

УДК 378.147.013

ЗАБЕЗПЕЧЕННЯ ВАЛІДНОСТІ ТА НАДІЙНОСТІ МОВНИХ ТЕСТІВ ЗГІДНО ВИМОГ СТАНДАРТІВ НАТО «STANAG 6001»

Ю. В. Грищук

Національний університет оборони України імені Івана Черняховського, Київ, Україна
julia_danyluk@ukr.net

В. Д. Галкіна

Національний університет оборони України імені Івана Черняховського, Київ, Україна
kryvoru4ko@rambler.ru

Стаття розкриває основні питання забезпечення валідності та надійності стандартизованих мовних тестів згідно вимог стандартів НАТО «STANAG 6001». У статті проаналізовано сутність понять тест, тестування та стандартизований мовний тест згідно вимог стандартів НАТО «STANAG 6001». Встановлено, що валідність та надійність є основними критеріями розробки мовних тестів. На основі аналізу наукової літератури надано рекомендації щодо забезпечення валідності та надійності мовних тестів згідно вимог стандартів НАТО «STANAG 6001».

Ключові слова: тест, тестування, стандартизований мовний тест згідно вимог НАТО «STANAG 6001», валідність, надійність, рівень володіння іноземною мовою, стандартизований мовленнєвий рівень, тестований, Збройні Сили України.

PROVIDING VALIDITY AND RELIABILITY OF LANGUAGE TESTS ACCORDING TO NATO «STANAG 6001»

J. Hryshchuk

Ivan Cherniakhovskyi National Defence University of Ukraine, Kyiv, Ukraine
julia_danyluk@ukr.net

V. Halkina

Ivan Cherniakhovskyi National Defence University of Ukraine, Kyiv, Ukraine
kryvoru4ko@rambler.ru

The article highlights the main issues concerning providing validity and reliability of NATO «STANAG 6001» standardized language tests. The article presents the essence of the following concepts such as: test, standardized language test NATO «STANAG 6001». The authors show that the validity and reliability are key criteria for language tests development. According to the analysis of scientific literature, recommendations for providing validity and reliability of NATO «STANAG 6001» language tests are given in the article.

Key words: test, standardized language test NATO «STANAG 6001», validity, reliability, level of foreign language proficiency, candidate, the Armed Forces of Ukraine.

ПОСТАНОВКА ПРОБЛЕМИ

На сьогодні Збройні Сили України (ЗСУ) взяли курс на покращення міжнародного військового співробітництва з країнами – членами ЄС та НАТО. Одним із пріоритетних напрямків співробітництва є забезпечення представництва ЗСУ в багатонаціональних штабах оперативно-стратегічного та оперативно-тактичного рівнів; участь підготовленого військового особового складу в міжнародних операціях з підтримання миру і безпеки; участь у багатонаціональних військових навчаннях за межами України та проведення їх на території України тощо [4; 8].

Ефективність досягнення взаємосумісності ЗСУ зі збройними силами країн-членів ЄС та НАТО значною мірою залежить від рівня володіння іноземною мовою військових фахівців. У ЗСУ

цей рівень оцінюється за допомогою стандартизованого мовного тесту, який розробляється відповідно до певних критеріїв, основними з яких є валідність та надійність [4; 8]. На сьогодні, забезпечення валідності та надійності стандартизованих мовних тестів згідно вимог стандартів НАТО «STANAG 6001» (мовний тест НАТО «STANAG 6001») не достатньо повно висвітлюється в науковій літературі.

Відповідно до вищезазначеного, *метою* статті є теоретичне обґрунтування сутності та значення валідності і надійності під час розробки мовних тестів згідно вимог стандартів НАТО «STANAG 6001». Відповідно до мети, визначено такі завдання: 1. Визначити особливості мовних тестів згідно вимог стандартів НАТО «STANAG 6001». 2. Визначити шляхи забезпечення валідності та надійності під час розробки стандартизованих мовних тестів згідно вимог стандартів НАТО «STANAG 6001».

