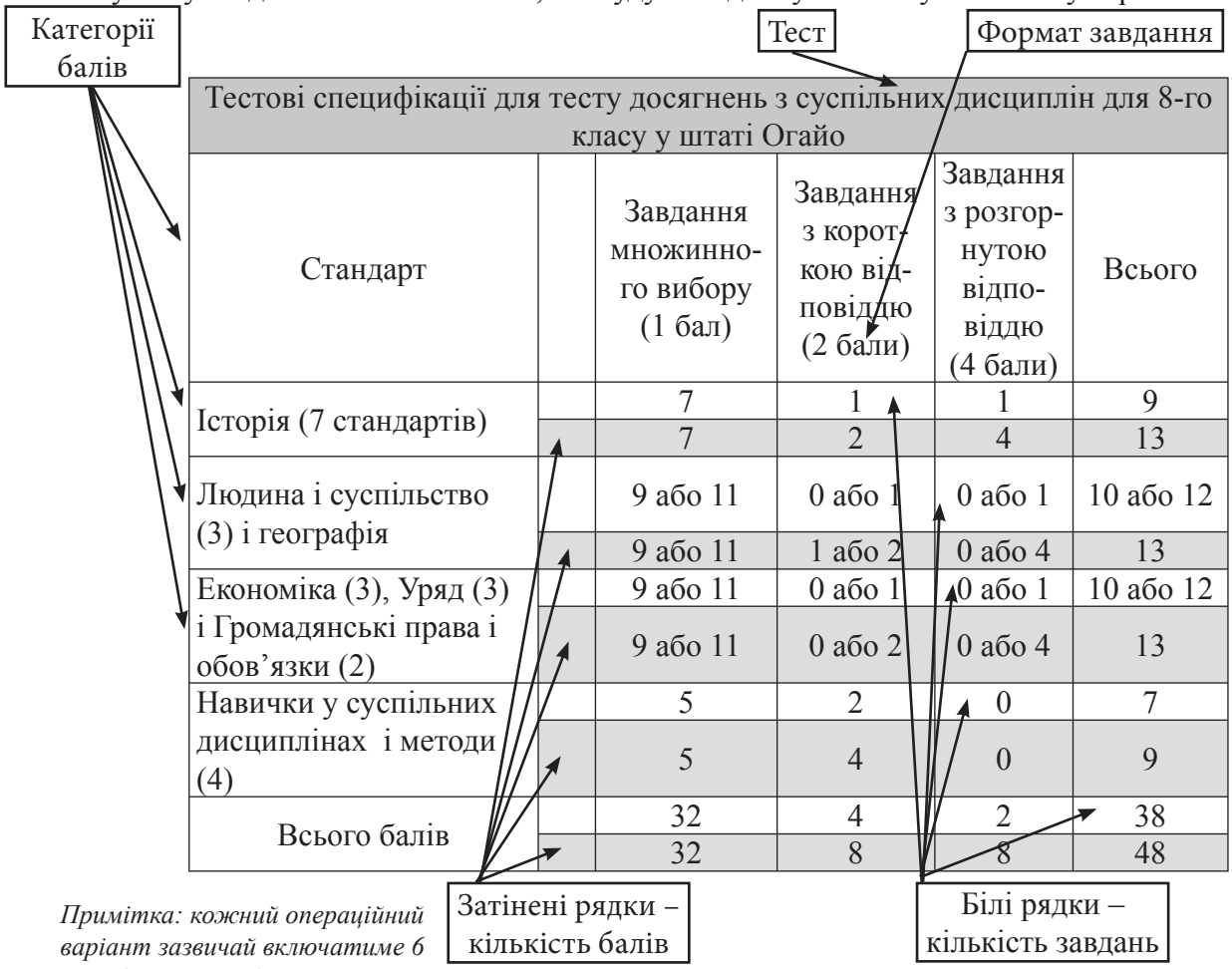
Бали по кожному стандарту	
Лексика	9
Процес читання	10 – 12
Інформативний текст	14 – 18
Літературний текст	12 – 18

ТЕСТОВІ СПЕЦИФІКАЦІЇ

Тестові специфікації є одним із способів кодифікації відповідності кожного варіанту тесту до конструкту, для оцінювання якого створений тестовий варіант. Зазвичай, типові тестові специфікації містять перелік категорій оцінювання разом із кількістю балів по кожній з категорій, а також типи завдань, які увійдуть до варіанту тестів по кожній категорії. Крім того, вказуються конкретні стандарти по кожній категорії, а також кількість і типи завдань для кожного стандарту.

ПРИКЛАД. ТЕСТОВІ СПЕЦИФІКАЦІЇ

Тестові специфікації описують по кожній категорії стандартів кількість завдань по кожному типу завдань і кількість балів, які будуть надані у кожному тестовому варіанті



Тестові специфікації для тесту досягнень з суспільних дисциплін для 8-го класу у штаті Огайо					
Стандарт		Завдання множинного вибору (1 бал)	Завдання з короткою відповіддю (2 бали)	Завдання з розгорнутою відповіддю (4 бали)	Всього
Історія (7 стандартів)		7	1	1	9
		7	2	4	13
Людина і суспільство (3) і географія		9 або 11	0 або 1	0 або 1	10 або 12
		9 або 11	1 або 2	0 або 4	13
Економіка (3), Уряд (3) і Громадянські права і обов'язки (2)		9 або 11	0 або 1	0 або 1	10 або 12
		9 або 11	0 або 2	0 або 4	13
Навички у суспільних дисциплінах і методи (4)		5	2	0	7
		5	4	0	9
Всього балів		32	4	2	38
		32	8	8	48

Примітка: кожний операційний варіант зазвичай включатиме 6 апробаційних завдань

СПЕЦИФІКАЦІЇ ТЕСТОВИХ ЗАВДАНЬ

Документи, що визначають специфікації тестових завдань, надають детальну інформацію про завдання, які будуть розроблені для оцінювання кожного стандарту або показника оцінювання. Документи, що визначають специфікації завдань, включають типи завдань та рамки змісту для кожного стандарту або показника, а також зразки завдань, щоб показати формулювання завдань по кожному з стандартів або показників. Цей документ використовується для надання рекомендацій для написання тестових завдань та для навчання розробників завдань.

ПРИКЛАД

***Стандарт*: Біологічні науки (Науки природознавчого циклу)**

Стандарт:

D. Пояснити як відбувається вимирання видів, коли відбуваються зміни навколишнього середовища та адаптивні характеристики є недостатніми для виживання.

Показники:

07-4 Визначте, як перенаселення впливає на екосистему.

07-5 Поясніть, що деякі зміни навколишнього середовища відбуваються швидко, а деякі – повільно (зміна лісів та водойм, пожежі і декомпозиція).

08-4 Поясніть, що різноманітність видів складалася завдяки поступовим процесам протягом багатьох поколінь.

08-5 Поясніть, як організм, адаптований до особливого навколишнього середовища, може зникнути, якщо середовище зміниться.

Рамки змісту:

- Різноманітність, адаптація, види та їх зв'язок зі змінами навколишнього середовища
- Зміни в організмі, обумовлені факторами навколишнього середовища
- Докази еволюційних процесів.
- Докази змін навколишнього середовища

Не оцінювати:

Щоб визначити верхні обмеження складності змісту, порівняйте з Класом 10 LS Шаблони E, H, I , що стосуються змін організму з часом та факторів і процесів, що обумовлюють ці зміни.

Типи завдань: Завдання множинного вибору, завдання на коротку відповідь та на розширену відповідь.

Атрибути стимулів:

Завдання для оцінки шаблону/індикатору можуть включати, але не обмежуються ними, наступні атрибути стимулів:

- Діаграми ланцюгів їжі, які демонструють зміни навколишнього середовища
- Графіки змін навколишнього середовища крізь час
- Сценарії
- Графіки змін копалин
- Зображення організмів

Атрибути відповідей:

MC Опції можуть включати, але не обмежуються ними, наступні атрибути:

- Характеристики/пояснення/передбачення змін в екосистемі, обумовлених змінами навколишнього середовища

Дистрактори

- Неправильні пояснення/характеристики, що репрезентують загальні неправильні тлумачення для дистракторів

SA/ER Відповіді, можуть включати, але не обмежуються ними, наступні атрибути:

- Характеристики і пояснення змін в екосистемі, обумовлених змінами навколишнього середовища
- Передбачення, яким чином організми і екосистема можуть реагувати довгостроково на специфічні зміни навколишнього середовища

РОЗРОБКА БАНКУ ТЕСТОВИХ ЗАВДАНЬ

При розробці банку тестових завдань важливо ретельно дотримуватися тестових специфікацій. Якщо, наприклад, у специфікаціях зазначено, що до кожного стандарту необхідно мати завдання на коротку відповідь, банк повинен містити завдання на коротку відповідь для кожного стандарту. Таким чином, якщо в певній специфікації зазначено, що для кожного стандарту потрібно мати 20% завдань в певній категорії балів, тоді банк повинен містити мінімум 20% завдань в цій категорії балів.

Також важливо включати весь діапазон рівнів складності у завданнях для певного стандарту. Наявність всього діапазону рівнів складності дозволяє бути більш гнучким при укладанні тестового варіанту. Також банк повинен містити завдання, що співпадають або перевищують рівень знань, передбачений для кожного стандарту. Це дозволяє забезпечити належну відповідність тестових варіантів.

РОЗДІЛ 3. ЗАГАЛЬНІ РЕКОМЕНДАЦІЇ ДО НАПИСАННЯ ТЕСТОВИХ ЗАВДАНЬ ТА ЕКСПЕРТНОЇ ОЦІНКИ

Метою будь-якого типу оцінювання є вимірювання певного конструкту та формулювання належних висновків з отриманих даних. Найбільш точні висновки можна зробити, коли поставлені питання є зрозумілими, чіткими та однозначними. Неоднозначних питань слід уникати, тому що вони вносять необ'єктивність, що знижує валідність висновків.

Перша перевірка, яка має бути зроблена при наданні експертної оцінки завданню, це визначити, чи дійсно завдання перевіряє саме те вміння, що передбачалося при написанні завдання. Таким чином, кожне завдання має також включати визначення вміння, стандарту або показника, який має оцінюватися завданням.

Якщо вміння входить до тестових специфікацій, наступним кроком для експерта буде перевірка специфікацій завдання, щоб визначити чи формат тесту (завдання множинного вибору, завдання з короткою відповіддю або завдання з розгорнутою відповіддю) відповідає цьому вмінню. Якщо формат завдання не підходить для даного вміння, експерт має відхилити завдання або попросити автора представити завдання у належному для цього вміння форматі.

Якщо формат завдання відповідає вказаному вмінню, експерт може перейти до розгляду того, чи дійсно основа завдання або його інструкція оцінює вказане вміння. Це знову може бути визначено перевіркою специфікацій завдань, що мають приводити перелік формулювань для оцінювання даного вміння. Якщо формулювання основи завдання або інструкції не відповідають затвердженому формулюванню, то завдання все ще може бути прийнятним для перевірки цього вміння. Усі можливі формулювання для завдання, що оцінюють вміння, можуть не бути включені до специфікацій завдання, таким чином, експерт повинен визначити на власний розсуд, чи оцінює основа завдання або інструкція відповідне вміння. Щоб зробити це, експерт має бути ознайомлений з призначенням вміння (стандарту, показника, тощо), яке оцінюється. Такі відомості можна отримати, ознайомившись із специфікаціями тесту та завдання, а також переглянувши затверджені завдання в банку та завдання, які знаходяться на апробації або завдання в операційних тестових варіантах.

Якщо експерт впевнений у тому, що завдання не вимірює потрібне вміння, тоді він або вона має визначити, чи оцінює це завдання інше вміння, зазначене в тестових специфікаціях. Якщо експерт вважає, що завдання не оцінює належним чином жодного з наведених вмінь, він або вона має відхилити завдання.

Якщо експерт вирішує, що завдання належним чином оцінює іншу навичку, він має надати завданню іншу класифікацію, що відповідає цьому вмінню. Крім того, експерт має визначити, чи потрібні у цьому банку тестових завдань додаткові завдання, що оцінюють відповідне вміння. Може бути, що банк містить надлишок завдань, що оцінюють це окреме вміння, просто тому, що ці завдання прості для написання. Однак, якщо банк завдань містить занадто багато завдань на одне й те саме вміння та недостатньо завдань до іншого вміння, тоді процес укладання остаточного варіанту тесту може опинитися під загрозою. Таким чином, експерт має визначити, чи потрібні у цьому банку завдання для вміння, якому відповідає це завдання, та, виходячи з цього, прийняти або відхилити перекласифіковане завдання.

Перед тим, як відхилити завдання, експерт має визначити, чи містить це завдання інший контекст або інший підхід, що відрізняється від сукупності існуючих завдань. Якщо завдання відрізняється від зібраних завдань, тоді його потрібно залишити. Якщо завдання відхилене, тоді автор завдань має надати додаткове завдання для вміння, що передбачалося спочатку.

В процесі створення і накопичення завдань проводяться розширені експертизи на відповідність і якість завдань. Чим більш чисельними є перевірки, тим кращими виходять тестові завдання. Важливість процесу перегляду не можна переоцінити – принципово важливо, щоб усі аспекти потенційного тестового завдання було ретельно вивчено. Після того, як завдання пройшли через прискіпливий аналіз внутрішніх перевірок розробником та клієнтом, комісія зовнішніх експертів розглядають завдання на об'єктивність, відповідність навчальним програмам та року навчання. Як тільки завдання затверджені усіма експертами, вони приймаються в остаточному вигляді та готові для апробації.

Результати апробації надають дані щодо виконання учнями кожного із завдань. Після аналізу даних встановлюються параметри завдання. Розробник і клієнт вивчають параметри завдань для того, щоб встановити їх придатність до використання у тестових варіантах. Всі характеристики завдань, описані у Розділі 7, мають бути перевірені та розглянуті. У багатьох випадках рішення стосовно відхилення або включення завдання у банк приймаються на основі врівноваження всіх параметрів завдання. Завдання, які вносяться до банку, повинні мати діапазон значень складності, розподільчої здатності та рівня досягнень, і, крім того, залишатися у встановлених межах. Усі разом завдання повинні задовольняти усім стандартам, які вимірюються тестом, і в ідеалі для кожного стандарту має бути весь діапазон визначених показників.

Після того, як завдання пройшли апробацію та визначено, що їхні статистичні дані задовольняють встановленим для тесту межам, завдання розміщуються у банку та можуть бути використані для укладення операційних тестових варіантів. Під час укладання тестового варіанту важливо пильно дотримуватися документів, що визначають структуру тесту, щоб пересвідчитися, що кожний варіант відповідає встановленій схемі. Кожному стандарту в тесті має відповідати вірна кількість завдань і балів за завдання, а також вірний формат завдань. Також важливо пересвідчитися, що обрані для варіанту завдання не є підказкою одне для одного, або що одне завдання не надає відповіді на інше завдання в тому ж варіанті тесту.

Після того, як попередній тестовий варіант укладено та узгоджено спеціалістом із змісту від розробника, він має бути надісланий спеціалісту із статистичних характеристик, щоб пересвідчитися у тому, що варіант відповідає статистичним вимогам до тесту. Якщо варіант відповідає цим вимогам, він готовий для відправки на перегляд клієнту. Варіант, готовий для відправки, має супроводжуватися зведеною таблицею із статистичною інформацією про тестовий варіант та попередньо укладеною таблицею конвертації балів.

Клієнт має перевірити варіант на відповідність тестовим специфікаціям, статистичним очікуванням, та на правильність загального формату. Технічна якість, представлення завдань та інструкцій у тестовому варіанті мають бути перевірені на точність та повноту. Клієнт має звірити завдання із варіантом для апробації, щоб пересвідчитись у тому, що завдання не було змінене ніяким чином. Таблиця конвертації балів має бути звірена із граничними балами встановлених стандартів і попередніми таблицями конвертації балів, щоб пересвідчитися у тому, що граничні бали відповідно до рівнів успішності не мають неналежних відхилень.

РОЗДІЛ 4. НАПИСАННЯ ТА ЕКСПЕРТНА ОЦІНКА СТИМУЛІВ

Стимулами є уривки текстів або графічні елементи, до яких має звертатися учень при відповіді на одне чи кілька завдань у тесті. Важливо пам'ятати, що стимули додають навантаження до читання, а це означає час на проведення тестування. Таким чином, стимули мають бути стислими і містити важливу інформацію, яку учасник тестування потребує для того, щоб відповісти на пов'язане з цим стимулом питання (кілька питань). Якщо на питання, пов'язані зі стимулом можна відповісти без звертання до нього, тоді цей стимул просто додає навантаження у вигляді читання і не додає ніякої вартості. При рецензуванні завдань, прив'язаних до стимулу, експерт має розглядати цінність, яку стимул вносить до тесту.

