

**НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ОСТРОЗЬКА АКАДЕМІЯ»
КАФЕДРА ПСИХОЛОГІЇ ТА ПЕДАГОГІКИ**

У. І. Нікітчук

«ПСИХОЛОГІЧНЕ ТЕСТУВАННЯ»

навчально-методичний посібник

до курсу «Психодіагностика»

для студентів спеціальності «Психологія»

факультету політико-інформаційного менеджменту

ОСТРОГ - 2015

*Рекомендовано до друку рішенням вченої ради
Національного університету «Острозька академія»
(протокол № від грудня 2015 року)*

Рецензенти:

Л. В. Засєкіна, доктор психологічних наук, професор, завідувач кафедри загальної та соціальної психології Східноєвропейського національного університету імені Лесі Українки.

О. В. Матласевич, кандидат психологічних наук, доцент, завідувач кафедри психології та педагогіки Національного університету «Острозька академія».

Нікітчук У. І.

Психологічне тестування: навчально-методичний посібник. – Острог: Видавництво Національного університету «Острозька академія», 2015.

У навчально-методичному посібнику подано інформацію щодо важливих аспектів психодіагностики, що стосуються розробки й адаптації психологічних тестів. Розглянуті питання проливають світло на основні змістові й формальні характеристики тестів, нормативи й рекомендації, що їх варто враховувати при конструюванні.

Навчально-методичний посібник адресовано студентам-психологам, які вивчають курс «Психодіагностика», а також тим, хто прагне зорієнтуватися у базових питаннях розробки й адаптації психологічних тестів і розпочати створення авторського психодіагностичного інструментарію.

ЗМІСТ

ПЕРЕДМОВА.....

ТЕСТ ЯК ОСНОВНИЙ ІНСТРУМЕНТ ПСИХОДІАГНОСТИКИ

Поняття та види тестів у психодіагностиці.....
Комп'ютеризовані та комп'ютерні тести.....
Види особистісних опитувальників.....
Тестування особистості з метою профорієнтації.....
Концепція факторів інтелекту як основа відповідних тестів.....
Діагностика спеціальних здібностей особистості.....
Контрольні запитання.....

ПІДГОТОВКА ДО КОНСТРУЮВАННЯ ТА АДАПТАЦІЇ ПСИХОЛОГІЧНОГО ТЕСТУ

Основні етапи конструювання тесту.....
Адаптація зарубіжних тестів.....
Характеристики досліджуваного, що можуть вплинути на результати
тестування.....
Контрольні запитання.....

ОСНОВНІ ХАРАКТЕРИСТИКИ ЕФЕКТИВНИХ ПСИХОЛОГІЧНИХ ТЕСТІВ

Особливості різних типів шкал вимірювання.....
Поняття надійності тесту та її види.....
Валідність як комплексний показник ефективного тесту.....
Дискримінативність тесту та його окремих частин.....
Стандартизація та норми.....
Контрольні запитання.....

РОЗРОБКА ЗАВДАНЬ ПСИХОЛОГІЧНОГО ТЕСТУ

Розробка теоретичного конструкту та специфікація майбутнього психологічного тесту.....	
Формулювання завдань особистісного опитувальника.....	
Розробка завдань для тестів здібностей.....	
Формулювання завдань для тестів досягнень.....	
Пілотажне дослідження й математичний аналіз завдань тесту.....	
<i>Контрольні запитання.....</i>	

КОРОТКИЙ ТЕРМІНОЛОГІЧНИЙ СЛОВНИК.....

СПИСОК ВИКОРИСТАНОЇ ТА РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ..

ДОДАТКИ.....

Додаток А. План лекційних та практичних занять з курсу «Психодіагностика»	
Додаток Б. Рекомендації щодо структури звітної роботи з конструювання тесту.....	
Додаток В. Зразок звіту про апробацію психодіагностичної розробки студентів На УОА	
Додаток Г. Зразок <u>елементів</u> звіту про апробацію Опитувальника особливостей харчової поведінки у жінок.....	

ПЕРЕДМОВА

Психодіагностика відноситься до тих дисциплін, вивчення яких є обов'язковим для здобуття диплому про вищу психологічну освіту. Проте, у більшості вузів України вона викладається як історія і практикум використання різних видів психодіагностичних методик. За кордоном психодіагностика досі вважається теорією і практикою проведення оцінки особистості чи діагностування психічних порушень засобами клінічної психології. Еквівалентом традиційного для нас розуміння психодіагностики є поняття «психологічне тестування» (psychological testing), вивчення якого передбачає оволодіння основами розробки психодіагностичного інструментію.