РЕЗУЛЬТАТИ ТЕОРЕТИЧНОГО АНАЛІЗУ ПРОБЛЕМИ

Проблема розробки мовних тестів стала предметом розгляду багатьох вітчизняних і закордонних учених [1; 6; 9; 12]. З одного боку, науковці вважають, що тест – фіксоване в часі випробування, призначене для встановлення кількісних та якісних індивідуально-психологічних якостей [2; 3]. З іншого боку, тест – це науково обґрунтований метод вимірювання досліджуваних особливостей особистості; метод діагностики, за допомогою якого вибір поведінки, що презентує передумови чи результати навчального процесу, повинен максимально відповідати принципам зіставлення, об'єктивності, надійності та валідності вимірів, повинен пройти опрацювання й інтерпретацію та бути прийнятним для застосування [7; 8]. На думку американської вченої А. Анастасі [10] тести відповідають таким вимогам: однозначності у формулюванні цілі, предмету і галузі застосування; однозначності алгоритму процедури проведення мовного тестування (МТ), зрозумілої його учасникам; наявності обґрунтованої процедури та алгоритму підрахунку й стандартизації тестових балів; наявності тестових шкал, що перевіряють на репрезентативність, надійність і валідність у заданій сфері застосування.

Таким чином, на основі аналізу психолого-педагогічної літератури з порушеного питання, ми зробили висновок, що такі терміни як тест, тестування у вузькому значенні означають використання і проведення тесту, а в широкому – сукупність процедурних етапів планування, розробки і випробування тестів, обробки та інтерпретації результатів проведеного тесту [5].

Мета стандартизованого тесту НАТО «STANAG 6001» полягає у визначенні загального рівня володіння іноземною мовою [8; 13], що є сукупністю мовних знань, мовленнєвих навичок та вмій, що визначають здатність особи до її практичного використання. У ЗСУ цей рівень оцінюється за шкалою стандартизованих мовленнєвих рівнів (далі – СМР), які відповідають рівням мовленнєвої компетенції за мовним стандартом НАТО «STANAG 6001».

Мовні тести НАТО «STANAG 6001» визначають відповідний рівень володіння іноземною мовою тестованого. Характерною особливістю, яка відрізняє даний вид тесту від інших є те, що до уваги не береться попередній досвід вивчення іноземної мови. Мета мовного тесту НАТО «STANAG 6001» полягає в наданні оцінки наскільки добре тестований може використовувати іноземну мову в іншомовному середовищі [8; 13]. Мовний тест НАТО «STANAG 6001» не базується на програмі навчальних курсів, або інших освітніх програмах, оскільки його основна мета полягає не в перевірці знання окремих аспектів мови, наприклад, граматики, лексики тощо, а у визначенні рівня володіння іншомовною компетентністю в цілому [4].

Отже, розглянувши особливості мовних тестів НАТО «STANAG 6001», перейдемо до опису таких основних критеріїв, які висуваються до розробки мовних тестів як валідність та надійність.

Термін «валідність» (англ. valid) означає обґрунтований, дійсний, придатний, той, що має силу [2; 8; 9]. Проблема визначення поняття валідності стала предметом розгляду багатьох вітчизняних та закордонних учених [7; 9–11]. Вони вважають, що валідність — це комплексна характеристика тесту, що відображає інформацію про коло досліджуваних явищ, а також міру репрезентативності процедури дослідження по відношенню до них [1; 11]. У спрощеному та узагальненому формулюванні валідність тесту — це поняття, яке вказує на те, що тест вимірює і наскільки успішно він це вимірює [7; 10]. Поняття валідності використовується для загальної характеристики тесту щодо відповідності одержаних результатів меті та вимогам оцінювання. Якщо тест не валідний, то отримані результати не є коректними, оскільки розроблений тест не відповідає визначеній меті тестування [2].

Учені розрізняють різні типи класифікації валідності залежно від критеріїв, покладених в їх основу. Так, зокрема, у психодіагностиці, основними видами валідності вважають змістову, критеріальну і конструктну валідності, а усі інші – похідними від них [3]. Суттєвими для нашого дослідження стали погляди науковців, які у своїх працях зазначають, що термін “конструктна валідність” використовується для трактування загального та всеохоплюючого поняття валідності, яке в свою чергу містить похідні, тобто інші підвиди валідності (змістову, прогностичну та ін.). Сфера

поширення змістової валідності – це тести навчальних досягнень, критеріально-орієнтовані тести, спрямовані на оцінювання якості засвоєння матеріалу, сформованості навичок та тести для здійснення професійного відбору кадрів [10]. Окрім того, багато науковців окремо відзначають роль очевидної валідності: уявлення про тест, сферу його застосування, результативність та прогностичну цінність, що виникає у досліджуваного або іншої особи, яка не володіє спеціальними знаннями про справжню валідність. Очевидна валідність зазвичай демонструється через назву та інструкцію до тесту і допомагає налагодити співробітництво з досліджуваними об'єктами [10].