РІВЕНЬ СКЛАДНОСТІ ПРИ ЧИТАННІ

Завжди необхідно визначати рівень складності тексту-стимулу перед тим, як включати його до варіанту тесту. Для тесту з читання рівень складності має відповідати або бути дещо вищим (на один або два класи) за рівень, визначений для року навчання учасників тестування. Для предметних тестів (з природничих дисциплін, математики, суспільствознавства) рівень складності читання текстів має бути принаймні на один клас (а бажано на два класи) нижчим за рік навчання учнів, для яких укладається тест.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Декілька досліджень було присвячено вивченню впливу рівня складності читання текстів на здатність учнів зрозуміти поняття, які пояснюються у тексті. Було виявлено наступне: якщо учні отримують текст з певної предметної галузі того ж рівня складності читання як і тексти, призначені для навчання читання (зазвичай такий рівень відповідає року навчання), то це перешкоджає їх спроможності засвоїти представлені поняття.

Чому це так? Рівень матеріалів, які використовуються для навчання читання, є достатньо складним для учнів через те, що такі тексти містять незнайому лексику та складні граматичні структури, які є новими для учнів. Тому вся увага та зосередженість при читанні витрачається на правильне розуміння слів у тексті. Через те, що така велика частина уваги іде на процес декодування тексту, в учнів залишається тільки невелика частина уваги, яку вони можуть приділити розумінню інформації, представленої у тексті.

Таким чином, у тесті, який НЕ призначений оцінювати навички читання, важливо, щоб рівень складності текстів був комфортним для учнів з тим, щоб надати їм змогу зосередити всю свою увагу на змісті. Рівень складності тексту, яка на два рівня нижча за рівень складності для року навчання учнів, що проходять тестування, буде комфортною для більшості учнів. Це у свою чергу забезпечить те, що учні матимуть максимальну вірогідність швидко й легко зрозуміти поняття, представлені у стимулі.

УПЕРЕДЖЕНІСТЬ ТА МОВНА ДОСТУПНІСТЬ

Важливо, щоб стимул відповідав вимогам щодо упередженості та мовної доступності. Стимул не повинен містити інформацію, яка б могла засмутити учня і через це відволікти його та завадити його спроможності продемонструвати свій рівень володіння конструктом, який вимірюється. Рекомендації щодо предметних тестів (природничі дисципліни та суспільствознавство) мають незначні відмінності від рекомендацій до тестів на перевірку вмінь та навичок (читання, письмо, математика). Оцінювання з читання, письма або математики НЕ повинні торкатися емоційних тем. Для тестів з природничих дисциплін та

суспільствознавства може бути необхідність включити теми, які можуть бути емоційними, такі як тема рабства при тестуванні з історії або тема еволюції у природничих тестах. Це тому, що такі ідеї або поняття є складовими знання з цих предметів, проте ці поняття не повинні бути присутніми у тестах з навичок та вмінь.

ЗАГАЛЬНІ РЕКОМЕНДАЦІЇ

1. ПРИСУТНІСТЬ СТЕРЕОТИПІВ

Матеріали для тестування не повинні представляти стереотипи людей за такими ознаками:

- Вік
- Інвалідність
- Стать
- Расова/етнічна приналежність
- Сексуальна орієнтація

2. ЧУТЛИВІ АБО СУПЕРЕЧЛИВІ ТЕМИ

Чутливі теми або ті, які потенційно можуть викликати почуття болю або образи, необхідно уникати в усіх випадках або представляти тактовно. Наприклад, уривок, в якому йдеться про історичне значення битви є прийнятним, в той час як графічне зображення битви не є. Суперечливими темами вважаються:

- Смерть та хвороби
- Азартні ігри*
- Політика (сьогодення)
- Расові стосунки
- Релігія
- Сексуальність
- Забобони
- Війна

*(Посилання на азартні ігри потрібно уникати в завданнях з математики на ймовірність.)

3. ПОРАДА

Матеріали для тестування не повинні пропагувати певні стилі життя або поведінку крім таких, які є найбільш загальними та стосовно яких існує загальна погодженість. Наприклад, рецепт корисної для здоров'я страви є прийнятним, але уривок тексту, в якому рекомендують конкретну дієту, не є прийнятним. Слід повністю виключити поради, що стосуються:

- Релігії
- Сексуальних уподобань

4. НЕБЕЗПЕЧНА ДІЯЛЬНІСТЬ

Слід приділяти увагу тому, щоб не представити небезпечні види діяльності таким чином, щоб вони здавалися привабливими або прийнятними.

5. РОЗМАЇТТЯ, РЕПРЕЗЕНТАТИВНІСТЬ ТА ЕТНОЦЕНТРИЗМ

Матеріали для тестування мають:

- Відображати розмаїття учасників тестування
- Містити стимули (такі як літературні твори), написані представниками меншин
- Використовувати особисті імена, що походять з різних етнічних громад
- Використовувати зображення людей, які представляють різні етнічні громади
- Уникати етноцентризму (ставлення, яке ґрунтується на переконанні, що всі люди мають розділяти мову, вірування, культуру або релігію певної групи)

6. ДИФЕРЕНЦІАЛЬНА ОБІЗНАНІСТЬ: ЕЛІТНІСТЬ ТА УПЕРЕДЖЕНІСТЬ ЗАВДАННЯ

Слід уникати використання спеціалізованих понять та термінів, які є сторонніми по відношенню до основного змісту тестових питань. Це застереження стосується термінології з таких галузей як:

- Будівництво
- Фінанси
- Спорт
- Право
- Техніка
- Військова тематика
- Політика
- Наука
- Технологія
- Сільське господарство

7. ДОСТУПНІСТЬ МОВИ

Мова повинна бути максимально чіткою, недвозначною і зрозумілою для всіх. Наступних елементів слід уникати або використовувати із застереженням:

- Пасивні конструкції
- Ідіоми
- Речення з множинним підпорядкуванням
- Займенники з нечітко вказаним словом або поняттям, які вони представляють
- Багатозначні слова
- Нестандартні граматичні конструкції
- Діалектизми
- Жаргонізми

8. ГРАФІЧНІ ЗАСОБИ

Усі відповідні зазначені вище вимоги також стосуються графічних засобів.

УНІВЕРСАЛЬНІСТЬ СТРУКТУРИ

Універсальність структури означає рівний доступ до тесту для всіх учнів, включаючи учнів з обмеженими можливостями, як наприклад вади зору або обмежені можливості щодо навчання. Проте дуже часто розробники тестів не враховують учнів з такими вадами при написанні та експертній оцінці завдань. Розробники тестів часто усвідомлюють неадекватність своїх завдань, коли необхідно створити варіант тесту в системі Брайля. Хоча перевірка засвоєння деяких понять можливо потребує використання складних графічних елементів, при розробці тестових завдань розумним було б враховувати принцип універсальності структури.

РОЗДІЛ 5. НАПИСАННЯ ТА ЕКСПЕРТНА ОЦІНКА ЗАВДАНЬ МНОЖИННОГО ВИБОРУ

Мета завдань множинного вибору – визначити, чи опанував учень певний рівень знань та навичок. Якщо учасник тестування опанував певний розділ змісту, але не може вибрати правильну відповідь через те, що питання нечітке, це завдання потрібно видалити або переписати. Якщо учасник тестування не опанував певний розділ змісту, але може легко вгадати правильну відповідь завдяки тому, як написане завдання, воно також є незадовільним. Слідуючи певним принципам, розробник завдань може навчитися писати завдання множинного вибору, які б точно вимірювали відповідний зміст.

У даному розділі надаються визначення та приклади складових завдання множинного вибору. У наступних підрозділах надаються принципи для написання та рецензування трьох головних складових завдань множинного вибору: основи, ключа та дистракторів.

ПОНЯТТЯ ТЕСТОВОГО ЗАВДАННЯ

Завданням є окреме питання і для деяких типів завдань набір потенційних відповідей на завдання. Термін «завдання» використовується для позначення питання і відповідей при тестуванні з використанням множинного вибору. Термін «завдання» також використовують для позначення питань, які вимагають написання відповіді, хоча відповідь учасника тестування не вважається частиною завдання.

Завдання множинного вибору містить питання або твердження, яке називається основою, після якого пропонується ряд *варіантів відповіді*. Тільки один з варіантів відповіді є правильним і називається *ключем*. Решта варіантів називають *дистракторами*, тобто вони є правдоподібними, але невірними варіантами відповіді.

Приклад

1. Дженіфер хоче відгородити на галявині прямокутну ділянку, щоб залишати там собаку. Вона купила достатньо огорожі, щоб відгородити 252 квадратні фути галявини. Якщо довжина ділянки 21 фут, то яка її ширина?

Основа

A. 12 футів ← Ключ
B. 14 футів
C. 231 фут
D. 273 фути

Варіанти

Дистрактори

ОСНОВА ТЕСТОВОГО ЗАВДАННЯ

♦ **Закрита основа:**

Закрита основа – це завершене питання або твердження, яке закінчується крапкою або знаком питання.

Приклад

2. Для того, щоб з'ясувати, коли фраза "give me a break" з'явилася в англійській мові, який тип інформації ви будете шукати в словнику?

Закрита основа

A. Походження і етимологію ← Ключ
B. Використану абревіатуру
C. Похідні слова
D. Варіант вимови

Варіанти

Дистрактори

♦ **Відкрита основа:**

Відкрита основа – це незавершене питання або твердження; це часто фраза, яку можна завершити правильно, якщо вибрати один з варіантів відповіді.

Приклад

3. Автор використовує порівняння "Носитися на інвалідному візку – все одно, що везти мотоцикл по звивистому путівцю" для того, щоб описати

Основа

A. Хвилювання від швидкості крісла ← Ключ
B. Витрати на утримання крісла
C. Вигляд моторизованого крісла
D. Простоту маневрування крісла

Варіанти

Дистрактори

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Декілька дослідників вивчали відмінності у когнітивних процесах учасників тестування при роботі з завданнями множинного вибору з основами відкритого і закритого типів. Загалом було виявлено, що добре сформульовані завдання з закритою основою учасники сприймають швидше і швидше надають на них відповіді, у порівнянні з завданнями з відкритою основою.

Чому це так? Завдання з закритою основою зі знаком питання є завершеною думкою. З когнітивної точки зору, для нас відносно легше утримувати у робочому сегменті пам'яті лінгвістичну подію із завершеним змістом як повне твердження або питання. Потім ми легко можемо шукати і вибирати відповідь серед поданих варіантів без необхідності кілька разів перечитувати питання.

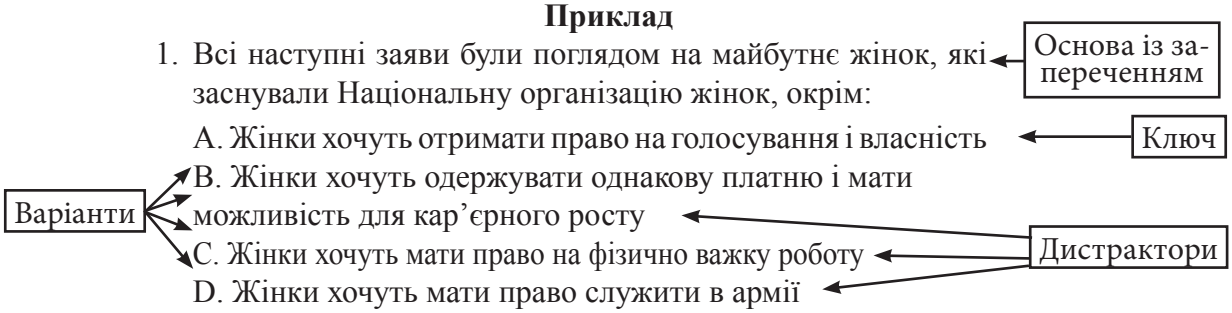
Відкрита основа містить неповну лінгвістичну подію, яка стає повною тільки з додаванням варіанта відповіді. Розглядаючи кожний варіант, учень має повертатися знову до незавершеного висловлювання,перечитувати його та додавати текст варіанта для того, щоб отримати повне висловлювання і таким чином визначити його зміст.

Учню може знадобитися прочитати закриту основу тільки один раз, звичайно в залежності від складності основи, але він буде змушений читати відкриту основу кілька разів для визначення правильності або неправильності кожного варіанту. Для обмеженого в часі стандартизованого оцінювання час тестування є цінним і, таким чином, відкриті основи бажано не використовувати.

♦ **Основа із запереченням:**

Завдання множинного вибору із запереченням в основі містить слово, яке позначає, що учасник тестування має обрати правильну відповідь як один із варіантів, який не відповідає або є протилежним по відношенню до інших варіантів. Основа із запереченням може бути сформульована як відкрита або закрита. Слова, які часто використовуються в запереченні, це «не», «ніколи», «не може», «крім».

Приклад



РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Декілька дослідників вивчали відмінності у когнітивних процесах учасників тестування при роботі з завданнями множинного вибору з основами, які містять або не містять заперечення. Загалом було виявлено, що добре сформульовані завдання з позитивною основою учасники сприймають швидше і швидше надають на них відповіді, у порівнянні з завданнями з основою, що містить заперечення. Також було виявлено, що при читанні завдання учні частіше невірно розуміють основи з запереченням у порівнянні з основами без заперечення.

Чому це так? Є декілька причин. Однією з них є той факт, що заперечення – це лінгвістична подія, яка потребує більше часу для обробки. Дослідження з психолінгвістики виявили, що ми потребуємо більше часу на сприйняття та обробку заперечення у будь-якому дискурсі, ніж на сприйняття і обробку стверджувального речення або питання.

Інша причина полягає в тому, що заперечення залежить від одного слова в основі або завданні; учасник тестування легко може пропустити одне слово при читанні і таким чином зрозуміти основу як стверджувальне речення. Третьою причиною є те, що при складанні стандартизованого тесту з завданнями множинного вибору учасники тестування схильні налаштовуватися на режим вибору позитивних відповідей. Вони просто не очікують заперечення в основі і через недогляд пропускають її, тому що режим пошуку відповідей не сигналізує про таку можливість, і учасники не змінюють свою стратегію пошуку.

Що робить учасник тестування, коли зустрічає завдання із запереченням в основі? Учасник може правильно прочитати основу, а потім вибрати правильну відповідь на негативне питання, але учасник також може неправильно прочитати основу, розуміючи її як позитивну, і вибрати перший варіант, який відповідає на питання. Якщо учасник тестування не прочитає всі інші варіанти, він можливо так і не зрозуміє, що інші кілька варіантів також надавали відповідь на це питання. Якщо учасник тестування таки пропустив заперечення в основі і потім читає всі варіанти, його спантеличують кілька привабливих відповідей. Такий учасник тестування може повернутися до питання, щоб перечитати його, і таким чином виявить заперечення і надасть правильну відповідь, але він також може вибрати один з правильних варіантів і продовжувати тест далі. Головна думка полягає у тому, що існує велика вірогідність того, що учасник тестування надасть неправильну відповідь на питання не тому, що в нього немає знань або навичок, а тому що він пропустив заперечення в основі. Таким чином, це відносить питання із запереченням до категорії завдань з «хитрощами».

Впроваджуючи програми стандартизованого оцінювання, прагнуть оцінити розуміння учнями певного обсягу знань. Якщо серед навичок, що підлягають тестуванню, не є вміння правильно читати заперечення, існує надто велика ймовірність змішання заперечення з навичками, які ви хочете оцінити. Таким чином, бажано не використовувати основи із запереченням.

В стандартизованих оцінювальних програмах, ми намагаємось оцінити розуміння студентів основного знання. Якщо ви не перевіряєте тільки одну з навичок, а саме - правильне читання негативної форми, вірогідність того, що ви поєднуєте негативну форму із навичками, які ви намагаєтеся оцінити, є дуже високою. Таким чином, використовування негативних основ розхолоджує.

ПРИНЦИПИ НАПИСАННЯ ТА ЕКСПЕРТНОЇ ОЦІНКИ ОСНОВИ ЗАВДАНЬ МНОЖИННОГО ВИБОРУ:

- Якісні основи завдань множинного вибору можна пов’язати з вмінням (вміннями), які вони оцінюють.
- Якісні основи завдань множинного вибору оцінюють істотну інформацію.
- Якісні основи завдань множинного вибору є стислими і містять тільки ту інформацію, яка необхідна учасникам тестування для надання відповіді на питання.
- В якісних основах завдань множинного вибору уникають матеріалу, що немає відношення до предмету оцінювання.
- В якісних основах завдань множинного вибору уникають багатослівності.
- Якісні закриті основи завдань множинного вибору мають надавати достатньо інформації для відповіді на питання без варіантів відповіді.
- Якісні відкриті основи завдань множинного вибору містять достатньо інформації для того, щоб надати учаснику тестування орієнтир до пошуку правильної відповіді.
- В якісних основах завдань множинного вибору уникають використання дієслів пасивного стану.

- В якісних основах завдань множинного вибору технічна термінологія використовується тільки при перевірці знання саме цієї термінології.
- В якісних основах завдань множинного вибору правильно використовується фраза «згідно з».
- В якісних основах завдань множинного вибору правильно використовується фраза «що із наведеного нижче».