Прикра ситуація з українськими психологічними тестами, на якій особливо хочеться наголосити у передмові до цього видання, може бути пов'язана зі специфікою викладання курсу «Психодіагностика» у вузах. Під прикрою ситуацією мається на увазі те, що у всьому світі накопичено більш як столітній досвід розробки психодіагностичних інструментів, є ціла низка розробок, які можна вважати взірцевими, еталонними, досвід створення яких може бути взятий за основу розробки не менш потужних і потрібних методик, а в Україні знайти апробований тест рідною мовою, про який є прозорі відомості щодо його надійності, дискримінативності, валідності (не говорячи вже про нормативні показники) і досі важко.

Тим не менше, інтернет-сайти наповнені цілим морем незграбно перекладених українською неапробованих методик, які знаходяться у легкому доступі і настільки широко використовуються навіть серйозними організаціями, що це просто таки лякає. Окремою болючою темою психодіагностики залишається й етика психодіагностичних досліджень, й відсутність будь якого контролю над використанням і розповсюдженням психологічних тестів, та все ж найбільш прикро за те, наскільки далеко знаходиться пересічний український психолог від розуміння того, що є

ефективним психодіагностичним інструментом, як можна (і як не можна) розробляти авторські методики.

Варто зазначити, що програмою викладання дисципліни «Психодіагностика» у Національному університеті «Острозька академія» передбачено теоретичне й практичне засвоєння студентами основних засад конструювання психологічних тестів. Цей посібник укладено з метою розкриття тих базових питань діагностики, які безпосередньо пов'язані з теорією і практикою розробки методик для психологічного тестування.

Навчально-методичний посібник «Психологічне тестування» адресований, насамперед, студентам-психологам. Сподіваюся, що зацікавлений студент зверне увагу на посилання, що їх містить це видання, до більш фундаментальної літератури і продовжуватиме свої наукові розвідки у цьому напрямку, поглиблюючи розуміння матеріалу, запропонованого на сторінках цього посібника. Також видання може бути корисне тим, хто прагне зорієнтуватися у базових питаннях розробки формалізованих тестів, зробити перші кроки до створення авторського психодіагностичного інструментарію.

Для ілюстрації деякого теоретичного матеріалу посібника використано приклади із діагностичних розробок студентів-психологів НаУОА, що були здійснені протягом останніх двох років. Сподіваємося, що це надихне читача спробувати й себе у ролі відповідального автора!

Хочу зазначити, що під час роботи над посібником виникали все нові й нові ідеї щодо розширення його змісту, тож сподіваюся, що це видання буде тільки першим і в подальшому вдосконалюватиметься. Очікую і наперед вдячна за зауваження та рекомендації до структури і змісту цієї навчально-методичної розробки. Висловлюю також подяку студентам за їх ідеї та прицільні запитання викладачу, - всі вони підштовхували до створення посібника, - а також колегам по кафедрі, які надихали і стимулювали до реалізації цього «проекту».

ТЕСТ ЯК ОСНОВНИЙ ІНСТРУМЕНТ ПСИХОДІАГНОСТИКИ

- ✓ Поняття та види тестів у психодіагностиці.
- ✓ Комп'ютеризовані та комп'ютерні тести.
- ✓ Види особистісних опитувальників.
- ✓ Тестування особистості з метою профорієнтації.
- ✓ Концепція факторів інтелекту як основа відповідних тестів.
- ✓ Діагностика спеціальних здібностей особистості.

Поняття та види тестів у психодіагностиці

За Л. Бурлачуком *тест* (англ. test— проба, випробування, перевірка) - *ансамбль стандартизованих, стимулюючих певну форму активності, часто обмежених у часі виконання завдань, результати яких піддаються кількісній (і якісній) оцінці і дозволяють встановити індивідуально-психологічні особливості людини.*

Згідно цього визначення стає очевидним, що тестом можна назвати лише той діагностичний інструмент, який відповідає ряду цілком визначених параметрів: узгодженість завдань, відповідність їх змісту вимірюваному конструкту, адекватна завданням вимірювання форма та спосіб пред'явлення завдань, здатність диференціювати досліджуваних, - все це складає психометричні основи психодіагностики. Поетапна перевірка відповідності психодіагностичної методики згаданим параметрам дає змогу оцінити її якість та рекомендувати її для використання на практиці.