На основі аналізу іноземних наукових джерел [9–12], ми вважаємо, що валідність мовних тестів НАТО «STANAG 6001» можна втілити за допомогою дотримання таких рекомендацій:

1. Забезпечення наявності детального опису тесту перед набуттям його чинності з метою розробки єдиного тесту для різних навчальних військових закладів;
2. Надання в буклеті зразків виконання тестових завдань;
3. Забезпечення відповідності тестових завдань меті тесту. Під час розробки мовних тестів слід пам'ятати про те, що тест буде вважатися валідним тільки тоді, коли він вимірюватиме саме ті мовленнєві компетенції та можливості, для яких він власне призначений. Наприклад, під час виконання субтесту з аудіювання не повинні враховуватися граматичні та орфографічні помилки тестованого. У протилежному випадку, даний тест не вважатиметься валідним;
4. Мінімізування ймовірності вгадування правильних відповідей тестових завдань. Дистрактори повинні бути належно підібраними;
5. Тестологи мають бути високоосвіченими і теоретично “підкованими” фахівцями, які враховують увесь спектр параметрів, необхідних для створення валідного тесту.

У психолого-педагогічних розвідках без уваги не залишився такий аспект як забезпечення надійності тестів [3; 10; 12]. Надійність методу вимірювання – це міра стійкості результатів, що впливає на точність з якою можна виміряти ту чи іншу конкретну ознаку [1; 5]. Перевірка надійності методу вимірювання стосується насамперед відновлення результатів при повторних вимірах. Тест вважається надійним, якщо він є внутрішньо узгодженим. Також тест вважається надійним, якщо він дає ті самі показники для кожного тестованого при повторному тестуванні (при умові, що тестований не змінився). Надійність при повторному тестуванні через певний час називають ретестовою надійністю [9–11].

На думку науковців, існує декілька шляхів забезпечення надійності мовних тестів, а саме: повторне тестування, застосування паралельних форм (варіантів) та забезпечення незалежності прийняття рішень екзаменаторів. Проведення повторного тестування кандидатів здійснюється протягом короткого періоду часу. Під час цього періоду тестовані не залучаються у процес навчання. Тест можна вважати надійним, якщо результати обох тестувань є однаковими, або майже однаковими. Якщо результати тестів дуже відрізняються, то тест вважається ненадійним [9].

Використання паралельних форм (варіантів) мовних тестів підвищує їхню надійність, оскільки проблеми щодо забезпечення рівня захищеності тестів зводяться до мінімуму. Ще одним, проте не менш важливим позитивним наслідком використання паралельних форм тестів, є зменшення ймовірності автоматичного вирішення тестових завдань. Це означає, що при постійному вирішуванні одних й тих самих завдань тестований вже починає виконувати їх автоматично, що, у свою чергу, призводить до втрати валідності тесту.

Незалежність прийняття рішень екзаменаторів означає, що екзаменатори оцінюють тестованого незалежно один від одного. Оцінки рівня володіння іноземною мовою тестованого не повинні суттєво відрізнятися, а в ідеальному варіанті повинні бути ідентичними. Якщо оцінка двох екзаменаторів співпадає, тоді це означає, що вони однаково розуміють та застосовують систему оцінювання.