ПРИНЦИПИ НАПИСАННЯ ТА ЕКСПЕРТНОЇ ОЦІНКИ КЛЮЧА ДО ОСНОВИ ЗАВДАНЬ МНОЖИННОГО ВИБОРУ:

- Якісні завдання множинного вибору мають відповідність між основою та правильною відповіддю.
- Якісні завдання множинного вибору мають тільки одну правильну відповідь.
- В якісних завданнях множинного вибору уникають правильної відповіді, яка б виділялася певним чином серед варіантів:
  - Правильна відповідь не повинна бути довшою за дистрактори;
  - Правильна відповідь не повинна граматично відрізняться від дистракторів;
  - Правильна відповідь не повинна бути єдиним варіантом, який містить суттєву лексику, використану в основі;
  - Правильна відповідь не повинна бути єдиним позитивним реченням або фразою серед негативних і навпаки;
  - Правильна відповідь не повинна містити граматичної підказки, як, наприклад, при використанні службових слів.
- В якісних ключах завдань множинного вибору уникають багатослівності.
- В якісних ключах завдань множинного вибору уникають використання дієслів пасивного стану.

ПРИНЦИПИ НАПИСАННЯ ТА ЕКСПЕРТНОЇ ОЦІНКИ ДИСТРАКТОРІВ ДО ОСНОВИ ЗАВДАНЬ МНОЖИННОГО ВИБОРУ:

- Якісні дистрактори завдань множинного вибору не містять очевидно неправильних відповідей.
- В якісних завданнях множинного вибору уникають дистракторів, які занадто наближені до правильної відповіді.
- Якісні дистрактори завдань множинного вибору представляють неправильне логічне розуміння або типові хибні уявлення, характерні для учасників тестування.
- Якісні дистрактори завдань множинного вибору граматично паралельні один з іншим та з правильною відповіддю.
- Якісні дистрактори завдань множинного вибору мають приблизно однакову довжину.
- Якісні дистрактори завдань множинного вибору не представляють як позитивно, так і негативно оформлені висловлювання.
- В якісних завданнях множинного вибору уникають дистракторів, які б були протилежними висловлюваннями.
- Якісні дистрактори завдань множинного вибору завжди уникають варіанту «все з вищенаведеного» або «жодне з вищенаведеного».
- В якісних дистракторах завдань множинного вибору уникають багатослівності.
- В якісних дистракторах завдань множинного вибору уникають використання дієслів пасивного стану.
- В якісних дистракторах та ключах завдань множинного вибору уникають використання таких слів як «всі», «завжди», «ніколи», але використовують «зазвичай» або «іноді».
- Якісні дистрактори та ключі завдань множинного вибору не повторюють в кожному варіанті інформацію, яку можна розмістити в основі.

НАПИСАННЯ ТА ЕКСПЕРТНА ОЦІНКА ОСНОВИ ЗАВДАННЯ МНОЖИННОГО ВИБОРУ

Якісні основи завдань множинного вибору можна пов'язати з вмінням (вміннями), які вони оцінюють

При написанні або винесенні експертного висновку стосовно завдання першим і найголовнішим кроком є упевнитись в тому, що завдання відповідає вмінню, яке необхідно оцінити.

Приклад 1:

ВМІННЯ: Відрізняти комах від інших членистоногих.

ЗАВДАННЯ: Як комахи відрізняються від інших членистоногих?

A. У комах немає спинного хребта.
B. У комах є зовнішня тверда оболонка
C. У комах зчленовані ноги.
D. У комах шестеро ніг.*

ОБГОВОРЕННЯ: У всіх членистоногих немає спинного хребта, є зовнішня тверда оболонка та зчленовані ноги. Комахи відрізняються від інших членистоногих тим, що мають шість ніг. Завдяки тому, що наведене вище завдання зосереджено на цій суттєвій різниці між комахами та іншими членистоногими, воно відповідає зазначеному вмінню.

Приклад 2:

ВМІННЯ: Визначати міста - столиці африканських країн.

ЗАВДАННЯ: Які з наведених африканських міст-столиць є також головним портом у Атлантиці?

A. Кіншаса
B. Лагос
C. Луанда*
D. Лусака

ОБГОВОРЕННЯ: Це завдання **не спрямоване** на встановленні столиці країни, як це передбачено у вмінні. Натомість, йдеться про точне розташування певних міст, що не є доречним для зазначеного вміння.

Приклад 2: Перероблений

ЗАВДАННЯ: Яке місто є столицею Анголи?

A. Хараре
B. Кіншаса
C. Луанда*
D. Лусака

Приклад 3:

ВМІННЯ: Визначати та вибирати відповідні одиниці для вимірювання кутів.

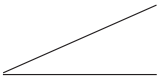
ЗАВДАННЯ: Який інструмент буде найкращим для вимірювання довжини скріпки для паперів?

A. компас
B. транспортир
C. лінійка
D. градуйована лінійка для вимірювання лісоматеріалів

ОБГОВОРЕННЯ: Це завдання **не оцінює** зазначене вміння. У визначенні вміння чітко вказано: «…вибирати відповідні одиниці». В цьому завданні запитується про відповідний інструмент.

Приклад 3: Перероблений

ЗАВДАННЯ: Яка одиниця виміру буде правильною для визначення розміру кута?



A. градус*
B. грам
C. фунт
D. квадратний дюйм

Якісні основи завдань множинного вибору оцінюють істотну інформацію

Крім відповідності вмінню, що передбачається перевірити, завдання повинно оцінювати істотну інформацію.

Приклад 1:

ВМІННЯ: Назвати англійських поетів - романтиків та їх роботи.

ЗАВДАННЯ: Яка з наведених далі поем належить Вільяму Вордсворту?

A. "Ода Соловейкові"
B. "На згасання Венеціанської Республіки"
C. "До руїни у Романтичній країні"
D. "Що ж, ми не підемо більше вештатись"

ОБГОВОРЕННЯ: Вільям Вордсворт був провідним англійським поетом - романтиком, то ж завдання про нього повинно належним чином відповідати зазначеному вмінню. Проте, Вордсворт написав сотні віршів. Розробник завдання має запитати про одну з найбільш відомих робіт Вордсворта, таку як «Прелюдія». Завдання на маловідомі і незначні вірші не відповідає критерію важливості.

Приклад 2:

ВМІННЯ: Назвати деталі з уривка.

ЗАВДАННЯ: Який колір був знаком на вікні магазину?

A. блакитний
B. зелений
C. червоний
D. жовтий*

ОБГОВОРЕННЯ: Це завдання **не оцінює** вміння, яке передбачено оцінити. Велика ймовірність того, що незалежно від того, з чим це завдання асоціюється, колір знаку на вікні магазина не є істотним, і таким чином, це питання не варто ставити.

Якісні основи завдань множинного вибору є стислими і вміщують тільки ту інформацію, яка необхідна учасникам тестування для надання відповіді на питання

Якісно складена основа містить всю інформацію, яка потрібна учаснику тестування для того, щоб відповісти на питання.

Приклад 1:

ВМІННЯ: Назвати причини і наслідки міграції корінних американців в Північній Америці.

ЗАВДАННЯ: Чому делавари або ленні ленапе виїхали зі своєї батьківщини?

A. Англійські поселенці витіснили ленні ленапе на захід до річки Огайо.
B. Худоба європейців знищила кукурудзяні поля ленні ленапе.
C. Шведські та голландські торговці переповнили землі ленні ленапе у пошуках хутра.
D. Ленні ленапе продали свою землю голландцям, яку вони хотіли тільки орендувати.

ОБГОВОРЕННЯ: По тому, як написане завдання, основа має на увазі, що ленні ленапе

поїхали тільки один раз і тільки з однієї причини. Насправді, ця народність, подібно до багатьох корінних американців, мігрувала кілька разів з багатьох зовнішніх причин. Більш уточнююча основа надає контекст для конкретного переміщення.

Приклад 1: Перероблений

ЗАВДАННЯ: Як відреагували ленні ленапе, що жили уздовж бухти Делавера, коли шведські торговці прибули до їх краю на початку 1600-х?

A. Вони оголосили війну шведам.
B. Вони поїхали подалі від шведів.
C. Вони залишилися і торгували із шведами.
D. Вони продали свої землі шведам.

Приклад 2:

ВМІННЯ: Називати американських президентів

ЗАВДАННЯ: Хто був президентом Сполучених Штатів з 1797 по 1801 роки, залишаючись при владі тільки чотири роки?

A. Ендрю Джексон
B. Джордж Вашингтон
C. Джеймс Полк
D. Джон Адамс*

ОБГОВОРЕННЯ: Основа містить непотрібну інформацію «залишаючись при владі тільки чотири роки». Ця додаткова інформація не потрібна, щоб відповісти на завдання.

Приклад 2: **Перероблений**

ЗАВДАННЯ: Хто був президентом Сполучених Штатів з 1797 по 1801 роки?

A. Ендрю Джексон
B. Джордж Вашингтон
C. Джеймс Полк
D. Джон Адамс*

Автори завдання повинні також дотримуватись загальноприйнятого правила уникати невизначеності, включаючи випадки вживання займенників, коли незрозуміло, до якого іменника вони відносяться.

ПРИКЛАД: Коли Мао Цзедун змістив Чан Кай-ші у 1949 році, куди він поїхав?

ОБГОВОРЕННЯ: Кого стосується «він»: Мао чи Чана?

КРАЩИЙ ВАРІАНТ: Коли Мао Цзедун змістив Чан Кай-ші у 1949 році, куди поїхав Чан?

В якісних основах завдань множинного вибору уникають матеріалу, що не відноситься до предмету тестування

Основа, яка містить сторонній матеріал, може вносити плутанину або матеріалу може бути недостатньо.

Приклад:

ЗАВДАННЯ: Потужна боротьба ослабила імперію інків якраз перед прибуттям іспанців. Хто став верховним правителем інків тільки для того, щоб бути викраденим і вбитим від руки Франціско Пісарро у 1533 році?

A. Атауальпа*
B. Гуаскар
C. Гуаяна Капак
D. Топа Інка

ОБГОВОРЕННЯ: Це завдання вимагає встановити єдиний факт: ім'я правителя інків, убитого Пісарро. Факт, що у Атауальпи була боротьба зі своїм братом до прибуття іспанців, є окремою інформацією, яка не належить цьому завданню.

Приклад: Перероблений

Завдання: Хто був верховним Інкою, вбитим від руки Франціско Пісарро у 1533 році?

A. Атауальпа*
B. Гуаскар
C. Гуаяна Капак
D. Топа Інка

ОБГОВОРЕННЯ: Основа, наведена вище, є більш стислою, ніж перший її варіант. Проте, перероблена основа все ще містить непотрібний матеріал. Уявіть собі, що ви складаєте тест. Ви достатньо впевнені, що Атауальпа був правителем інків, убитим Пісарро. Проте, а що коли ви не знаєте точну дату його смерті? Посилання на 1533 рік в основі може спантеличити вас. Вам може стати цікаво, чи може не один верховний правитель інків пав жертвою Пісарро, і чи знання дати не є важливим для надання правильної відповіді на питання.

Будь-яка інформація, яка не потрібна для формулювання питання, повинна бути вилучена із основи, тому що це може заплутати учасника тестування.

В якісних основах завдань множинного вибору уникають багатослівності

Чим довша основа, тим довше її читати.

БАГАТОСЛІВНА: Який з наведених американських композиторів написав «Весну в Апалачах»?

СТИСЛА: Хто написав «Весну в Апалачах»?

БАГАТОСЛІВНА: Яку тему автор цієї роботи намагався донести?

СТИСЛА: Яка тема цієї роботи?

Кожного разу, коли ви пишете завдання, переконайтеся, що ви представили найважливішу інформацію у найбільш стислий спосіб.

Приклад:

ЗАВДАННЯ: Яку тему автор цієї роботи намагався донести?

A. якою нудною може бути робота
B. які складні люди*
C. яке складне життя
D. як все дорого

ОБГОВОРЕННЯ: Основа, наведена вище є багатослівною, її можна легко спростити.

Приклад: Перероблений

Яка тема цієї роботи?

A. якою нудною може бути робота
B. які складні люди
C. яке складне життя
D. як все дорого

Крім того, завдання множинного вибору може мати відкриту чи закриту основу. Зазвичай закрита основа має перевагу, тому що завершена думка є легшою для читання й сприйняття учасником тестування, ніж незавершена думка. Проте іноді закрита основа призводить до багатослівності у формулюванні питання.

Якісні закриті основи завдань множинного вибору мають надавати достатньо інформації для того, щоб відповідь на питання можна було надати без варіантів відповіді

Приклад:

ЗАВДАННЯ: Яка із наведених країн не знаходиться в Європі?

A. Сенегал
B. Греція
C. Польща
D. Фінляндія

ОБГОВОРЕННЯ: Основа повинна містити достатньо інформації, щоб відповісти на питання без варіантів відповіді.

Приклад: Перероблений

ЗАВДАННЯ: На якому континенті знаходиться Сенегал?

A. Африка
B. Азія
C. Європа
D. Південна Америка

Якісні відкриті основи завдань множинного вибору містять достатньо інформації для того, щоб надати учаснику тестування орієнтир до пошуку правильної відповіді

Приклад 1:

ЗАВДАННЯ: Вашингтон

A. увійшов до складу країни в 1889 році.
B. входив до складу території Орегон до 1912 року.
C. став штатом в 1811 році.
D. входив до складу Північних Територій до 1848 року.

ОБГОВОРЕННЯ: Основа в цьому завданні є двозначною і могла б мати стільки правильних продовжень, що учневі доведеться читати спочатку варіанти відповідей, перш ніж він/вона зможе зрозуміти, про що питають в завданні. «Вашингтон» міг би бути Джорджем Вашингтоном; Вашингтон, Округ Колумбія; штат Вашингтон; або якийсь інший Вашингтон.

Приклад 1: **Перероблений**

ЗАВДАННЯ: У якому році штат Вашингтон приєднався до Сполучених Штатів?

A. 1889
B. 1811
C. 1848
D. 1912

Приклад 2:

ЗАВДАННЯ: Однорічник

A. є історією передгір'я Озарк.
B. описує життя на лісопилці у Вашингтоні.
C. є історією однорічної дитини.
D. був написаний Д. Селінджером

ОБГОВОРЕННЯ: Це завдання не формулює жодної проблеми і таким чином насправді є низкою питань «вірно» - «невірно», що стосуються однієї загальної теми.

Приклад 2: Перероблений

ЗАВДАННЯ: Місцевість Однорічник є

A. у передгір'ї Озарк.
B. лісопилці у Вашингтоні.
C. Гори Алегані.
D. лесовий район у Флориді.

В якісних основах завдань множинного вибору уникають використання дієслів пасивного стану

Автори повинні зазвичай уникати пасивної форми дієслів.

Перш за все, конструкції з дієсловами у пасивній формі є більш багатослівні, ніж інші.

ПАСИВНА ФОРМА: Яким футболістом був забитий гол?

АКТИВНА ФОРМА: Який футболіст забив гол?

По-друге, в пасивній конструкції, взаємовідношення між темою і дієсловом ослаблено. Саме футболіст, а не м'яч, виконав дію.

Тільки тоді віддається перевага пасивній конструкції, коли речення зосереджено на особі або об'єкті, з яким ця дія відбувається:

Приклад: Наполеон був остаточно розбитий Веллінгтоном у битві при Ватерлоо.

Автор речення прагнув зосередити увагу читача на Наполеоні, а не на Веллінгтоні.

В якісних основах завдань множинного вибору технічна термінологія використовується тільки при перевірці знання саме цієї термінології

Якщо перевіряється саме знання технічної термінології, тоді ця термінологія має використовуватися у завданнях.

Приклад:

ВМІННЯ: Називати приклади літературних образних висловів таких, як, наприклад, персоніфікація, метафора, порівняння та гіпербола.

ЗАВДАННЯ: Який з наведених образних висловів є прикладом персоніфікації?

A. Той сандвіч є великим як матрац!
B. Небо плакало над хворою землею.*
C. Річка стала киплячим казаном.
D. Вітер був як вологе полотно.

Проте, якщо ви не перевіряєте саме формальну термінологію, користуйтеся словами, які є прості, конкретні та близькі вашій аудиторії.

Приклад 1:

ВМІННЯ: Визначати настрій художнього твору

ЗАВДАННЯ: Яке емоційне середовище створює Шекспір у сцені на болоті у Макбеті?

A. холодно раціональне
B. похмуре*
C. ніжно меланхолійне
D. безтурботне

ОБГОВОРЕННЯ: Варіант основи вимагає знання терміну «емоційне середовище» і не є тим, що передбачає вміння.

Приклад 1: Перероблений

ЗАВДАННЯ: Який настрій створює Шекспір у сцені на болоті у Макбеті?

A. холодно раціональний
B. похмурий*
C. ніжно меланхолійний
D. безтурботний

В якісних основах завдань множинного вибору правильно використовується фраза «згідно з»

Іноді хочеться написати завдання, в якому питання стосується позицій або переконання окремої особи або організацій, а не загальноприйнятих фактів. У цьому випадку буде прийнятним почати основу із «Згідно з….»

Приклад: Згідно з концепцією Жана Піаже, перша стадія розвитку пізнання індивідуума є

A. конкретно - операційна стадія.
B. формально-операційна стадія.
C. передопераційна стадія.
D. сенсорно-моторна стадія.*

Уникайте використовувати цю фразу у завданням для учнів, які за віком можуть її не розуміти.

НЕВДАЛИЙ ВАРІАНТ: Згідно з Джоном Трімейном, яких солдатів постачає Нова Англія?

КРАЩИЙ ВАРІАНТ: В оповіданні Джона Трімейна, якої думки Джон про солдатів Нової Англії?

В якісних основах завдань множинного вибору правильно використовується фраза «що із наведеного нижче».

