УДК 378.146

Малік Олександр Іванович
начальник Навчально-наукового центру мовної підготовки
Академія сухопутних військ імені гетьмана Петра Сагайдачного, м. Львів, Україна
malik_alex@mail.ru

ОГЛЯД ДІЄВИХ МЕТОДІВ СТВОРЕННЯ ТЕСТОВИХ ЗАВДАНЬ І ПРОЕКТУ ТЕСТУ, ПРИЗНАЧЕНИХ ДЛЯ ВИЗНАЧЕННЯ РІВНЯ ПРОФЕСІЙНОЇ ПІДГОТОВКИ ВІЙСЬКОВОСЛУЖБОВЦІВ

Анотація. Актуальність викладеного матеріалу обумовлена тим, що в Україні триває процес упровадження тестування як основного засобу контролю засвоєння знань. Стаття присвячена теоретико-практичним аспектам проблеми укладання тестових завдань, призначених для визначення рівня професійної підготовки. Розглянуті базові елементи процесу створення тестових завдань і проекту тесту, дотримання яких дає основу для створення надійного засобу контролю. Сформульовано перелік принципово важливих критеріїв укладання тестових завдань. Висунуті основні вимоги до структурних складових тестових завдань з метою отримання високоякісних тестів. Вказані основні параметри, що впливають на ефективність тесту в цілому.

Ключові слова: тестування; критерії відбору тестового матеріалу; шкала оцінювання; рівень професійної підготовки.

1. ВСТУП

Постановка проблеми. На сучасному етапі розвитку суспільства Україна знаходиться у процесі реформування, розбудови, а подекуди й створення нових моделей і форм багатьох своїх інституцій. У цьому широкомасштабному процесі чи не головною ланкою слід вважати професійну підготовку національних кадрів, спроможних брати участь у вирішенні різноманітних завдань, які постають перед державою. Саме тому зараз актуальною і життєво необхідною для вирішення стала проблема розробки системи оцінки якості професійної підготовки усіх категорій спеціалістів, яка б давала об'єктивну порівняльну інформацію на різних рівнях цього процесу. Тільки так можна буде оцінити ефективність діяльності як окремого навчального закладу, так і всієї професійної освіти.

Важливою складовою системи оцінювання якості професійної підготовки виступає педагогічне вимірювання рівня знань, що визначає суть педагогічної діагностики, від якості якої великою мірою залежить як ефективність організації процесу підготовки фахівця, так і рівень його професійної компетенції. В Україні, як і в розвинених країнах світу, триває процес упровадження тестування як основного засобу педагогічного вимірювання. Наразі констатується незворотний і динамічний перехід до комп'ютерних технологій, що свідчить про розширення інноваційного технологічного етапу в галузі педагогічної діагностики [3, с. 6].

Упровадження тестування в систему професійної підготовки фахівців неможливе без жорсткої стандартизації процесу створення тестових завдань і проекту тесту. Здійснення педагогічного вимірювання має базуватись на відповідному методичному обґрунтуванні для розроблення і застосування сучасних засобів тестування й обробки отриманих результатів з урахуванням основних психолого-педагогічних принципів професійної підготовки. Тільки валідні й надійні тести можуть адекватно визначити рівень набутих професійних компетенцій.

© О. І. Малік, 2013.

У рамках даного дослідження нами були проаналізовані наукові праці А. Б. Андрєєва, А. О. Бєльдягіної, І. М. Зварича, Л. Інгекамп, Н. Н. Самилкіної, Г.О. Хубаєва, Н. В. Чорної, M. Brichcin, H. Madsen, G. Rasch та інших дослідників, у працях яких розглядається методологія виміру навчальних досягнень. Критерії відбору тестового матеріалу розглядалися з опорою на дослідження В. С. Аванесова, І. Є. Булах, В. І. Васильєва, П. Клайна, М. Б. Мельникової та ін. Аналіз наукових праць свідчить про неабияку актуальність досліджуваної проблеми у зв'язку зі стрімкими темпами впровадження тестування як методу визначення рівня компетентності у процесі професійної підготовки фахівців.

Метою даної статті є висвітлення основних складових процесу створення тестових завдань і проекту тесту для визначення рівня професійної підготовки.