Отже, на основі поглядів зарубіжних спеціалістів, які займаються розробкою мовних тестів НАТО «STANAG 6001» [5; 9–12], ми вважаємо, що надійність мовних тестів НАТО «STANAG 6001» забезпечується за допомогою дотримання таких рекомендацій:

1. Забезпечення наявності достатньої кількості завдань з метою перевірки знань, інтелектуальних і практичних умінь та навичок. На думку науковців, завдання та питання у тестах повинні бути максимально незалежними один від одного. У випадку, коли тестований недостатньо добре зрозумів одне питання, тоді у нього є можливість вірно відповісти на друге, оскільки воно є незалежним від першого питання. Якщо взяти за приклад тест з говоріння, то екзаменатор повинен задавати питання різної тематики з тією ж самою метою, яка зазначалася вище. Екзаменатору слід брати до уваги той факт, що тестований не завжди може орієнтуватися в усіх темах, особливо, якщо мова йде про СМР 1.
2. Надання такої кількості завдань, яка дозволить не тільки вірно оцінити тестованого, але й не виснажить його. Але з іншої сторони, у разі включення незначної кількості завдань, можна помилитися в оцінюванні рівня володіння іноземною мовою тестованого;
3. Забезпечення розташовування менш складних завдань на початку тесту. Така стратегія

підвищить впевненість тестованого у своїх силах та зменшить його переживання;

4. Виключення з тесту альтернативних завдань, оскільки практика надання тестованому можливості вибору тестових завдань негативно впливає на надійність тесту;

5. Уникнення двозначного тлумачення тестових завдань. Тестологи, обговорюючи завдання, повинні максимально з усіх сторін їх розглянути. Але, без сумніву, найкращий спосіб перевірки тестового завдання – це його апробація;

6. Наявність інструкцій щодо процедури виконання тестових завдань, які повинні бути чіткими та лаконічними; бажано у письмовій формі, щоб не упустити важливу інформацію. Екзаменатор не повинен покладатися на те, що деякі елементи процедури проходження тестування будуть самі собою зрозумілі для тестованого. Екзаменатор повинен озвучити усі вимоги, які висуваються до тестованого до початку та під час проходження мовного тестування;

7. Забезпечення високої якості тестових матеріалів: буклетів та аудіозаписів. Буклети повинні бути надрукованими зручним для читання шрифтом, а аудіозапис повинен бути якісним. В іншому випадку, тестовані зіштовхнуться з непередбачуваними труднощами, які негативно вплинуть на результат тестування;

8. Наявність чітко визначеної системи оцінювання;

9. Наявність двох, або більше екзаменаторів, що забезпечить об'єктивне оцінювання тестованого;

10. Забезпечення конфіденційності процедури проходження мовного тестування. Слід надати тестованому особистий номер, щоб не зазначати на бланку відповідей його особисті дані. Така процедура сприятиме конфіденційності, яка у свою чергу підвищить надійність тесту, оскільки екзаменатор не буде знати чію саме роботу він оцінює.

ВИСНОВКИ

Отже, можна зробити висновок, що сутність мовних тестів НАТО «STANAG 6001» полягає у наявності максимально наближених до реальності завданнях та оцінюванні володіння чотирма видами мовленнєвої діяльності, що є необхідними для функціонування в іншомовному середовищі. Особливістю мовного тесту НАТО «STANAG 6001» є те, що він визначає загальний рівень володіння іноземною мовою, не беручи до уваги попередній досвід навчання та не прив'язаний до навчальної програми курсів з іноземної мови.

Проведений нами аналіз літературних джерел за порушеною проблемою дозволяє нам стверджувати, що валідність та надійність є основними критеріями під час розробки мовних тестів НАТО «STANAG 6001». Тестологи повинні прикладати якомога більше зусиль, щоб зробити тест максимально надійним. Якщо тест не буде відповідати основним вимогам надійності, то він не буде вважатися валідним. Тест буде вважатися валідним лише у тому випадку, коли він вимірюватиме саме ті мовленнєві компетенції, для яких він власне призначений. У свою чергу, тест можна буде назвати надійним, якщо він буде внутрішньо узгодженим та надасть ті самі показники для кожного тестованого при повторному тестуванні (при умові, що випробуваний не змінився).