«Що із наведеного нижче» не є іменником; тож виходить, воно не може бути підметом у реченні. Радимо вживати після цієї фрази іменник.

НЕВДАЛИЙ ВАРІАНТ: Що із наведеного нижче є сумчастими тваринами?

КРАЩИЙ ВАРІАНТ: Які з наведених тварин є сумчастими?

НАПИСАННЯ ТА ЕКСПЕРТНА ОЦІНКА КЛЮЧІВ ДО ОСНОВ ЗАВДАНЬ МНОЖИННОГО ВИБОРУ

Якісні завдання множинного вибору мають відповідність між основою та правильною відповіддю

Дуже важливо декілька разів перевірити, щоб пересвідчитись, що ключ є правильною відповіддю. Це означає, що необхідно виконати математичні обчислення більше одного разу, перевіряти всі факти в численних довідкових джерелах і використовувати сучасні джерела стосовно інформації, яка може змінюватися.

НЕПРАВИЛЬНИЙ КЛЮЧ: Який із наведених генералів Конфедерації був з Огайо?
A. Джордж Арчібальд
B. Жордж МкКлілан*
C. Джон Александр
D. Улісс С. Грант

ОБГОВОРЕННЯ: Є дві проблеми у цьому завданні. Перша, основа є двозначною. Чи означає фраза «з Огайо», «народжений в Огайо,» «жив у Огайо» або «представляв Огайо»? МкКлілан, хоч і очолював перші польові загони з Огайо у Громадянській Війні, не був народжений в Огайо; і навіть ніколи там не жив. Він народився в Пенсільванії, відвідував коледж у Вест Пойнті і не мав ніякого зв’язку з Огайо до 1860 року, коли він став президентом Залізниці Огайо та Міссісіпі.

Якісні завдання множинного вибору мають тільки одну правильну відповідь

Переконайтеся, що правильна відповідь тільки одна.

ЗАВДАННЯ: Який генерал Конфедерації народився у Огайо?

A. Джозеф Хукер
B. Улісс С. Грант*
C. Вильям Текумзе Шерман*
D. Вінфілд Скотт

ОБГОВОРЕННЯ: І Грант, і Шерман народилися в Огайо. Треба замінити або В або С.

В якісних завданнях множинного вибору уникають правильної відповіді, яка б виділялася певним чином серед варіантів

Ключ не повинен відрізнятися від інших варіантів відповіді. Суттєва ймовірність того, що ключ, який виділяється серед варіантів відповіді, буде вибраний тими учнями, які насправді не знають, що це є правильна відповідь.

- Не робіть ключ довшим за дистрактори.
  Учений Майкл Фарадей відкрив

  A. електромагнітну індукцію.*
  B. гравітацію.
  C. швидкість світла.
  D. рентгенівське випромінювання.

- Не робіть ключ граматично відмінним від дистракторів.
  Що таке ліпід?

  A. жирова речовина*
  B. на ранній стадії розробки
  C. взаємовигідно
  D. речовина, пов'язана з особливостями організму

  ОБГОВОРЕННЯ: Ключем є іменникова конструкція, а дистракторами є інші конструкції.

- Не робіть ключ єдиним варіантом, який містить суттєву лексику з основи.
  Яка функція державного **нотаріуса (notary)**?

  A. забезпечувати присутність обвинуваченого в суді
  B. розслідувати причину смерті
  C. захищати інтереси підлітків
  D. **посвідчувати (notarize)** офіційні документи*

- Не надавайте ключ у вигляді позитивного висловлювання на фоні трьох негативних варіантів або навпаки.
  Згідно з формальним етикетом, весільне посипання може робити

  A. будь-хто, хто зараз разом працює з нареченою.
  B. будь-хто, хто знав наречену протягом щонайменше десяти років.
  C. будь-хто, хто є близьким другом нареченої та дружки.
  D. будь-хто, хто не **є** близьким членом родини нареченої.*

- Не надавайте підказки до ключа, використовуючи службові слова, як наприклад «ці», «спеціальна» або «певні».

  Що із наведеного, можна знайти у системі управління нацією, керованій диктатурою?

  A. вільна преса
  B. спеціальна таємна поліція*
  C. належне дотримання закону
  D. активні профспілки

В якісних ключах завдань множинного вибору уникають багатослівності

Уникайте зайвих слів в ключі так само, як і в основі.

БАГАТОСЛІВНА: Якої форми орбіта комети?

A. Вона знаходиться у формі еліпса, навколо сонця.

ПОЛІПШЕНА / СТИСЛА:

A. еліптична

В якісних ключах завдань множинного вибору уникають використання дієслів пасивного стану

Уникайте пасивної форми в ключі так само, як і в основі.

ПАСИВНА ФОРМА В КЛЮЧІ:

Хто здійснив дуже успішне зимове завоювання Русі?

A. Південна Русь була спустошена Батиєм, онуком Чингісхана.

АКТИВНА ФОРМА В КЛЮЧІ:

A. Батий, онук Чингісхана

ОБГОВОРЕННЯ: Варіант ключа з дієсловом в пасивному стані більш багатослівний, ніж варіант з активним станом. До того ж, основа містить дієслово в активному стані; запитання ставиться про особу, яка очолила завоювання. Варіант ключа з пасивним станом дієслова не відповідає основі граматично, і це вносить плутанину, тому що він починається з назви країни, а не з імені особи.

НАПИСАННЯ ТА ЕКСПЕРТНА ОЦІНКА ДИСТРАКТОРІВ ДО ОСНОВИ ЗАВДАННЯ МНОЖИННОГО ВИБОРУ

Якісні дистрактори завдань множинного вибору не містять очевидно неправильних відповідей

Кожний дистрактор у завданні множинного вибору повинен виглядати як правдоподібна альтернатива для будь-кого, хто не володіє знаннями або навичками, які необхідно оцінити. Жоден з дистракторів не повинен бути відверто неправильною відповіддю, яку можна виключити без жодного напруження з боку учня.

ПРИКЛАД НЕВДАЛОГО НАПИСАННЯ ДИСТРАКТОРІВ:

У книзі Рамона Назавжди, кому з героїв приходиться чекати наодинці у прийомному відділенні лікарні, після того, як народжується дитина?

A. Пану Квімбі
B. Пані Квімбі
C. Рамоні*
D. Новонародженій дитині

ОБГОВОРЕННЯ: Ключем є C: Рамоні. «Новонароджена дитина» - це відверто неправильна відповідь, яку учень може ЛЕГКО виключити, навіть якщо і не читав оповідання. Коли це новонароджений малюк міг чекати наодинці в прийомному відділенні лікарні?

ПРИКЛАД НЕВДАЛОГО НАПИСАННЯ ДИСТРАКТОРІВ:

Чарлі Паркер був одним з музикантів, які створили музичний стиль, відомий як

A. бібоп*
B. блуграсс
C. «холодний» джаз
D. свінг

ОБГОВОРЕННЯ: Учень, який знає, що A, C та D є стилями джазу, а от B є стилем музики кантрі, може виключити варіант B і без будь-якого знання про Чарлі Паркера.

В якісних завданнях множинного вибору уникають дистракторів, які занадто наближені до правильної відповіді

Вам не хочеться створювати дистрактори, які занадто очевидно є помилковими, але хотілося б уникнути дистракторів, які є занадто близькими до ключа. Ви ж не пробуєте обдурити учасника тестування.

ПРИКЛАД НЕВДАЛОГО НАПИСАННЯ ДИСТРАКТОРІВ:

Що є найпоширенішим симптомом собачої чуми?

A. слинотеча
B. випадіння волосся
C. потовщання подушок на лапах
D. сильна діарея*

ОБГОВОРЕННЯ: Сильна діарея є загальним симптомом даної хвороби. Варіанти A та B є гарними дистракторами, оскільки вони є симптомами інших собачих хвороб. Варіант C є дуже близьким до ключа. Потовщання подушок на лапах буває іноді симптомом собачої чуми, хоча воно не таке вже й поширене, як діарея.

Якісні дистрактори завдань множинного вибору представляють неправильне логічне розуміння або типові хибні уявлення, характерні для учасників тестування

Дистрактор повинен представити результат правдоподібного, але помилкового процесу мислення учня (невипадкова помилка).

ВДАЛИЙ ПРИКЛАД:

Као заплатив $8.00 за легку закуску і хоче залишити 20% «на чай». Скільки повинно бути чайових?

A. $0,16
B. $0,40
C. $1,60*
D. $4,00

ОБГОВОРЕННЯ: C є ключем, тому що щоб вирахувати 20% чайових, учень повинен помножити суму за рахунком $8.00 на 0,20. Учень, який вибрав варіант A, помножив на 0,02. Учень, який вибрав варіант B, поділив на 20. Учень, який вибрав варіант D, поділив на 2.

ВДАЛИЙ ПРИКЛАД:

Хто прославився у якості медичної сестри під час Кримської Війни?

A. Клара Бартон
B. Доротея Дікс
C. Флоранс Найтінгейл*
D. Мері Ен Бікердайк

ОБГОВОРЕННЯ: Клара Бартон є вдалим дистрактором, тому що вона прославилася як медична сестра під час Громадянської Війни. Доротея Дікс та Мері Ен Бікердайк також були медсестрами під час Громадянської Війни.

Якісні дистрактори завдань множинного вибору граматично паралельні один з іншим та з правильною відповіддю

Дистрактори повинні бути граматично паралельні один до одного. Наприклад, всі дистрактори мають бути іменниками або прикметниками. Дистрактори, які не є граматично паралельними, важко читати.

ПРИКЛАД НЕПАРАЛЕЛЬНИХ ДИСТРАКТОРІВ:

Яке з наведених положень міститься у Законі про відміну податку на нерухомість, прийнятому у США палатою представників у 2001 році?

A. поступове зниження податкової ставки, починаючи з 2010 року
B. положення, що передбачає повну відміну негайно
C. збільшення звільнення до $2 мільйонів у 2002 році
D. Податок буде повністю відмінено до 2011 року.*
ОБГОВОРЕННЯ: Три дистрактори є фрагментами речення, а ключ - повним реченням.

Якісні дистрактори завдань множинного вибору мають приблизно однакову довжину

Дистрактори повинні бути приблизно однакової довжини.

ПРИКЛАД ДИСТРАКТОРІВ, ЯКІ НЕ Є ПАРАЛЕЛЬНИМИ ЗА ДОВЖИНОЮ:

ЗАВДАННЯ**:** Ключовою подією мітозу є:

A. схрещення генів однієї хромосоми з генами іншої.*
B. подвоєння ДНК.
C. вихід гаплоідних клітин.
D. відокремлення хромосом.

ОБГОВОРЕННЯ: Необхідно, щоб всі варіанти відповіді були однакової довжини. Якщо неможливо витримати однакову довжину для всіх варіантів відповіді, слід спробувати створити два варіанти короткі та два довгі, щоб таким чином уникнути одного варіанта, який би відрізнявся від інших.

Якісні дистрактори завдань множинного вибору не представляють як позитивно, так і негативно оформлені висловлювання

ПРИКЛАД НЕВДАЛОГО НАПИСАННЯ ДИСТРАКТОРІВ:

Як необхідно обробляти чорнильну пляму на одязі, що линяє?

A. Обробіть пляму крохмалем або дитячою присипкою, а потім зітріть крохмаль чи присипку.
B. Не **періть** одяг, поки не перевірите чорнило на шматку ганчірки, щоб побачити, як прання впливає на пляму.*
C. Покладіть одяг в морозилку в прозорому пластиковому кульку на годину, а потім спробуйте вивести пляму лаком для нігтів.
D. Сильно натріть пляму шматком білого хліба, а потім замочіть одяг в розчині прального порошку та кип'яченої води.
ОБГОВОРЕННЯ: Варіант B, який є ключем, виділяється через те, що він один містить заперечення.

В якісних завданнях множинного вибору уникають дистракторів, які б були протилежними висловлюваннями

ПРИКЛАД НЕВДАЛОГО НАПИСАННЯ ДИСТРАКТОРІВ:

Чорниці виростають краще в яких умовах?

A. **кисла земля***
B. **лужна земля**
C. густа тінь
D. вологі умови

ОБГОВОРЕННЯ: Земля є або лужним або кислотним середовищем. Це є протилежні якості. Саме через те, що більшість рослин вимагає або того, або іншого (неважливо, що їм потрібно ще), учень може визначити, що ключем є або варіант A або B і без читання варіантів C або D.

Якісні дистрактори завдань множинного вибору завжди уникають варіанту «все з вищенаведеного» або «жодне з вищенаведеного»

Необхідно завжди уникати таких варіантів.

ПРИКЛАД НЕВДАЛОГО НАПИСАННЯ ДИСТРАКТОРІВ:

Два процеси, які дозволяють клітинам вивільняти енергію від їжі, є

A. виведення та дифузія.
B. бродіння та клітинне дихання.
C. осмос та мимовільна генерація.
D. жоден із наведених варіантів.

ОБГОВОРЕННЯ: У більшості завдань множинного вибору учень може відкинути або прийняти кожний варіант окремо. Проте завдання, що містять варіанти «все з вищенаведеного» або «жодне з вищенаведеного», вимагають від учня розглядати всі варіанти в сукупності, що є більш складним завданням. До того ж, завдання з ключем «жодне з вищенаведеного» скоріше перевіряє те, що не відповідає дійсності, ніж навпаки, і загалом є небажаною стратегією.

В якісних дистракторах завдань множинного вибору уникають багатослівності

Уникайте зайвих слів у дистракторах, як і в інших компонентах завдання.

ПРИКЛАД НЕВДАЛОГО НАПИСАННЯ ДИСТРАКТОРІВ:

Яке перетворення енергії відбувається в зелених рослинах під час фотосинтезу?

A. Теплова енергія сонця перетворюється хлоропластами в рослині на електричну енергію.
B. Теплова енергія сонця перетворюється хлоропластами в рослині на енергію світла.
C. Хімічна енергія від переробки поживних речовин в рослині перетворюється на механічну енергію.
D. Теплова енергія сонця перетворюється хлоропластами в рослині на хімічну енергію.

ОБГОВОРЕННЯ: Доповнення, що складається із слів «від сонця» та «хлоропластами в рослині» повторюється в трьох варіантах, що робить їх зайвими. Це також робить варіант С радикально відмінним через те, що в ньому не використовуються ці слова.

ВДАЛИЙ ПРИКЛАД:

Яке перетворення енергії відбувається в зелених рослинах під час фотосинтезу?

A. Теплова енергія перетворюється на електричну енергію.
B. Теплова енергія перетворюється на енергію світла.
C. Хімічна енергія перетворюється на механічну.
D. Енергія світла конвертується у хімічну енергію.

В якісних дистракторах завдань множинного вибору уникають використання дієслів пасивного стану

ПРИКЛАД НЕВДАЛОГО НАПИСАННЯ ДИСТРАКТОРІВ:

Як вирішується проблема морської черепахи?

A. Старші нагадують морській черепасі, що вона завжди була красивою.
B. Морська черепаха вирішує жити без панцирю.
C. Пакувальний щур повертає коштовні камені, які він узяв.
D. Більше коштовних каменів знайдено для морської черепахи монстром Джіла.

ОБГОВОРЕННЯ: Уникайте непотрібних слів у відволікаючих варіантах, як ніде в іншому місці завдання.

ВДАЛИЙ ПРИКЛАД:

Як вирішується проблема морської черепахи?

A. Старші нагадують морській черепасі, що вона завжди була красивою.
B. Морська черепаха вирішує жити без панциру.
C. Пакувальний щур повертає коштовні камені, які він узяв.
D. Джіла-монстр знаходить більше коштовностей для черепахи.

В якісних дистракторах та ключах завдань множинного вибору уникають використання таких слів як «всі», «завжди», «ніколи», але використовують «зазвичай» або «іноді»

ПРИКЛАД НЕВДАЛОГО НАПИСАННЯ ДИСТРАКТОРІВ:

Закони Джима Кроу, прийняті у Південних штатах після Реконструкції, були покликані:

A. підтримати цілі 13-ої, 14-ої та 15-ої поправок.
B. заохочувати всіх афроамериканців купувати свої власні ферми.
C. надавати початкову освіту колишнім рабам.
D. заснувати расову ізоляцію.

ОБГОВОРЕННЯ: Багато натренованих учнів знають, що використання слів «все», «завжди», і «ніколи» у дистракторах ще більше підкреслює неправильність варіанту. В деяких матеріалах для підготовки тестів учням рекомендують відхиляти варіанти, які містять такі слова. Проте в деяких випадках обмежуючі слова можуть вказувати на правильність варіанту відповіді.

Якісні дистрактори та ключі завдань множинного вибору не повторюють в кожному варіанті інформацію, яку можна розмістити в основі

Приклад:

Завдання: Член ради директорів корпорації зазвичай вибирається ким із наведеного списку?

A. Власниками облігацій корпорації
B. Акціонерами корпорації
C. Президентом корпорації
D. Працівниками корпорації

ОБГОВОРЕННЯ: Слово «корпорація» повинно бути в основі. Завдання перш за все повинно бути чітким і зрозумілим. Через це не завжди можна включати в основу матеріал, що повторюється.

ПОКРАЩЕНЕ ЗАВДАННЯ:

Які люди, пов’язані з корпорацією, зазвичай вибирають директорів?