Питання про види психологічних тестів широко висвітлене у доступній літературі з психодіагностики. Воно, як і багато інших, відноситься до розроблених питань, що не потребують перегляду. Втім, коли перед дослідником постає завдання розробки або адаптації методики, він з самого початку повинен визначитися з тим, який вид прагне отримати

на виході, адже від цього залежатиме і форма постановки завдань, і процедура діагностичного обстеження, і спосіб визначення змісту завдань, і багато іншого. Отже, вважаємо за доцільне надати деякі узагальнені відомості про види тестів у психодіагностиці.

Існують різні критерії для класифікації психологічних тестів, зокрема:

1) особливості тестових завдань (вони можуть бути вербальними та невербальними, тож і тести у цьому сенсі поділяють на *вербальні* та *практичні*);

2) спосіб проведення тестування (*індивідуальні, групові, комп'ютерні* тести) [6];

3) спрямованість тесту (залежно від предмету вимірювання прийнято виокремлювати *тести здібностей, особистісні тести, тести окремих психічних функцій*);

4) принцип, що лежить в основі конструювання тесту (за цим критерієм тести поділяються на *факторно-аналітичні* та *критеріально-орієнтовані (емпіричні)*; саме на цьому критерії ми зупинимося детальніше на сторінках посібника, тож детальніший опис буде подано нижче);

5) наявність часових обмежень: (тести, для яких часові обмеження відсутні або виконують допоміжну функцію, називаються *тестами результативності*; за наявності часових обмежень мова йде про *тести швидкості*) [9];

6) кількість вимірюваних властивостей (тест, що вимірює лише одну змінну називається *простим*, за наявності двох і більше змінних (частіше за все незалежних) мова йде про *комплексний тест*)

7) психічна активність досліджуваних (залежно від того, яку психологічну активність передбачають завдання, тести можуть бути інтроспективні, екстроспективні, проєктивні, виконавчі) [35];

8) якість тесту (за наявності уніфікованої й регламентованої процедури дослідження мова йде про *стандартизований тест*, якщо ж

використання тесту здійснюється без строгих правил та інструкцій, цей інструмент є *нестандартизованим*, що, звичайно, ставить під питання його психометричні основи);

9) матеріал, яким оперує досліджуваний (тести можуть бути бланкові, апаратурні чи предметні) [14];

10) методичний принцип, що його покладено в основу розробки психодіагностичного інструменту (так, якщо виконання завдань оцінюється як правильне чи неправильне, мова йде про *об'єктивні тести*, а якщо неправильних відповідей немає, то це – *стандартизовані самозвіти*: тести опитувальники, відкриті опитувальники, шкальні техніки, індивідуально-орієнтовані ідеографічні техніки, проєктивні, діалогічні техніки) [30].

Допитливий студент може задатися запитаннями:

- ÷ як розмежувати часто вживані у літературі як синонімічні терміни «тест», «опитувальник», «методика»?
- ÷ як правильно назвати те, що я конструюю?

Відповіді на ці питання пропонує Л. Бурлачук, пояснюючи, що суміжне використання термінів «тест» і «методика» було зумовлене історично, хоча з точки зору психодіагностики термін «методика» коректніше вживати тоді, коли мова йде про нестандартизовані діагностичні інструменти або такі, що не вимірюють, а оцінюють особистість (наприклад, *проєктивні методики*). *Опитувальниками* традиційно називають такі психодіагностичні інструменти, які, на відміну від інших тестів, спрямовані на суб'єктивну оцінку досліджуваним самого себе або інших людей. Більшість особистісних тестів мають саме такі завдання, тому і в Україні, і за кордоном називаються опитувальниками ('inventory', 'questionnaire').

Комп'ютеризовані та комп'ютерні тести

Сучасна психодіагностика відрізняється від усталених традицій тим, що **роль експериментатора все частіше виконує комп'ютер**. Цей факт потребує уваги, адже саме з ним пов'язані найбільш значні перспективи і найактуальніші завдання галузі. Якщо раніше процес дослідження надійності, валідності тесту, а надто збору нормативних показників був тривалим і дорогим (у зв'язку з чим кількість нестандартизованих тестів і опитувальників, адаптація яких не завершена, просто вражає), то сьогодні існують всі передумови для того, щоб скорочувати ці витрати. З іншого боку, перед розробниками психодіагностичного інструментарію стоять завдання, пов'язані з володінням відповідним програмним забезпеченням, що передбачає як розробку комп'ютерного чи комп'ютеризованого тесту, так і математичне опрацювання значного масиву даних.