Перспективи подальшого наукового пошуку пов'язуємо з дослідженням та обґрунтуванням інших критеріїв, що висуваються до мовних тестів НАТО «STANAG 6001»: диференційна здатність, практичність та економічність.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Банкевич Л. В. Тестирование лексики иностранных языков / Л. Банкевич. – М. : 1981. – С. 56.
2. Булах І. Є., Мруга М. Р. Створюємо якісний тест / І. Булах, М. Мруга // Посібники для викладачів, вчителів та методистів. – К. : Майстер-клас, 2006. – 160 с.
3. Галян І. М. Психодіагностика / І. Галян. – К. : Академвидав, 2011. – 459 с.
4. Концепція мовної підготовки особового складу Збройних Сил України. Наказ Міністра оборони України від 01.06.2009 р. №267.
5. Клайн П. Справочное руководство по конструированию тестов. Введение в психометрическое проектирование / П. Клайн. – К. : Майстер-клас, 1994. – 238 с.
6. Мильруд Р. П., Матиенко А. В., Максимова І. Р. Зарубежный опыт языкового тестирования и оценки качества обучения иностранным языкам // Иностр. языки в школе. – 2005. – № 7. С. 32–41.
7. Ніколаєва С. Ю. Тестовий контроль лексичних навичок аудіювання (англійська мова) / С. Ніколаєва. – К. : Ленвіт, 1997. – 91 с.
8. Осітров В. В. Умови та дидактичні основи контролю рівня володіння іноземною мовою військовослужбовців з використанням сучасної технології комп'ютерно-адаптованого

- тестування / В. Осітров // Вісник НАОУ. – 2009. – № 2 (10). – С. 56–62.
9. Anderson, J.C., Buck, G. (1993) Standards in testing: a study of the practice of UK examination boards in EF/ESL testing. Language Testing. Cambridge.
 10. Anastasi, A. Urbina, S. (1997) Psychological Testing (7th edition). Upper Saddle River, N.J: Prentice Hall, 95–110.
 11. Buck, G. (2001) The testing of listening comprehension: an introspective study. Cambridge: Cambridge University Press, 35–67.
 12. Hughes Arthur (2003) Testing for Language Teachers. Cambridge, 11–36.
 13. NATO STANDARTIZATION AGREEMENT 6001 (Eition 4). Access Mode : <http://www.stanag6001.com>

REFERENCES

1. Bankevich, L. V. (1981) Testirovanie leksiki inostrannyh jazykov. Moskva.
2. Bulah, I. Je., Mruga, M. R. (2006) Stvorjujemo jakisnyj test. Kyi'v : Majster-klas.
3. Galjan, I. M. (2011) Psyhodiagnostyka. Kyi'v : Akademvydav.
4. Konceptija movnoi' pidgotovky osobovogo skladu Zbrojnyh Syl Ukrai'ny. Nakaz Ministra oborony Ukrai'ny vid 01.06.2009 p. № 267.
5. Klajn, P. (1994) Spravochnoe rukovodstvo po konstruirovaniju testov. Vvedenie v psihometricheskoe proektirovanie. Kiiv.
6. Mil'rud, R. P., Matienko, A. V., Maksimova, I. R. (2005) Zarubezhnyj opyt jazykovogo testirovanija i ocenki kachestva obuchenija inostrannym jazykam. Inostr. jazyki v shkole, 7, 32–41.
7. Nikolajeva, S. Ju. (1997) Testovyj kontrol' leksychnyh navychok audijuvannja (anglijs'ka mova). Kyi'v : Lenvit.
8. Ositrov, V. V. (2009) Umovy ta dydaktychni osnovy kontrolju rivnja volodinnja inozemnoju movoju vijs'kovosluzhbovciv z vykorystannjam suchasnoi' tehnologii' komp'juterno-adaptovanogo testuvannja. Visnyk NАОU, 2 (10), 56–62.
9. Anderson, J. C., Buck, G. (1993) Standards in testing: a study of the practice of UK examination boards in EF/ESL testing. Language Testing. Cambridge.
10. Anastasi, A. Urbina, S. (1997) Psychological Testing (7th edition). Upper Saddle River, N.J: Prentice Hall, 95–110.
11. Buck, G. (2001) The testing of listening comprehension: an introspective study. Cambridge: Cambridge University Press, 35–67.
12. Hughes Arthur (2003) Testing for Language Teachers. Cambridge, 11–36.
13. NATO STANDARTIZATION AGREEMENT 6001 (Eition 4). Access Mode : <http://www.stanag6001.com>

Стаття надійшла до редакції 31.03.2015 р.