A. Власники облігацій
B. Акціонери
C. Офіційні особи
D. Службовці

РОЗДІЛ 6. НАПИСАННЯ, ЕКСПЕРТНА ОЦІНКА ТА ОЦІНЮВАННЯ ЗАВДАНЬ ВІДКРИТОГО ТИПУ

Створення відкритих завдань вимагає уважності й майстерності. Оцінювання таких завдань потребує значних витрат і багато часу через те, що екзаменатори мають прочитати кожну відповідь, крім того учасникам тестування потрібно більше часу для написання відповіді на відкриті завдання, ніж на завдання множинного вибору. Таким чином, збільшуючи кількість відкритих завдань в тесті, ми поставимо менше запитань і охопимо менший обсяг змісту/стандартів/навичок.

Завдання відкритого типу мають чітко інформувати учасників тестування про те, яка відповідь від них очікується. Вони також мають чітко вказувати, який саме аспект буде оцінюватися (наприклад, орфографія, граматика тощо).

Завдання відкритого типу – це ті, які вимагають від учасника тестування написати відповідь без можливості вибору із запропонованих варіантів. В залежності від предметної галузі, відповіддю на такі питання можуть бути одне слово, кілька слів, письмова відповідь обсягом у кілька абзаців або складне рівняння. Існують два загальні типи відкритих завдань: завдання з короткою відповіддю та завдання з розгорнутою відповіддю.

- **Завдання з короткою відповіддю:**

Завдання з короткою відповіддю зазвичай вимагають від студента надати менше інформації, ніж завдання з розгорнутою відповіддю. Зазвичай на виконання такого завдання учню потрібно від 2 до 3 хвилин.

- **Завдання з розгорнутою відповіддю:**

Завдання з розгорнутою відповіддю зазвичай вимагають від учня надати більше інформації або подробиць (наприклад, надати стислий огляд певного матеріалу). Зазвичай студенту потрібно від 5 до 6 хвилин, щоб виконати такі завдання. Деякі завдання з розгорнутою відповіддю можуть вимагати навіть більше часу, можливо від 10 до 15 хвилин.

СКЛАДОВІ ЗАВДАНЬ ВІДКРИТОГО ТИПУ

Всі завдання відкритого типу складаються з двох частин: інструкції та схеми оцінювання.

Інструкція: частина завдання, яка представляє запитання або завдання для відповіді.

Схема оцінювання: критерії, які використовуються для оцінювання відкритого завдання. Крім критеріїв схема оцінювання повинна включати зразки відповіді і/або правильну відповідь.

НАПИСАННЯ ТА ЕКСПЕРТНА ОЦІНКА ІНСТРУКЦІЙ ДО ЗАВДАНЬ ВІДКРИТОГО ТИПУ

ПРИНЦИПИ НАПИСАННЯ ТА ЕКСПЕРТНОЇ ОЦІНКИ ІНСТРУКЦІЙ ДО ЗАВДАНЬ ВІДКРИТОГО ТИПУ:

- Якісні завдання відкритого типу можна пов’язати з вмінням (вміннями), які ці завдання оцінюють.
- На якісні завдання відкритого типу не можна надати відповідь просто «так» або «ні».
- Якісні завдання відкритого типу побудовані на темах або добірках матеріалів, які є доступними учасникам тестування і не вимагають від учасників розкриття особистої інформації.
- Якісні завдання відкритого типу не містять зайвої багатослівності, непотрібних

заперечень або пасивних конструкцій.

- Якісні завдання відкритого типу побудовані таким чином, що в результаті їх виконання можна отримати відповіді, які охоплюють усю шкалу балів.
- Якісні завдання відкритого типу містять питання, в якому чітко вказані обсяг і тип інформації, яку необхідно надати при відповіді для того, щоб отримати найвищий бал.

НАПИСАННЯ ТА ЕКСПЕРТНА ОЦІНКА СХЕМИ ОЦІНЮВАННЯ ЗАВДАНЬ ВІДКРИТОГО ТИПУ

ПРИНЦИПИ НАПИСАННЯ АБО ЕКСПЕРТНОЇ ОЦІНКИ СХЕМИ ОЦІНЮВАННЯ ЗАВДАНЬ ВІДКРИТОГО ТИПУ:

- Якісні завдання відкритого типу мають такі схеми оцінювання, що чітко встановлюють бали.
- Якісні завдання відкритого типу з предметів інших, ніж писемне мовлення, мають надавати чіткі зразки відповідей для кожного балу, який можна отримати за завдання.
- Якісні завдання відкритого типу мають такі схеми оцінювання, що містять паралельні описові твердження для шкали балів за завдання.
- Якісні завдання відкритого типу мають такі схеми оцінювання, які містять паралельні описові твердження, що стосуються навичок або стандартів виконання, які оцінюються, і не містять тверджень стосовно навичок, що не мають відношення до навичок або стандартів виконання, які оцінюються.
- Якісні завдання відкритого типу мають такі схеми оцінювання, що повністю відповідають поставленим запитанням.

Далі наводимо в якості прикладу уривок для читання для 7-го класу та завдання відкритого типу до нього. Ці приклади завдань надають зразок інструкцій до завдань та схеми оцінювання та містять необхідні складові якісного завдання відкритого типу.

Приклад 1: Безпосадочний політ з одного узбережжя до іншого

Ричард Бауман

Це уривок із статті пілотів Оуклі Келлі і Джона Макреді з Огайо, які першими здійснили безпосадочний переліт з одного узбережжя до іншого.

Келлі керував літаком під час зльоту, а Макреді мав посадити літак, сподіваємося, в Сан Дієго. Фоккер, з двигуном Ліберті потужністю 400 кінських сил, заревів з усіх сил, і величезний літак підстрибнув і помчався вдовж злітної смуги. “Перша спроба злетіти була невдалою”, - писав Макреді у виданні журналу “National Geographic” в липні 1924 року. Після того, як літак пробіг по смузі близько милі, цей величезний, важкий транспорт не виявив ніякого наміру злетіти у повітря”.

Літак знов був заправлений пальним і була зроблена друга спроба злетіти. Келлі знов повів п’ятитонний літак по злітній смузі. Хоча літак ще не досяг злітної швидкості, тобто “точки неповернення”, Келлі повністю відкрив дроселі, двигуни трусило, а літак котився по смузі. В кінці смуги було захисний вал довжиною близько двадцяти футів, і Келлі буквально спрямував літак на невеличкий пагорб.

“Над краєм... ми злетіли, але були все ще близько від землі”. – розказував Макреді про цей захоплюючий дух зліт. “Т-2 летів, але не було ніяких ознак набору висоти”. Літак погойдувався так, начебто він збирався розбитися, але Келлі, який задіяв всю свою майстерність пілота, вдалося набрати швидкість, необхідну для польоту Т2.

“Келлі зробив чудо, піднявши в повітря величезний важкий літак”- зауважив Макреді. – “20 хвилин нашого зльоту над Лонг Айлендом навряд чи хтось помітив. Насправді протягом перших п’яти миль ми просто уникали стовпів і дротів. Дуже повільно Фоккер

набирав висоту - двадцять, тридцять, сорок футів, його шасі зрізало верхівки дерев... Повернувши на захід, тяжкому літаку насилу вдалося набрати висоту 500 футів в той час, коли Келлі і Макреді були над пагорбами Пенсільванії.

Відповідно до конструкції літака, пілот сидів з лівого боку відкритої кабіни, прямо біля двигуна. Другий пілот сидів за 10 футів позаду від пілота, але усередині літака. Він мав ті ж важелі, що й пілот, але у нього не було переднього огляду і огляду справа. Келлі і Макреді регулярно мінялися місцями протягом польоту. Це було нелегко. Для цього другий пілот мав взяти на себе керування літаком, в той час як перший пілот мав встати, скласти своє сидіння, а потім перелізти через невеличкий отвір до середини літака. Після того, як другий пілот складав з себе управління , він перелізав в кабіну і брав на себе керування літаком.

Єдиною нагодою протягом польоту, коли два пілоти могли розмовляти один з одним, було кілька хвилин під час зміни місць. На літаку не було внутрішнього зв'язку, а вітер і шум двигунів заважали розмові.

У Макреді та Келлі були проблеми щодо керування літаком по місцевим орієнтирам. Найбільші міста Пенсільванії були скриті в густому тумані, в той час як ріки, автостради і залізничні дороги, по яким вони мали намір орієнтуватися, не можна було чітко розпізнати з літака. Тому пілоти розраховували своє місцеположення за допомогою компасу і швидкості вітру.

Потім почалися технічні проблеми. Електрична система літака припинила виробляти енергію для зарядки батарей. Без підзарядки батареї будуть швидко виходити з ладу, і вони будуть змушені зробити посадку. Келлі знаками показав Макреді взяти на себе керування літаком на місці другого пілота. Так як Келлі сидів як раз біля двигуна, він мав змогу зняти кожух з двигуна і знайти неполадку.

Без будь-якої видимості праворуч або попереду Макреді мав вести літак, покладаючись виключно на власну інтуіцію. Тим часом Келлі зміг дістатися моторного відсіку і виявив проблему, якою, як з'ясувалося, був несправний стабілізатор напруги. Незважаючи на задуху, шум і рухливі частини двигуна, він зняв стабілізатор , щоб полагодити його.

Полагодити стабілізатор літака, який стоїть на стоянці, справа нелегка, а зробити це під час польоту рівнозначно чуду.

Приклад 2: Відкрите завдання з короткою відповіддю

ВМІННЯ: Робити порівняння по тексту, звертаючи увагу на стиль автора, а також буквальний зміст тексту та той, що мається на увазі.

ОСНОВА: Визначте дві відмінності між першою і другою спробою підняти літак, які описані в уривку.

СХЕМА ОЦІНЮВАННЯ:

Бали (0-2):	Відповідь учня
2	Відповідь включає одну з наступних правильних відмінностей: • Перша спроба "була невдалою", в той час як друга спроба була вдалою. • Пілот просто вів літак по злітній смузі при першій спробі, тоді як під час другої спроби він перелетів захисний вал в кінці злітної смуги.
1	Відповідь містить одну з правильних вищенаведених відмінностей з неправильною або не вказаною відмінністю.
0	Відповідь, що не містить жодної з правильних вищенаведених відмінностей, відповідь неповна, невідповідна або відсутня взагалі.

ОБГОВОРЕННЯ: Зверніть увагу на те, що і питання в основі завдання і схема

оцінювання належним чином стосуються вміння, що перевіряється. Схема оцінювання містить всі необхідні компоненти, включаючи чітке пояснення очікування правильної відповіді при наданні трьох балів, прийнятні приклади і схожі формулювання. Зверніть увагу на те, що бал 1 у схемі оцінювання відсилає до прикладів, наданих для здобуття 2 балів, замість того, щоб їх повторювати.

ПРИКЛАД 3: Відкрите завдання з короткою відповіддю

ВМІННЯ: Оцінювати доречність і точність подробиць автора, що визначають стратегію аргументації і приклади упередженості і стереотипів.

ОСНОВА: Визначте дві конкретні подробиці в уривку, які ілюструють ту небезпеку, з якою зіткнулися Келлі і Макреді до того, як вони дісталися Пенсільванії.

СХЕМА ОЦІНЮВАННЯ:

Бали (0-2):	Відповідь учня
2	Відповідь включає дві з наступних правильних деталей: • "... Келлі буквально перетащив літак через невеличкий пагорб" • "Літак так трясся, начебто збирався розбитися...." • Келлі і Макреді "заледве уникали стовпи і дроти". • "... Фоккер дуже повільно набирав висоту – двадцять, тридцять, п'ятдесят футів, його шасі зрізало верхівки дерев". • "... важкий літак з трудом зміг піднятися на висоту 500 футів...."
1	Відповідь включає одну правильну з вищевказаних деталей, з одною неправильною або відсутньою.
0	Бали не надаються, якщо твердження не існує, або повністю неправильне. Відповідь не містить жодної правильної деталі, з тих, що були наведені вище, відповідь неповна, невідповідна або відсутня взагалі.

ОБГОВОРЕННЯ: Зверніть увагу на те, що і питання в основі завдання і схема оцінювання належним чином стосуються вміння, що перевіряється. Схема оцінювання містить всі необхідні компоненти, включаючи чітке пояснення очікування правильної відповіді при наданні двох балів, прийнятні приклади і схожі формулювання. Зверніть увагу на те, що бал 1 у схемі оцінювання відсилає до прикладів, наданих для здобуття 2 балів, замість того, щоб їх повторювати. Так як існує більш ніж дві можливі відповіді, всі можливі відповіді включені до схеми оцінювання.

Приклад 4: Відкрите завдання з розгорнутою відповіддю

ВМІННЯ: Оцініть відповідність, точність і доречність подробиць автора, що визначають стратегію аргументації і приклади упередженості і стереотипів.

ОСНОВА: Події, описані в уривку, підкреслюють позитивні риси характеру, які продемонстрували Келлі і Макреді. Визначте дві з цих видатних рис, зазначених в уривку, та надайте підтвердження кожній з них різними деталями з уривку. Відповіді можуть відноситися до одного або до обох чоловіків.

СХЕМА ОЦІНЮВАННЯ:

Принципи оцінювання	
Бали (0-4):	**Відповідь студента**
4	*Відповідь включає дві з наступних рис і одну підкріплюючу деталь для кожної риси:* • ***(Риса № 1)*** *відвага або сміливість* ***(Підтверджуючі деталі № 1)*** • *Вони були першими, хто здійснив безпосадочний політ через усю країну.* • *Вони злетіли, скотившись зі схилу.* • *Вони витримали важкий і небезпечний зліт.* • *Вони помінялися місцями під час польоту.* • *Вони відремонтували літак під час польоту.* • *Другий пілот вів літак з заблокованою видимістю в той час, як пілот ремонтував літак.* • ***(Риса № 2)*** *впевненість, майстерність або винахідливість* ***(Підтверджуючі деталі № 2)*** • *У них було достатньо майстерності (професіоналізму) підняти літак в повітря і утримати його в польоті...* • *Вони змогли використати навички визначення місцеположення тільки по компасу для того, щоб вести літак.* • *Вони змогли виявити неполадки і відремонтувати літак під час польоту.* • ***(Риса № 3)*** *спокій, терпіння або рішучість* ***(Підтверджуючі деталі № 3)*** • *Вони зробили дві спроби злетіти і не здалися.* • *Вони змогли витримати дуже повільний зліт на початку їх польоту.* • *Вони витримали незручну конструкцію літака, яка вимагала від них мінятися місцями (і обмежену видимість у хвості літака).* • *Вони використали навички визначення місцеположення тільки по компасу, коли не змогли орієнтуватися за об'єктами на місцевості.* • *Вони відремонтували літак під час польоту.* *ПРИМІТКА: Бал надається за будь-яку переконливу рису або підтверджуючу деталь. Учні можуть посилатися на обох або на одного пілота, описуючи їх риси.*
3	Відповідь включає дві з вищенаведених рис, одну з правильними деталями, а іншу з неправильними або відсутніми деталями. *ПРИМІТКА: Бал надається за будь-яку переконливу рису або підтверджуючу деталь. Бали не надаються за правильні деталі, якщо риса, до якої вони належать, неправильна або відсутня.*
2	Відповідь включає дві з вищенаведених правильних рис, обидві з неправильними або відсутніми деталями. - АБО – Відповідь включає одну з вищенаведених правильних рис з правильними деталями - ТА – неправильну або відсутню рису. ПРИМІТКА: Бал надається за будь-яку переконливу рису або підтверджуючу деталь. Бали не надаються за правильні деталі, якщо риса, до якої вони належать, неправильна або відсутня.
1	Відповідь включає одну з вищенаведених правильних рис з неправильними або відсутніми деталями. - ТА – неправильну або відсутню рису. ПРИМІТКА: Бал надається за будь-яку переконливу рису або підтверджуючу деталь. Бали не надаються за правильні деталі, якщо риса, до якої вони належать, неправильна або відсутня.

0	Бали не надаються, якщо твердження не існує або повністю неправильне. Відповідь не містить одну з вищенаведених правильних рис, вона неповна, невідповідна або відсутня. ПРИМІТКА: Бал надається за будь-яку переконливу рису або підтверджуючу деталь. Бали не надаються за правильні деталі, якщо риса, до якої вони належать, неправильна або відсутня.

ОБГОВОРЕННЯ: Зверніть увагу на те, що і питання в основі завдання і схема оцінювання належним чином стосуються вміння, що перевіряється. Ця схема оцінювання розроблена дуже вдало, тому що спочатку надає обидві частини інформації, необхідні для здобуття кожного балу, а потім включає перелік різних можливих відповідей. Коли бал можна здобути у різний спосіб, наприклад, 2 бали, чітко пропонуються всі можливі варіанти відповіді . Кожний бал також містить «ПРИМІТКУ», яка конкретно пояснює, що НЕ є прийнятним. Така примітка надає екзаменатору і експерту завдання чітке розуміння того, що є прийнятним і що є неприйнятним, не залишаючи простору для двозначності.

Приклад 5: Відкрите завдання з розгорнутою відповіддю

ВМІННЯ: Проаналізувати приклади причини і наслідків, фактів і думки.