2. МЕТОДИ ДОСЛІДЖЕННЯ

Дослідження проводилось в рамках НДР «Удосконалення методики тестування особового складу Збройних сил України з використанням новітніх комп'ютерних технологій з урахуванням видової специфіки змісту мовної підготовки» Навчально-наукового центру мовної підготовки Академії сухопутних військ імені гетьмана Петра Сагайдачного (Свідоцтво про реєстрацію авторського права № 30458 від 28.09.2009 р.). Під час дослідження використовувався комплекс методів, зокрема: науковий аналіз, порівняння і синтез даних; узагальнення і систематизація вітчизняного й зарубіжного досвіду використання комп'ютерних технологій у діагностиці професійної підготовки.

3. РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

У наш час існує багато класифікацій видів тестів, тому дати універсальне визначення для всіх видів навряд чи можливо і потрібно. У даній роботі ми говоритимемо лише про проектування тих видів тестів, котрі використовуються як вимірники рівня професійної підготовки і стали складовою комп'ютерних програм визначення рівня професійної підготовки військовослужбовців. Під проектуванням тесту розуміється педагогічне проектування, тобто всестороння розробка змістових і технологічних характеристик тесту.

Для визначення рівня професійної компетентності в певній галузі людської діяльності застосовують діагностичні тести, призначені для вимірювання рівня набутих компетенцій у певній навчальній дисципліні, котра формує фахову компетентність спеціаліста.

Одна з класифікацій видів тестів поділяє їх за принципом укладання на *стандартизовані* і *не стандартизовані*.

Швидке поширення тестових методик, зокрема у процесі діагностики рівня фахової компетентності, яке спостерігається останнім часом в Україні, відбувається здебільшого за рахунок використання саме нестандартизованих тестів. Причина полягає у відносній простоті розробки таких тестів, які укладаються самим викладачем (інструктором) для своїх слухачів, складаються на невеликому за обсягом матеріалі конкретної теми або дисципліни, не вимагають довготривалих процедур експертизи й узгодження.

Стандартизація тесту потребує копіткої роботи і тривалого часу, оскільки стандартизований тест є таким, який пройшов попереднє випробування на великій кількості учасників тестування і має кількісні показники якості. Стандартизовані тести

супроводжуються паспортом, у якому містяться норми, умови та інструкції для багаторазового використання тесту в різних умовах.

Разом з тим, дослідження результатів численних тестувань, проведених ННЦМП, дають підстави стверджувати, що діагностика рівня професійної підготовки військовослужбовців має проводитись на основі саме стандартизованих тестів за умови надання їм таких властивостей:

– вони ґрунтуються на фіксованому (стандартному) змісті, який, у свою чергу, заснований на затверджених нормативах фахової компетенції;

– дані тести базуються на стандартних процедурах проведення тестування і оцінки результатів, які мають виконуватися за кожного адміністрування (проведення тесту);

– вони ретельно перевіряються в ході емпіричного дослідження, піддаються експертизі й сертифікації і мають надійно визначені статистичні характеристики вимірювальних якостей.

Порівнюючи стандартизовані і не стандартизовані тести, слід взяти до уваги, що попри простоту застосування останніх, вони демонструють недоліки, суттєві у процесі визначення рівня професійної компетенції. Укладення тесту на власний розсуд викладача породжує неоднозначність результатів тестування і порушує принцип об'єктивності, яка є основною перевагою даного методу контролю. Така ситуація зумовлюється, передусім, розбіжністю вимог викладачів різних кафедр і навчальних закладів, відмінністю в їхньому рівні вибагливості в оцінюванні тієї самої відповіді, різницею в професійній кваліфікації. Об'єктивність результатів тестування не стандартизованими тестами може бути порушена і внаслідок суто психологічного фактору, коли закладена в тесті шкала оцінок свідомо завищується викладачем через побоювання, що результати тестування будуть використані для оцінювання його роботи, або занижується ним з інших міркувань.

Ще одним визнаним недоліком застосування не стандартизованих тестів є розбіжності між різними викладачами, кафедрами, ВВНЗ стосовно обсягу, структури і відбору того матеріалу, на основі якого буде зроблено висновок про рівень набутих слухачем професійних компетенцій. Унаслідок такої неузгодженості ключових питань визначення рівня фахової компетентності може виникнути парадоксальна ситуація, коли відмінник з певної дисципліни в одному ВВНЗ покаже незадовільний результат з тієї ж дисципліни в іншому навчальному закладі.