До переваг сучасних комп'ютеризованих та комп'ютерних методик психодіагностики відносяться можливості:

- забезпечувати постійність умов тестування;
- точно реєструвати реакції досліджуваного;
- відновлювати і відслідкувати послідовність дій досліджуваного;
- формувати банки психодіагностичних даних;
- встановлювати тестові норми для різних груп досліджуваних;
- полегшити найбільш рутинні й трудомісткі аспекти діагностичної роботи психолога;
- забезпечувати конфіденційності результатів тестування;
- знижувати собівартість психологічних досліджень;
- мінімізувати негативний впливу, що він часто-густо виникає у ситуації міжособистісної взаємодії між експериментатором і досліджуваним;
- актуалізувати «ігрову» мотивацію досліджуваного, коли процес тестування стає привабливішим (а це, у свою чергу, підвищує достовірність результатів);

- забезпечувати швидкий зворотній зв'язок досліджуваному щодо результатів тестування;
- розробляти динамічні стимули для діагностики.

Варто зазначити, що використання комп'ютерних технологій у тестуванні разом із дотриманням норм законодавства щодо авторських прав здатне ефективно вирішувати також і проблему безконтрольного поширення психологічних тестів, яка сьогодні гостро постала перед дослідниками. Про це деякі автори неодноразово наголошують на сторінках своїх праць з діагностики [5, 6, 9] і це питання дійсно потребує сьогодні пильної уваги дослідників.

Звичайно, існують тести, які не можуть бути комп'ютеризовані. Крім того, «осучаснення» й переведення в цифровий варіант вже існуючої методики обов'язково передбачає її подальшу емпіричну перевірку: обґрунтування надійності комп'ютеризованого варіанту методики, ревалідацію, рестандартизацію на відповідних вибірках.

Серед «слабких місць» комп'ютеризованих методик вказують, здебільшого, на недоліки інтерпретації тестових результатів, отримані шляхом виключно комп'ютерних алгоритмів, а також на те, що у певній кількості досліджуваних (у 1993 році їх було близько 30%[9]) спостерігаються прояви комп'ютерної тривожності. Варто зазначити, що за кордоном серед висококваліфікованих професіоналів, орієнтованих на індивідуальну діагностику, усе більш виразно виявляється тенденція неприйняття будь-яких автоматичних алгоритмів у інтерпретації тестових результатів.

Види особистісних опитувальників

Особистісні опитувальники є переліком різної кількості запитань чи тверджень, на які досліджуваний відповідає у відповідності із інструкцією і типом завдання. У психодіагностиці існує досвід створення абсолютно

різних за довжиною опитувальників (від 20 до 550), що підтверджує тезу про те, що кількість завдань може значно варіювати. Не варто довіряти опитувальникам, у яких запитань менше, ніж 20, адже їх спроможність забезпечити надійне вимірювання прямує до нуля. На сьогодні у психодіагностиці накопичено великий досвід конструювання особистісних опитувальників. Їх класифікують за діагностичною спрямованістю так:

- *опитувальники рис особистості* (наприклад, багатовимірні опитувальники «Шістнадцять особистісних факторів» Р. Кеттелла, опитувальник Е. Шострома, Нео РІ, та одновимірні, такі як, наприклад, шкала тривожності Дж. Тейлора та ін.);
- *опитувальники типологічні* (наприклад, «ЕРІ» Г.Айзенка, ММРІ, опитувальник Леонгарда-Шмішека, індикатор типів Маєрс-Бріггс);
- *опитувальники мотивів* (наприклад, «Перелік особистісних переваг» А. Едвардса, опитувальники А. Мехрабіана, Ю.Орлова, Шкала соціального схвалення Марлоу-Крауна);
- *опитувальники інтересів* (ця група опитувальників детально розглядатиметься у наступному пункті);
- *опитувальники цінностей* (наприклад, опитувальники Д. Супера, М. Рокіча);
- *опитувальники установок* (наприклад, шкала атитюдів Л. Терстоуна).

За сферою застосування можна виокремлювати клінічні, профорієнтаційні опитувальники, опитувальники, що розроблені для навчальних цілей, для сфери професійного добору та роботи з персоналом тощо.