ОСНОВА: Опишіть чотири факти, які трапилися, коли відмовив стабілізатор напруги на Фоккер Т-2.

СХЕМА ОЦІНЮВАННЯ:

Бали (0-4):	**Відповідь учня**
4	Відповідь містить чотири з наступних правильних наслідків: • Електрична система перестала виробляти електрику для зарядки батарейок. • Макреді взяв на себе управління літаком з місця другого пілота. • Келлі перевірив двигун і з'ясував в чому проблема. • Келлі зняв стабілізатор, щоб його відремонтувати.
3	Відповідь включає три з правильних вищенаведених наслідків, а один наслідок є неправильним або відсутній.
2	Відповідь включає дві з правильних вищенаведених наслідків, а два є неправильними або відсутні.
1	Відповідь включає один з правильних вищенаведених наслідків, а три є неправильними або відсутні.
0	Бали не надаються, якщо твердження не існує або є помилковим. Відповідь не містить жодного з правильних наслідків, зазначених вище, відповідь неповна, не має відношення до цього питання або відсутня.

ОБГОВОРЕННЯ: Зверніть увагу на те, що і питання в основі завдання і схема оцінювання належним чином стосуються вміння, що перевіряється. Схема оцінювання містить всі основні компоненти, включаючи чітке пояснення очікування правильної відповіді на всі три бали, позначені прийнятні приклади і паралельні формулювання. Зверніть увагу на те, що для балів 3, 2 і 1, схема оцінювання містить посилання на приклади, надані за 4 бали, замість того, щоб повторювати їх. Так як є більше ніж чотири можливих відповіді на це питання, всі можливі відповіді включені до схеми оцінювання.

РОЗДІЛ 7. ТЕХНІЧНІ ХАРАКТЕРИСТИКИ ТЕСТОВИХ ЗАВДАНЬ

До цього моменту перевірка тестових завдань проводилася тільки дорослими. Відповіді учнів на кожне питання потрібно фіксувати і аналізувати, тому що учні не завжди відповідають так, як очікувалося - вони часто мають інше розуміння та сприйняття, яке відрізняється від точки зору та розуміння дорослих. Таким чином, для проведення окремої апробації завдання укладаються у варіанти подібно до тестових зошитів в операційних тестах, і ці варіанти надаються випадковій та репрезентативній вибірці учнів. Це ідеальний метод для апробації великої кількості завдань або для апробації завдань з розгорнутою відповіддю з метою апробувати інструкції до завдання. Одним із недоліків проведення такої окремої апробації є те, що учні можуть не продемонструвати максимальний рівень успішності при виконанні тесту, який для них не має особливого значення. Це особливо стосується учнів старших класів. Крім того, до впровадження системи стандартизованого оцінювання навчальний процес у школі може не бути настільки узгодженим із загальними стандартами щодо змісту, як це може бути, коли система стандартизованого оцінювання діє.

Іншим методом проведення апробації є апробація шляхом включення апробаційних завдань в операційні варіанти тесту. Операційними варіантами тесту є ті, що призначені для всіх учнів і є джерелом тестових балів. Це є ідеальним методом для поповнення банку завдань множинного вибору та завдань з короткою відповіддю, оскільки проведення апробації завдань проходить в процесі реального тестування, коли учні намагаються продемонструвати якомога вищий рівень успішності. Відповіді учнів на такі завдання використовуються тільки для оцінки завдань з метою їх можливого подальшого включення до операційних тестів та не впливають на отримані учнями тестові бали. Крім того, аналіз завдань проводиться на рівні групи, і особистість учнів ніколи не ідентифікується. Дані відповідей аналізуються, та встановлюються різні статистичні характеристики завдань. Статистичні характеристики завдань використовують для оцінювання таких аспектів як складність та розподільча здатність завдань. Статистичними характеристиками завдань за класичною теорією тестів є бісеріальна кореляція, бісеріали для кожного варіанту відповіді, показник **p**, квантиль-аналіз, та диференційне функціонування завдання (DIF).

КОЕФІЦІЄНТ БІСЕРІАЛЬНОЇ КОРЕЛЯЦІЇ

Класичною статистичною характеристикою, що має розглядатися, є коефіцієнт бісеріальної кореляції. Ця характеристика визначає, наскільки добре завдання розподіляє учнів на тих, хто має обмежене володіння матеріалом, від тих, хто відмінно опанував матеріал. Показник бісеріальної кореляції може коливатися від -1.0 до 1.0. Висока позитивна бісеріальна кореляція вказує на те, що учні з високим рівнем досягнень з більшою вірогідністю вибирають правильні відповіді, ніж учні з низьким рівнем. Низька або нульова бісеріальна кореляція вказує на те, що учні з високим рівнем знань вже не вибирають правильні відповіді з більшою вірогідністю, ніж учні з низьким рівнем. Завдання з бісеріальною кореляцією .25 або менше вважаються неприйнятними для використання у тестах, а завдання з негативною бісеріальною кореляцією ніколи не можна використовувати в тесті. Якщо завдання не диференціює належним чином, воно не надає додаткової інформації про успішність учнів, а якщо бісеріальна кореляція надто низька, завдання не надає ніякої інформації взагалі. Крім того, такі завдання можуть в дійсності оцінювати зовсім інші аспекти конструкту, що вводить похибку у вимірювання того конструкту, для оцінювання якого розроблений тест.

БІСЕРИАЛЬНА КОРЕЛЯЦІЯ ДИСТРАКТОРІВ

Бісеріальна кореляція для кожного варіанту відповіді у завданнях множинного вибору показує, як спрацьовують дистрактори та ключ. Якщо правильна відповідь повинна мати позитивну бісеріальну кореляцію, показник бісеріальної кореляції для дистракторів має бути від'ємним. Від'ємна бісеріальна кореляція для дистракторів вказує на те, що учні з високими досягненнями з меншою вірогідністю вибирають неправильні відповіді, ніж учні з низькими досягненнями. Дистрактор, що має позитивну бісеріальну кореляцію, не є достатньо відволікаючим варіантом, тому що учні з високими досягненнями вибирають варіант з неправильною відповіддю з такою ж або більшою вірогідністю, ніж учні з низькими досягненнями. Такий дистрактор може мати певний недолік, наприклад, є надто близьким до ключа, або містить інформацію, що провокує сильних учнів надто довго думати над відповіддю, або надати ще одну правильну відповідь. Такі завдання мають бути ретельно перевірені на наявність таких недоліків. Будь-яке завдання, в якому один з дистракторів має позитивною показник бісеріальної кореляції, що є більшим за бісеріальну кореляцію для правильної відповіді, є поганим завданням. Необхідно проаналізувати завдання з тим, щоб побачити, чи можна визначити і виправити проблему. Якщо крім незадовільного дистрактора це завдання є вдалим, то цей дистрактор потрібно замінити, а завдання має знову пройти апробацію.

ПОКАЗНИК Р

Показником складності завдання є статистичне значення р, яке вказує на відсоток учнів, які правильно відповіли на завдання. Завдання, в якому значення р перевищує .90 загалом є надто легкими для тієї групи, в якій воно було застосовано, це також може вказувати на те, що завдання оцінює вміння і навички, які були сформовані у попередньому навчальному році. І навпаки, завдання із значенням р нижче за рівень випадковості (наприклад, .25 у тестовому завданні з чотирма можливими варіантами відповіді) також будуть позначені як проблемні. Дуже низьке значення р може означати, що ключ до завдання вказаний невірно, або що завдання оцінює знання або вміння, які учень ще не опановував, або що завдання надто складне для даного року навчання.

КВАНТИЛЬ-АНАЛІЗ

Квантиль-аналіз надає більше інформації про те, як ключ та дистрактори діють у завданнях множинного вибору. Квантиль-аналіз дозволяє користувачеві оцінити розподільчу здатність та складність завдання водночас. В таблиці нижче наведений приклад квантиль-аналізу для завдання множинного вибору. Показники успішності складання тесту по всій вибірці розділені на п'ять рівних частин, де учні з найкращими результатами розташовані у Квантилі 1, а учні з найгіршими результатами – у Квантилі 5.

	A	B	C	D (Ключ)
Квантиль1	4,6	17,9	7,8	69,6
Квантиль 2	8,9	30,2	13,8	47,0
Квантиль 3	12,0	36,0	17,2	34,5
Квантиль 4	14,4	40,0	19,3	25,7
Квантиль 5	17,5	40,3	22,4	17,0
Загалом	11,6	33,1	16,2	38,3

Ключ і дистрактори для цього завдання безумовно працюють добре. Як показує колонка ключа, відсоток учнів, які обирали правильну відповідь, зменшується з верхнього рядка до нижнього. 69,6% учнів з найкращими результатами вибрали правильну відповідь, у той час як з квантиля 5 правильну відповідь вибрали тільки 17,0%. Відсоток по кожному з

дистракторів зростає від Квантиля 1 до Квантиля 5, а відсоток у загальному нижньому рядку показує, що дистрактор В є найсильнішим для всіх учнів. У цьому завданні бісеріальну кореляцію для дистрактора В потрібно перевірити, щоб впевнитися, що вона від'ємна.

Квантиль-аналіз є також інформативним для завдань відкритого типу. Таблиця нижче містить такий аналіз для завдання відкритого типу з максимальним балом 2.

	0 балів	1 бал	2 бали	Відповіді немає
Квантиль1	2,4	5,7	91,8	0.1
Квантиль 2	5,8	7,1	87,0	0,2
Квантиль 3	9,6	13,2	76,7	0,4
Квантиль 4	13,4	17,9	68,1	0,6
Квантиль 5	21,1	26,6	48,5	3,4
Загалом	10,2	13,8	75,0	0,9

Як можна побачити, це завдання добре працює. Відсоток учнів, які отримали 0, 1 та 2 бали, зростає від Квантиля 1 до Квантиля 5, а відсоток для кожного квантиля зростає з 0 балів до 2.

ДИФЕРЕНЦІЙНЕ ФУНКЦІОНУВАННЯ ЗАВДАННЯ (DIF)

Статистичний показник DIF вказує на те, чи завдання спрацьовує однаково у різних групах. Іноді DIF-аналізу розуміють невірним чином. Такий аналіз проводиться для того, щоб визначити потенційну необ'єктивність завдання і ні в якому разі не стосується різниці в успішності між різними групами. Дані, що розглядаються, стосуються окремого завдання, і визначають, чи є це завдання таким, що потенційно може надати перевагу або, навпаки, бути несприятливим фактором для будь-якої з груп учнів, які мають однаковий рівень академічних досягнень. Потенційна необ'єктивність присутня у випадку, коли учасники тестування, які мають однаковий рівень досягнень та належать до різних груп мають різну вірогідність надати правильну відповідь на завдання. Незважаючи на те, що завдання, які містять потенційну необ'єктивність, відмічаються як проблемні, потенційна необ'єктивність не може бути визначена аналізом DIF, оскільки може існувати реальна різниця у різних групах через відмінності у викладанні. Завдання мають бути ретельно перевірені, щоб визначити, чи дійсно вони містять якийсь чинник необ'єктивності. Зазвичай завдання перевіряються незалежними експертами, група яких представляє розмаїття вибірки учасників тестування.

Таблиця нижче надає приклади аналізу по квантилях. Перша таблиця показує відсутність різниці між двома етнічними групами: білими та чорношкірими/ афроамериканцями. Друга таблиця показує розбіжність між цими двома групами.

	Складність завдання (p-показник) відповідно до рівня знань і умінь				
Етнічна приналежність	Нижчий квантиль	2 Квантиль	3 Квантиль	4 Квантиль	Вищий квантиль
Чорношкірі/ афроамериканці	.20	.35	.50	.65	.80
Білі	.20	.35	.50	.65	.80

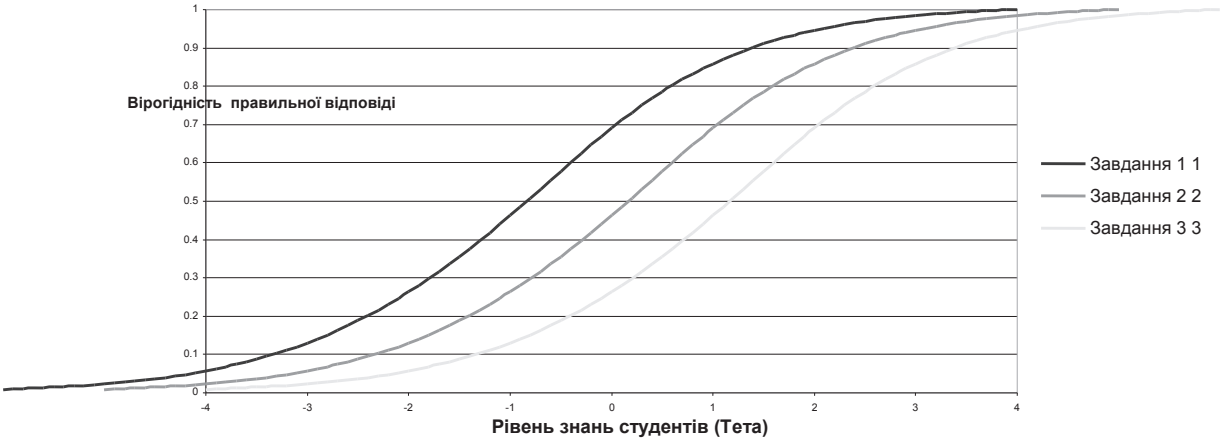
	Складність завдання (p-показник) відповідно до рівня знань і умінь				
Етнічна приналежність	Нижчий квантиль	2 Квантиль	3 Квантиль	4 Квантиль	Вищий квантиль
Чорношкірі/ афроамериканці	.05	.20	.35	.50	.65
Білі	.20	.35	.50	.65	.80

Для експертів значення DIF виражено літерами A, B і C, де A вказує на відсутність DIF, а C – на потенційно високе DIF. Позначки плюс (+) чи мінус (-) показують, до якої групи відносяться ці літери. Плюс показує потенційну необ'єктивність на користь більшості, у той час як мінус показує потенційну необ'єктивність на користь меншості.

ТЕОРІЯ ВІДПОВІДІ НА ЗАВДАННЯ (IRT)

Теорія відповідей на завдання (IRT) включає математичний аналіз даних про відповіді на завдання для того, щоб охарактеризувати відповіді учнів на тестові завдання. Моделі IRT можуть оцінювати складність завдання, розподільчу здатність, а також показник вгадування правильної відповіді у завданнях множинного вибору. В штаті Огайо використовують однопараметричну модель Раша, яка оцінює тільки параметр складності завдання. Модель часткового заліку Мастерса (Masters' Partial Credit Model) використовують для оцінки параметрів для завдань відкритого типу. Складність завдань виражається у величинах тета. Середнє значення по шкалі тета дорівнює нулю, стандартне відхилення – одиниці, так щоб значення тета, для зручності, виражалося в числах від -3 до 3, де більші числа означають вищі академічні досягнення. Величина тета обчислюється як точка, в якій для учня з даним рівнем успішності вірогідність надання правильної відповіді на завдання становить 0,5. Банк тестових завдань відцентрований по середньому рівню складності першого операційного тесту, таким чином середнє значення складності першого операційного тесту дорівнює нулю. Простіші завдання мають менші значення тета, тоді як більш складні завдання мають більші значення тета. Моделі IRT особливо корисні для укладання та урівнювання тестових варіантів, тому що параметри завдань є відносно незалежними від вибірки, що дозволяє розташовувати їх на одній шкалі.

Величина p не залежить від вибірки учнів, які проходять тестування. У моделях IRT складність завдань калібрується по відношенню до інших завдань, так щоб визначення складності не залежало від вибірки, за допомогою якої визначаються рівні знань та навичок.



Необхідно проаналізувати весь набір статистичних показників завдань, перед тим як заносити їх до операційної частини банку тестових завдань, тому що кожний з показників надає різну інформацію про те, як учні працюють із завданням. Чим більше інформації перевіряє експерт, тим краще він/вона проінформовані стосовно завдання, що в результаті дозволяє приймати кращі рішення. Необхідно витримувати баланс при вирішенні, коли завдання потрібно залишити, а коли відхилити. Важливо пам'ятати, що при створенні високоякісних тестових варіантів необхідно включати завдання з гарними характеристиками вимірювання – включення низькоякісних завдань призводить до укладання тестових варіантів з поганими психометричними характеристиками.

РОЗДІЛ 8. УКЛАДАННЯ ТА ЕКСПЕРТНА ОЦІНКА ТЕСТОВИХ ВАРІАНТІВ

УКЛАДАННЯ ТЕСТОВИХ ВАРІАНТІВ

Для будь-якого тесту в рамках програми оцінювання на рівні штату потрібна висока технічна якість. Якість тесту включає надійність, послідовність вимірювання, валідність, тобто ступінь того, наскільки зібрані докази та теорії підтверджують спеціальні інтерпретації тестових балів визначеного використання результатів тесту, а також точність змісту та ключів до тесту.

Було б відносно легко укладати тест, використовуючи тільки статистичні дані і не беручи до уваги питання змісту та самого навчального процесу. Так само було б відносно легко укладати тест, зважаючи тільки на питання змісту та навчального процесу і не беручи до уваги технічні аспекти. Коли дві вимоги – технічна якість та якість змісту – накладаються одна на одну, укладання тесту стає складнішим і в багатьох випадках стає мистецтвом.