Попри це, за відсутності стандартизації спеціалісти відмічають також недоліки у структурі тестів і тестових завдань, формулюванні стимулів та інструкцій до завдань; підборі варіантів відповідей, інтерпретації результатів, а також відсутність апробації тестових матеріалів, порушення технології і вимог до процедури тестування. [4, 5]

На сьогоднішній день перелік сформульованих спеціалістами правил укладання тестів, у тому числі комп'ютерних, і тестових завдань до них є практично неосяжним, що унеможливлює їх дотримання у повному обсязі. Отже, перед практиком-тестологом постає завдання виокремити з усього різноманіття такі правила і рекомендації, дотримання яких здатне кардинально впливати на показники надійності і валідності тесту.

Результати дослідницької роботи автора стосовно діагностування професійної компетенції військовослужбовців дали підстави стверджувати, що перелік нижче вказаних правил можна вважати найбільш значущим і впливовим у процесі проектуванні валідних і надійних тестових вимірників.

Аналіз результатів апробаційних тестувань для визначення рівня підготовки військовослужбовців з англійської мови, проведених ННЦ МП з використанням розроблених ним електронних засобів діагностики підтверджує висновок про те, що

одним із найсуттєвіших чинників забезпечення валідності тесту, і комп'ютерного в тому числі, є правильність обраного принципу відбору тестового матеріалу. Отже, найважливішими критеріями, які забезпечують змістову валідність тесту, виявилися такі характеристики змісту тестових завдань (ТЗ):

Значимість. За цим критерієм до змісту ТЗ включаються лише ті елементи дисципліни, які можна віднести до найбільш суттєвих, ключових, на основі яких формується професійна компетентність майбутнього спеціаліста. Такі елементи можна назвати структурними. Отже, у ТЗ слід включати тільки ті матеріали, що відіграють роль структурних елементів в індивідуальній фаховій компетентності і є стрижнем загального змісту професійної підготовки. Тільки за такої умови результати тестування будуть адекватно відображати рівень набутих фахових компетенцій, а не випадкове знання другорядних елементів дисципліни.

Репрезентативність. Даний критерій доповнює й уточнює попередній, оскільки необхідність включення у ТЗ лише значимих елементів змісту дисципліни не означає обов'язкового включення в тест усіх таких елементів. Багато які з них пов'язані між собою в загальній структурі знань або включені один в інший повністю або частково. Намагання відобразити у тесті всі без винятку важливі елементи контрольованого матеріалу призводить до надмірної довжини тесту, що знижує його якість за багатьма показниками.

Наукова вірогідність. Цей критерій особливо важливий для тестування з теоретичних дисциплін. У науці часто існують дискусійні точки зору або не апробовані висновки, які може поділяти або не поділяти укладач тесту. Такі не підтверджені гіпотетичні положення не слід включати до ТЗ, оскільки їх знання або незнання мало характеризує рівень фахової підготовки тестованої особи. Отже, у ТЗ слід включати тільки той зміст навчальної дисципліни, що піддається раціональній аргументації і є визнаним офіційною наукою.

Відповідність рівню сучасного стану розвитку теорії і практики дисципліни. Цей критерій випливає з необхідності готувати фахівців і перевіряти їхні знання на сучасному матеріалі і за вимогами останніх досягнень науково-технічної та гуманітарної думки. Стосовно підготовки з мовних дисциплін цей критерій зобов'язує слідкувати за сучасними змінами у розвитку мови, що вивчається, і дозволяє забезпечити діагностику комунікативних компетенцій тестованого, тобто здатність здійснювати комунікацію з реальними, а не гіпотетичними носіями мови.

Варіативність змісту завдань тесту створює можливість адаптації до змін у розвитку науки, науково-технічного прогресу і до змін у змісті навчальної дисципліни або до нових підручників. Цей критерій дозволяє також враховувати рівень підготовки контингенту, який проходить тестування, для забезпечення відповідних параметрів якості тесту.

Системність змісту. Це означає підбір такого змісту ТЗ, котрий відповідає вимогам системності знань. Крім підбору завдань із системним змістом важливо мати завдання, пов'язані між собою загальною структурою знань. Це можливо лише в тих випадках, коли завдання тесту пов'язані із загальною факторною структурою знань.

Комплексність і збалансованість змісту ТЗ. Тести для контролю фахової компетентності повинні відображати основні теми навчального курсу комплексно, з урахуванням взаємозв'язків між окремими темами, а кількість ТЗ з кожної теми проектується у чіткій відповідності до її значущості. Сукупність ТЗ має збалансовано відображати основний теоретичний матеріал: поняття, закони і закономірності, гіпотези, факти, структурні компоненти теорії разом з методами наукової і практичної діяльності, навичками ефективно розв'язувати типові професійні завдання.

Відповідність заданим параметрам. Для отримання валідних результатів тестування зміст ТЗ необхідно ставити у пряму залежність від заданих замовником параметрів і, зокрема, від мети контролю. Як відомо, рівень складності змісту ТЗ обирається відповідно до рівня підготовки контингенту, що підлягає тестуванню: переважання легких ТЗ покаже неадекватно високі результати, а перебільшена кількість складних ТЗ їх безпідставно занизить – в обох випадках тест виявиться невалідним. Однак, якщо замовник тестування визначає за мету відбір невеликої кількості найбільш підготовлених слухачів, слід включити до тесту завдання високого рівня складності, тому що лише за допомогою таких завдань можна відібрати найкращих. Якщо ж, навпаки, потрібно відсіяти найслабших, то зробити це краще за допомогою відносно легких завдань; ті, хто не виконають ці завдання, і є найслабші. За необхідності оцінити слухачів у всьому діапазоні знань – від нижчого до вищого – тест має містити ТЗ мінімального, середнього і вищого рівнів складності. Якщо замовник поставив завдання оцінити відповідність знань випускників декількох освітніх закладів вимогам мінімуму змісту навчальних дисциплін, зафіксованого в державних освітніх стандартах, то у структуру тесту слід включити тільки ТЗ мінімального рівня складності, які більшість спроможна вирішити. Наразі слід бути готовим до вирішення таких питань, як визначення мінімального „прохідного” відсотка правильного виконання ТЗ і визначення параметрів оцінювання складання тесту в цілому залежно від відсотка правильних відповідей [6, с. 45].

Вважаємо, що розповсюджена донедавна, а останнім часом піддавана критиці [11, с. 58], рекомендація включати до тесту більше завдань середньої складності виправдана з погляду визначення надійності виміру за формулами класичної теорії тестів. Існуючі в цій теорії методи оцінювання надійності тесту дають зниження надійності у разі включення до тесту помітно легких і складних завдань.

Розташування ТЗ у тесті в порядку зростаючої складності. Цей критерій означає, що кожен навчальний елемент у процесі контролю викликає певний рівень складності, який має враховувати укладач під час проектування ТЗ. У таких навчальних дисциплінах, як логіка, математика, іноземні мови, статистика, філософія й ін., знання наступних елементів курсу залежить від знання попередніх навчальних елементів. Тому вивчати ці дисципліни можна тільки із самого початку і без пропусків. За цим же принципом побудовані класичні тести.

Стосовно ж комп’ютерних тестів на визначення рівня професійної компетентності даний принцип не завжди обов’язковий. Практика показує, що надійних і ефективних результатів у тестуванні дає принцип випадкового вибору комп’ютером тестових завдань з банку ТЗ. Слід наголосити, що такий принцип генерування тесту, однак, допустимий лише в рамках варіативного вибору з банку ТЗ одного рівня складності.

Найефективніші результати застосування перелічених вище критеріїв відбору тестового матеріалу дає багаторівневий тест як комплексний засіб всебічного контролю оцінювання професійної компетентності за заданими замовником параметрами. Такий тест включає ТЗ з двох або більше рівнів фахової компетенції і дозволяє оперативно і комплексно визначати рівень професійної компетентності. Багаторівневий тест використовує переваги адаптивного методу тестування і передбачає перехід до складання завдань наступного рівня компетентності залежно від позитивного (мінімум 70 % правильних відповідей) результату складання завдань попереднього рівня компетентності.

Валідність ТЗ і тесту в цілому напряму залежить не тільки від змістового наповнення, але й коректності структурних елементів ТЗ, якими є інструкція, власне завдання і відповіді до завдання.