Механічні аспекти тесту – організація, вказівки та його друк – такі ж важливі, як і якість завдань. Багато факторів безпосередньо впливають на здатність студента проявити володіння матеріалом під час оцінювання, в той час як інші впливають на можливість точно та ефективно обчислювати тестові бали. Ще інші фактори впливають на висновки, які можна зробити на основі результатів тестування.

Для забезпечення порівнюваності тестових варіантів, кожний варіант повинен відповідати тестовим специфікаціям. Це означає, що необхідно дотримуватися постійної кількості балів у кожній категорії балів, а також кількості завдань та їх типів по кожному стандарту/показнику. Крім відповідності вимогам до змісту, зазначеним у тестових специфікаціях, операційні варіанти тестів повинні відповідати статистичним вимогам до тестового варіанту. Таким чином, після того укладання тесту і проведення експертної оцінки стосовно відповідності його змісту, тестові варіанти мають також проходити експертизу на статистичну прийнятність.

Коли розробник тесту надає експертну оцінку змісту тестового варіанту, крім підтвердження того, що дотримані всі вимоги стандарту, експерт має також перевірити, чи не містять тестові завдання підказки один до одного. Це означає, що в одному завданні не повинно бути матеріалу, який є відповіддю до іншого завдання. Це також означає, що відповідь учня на завдання не повинна залежати від відповіді на будь-яке інше завдання і має бути статистично незалежною. Це поняття називається «локальною незалежністю». Така незалежність найбільш типово порушується, якщо одне завдання містить натяк на відповідь до іншого завдання. У такому випадку, справжній рівень знань учня при відповіді на друге завдання не можна виміряти.

Зазвичай легко помітити, що одне завдання містить підказку до відповіді на інше завдання. Більш складно побачити, у якому випадку учня буде винагороджено або покарано двічі при наявності чи відсутності необхідних знань або навичок. Це також є порушенням локальної незалежності і трапляється зазвичай, коли два чи більше завдання пов’язані із загальним стимулом.

Крім цього, розробникам при укладанні та експертній оцінці тестів необхідно зважати на моменти загального характеру. Багато з них здаються очевидними, але їх можна легко випустити з уваги. В процесі розробки тестові варіанти звичай перевіряють на наступне:

- Завдання розташовуються у правильній послідовності.
- Кожне з завдань одного типу представлено послідовно:

- Перелік варіантів відповіді в завданнях множинного вибору є послідовним.
- Варіанти відповідей позначені літерами A, B, C і D.
- Всі графічні елементи ілюстративного матеріалу представлені послідовно.

- Всі таблиці та діаграми повинні мати назви та представлені у однаковому форматі.
- Кількість варіантів відповідей є приблизно однаковою у всьому тестовому варіанті.
- Ключі до завдань повинні бути перевірені першим експертом, а також додатково одним незалежним експертом.
- Всі стимули мають саме ті завдання, що пов'язані з ними.
- Тематика завдань, уривків або стимулів не є дуже подібною.
- Варіант тесту не містить орфографічних або граматичних помилок, або помилок в ілюстративному матеріалі.
- Формулювання, структура та загальний вигляд завдання є такими, як і при апробації цього завдання.
- Витримується гендерний та етнічний баланс.
- Немає випадків, коли два відкриті завдання розташовані одне за одним.
- Блоки завдань до уривку не починаються або не закінчуються відкритим завданням.
- Питання, що відносяться до одного уривку /стимулу розташовані на тій самій сторінці або на сторінці поруч.
- Нумерація сторінок та її розташування є правильними.
- Розташування літер або назв стосовно змісту є правильним.
- Верхні та нижні колонтитули на всіх сторінках є вірними з дотриманням однакового інтервалу.
- На кожній сторінці є належні позначки «Перейдіть на наступну сторінку» або «Кінець».
- Форматування всіх текстів та тестових завдань перевірено на належне використання інтервалів, жирного шрифту та курсиву.
- Завдання, включені до тесту з метою апробації, не розташовані на одній сторінці разом з операційними завданнями.
- Кожне завдання та тестовий варіант в цілому перевірені згідно з відповідними рекомендаціями щодо стилю.
- Інструкції до завдань є послідовними і точними.

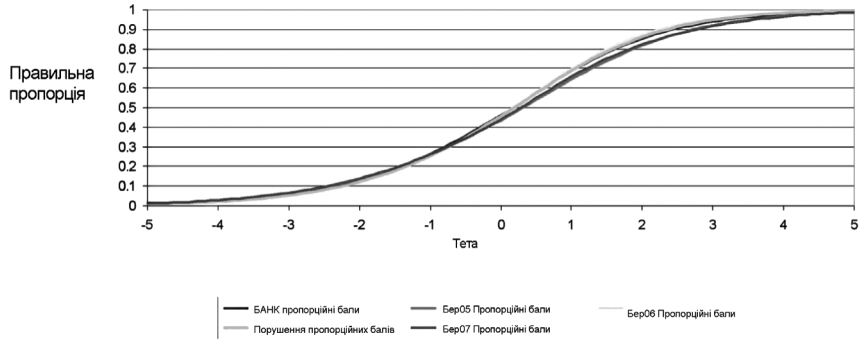
ТЕХНІЧНА ЯКІСТЬ ТЕСТУ

Розробка завдань спрямована на вимірювання рівня підготовки з певної предметної галузі. На відміну від тестів, орієнтованих на норму, в яких розробка завдань спрямована на здатності середнього учня, рівень складності завдань у критеріально-орієнтованих тестах може не відповідати рівню підготовки середнього учня, тому що цільові стандарти успішності можуть бути вищими або нижчими, ніж існуючий на даний момент середній показник рівня досягнень. При калібруванні рівня складності завдання середня складність тестових завдань відцентрована по відношенню до першого операційного тесту таким чином, що середня складність першого операційного тесту дорівнює нулю. Всі наступні операційні варіанти тестів повинні бути настільки наближено еквівалентними до першого операційного варіанту тесту, наскільки це можливо; таким чином щоб цільовий рівень складності дорівнював нулю при всіх діях по укладанню тестів. Орієнтуючи складність всіх операційних тестових варіантів на перший операційний тест, розробники тестів також можуть забезпечити те, щоб сирі бали, прив'язані до границь рівнів успішності, залишалися відносно стабільними при проведенні всіх тестових сесій.

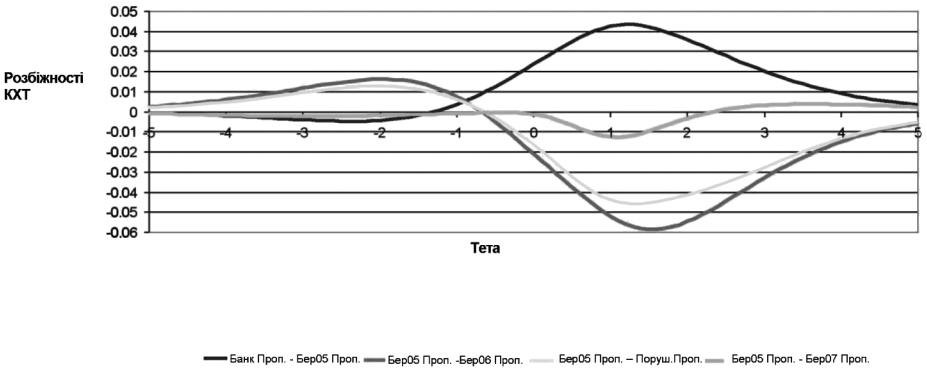
Щоб виміряти діапазон рівня досягнень, від базового рівня до просунутого, складність тестових завдань також повинна варіюватися. При укладанні першого операційного варіанту тесту розподіл складності тестових завдань оцінюється відносно до банку завдань. Завдяки цьому розподіл складності тестових завдань у першому операційному варіанті

тесту відображає розподіл складності завдань, який передбачався для розробки завдань. До того ж, цей підхід допомагає впевнитися, що банк має оптимальну конфігурацію для укладання тестових варіантів з порівнянним розподілом складності завдань.

Розподіл складності завдань оцінюється шляхом аналізу співвідносності характеристичних кривих тестів до запропонованих тестових варіантів. Характеристичні криві тесту показують відсоток загальних сирих балів, пов'язаних з кожним рівнем успішності. Якщо інформація, представлена певним тестовим варіантом, відрізняється від першого операційного варіанту більш ніж на 5% для будь-якого рівня успішності, то запропонований тестовий варіант вважається неприйнятним. Це допомагає забезпечити еквівалентність тестових варіантів як в плані середнього рівня складності тесту, так і в плані розподілу складності завдань у тестових варіантах. Нижче наведений приклад характеристичної кривої завдання. Зверніть увагу на те, як близько знаходяться криві одна до одної.



Наведений нижче графік використовується для оцінки розбіжностей між характеристичними кривими тестів. При аналізі розбіжностей кривих слід звертати увагу на декілька важливих аспектів. По-перше, розбіжності між характеристичними кривими тестів можуть виявлятися в певних ділянках кривих, що відповідають шкалі рівня успішності. Ці розбіжності відображають різні акценти у тестовій інформації варіантів тесту на шкалі рівнів успішності. Важливою причиною забезпечення того, що перший операційний варіант тесту відповідає банку тестових завдань, є факт, що розробка наступних операційних варіантів залежатиме від наявності завдань у тесті; і тому, якщо перший операційний варіант тестів відрізняється від банку розподілом тестової інформації, буде все складніше співвідносити тестові варіанти з першим операційним тестом через те, що вимоги до банку будуть непропорційними до наявності завдань в тих чи інших ділянках шкали рівня успішності. Важливим є те, що такі види проблем при укладанні тестів можуть бути вирішені шляхом одночасного створення певної кількості операційних варіантів тестів. Одночасне створення кількох тестових варіантів забезпечує більшу гнучкість при перестановці завдань та уривків у межах даних тестових варіантів з метою досягнення бажаних властивостей вимірювання. Як тільки ухвалюється остаточний варіант тесту, кількість наявних у банку завдань для подальшого укладання тестів зменшується, що набагато ускладнює подальше укладання психометрично виважених тестових варіантів.



Коли варіанти тестів укладаються вперше, розраховуються попередні таблиці конвертації балів. Ці таблиці ґрунтуються на параметрах завдань, отриманих під час апробації. Сирі граничні бали для кожного з рівнів виконання повинні бути якомога більше наближеними до граничних балів, визначених під час процедури встановлення стандартів. Це є найбільш важливим, коли йдеться про кваліфікаційні граничні бали. Якщо граничний бал для будь-якого рівня успішності змінився більше ніж на 1, тестовий варіант необхідно проаналізувати з тим, чи не можна поміняти або виключити деякі завдання.

КОНТРОЛЬНІ ПРИМІРНИКИ

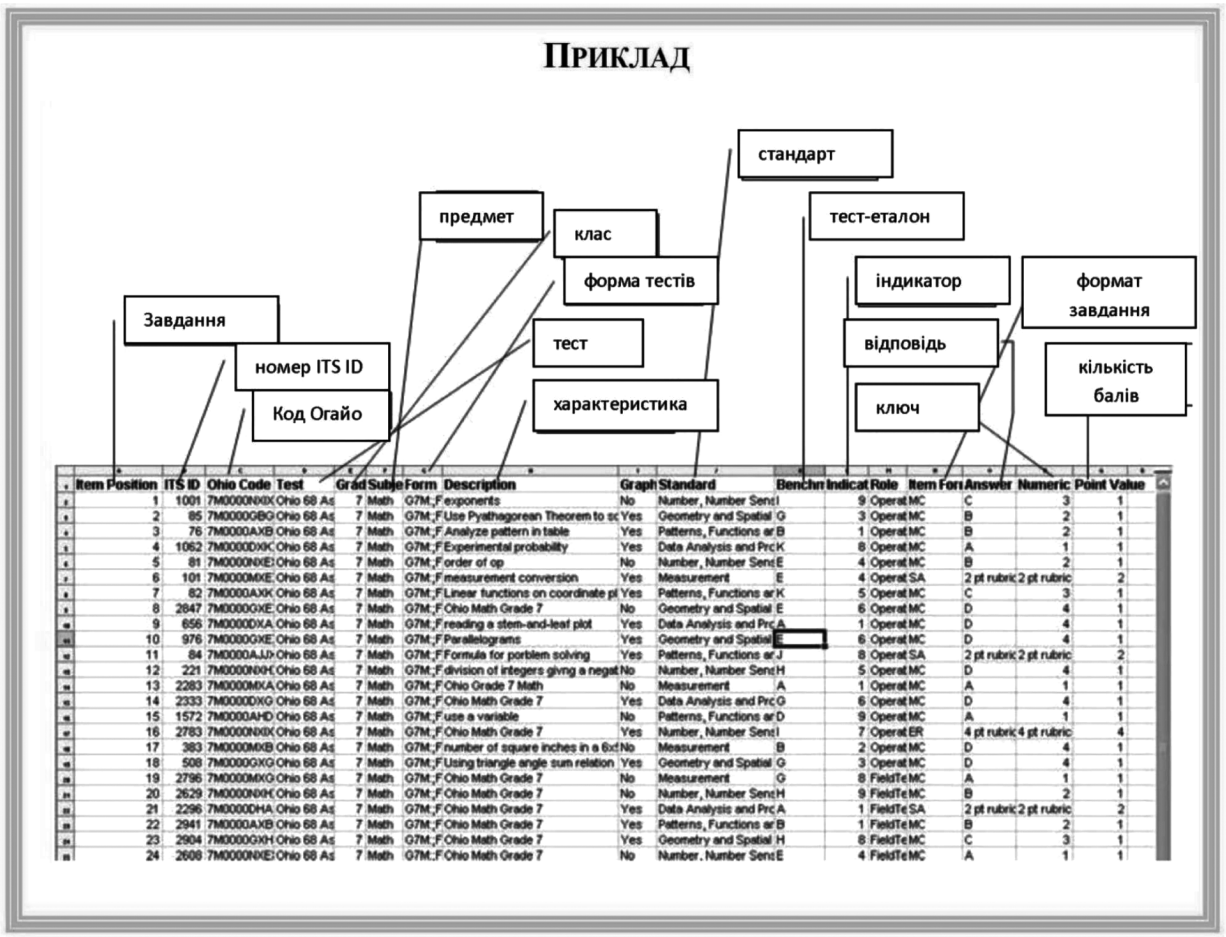
Після дотримання всіх вимог стосовно змісту та статистичних характеристик розробник тестів може перейти до створення контрольного примірника тестового варіанту. Такий контрольний примірник містить перелік включених тестових завдань у такій послідовності, в якій вони будуть розташовані в операційному тесті. Контрольний примірник містить зміст завдань та статистичну інформацію до завдань. Разом з мапою, розробник тестів роздруковує контрольні примірники всіх завдань тестового варіанту у такому форматі, в якому на одному аркуші друкується одне завдання (такі примірники тестових завдань позначають символом «1 х 1»).

Спочатку контрольні примірники тестових варіантів і окремих завдань проходять внутрішню експертизу для визначення їх змістової відповідності тестовим специфікаціям. Після такої експертизи контрольні примірники відправляють на затвердження клієнтові, перш ніж форматувати матеріали як тестовий варіант для друку. Це дозволяє клієнтові переглянути завдання на відповідність змісту та статистичним характеристикам. Якщо на будь-якому етапі експертизи завдання зазнають змін, оновлена версія варіанту та контрольний примірник знову надсилаються клієнтові.

Клієнт несе відповідальність за відповідність тестового варіанту тестовим специфікаціям: середнє значення р має бути в межах від .50 до .60; всі показники бісеріальної кореляції мають бути додатними і в ідеалі бути не менше значення .25; середнє значення тета має дорівнювати нулю; граничні бали повинні відповідати граничним балам, визначеним при встановленні стандартів. Клієнт повинен також перевірити, чи співпадає кодування завдань, наведене в електронних таблицях програми, з кодами представлених завдань.

Всі вибрані завдання повинні бути визначені як прийнятні в банку тестових завдань та в звітних документах про проведення експертних оцінок. Також має бути баланс між завданнями з використанням стимулів та завданнями без стимулів. Крім того, клієнт повинен перевірити, чи немає в одних завданнях підказок до інших, чи вірно вказані ключі до завдань, чи достатньо охоплено представлений у тесті зміст, чи немає граматичних або орфографічних помилок, чи витримується гендерний та етнічний баланс у тестовому варіанті, чи ілюстративний матеріал і графічні компоненти є точними та повними, чи є

точним зміст, що перевіряється завданнями, і чи завдання не зазнали змін після проведення апробації. Потрібно ретельно пересвідчитись у тому, що всі питання, до яких потрібно надати стимули, їх мають, що підряд не розташовано декілька завдань відкритого типу, і що уривки не починаються або не закінчуються завданнями відкритого типу. Заміна завдань в ідеалі відбувається саме на цьому етапі.