Інструкція до ТЗ визначає перелік дій учасника тестування. Якщо вона сформульована неадекватно формі й змісту завдання або не враховує особливості контингенту, для якого призначений тест, відповіді учасника тестування відображатимуть лише цю неадекватність, а не рівень набутих професійних компетенцій.

Змістова частина завдання має «валідоутворюючі» властивості не тільки з огляду на характер змісту ТЗ, який розглядався вище, але і на форму подання цього змісту. Розповсюджена донедавна заборона класичної теорії тестів на формулювання завдання у формі питань стає все менш суворою. Наш досвід підтверджує недоцільність обмеження форми завдання лише стверджувальним висловлюванням, тим більше, що педагогічний контроль у вигляді «запитання – відповідь» є звичним для вітчизняної системи освіти. Однак підкреслимо, що включення питальних форм у ТЗ допустиме лише за умови можливості дати беззаперечно однозначну відповідь на поставлене питання, що воно сформульоване чітко і коротко. Використання розширених синтаксичних конструкцій ускладнює розуміння змісту ТЗ і невиправдано відволікає увагу від виконання власне завдання. Зрозумілість ТЗ має бути абсолютною, якщо хоча б один з учасників апробаційного тестування його не зрозумів – воно не має включатись до тесту;

Ще однією дієвою рекомендацією щодо оформлення змістовної частини ТЗ вважаємо те, що вона має містити більше слів, ніж кожен з варіантів відповідей у закритих ТЗ або передбачувана відповідь у відкритому ТЗ. Така форма ТЗ змушує тестованого зосередитися саме на змісті ТЗ, а не на дистракторах, відтак пошук правильної відповіді стає більш осмисленим.

Усі варіанти *відповідей на завдання,* як дистрактори, так і правильні відповіді, повинні бути однаково привабливі і за змістом, і за структурою – це основний принцип підбору дистракторів, від коректності якого великою мірою залежить і валідність і надійність тесту. Остаточно правильність підбору дистракторів може бути перевірена лише емпіричним шляхом, за результатами апробаційних тестувань. Однак і на етапі проектування тесту можна підвищити вірогідність ефективної «роботи» дистракторів дотриманням перевірених практикою правил. По-перше, усі варіанти відповідей повинні бути граматично погоджені зі змістовною частиною ТЗ, однорідні за змістом, сформульовані коротко (не більше ніж 5–6 слів), мати ідентичну структуру і обсяг. Попри це, принципово важливо, щоби частота вибору того самого номера місця для правильної відповіді в різних завданнях була приблизно однакова. Якщо ТЗ містить серед інших альтернативні відповіді, їх не слід наводити відразу після правильної, тому що увага учасника тестування зазвичай зосереджується тільки на цих двох відповідях, а решта дистракторів «не працює».

Щодо підвищення валідності і надійності тестових завдань закритої форми (які, часто використовуються для тестування фахової компетентності) досвід тестування підтвердив суттєвий вплив правильного визначення необхідної кількості дистракторів на показники якості тесту. Цілком зрозуміле прагнення укладачів тестів до збільшення числа дистракторів задля зниження вірогідності вгадування призводить до подовження тесту і процедури тестування, що спричиняє включення психологічних факторів в результати тестування (розсіювання уваги, втома тощо). Відтак, незважаючи на відомий середній показник необхідної кількості дистракторів (4–6 на кожне ТЗ), оптимальна кількість варіантів відповідей повинна визначатись у кожному випадку окремо залежно від конкретних умов, параметрів і мети тестування. На етапі проектування комп'ютерного тесту дуже корисно правильно передбачити рівень його ефективності. Ефективним вважається тест, який краще, ніж інші тести, вимірює рівень

знань та вмінь з меншою кількістю завдань, якісніше, швидше, дешевше, і все це, за можливістю, у комплексі [2].

У визначенні ефективності тесту звертається увага на два ключових елементи – це кількість ТЗ у тесті і відповідність їх складності рівню підготовленості тих, хто складатиме цей тест. Показник складності ТЗ і самого тесту є змістовним і формальним одночасно. Змістовним показником тому, що складність великою мірою залежить від відповідності змісту рівневі підготовленості самих учасників тестування, у той час як у недосконалому тесті на результати починає помітно впливати форма завдань (особливо якщо вона неадекватна змісту).

Якщо з будь-якого тесту з великою кількістю завдань зробити оптимальний вибір меншої кількості, то може утворитися система, що не поступається за своїми властивостями тесту з порівняно великою кількістю завдань. Тест із меншою кількістю завдань у такому випадку слід вважати порівняно більш ефективним.

На ефективність тесту впливає і його форма. Порушення тестової форми завжди призводить до неправильного відтворення змісту і до гіршого розуміння завдання учасниками тестування. Це особливо відчутно у випадку комп'ютерної програми тестування, де форма пред'явлення ТЗ має бути максимально сприйнятливою. А тому розробник програмного забезпечення повинен отримати від укладача завдання такої форми і таких параметрів, щоби вони були розміщені на оглядовому полі максимально дохідливо, тобто, щоби форма представлення ТЗ в електронному засобі контролю не ускладнювала їх розв'язання учасником тестування.

Ефективність тесту інколи залежить також і від обраного принципу підбору завдань. Якщо підбирати завдання для виміру на всьому діапазоні зміни складності, то знижується точність вимірювання на окремій ділянці. І навпаки, якщо прагнути точно виміряти знання учасників тестування, наприклад, середнього рівня підготовленості, то це вимагатиме більше завдань саме даного рівня складності. Тому в завданні замовника на проектування комп'ютерного засобу визначення рівня професійної підготовки повинні міститись чітко сформульовані визначення тих рівнів професійної підготовки, які мають бути оцінені в ході тестування. Зауважимо, що заключний висновок про рівень ефективності ТЗ і тесту в цілому робиться лише на підставі статистичної обробки результатів апробаційного тестування.

Під час створення проекту тесту особливої уваги потребує формування ***шкали оцінок*** правильності виконання завдань. Насамперед важливо обґрунтовано визначити обраний тип шкали оцінок: дихотомічна (тільки за фактом правильності чи неправильності відповіді), або політомічна (визначення ступеня правильності відповіді): правильно, повно, точно, неточно, не зовсім правильно тощо. Правильний вибір шкали оцінювання напряму залежить від вимог замовника щодо параметрів тестування та його мети. Якщо метою тестування визначено отримання відомостей про кількість слухачів, які оволоділи тими чи іншими професійними компетенціями, дихотомічна шкала цілком достатня для отримання такої інформації. Однак для тестового контролю рівня оволодіння однією чи декількома професійними компетенціями більш доцільним буде використання політомічної шкали, яка стає просто необхідною у використанні ТЗ відкритого типу або у разі потреби оцінювати спосіб розв'язання проблеми.

Важливим питанням залишається і кількісний показник шкали оцінювання. Суб'єктивізм, присутній під час формування шкали оцінок, цілком зрозумілий, оскільки тут багато залежить від досвіду, інтуїції, компетентності, професіоналізму укладача тесту. Вимоги, що пред'являються до рівня фахової підготовки, коливаються в дуже широких межах і залежать від численних факторів. Практика свідчить, що, наприклад, у п'ятдесятибальній системі можливі різні підходи до визначення ціни

оцінки. Так, в одному випадку, якщо учасник тестування набирає від 31 до 50 балів, то він одержує оцінку «відмінно», від 25 до 30 балів — «добре», від 20 до 24 — «задовільно», менше 20 — «незадовільно». В інших випадках ця ціна вже має іншу вартість: 41–50 балів — «відмінно», 33–40 балів — «добре» і т. д. [12, с. 124].

На складність створення шкали оцінювання суттєво впливають також розбіжності у змісті, обсязі та технічному забезпеченні навчальних предметів, що викладаються в різних ВВНЗ. Фактична відсутність чітких рекомендацій зі складання шкал оцінок щодо конкретної навчальної дисципліни ускладнює цю ситуацію.

У практиці апробаційних, поточних, діагностичних і контрольних тестувань, які проводяться у ННЦ МП, успішно застосовується стобальна шкала оцінювання, рекомендована мовним стандартом НАТО СТАНАГ 6001 і використовувана у багатьох інших країнах.