Item Position	ITS ID	Ohio Code	Test	Grad	Subje	Form	Description	Graph	Standard	Benchm	Indicat	Role	Item For	Answer	Numeric	Point Value
1	1001	7M0000NXK	Ohio 68 As	7	Math	G7M_F	exponents	No	Number, Number Sen	I	9	Operat	MC	C	3	1
2	85	7M0000GBG	Ohio 68 As	7	Math	G7M_F	Use Pythagorean Theorem to sc	Yes	Geometry and Spatial	G	3	Operat	MC	B	2	1
3	76	7M0000AXB	Ohio 68 As	7	Math	G7M_F	Analyze pattern in table	Yes	Patterns, Functions ar	B	1	Operat	MC	B	2	1
4	1062	7M0000DXK	Ohio 68 As	7	Math	G7M_F	Experimental probability	Yes	Data Analysis and Prc	K	8	Operat	MC	A	1	1
5	81	7M0000NXE	Ohio 68 As	7	Math	G7M_F	order of op	No	Number, Number Sen	E	4	Operat	MC	B	2	1
6	101	7M0000MXE	Ohio 68 As	7	Math	G7M_F	measurement conversion	Yes	Measurement	E	4	Operat	SA	2 pt rubric	2 pt rubric	2
7	82	7M0000AXK	Ohio 68 As	7	Math	G7M_F	Linear functions on coordinate pl	Yes	Patterns, Functions ar	K	5	Operat	MC	C	3	1
8	2847	7M0000GXE	Ohio 68 As	7	Math	G7M_F	Ohio Math Grade 7	No	Geometry and Spatial	E	6	Operat	MC	D	4	1
9	656	7M0000DXA	Ohio 68 As	7	Math	G7M_F	reading a stem-and-leaf plot	Yes	Data Analysis and Prc	A	1	Operat	MC	D	4	1
10	976	7M0000GXE	Ohio 68 As	7	Math	G7M_F	Parallelograms	Yes	Geometry and Spatial	E	6	Operat	MC	D	4	1
11	84	7M0000AJJ	Ohio 68 As	7	Math	G7M_F	Formula for porblem solving	Yes	Patterns, Functions ar	J	8	Operat	SA	2 pt rubric	2 pt rubric	2
12	221	7M0000NXH	Ohio 68 As	7	Math	G7M_F	division of integers givng a negat	No	Number, Number Sen	H	5	Operat	MC	D	4	1
13	2283	7M0000MXA	Ohio 68 As	7	Math	G7M_F	Ohio Grade 7 Math	No	Measurement	A	1	Operat	MC	A	1	1
14	2333	7M0000DXG	Ohio 68 As	7	Math	G7M_F	Ohio Math Grade 7	Yes	Data Analysis and Prc	G	6	Operat	MC	D	4	1
15	1572	7M0000AHD	Ohio 68 As	7	Math	G7M_F	use a variable	No	Patterns, Functions ar	D	9	Operat	MC	A	1	1
16	2783	7M0000NXK	Ohio 68 As	7	Math	G7M_F	Ohio Math Grade 7	Yes	Number, Number Sen	I	7	Operat	ER	4 pt rubric	4 pt rubric	4
17	383	7M0000MXB	Ohio 68 As	7	Math	G7M_F	number of square inches in a 6x	No	Measurement	B	2	Operat	MC	D	4	1
18	508	7M0000GXG	Ohio 68 As	7	Math	G7M_F	Using triangle angle sum relation	Yes	Geometry and Spatial	G	3	Operat	MC	D	4	1
19	2796	7M0000MXG	Ohio 68 As	7	Math	G7M_F	Ohio Math Grade 7	No	Measurement	G	8	FieldTe	MC	A	1	1
20	2629	7M0000NXH	Ohio 68 As	7	Math	G7M_F	Ohio Math Grade 7	No	Number, Number Sen	H	9	FieldTe	MC	B	2	1
21	2296	7M0000DHA	Ohio 68 As	7	Math	G7M_F	Ohio Math Grade 7	Yes	Data Analysis and Prc	A	1	FieldTe	SA	2 pt rubric	2 pt rubric	2
22	2941	7M0000AXB	Ohio 68 As	7	Math	G7M_F	Ohio Math Grade 7	Yes	Patterns, Functions ar	B	1	FieldTe	MC	B	2	1
23	2904	7M0000GXH	Ohio 68 As	7	Math	G7M_F	Ohio Math Grade 7	Yes	Geometry and Spatial	H	8	FieldTe	MC	C	3	1
24	2608	7M0000NXE	Ohio 68 As	7	Math	G7M_F	Ohio Math Grade 7	No	Number, Number Sen	E	4	FieldTe	MC	A	1	1

ДІАЗОТИПНІ КОПІЇ ТЕСТУ

Після того, як контрольні примірники тестових варіантів і окремих завдань затверджені клієнтом, створюються діазотипні копії тестових варіантів. Діазотипна копія є друкованою версію розташування завдань у тому вигляді, в якому їх будуть бачити учасники тестування. Діазокопію необхідно ретельно перевірити на відповідність затвердженим контрольним примірникам.

Часто саме на цьому етапі помічають приховані підказки одного завдання до іншого. Це трапляється тому, що тільки на першій діазотипній копії можна побачити завдання в тому вигляді, в якому вони будуть представлені на сторінці. Саме тому іноді завдання переміщують на іншу сторінку у тестовому зошиті. Якщо вимагається зміна такого типу, створюється новий контрольний примірник тестового варіанту, який з точністю відображає новий порядок розташування завдань у тестовому варіанті. Діазокопію також переглядають для того, щоб упевнитися, що форматування завдань співпадає з специфікаціями у документах, що визначають особливості стилю для тестових варіантів. Слід зазначити, що для операційних тестових варіантів формулювання завдань не можна змінювати. Формулювання завдань, включених з метою апробації, можна змінити. Наступні контрольні листи, розроблені для тестів досягнень та випускних тестів в штаті Огайо ілюструють контрольні заходи на етапі першої діазокопії.

КОНТРОЛЬНИЙ ЛИСТ ДО ПЕРШОЇ ДІАЗОКОПІЇ ТЕСТУ АКАДЕМІЧНИХ ДОСЯГНЕНЬ

Варіант_____ Дата отримання _____

____ Обкладинка (Тест академічних досягнень на рівні штату Огайо, предмет, тестовий зошит, літера, якою позначено варіант тесту, місяць та рік проведення тестування, авторські права, колір)

____ Підтвердження (класи 3-6)

Рядки для оцінок (класи 7-8)

____ Звірити інструкції до завдань з рекомендаціями DFA

____ Верхні колонтитули

* Будь-ласка, користуйтесь олівцем (розташовано в лівій частині сторінки, тільки в зошитах з завданнями)

____ Нижні колонтитули («Перейдіть на наступну сторінку», стрілка присутня у всіх колонтитулах, крім сторінки з вказівками та останньої сторінки. Якщо в центрі сторінки є напис «Перейдіть на наступну сторінку», у колонтитулі розташована тільки стрілка.)

____ Нумерація сторінок

____ Інструкції до завдань (завдання xx-xx)

____ Нумерація завдань

____ Позначення варіантів відповідей (А, Б…)

____ Перевірити точність відсилання до всіх рядків, абзаців та ілюстративного матеріалу

____ Відповідність завдань множинного вибору та відкритих завдань документу, в якому надаються ключі (тільки для класів 5-8)

____ Внесені зміни до діазокопії (Потрібно виділити всі об’єкти на електронній таблиці правим кліком, обрати команду «Показати» для виводу всіх приміток)

____ Позначки зносок у тексті та зносках виділяються жирним шрифтом, двокрапки не виділяються жирним шрифтом, кожна зноска, позначена в уривку, має означення.

____ Перевірити курсив в уривках

____ Ізольовані строки / строки, що спливають

____ ITS верифікація (Під час друку обов’язково включіть наступне: «Завдання: чинний статус»)

Код Огайо на варіанті повинен співпадати з кодом Огайо в ITS

____ Коди Огайо на варіанті відповідають кодам Огайо на контрольних примірниках, розширені відповіді розташовані у правильних рядках

____Перевірте наявність та точність інформації стосовно апробації

Коментар :

Операції з контролю якості _____ Дата _____

Зміни, впроваджені до змісту _____ Дата _____

КОНТРОЛЬНИЙ ЛИСТ ДО ПЕРШОЇ ДІАЗОКОПІЇ ВИПУСКНОГО ТЕСТУ

Варіант _____ Дата отримання _____

____ Обкладинка (Випускний тест на рівні штату Огайо, логотип, предмет, тестування, місяць, рік, авторські права, колір)

____ Підтвердження /Рядки для оцінок

____Перевірити інструкції операційним тестом, що проводився весною 2007 року, що розташований на загальному хард-диску.

____ Колонтитули

____ Нижні колонтитули («Перейдіть на наступну сторінку», лінія виникає у всіх колонтитулах, окрім останньої сторінки. Якщо в центрі сторінки є напис «Перейдіть на наступну сторінку», у колонтитулі розташована тільки лінія.)

____ Нумерація сторінок

____ Вказівки (питання хх-хх)

____ Нумерація завдань

____ Варіанти відповідей (А, Б…)

____ Питання множинного вибору та відкриті питання відповідають Документу з відповідями

____ Впроваджено внутрішні зміни до діазокопії (Потрібно виділити всі об’єкти у динамічній таблиці правим кліком, обрати команду «Показати» для розгляду всіх рядків приміток)

____ Позначки зносок у тексті та зносках виділяються жирним шрифтом, двокрапки не виділяються жирним шрифтом, кожна зноска, позначена в уривку, має означення

____ Перевірити точність відсилання до всіх рядків, абзаців та ілюстративного матеріалу

____ Перевірити курсив в уривках

____ ITS верифікація (Під час друку обов’язково включіть наступне: «Завдання: чинний статус»)

Код Огайо на варіанті повинен співпадати з кодом Огайо в ITS

____ Коди Огайо на варіанті відповідають кодам Огайо на контрольних примірниках

____ Перевірте, чи використовується Avant Garde ITC – тільки для «4» та великої літери «I»

Коментарі

Операції з контролю якості _____ Дата _____

Зміни, впроваджені до змісту _____ Дата _____

Якщо необхідно внести зміни до діазокопії 1, клієнту надсилають діазокопію 2 та оновлений контрольний примірник варіанту тесту. Клієнт перевіряє діазокопію для того, щоб впевнитися, що необхідні зміни були внесені належним чином. Якщо так, тестовий варіант переводиться у формат «до друку». Цей формат перевіряють для того, щоб пересвідчитися, що він відповідає затвердженому формату або діазокопії 2. Якщо формат «до друку» та діазокопія співпадають, тестові варіанти можна друкувати.

ГЛОСАРІЙ

Завдання: окреме питання, а для деяких типів завдань, це також набір потенційних відповідей

Відкриті завдання/Завдання відкритого типу: завдання, які вимагають від учасника тестування надати або побудувати відповідь без наявності варіантів вибору відповіді

Завдання з розгорнутою відповіддю: завдання відкритого типу, які вимагають від учасника тестування надання достатньо великого обсягу інформації або деталей під час відповіді; для виконання таких завдань відводиться від п'яти до п'ятнадцяти хвилин

Завдання з розгорнутою відповіддю: завдання відкритого типу, які вимагають від учасника тестування надання меншого обсягу інформації або деталей, ніж в завданнях з розгорнутою відповіддю; для виконання таких завдань відводиться від двох до п'яти хвилин

Завдання множинного вибору: завдання, які мають від трьох до п'яти варіантів відповіді, лише один з яких є вірним

Варіанти відповідей: запропоновані варіанти відповідей

Дистрактори: правдоподібні, але неправильні варіанти відповіді

Ключ: правильний варіант відповіді

Інструкція: частина завдання, яке представляє питання або завдання, на яке потрібно надати відповідь

Основа: питання або твердження, яке часто представляє завдання множинного вибору

Локальна незалежність: характеристика, яку демонструє відповідь, на яку не впливають відповіді на інші завдання тесту і яка є статистично незалежною від них

Одномірність: перевірка тільки однієї здібності/вміння

Оцінювання виконання: складається із завдання або серії завдань, які студент повинен виконати

Психологічний конструкт: когнітивні здібності, якими користується учасник тестування при наданні відповіді на завдання

Схема оцінювання: критерії, за якими надаються бали при оцінюванні завдання відкритого типу

ДОДАТОК

ВПРАВА 1

1. Прочитайте вміння і відповідну пару завдань нижче. Тільки одне з завдань відповідає певному вмінню. В рядку після завдання вкажіть, яке із завдань відповідає вмінню і поясніть чому.

ВМІННЯ: Визначати барви веселки та порядок, в якому вони з'являються.

ЗАВДАННЯ А: Який колір йде попереду жовтого у веселці?

A. блакитний
B. помаранчевий*
C. червоний
D. фіолетовий

ЗАВДАННЯ В: Скільки кольорів у веселці?

A. 4
B. 5
C. 6
D. 7*

2–11. Вкажіть недоліки в побудові завдань, користуючись наданими нижче пунктами.

A. зайва деталь в основі завдання
B. пасивна форма
C. неоднозначна основа
D. ключові слова в основі і в ключі
E. варіанти відповіді не є паралельними граматично
F. ключ довший за дистрактори
G. негативна конструкція
H. дистрактор дуже близький до ключа
I. явний дистрактор
J. ключ, який ніколи не повинен використовуватися

1. В астрономії, що є *червоне зміщення* ?

A. рух об'єктів через всесвіт, що розширяється
B. зовнішня газова оболонка, що оточує зірку
C. зміщення світла у бік довгих червоних довжин хвилі*
D. зірка стала крихітною в результаті гравітаційного краху

2. Ким першим було зрозуміло кровообіг?

A. Бойль
B. Гален*
C. Гіппократ
D. Ньютон

3. Коли бог потойбічного світу вкрав доньку Деметри, Персефону, що вона робила?

A. вимагала допомоги від Зевса*
B. йшла від усіх на якийсь період часу
C. зверталася з проханням до бога потойбічного світу
D. перетворила землю на пустелю

4. Яка частина пасльону є отруйною?

A. листя
B. корені
C. зерна
D. всі з вищенаведеного *

5. Яке речення НЕ містить слово з орфографічною помилкою?

A. Анабель пише на папері кремового кольору.*
B. Всі вважають, що нова вистава причудова.
C. Я зашов забрати запрошення на весілля.
D. Сабаці дозволяють їсти зі столу.

6. Хто був першим Президентом Сполучених Штатів, від 1789-1797?

A. Джордж Вашингтон*
B. Джеймс Медисон
C. Джон Адамс
D. Томас Джеферсон

7. Яка з наступних порід є собакою для полювання ?

A. Гончак*
B. Шпіц
C. Пудель
D. Уельський тер'єр

8. 2 x 38 = ___

A. 40
B. 66
C. 76*
D. 79

9. Яка функція імунної системи?

A. Вона управляє запасом тіла повітря
B. передавати інформацію мозку
C. захищати тіло від патогенних мікроорганізмів*
D. розвиток швидкості і сповільнення процесів в організмі

10. Що таке північне сяйво?

A. смуги світла, іноді видного в деяких областях вночі, викликаних зарядженими частинками в магнітній поля землі*
B. вогні, що випускають світляки і інші тварини
C. віддзеркалення місячного світла від снігу
D. метеори

ВІДПОВІДІ І ПОЯСНЕННЯ:

1. (D) У правильному варіанті відповіді використовуються два ключові слова, які містяться в основі: «червоний» і «зміщення».

2. (B) Пасивна конструкція в основі; основа має формулюватися іншими словами: «Хто першим зрозумів кровообіг?»

3. (C) Наприкінці основи слово «вона» («що вона зробила»), може відноситися або до Деметри або до Персефони. Ця основа повинна бути виправлена, щоб виключити таку двозначність.

4. (J) «Все з вищенаведеного» не повинно використовуватися.

5. (G) Основа сформульована із запереченням.

6. (A) Знання дат, протягом яких Вашингтон займав пост, не є важливим для того, щоб знати, що він був першим президентом; насправді, дати можуть заплутати учасника тестування.

7. (H) Автор завдання склав його, думаючи що гончак – єдина з перелічених порід, що є собакою для полювання. Насправді, хоча і пудель і уельський тер'єр в основному утримуються як декоративні собаки, обидві породи все ще іноді використовуються для полювання.

8. (I) Знаючи, що будь-яке число при множенні на 2 в результаті надає парне число, варіант D можна виключити без виконання обчислення.

9. (E) Ключ і варіант B – сполучення слів, в яких головним словом є дієслово в неозначеній формі. Варіант A представлений повним реченням, а варіант D є іменниковим сполученням слів. Варіанти потрібно переписати таким чином, щоб вони були граматично паралельними.

10. (F) Такий варіант як A, який є набагато довшим (і стилістично більш технічним), ніж три дистрактори, виглядає правдоподібнішим і тому його більш ймовірно будуть вибирати учасники тестування.

ДЛЯ НОТАТОК:

USAID
ВІД АМЕРИКАНСЬКОГО НАРОДУ

USETI
ALLIANCE

АМЕРИКАНСЬКІ РАДИ
З МІЖНАРОДНОЇ ОСВІТИ

вул. Еспланадна, 20, 6-й поверх, Київ, Україна, 01001 Т 044 2893952 www.useti.org.ua

Це видання стало можливим завдяки підтримці американського народу відповідно до Договору про співпрацю #121-A-13-00001 з Агентством США з міжнародного розвитку (USAID). Думки, висловлені тут, належать авторам і не обов'язково відображають точку зору USAID чи Уряду США.