Для подолання розбіжностей і суб'єктивізму у процесі визначення рівня фахової компетенції шкала оцінювання виконання тесту з кожного конкретного навчального предмету повинна відображати основні вимоги освітньо-професійних стандартів та освітньо-кваліфікаційних характеристик підготовки спеціалістів. А тому вважаємо, що **системі тестування на визначення рівня професійної підготовки необхідно надати чіткого стандартизованого характеру**, вказавши при цьому основні вимоги щодо змісту, обсягу та форми конкретних компетенцій, які мають формувати єдиний стандарт професійної підготовки.

4. ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Тест є творчим документом, отже, завжди має ознаки оригінальності. Тим не менше, тест для визначення професійної компетентності військовослужбовців **повинен стати стандартизованим документом**, що має свої усталені конструкції, обсяги та цілі. Наведений перелік рекомендацій щодо механізму створення тестових завдань і проекту тесту представляється як базовий, перевірений практикою численних апробаційних та контрольних тестувань, здійснених ННЦ МП, дотримання яких дає основу для створення надійного засобу контролю.

Розглянуті вище елементи процесу укладання тестових завдань, проекту тесту і проведення тестування не претендують на вичерпне викладення рецептів вирішення всіх проблем і питань, які постають у практиці проведення тестової діагностики рівня професійної компетенції військовослужбовців. Пошук відповідей на все нові питання, що виникають в процесі розвитку, зокрема, комп'ютерного тестування, повинні стати предметом подальших досліджень.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Загальноєвропейські Рекомендації з мовної освіти: вивчення, викладання, оцінювання / перекл. Шерстюк О. — К. : Ленвіт, 2003. — 261 с.
2. Аванесов В. С. Композиция тестовых заданий / В. С. Аванесов. — М.: Центр тестирования Минобразования РФ, 2002. — 239 с.
3. Андреев А. Б. Компьютерное тестирование: системный подход к оценке качества знаний студентов / А. Б. Андреев. — М. : Педагогика, 2001. — 204 с.
4. Бельдягіна А. О. Автоматизований контроль знань з лінгвістичних дисциплін [Електронний ресурс] / А. О. Бельдягіна. — Режим доступу до документа : www.ev.nuos.edu.ua.
5. Интернет-образование: не миф, а реальность XXI века / [Зайцева Ж. Н., Рубин Ю. Б., Титарев Л. Г. и др.] ; под общ. ред. Тихомирова В. П. — М. : МЭСИ, 2000. — 189 с.
6. Булах І. Є. Педагогічне оцінювання і тестування: правила, стандарти, відповідальність / І. Є. Булах, М. Р. Мруга. — К. : Майстер-клас, 2007. — 272 с.

7. Булах І. Є. Основи педагогічного оцінювання. / І. Є. Булах. — К. : Майстер-клас, 2005. — 96 с.
8. Васильев В. И. Теория и практика формирования программно-дидактических тестов В. И. Васильев, Т. Н. Тягунова. — М. : МЭСИ, 2001. — 130 с.
9. Клайн П. Справочное руководство по конструированию тестов / П. Клайн. ; пер. Е. П. Савченко. — М. : ПАН Лтд, 1994. — 283 с.
10. Мельникова М. Б. Теория и практика конструирования педагогических тестов : учеб. пособие / М. Б. Мельникова, Г. С. Ковалева. — М. : Исслед. центр, 1996. — 47 с.
11. Олійник М. М. Тест як інструмент кількісної діагностики рівня знань в сучасних технологіях навчання / Олійник М. М. — Донецьк : ДонНУ, 2001. — 84 с.
12. Хубаев Г. О. О построении шкалы оценок в системах тестирования / Г. О. Хубаев // Высшее образование в России. — 1996. — № 1. — С. 122–125.

Матеріал надійшов до редакції 19.03.2013 р.

ОБЗОР ДЕЙСТВЕННЫХ МЕТОДОВ СОЗДАНИЯ ТЕСТОВЫХ ЗАДАНИЙ И ПРОЕКТА ТЕСТА, ПРЕДНАЗНАЧЕННЫХ ДЛЯ ОПРЕДЕЛЕНИЯ УРОВНЯ ПРОФЕССИОНАЛЬНОЙ ПОДГОТОВКИ ВОЕННОСЛУЖАЩИХ

Малик Александр Иванович
начальник Научно-учебного центра языковой подготовки
Академия сухопутных войск имени гетмана Петра Сагайдачного, г. Львов, Украина
malik_alex@mail.ru

Аннотация. Актуальность изложенного материала обусловлена тем, что в Украине продолжается процесс внедрения тестирования как основного средства контроля усвоения знаний. Статья посвящена теоретико-практическим аспектам проблемы создания тестовых заданий, предназначенных для определения уровня профессиональной подготовки. Рассмотрены базовые элементы процесса создания тестовых заданий и проекта теста, соблюдение которых дает основу для создания надежного средства контроля. Сформулирован перечень принципиально важных критериев создания тестовых заданий. Предъявлены основные требования к структурным составляющим тестовых заданий с целью получения высококачественных тестов. Указаны основные параметры, которые влияют на эффективность теста в целом.

Ключевые слова: тестирование; критерии отбора тестового материала; шкала оценивания; уровень профессиональной подготовки.

REVIEW OF EFFECTIVE METODS OF TEST TASKS AND TEST PROJECT CREATING FOR DETERMINING LEVELS OF PROFESSIONAL MILITARY TRAINING

Olexandr I. Malik
Head of Educational Scientific Language Training Centre
Army Academy named after hetman Petro Sahaidachyi, Lviv, Ukraine
malik_alex@mail.ru

Abstract. Actuality of the given material is specified by the fact that the process of adopting tests as the main means of knowledge control is going on in Ukraine. The article is dedicated to the theoretical and practical aspects of tests creating designated for determining levels of professional training. The article examines basic elements of the process of tests and test project creating following which gives the framework for creating reliable means of control. The article presents the list of the most important criteria of test tasks creating. The principal requirements to the structural elements of test tasks in order to create high quality tests have been suggested. The article also states the main features that can influence the test effectiveness in whole.

Keywords: testing; criteria for the selection of test material; assessment scale; professional training level.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Common European recommendations on language education: learning, teaching, assessment /Translated by Sherstiuk O. — K.: Lenvit, 2003. — 261 p. (in Ukrainian)
2. Avanesov V. S. The composition of test tasks / V.S. Avanesov. — M.: RF Ministry of Education's Testing Center, 2002. — 239 p. (in Russian)
3. Andreev A. B. Computer testing: a systematic approach for determining the quality of students knowledge / A.B. Andreev. — M.: Pedagogika, 2001. — 204 p. (in Russian)
4. Bel`diagina A. O. Automated knowledge control of linguistic disciplines [online] / A.O. Bel`diagina. — Available from: http//www.ev.nuos.edu.ua. (in Ukrainian)
5. Online education: not a myth but a reality of XXI century / [Zai`tceva Z.N., Rubin I.B., Titarev L.G. and others]; Edited by Tihomirova V. P. — M.: ME`SI, 2000. — 189 p. (in Russian)
6. Bulakh I. E. Educational assessment and testing: rules, standards, responsibility / I. E. Bulakh, M.R. Mruga. — K. : Mai`ster-class, 2007. — 272 p. (in Ukrainian)
7. Bulakh I. E. Fundamentals of educational assessment. / I. E. Bulakh. — K. : Mai`ster-class, 2005. — 96 p. (in Ukrainian)
8. Vasil`ev V. I. Theory and practice of forming the didactic tests / V.I. Vasil`ev, T.N. Tiagunova. — M. : ME`SI, 2001. — 130 p. (in Russian)
9. Clai`n P. Tests constructing reference guide / P. Clai`n.; translated by. E.P.Savchenko. — M.: PAN Ltd, 1994. — 283 p. (in Russian)
10. Mel`nikova M. B. Theory and practice of pedagogical tests constructing: Study guide. / M.B. Mel`nikova, G.S. Kovaleva. — M. : Issled. centr, 1996. — 47 p. (in Russian)
11. Olii`nik M. M. Test as a diagnostic tool for quantifying the level of knowledge in modern educational technologies / Olii`nik M. M. — Donetc`k: DonNU, 2001. — 84 p. (in Ukrainian)
12. Hubaev G. O. About preparing the assessments scale in testing systems / G. O. Hubaev // Vy`sshee obrazovanie v Rossii. — 1996. — № 1. — P. 122–125. (in Russian)