Актуальність дослідження різних аспектів професійної спрямованості особистості ніколи не вичерпується з часом, адже це та соціально зумовлена властивість особистості, що значною мірою визначає якість її

життя і життя суспільства в цілому. У наступному пункті посібника ми вважаємо за доцільне більш детально зупинитися на опитувальниках, що їх використовують для профорієнтації, оскільки ці інструменти відносяться до дуже вживаних, важливих, затребуваних, але, тим не менше, надто слабо розроблених для українського досліджуваного.

Тестування особистості з метою профорієнтації

Опитування, проведене нами у лютому 2015 року на вибірці із 75 респондентів, які здобули вищу освіту, (середній вік – 39,6 років, стандартне відхилення – 10,7 р.), показало, що 57% опитаних (43 досліджувані) не працюють за спеціальністю (рис. 1), з них 34,9% досліджуваних (15 чоловік) опановували новий фах у вищому навчальному закладі.

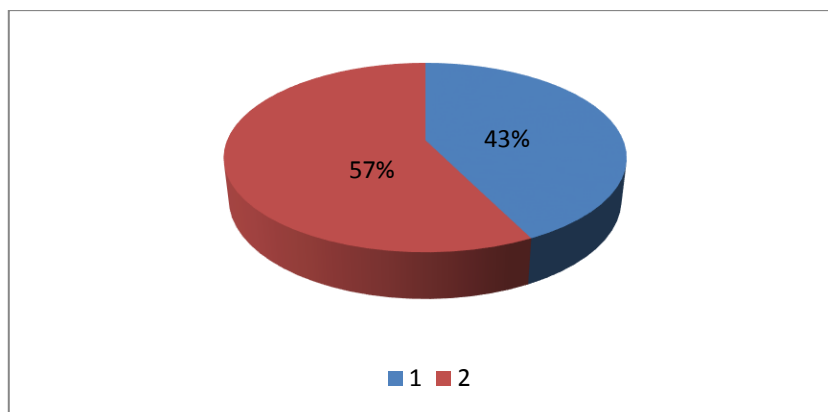


Рис. 1. Розподіл досліджуваних із вищою освітою залежно від того, чи працюють вони за фахом

Умовні позначення; 1 – досліджувані, що працюють за здобутою спеціальністю; 2 – досліджувані, що працюють не за фахом.

Найчисельніші групи серед тих, хто працює за здобутою першою спеціальністю стали група юристів (11 респ.), медиків (6 респ.) та учителів (6 респ.). До спеціальностей, що у досліджуваній вибірці виявилися найменш реалізованими, оскільки, здобувши їх, досліджувані або змінювали фах (здобували нову освіту і за нею працюють) або просто

працюють у іншому напрямку, потрапили психологи (в тому числі практичні психологи) (7 респ.), філологи (6 респ.) та представники технічних й економічних спеціальностей (по 5 респ. у кожній групі). Звичайно, трактувати ці дані варто обережно, адже і вибірка не надто чисельна для соціально-демографічних досліджень, і, наприклад, психологів за першим фахом у вибірці виявилось найбільше – 9 респондентів. До того ж, можна робити багато припущень щодо обумовленості таких результатів фактором неправильного вибору професії чи іншими (соціальними, історичними, економічними) факторами. Викладачам вузів знайома ситуація, коли студент часто не першого року навчання, важко здобуваючи фах, стверджує, що обрав не ту професію і працювати у цій галузі не хоче. Всі ці дані свідчать про те, що професійна орієнтація населення у період вибору професії потребує особливої уваги дослідників-психологів.

Звичайно, професійна орієнтація населення повинна базуватися на наукових дослідженнях, спиратися на положення міжнародних актів і практичний досвід розвинутих країн світу, тож ми порівняємо концептуальні засади доступних в Україні методик із визнаними закордонними розробками того ж спрямування.

Період вибору професії як етап розвитку професійної спрямованості особистості у моделі професійного розвитку Дональда Сьюпера входить до стадії пробудження, що характеризується розвитком «Я-концепції» через *фази інтересів* (11-12 років) та *здібностей* (13-14 років). За класифікацією стадій розвитку суб'єкта професійної діяльності й професійного становлення В. Бодрова, Є. Клімова, Т. Кудрявцева період вибору професії припадає на стадію оптимізації – підготовки до життя і праці, професії – і триває з 11-12 до 14-18 років. На цій стадії відбувається оволодіння системою соціально значущих ціннісних уявлень про життя і професійний шлях, засвоєння систем належних стосунків з товаришами і дорослими, активний самоаналіз, відносини у світі дорослих, формування

інформаційних основ моральної, соціальної і професійної спрямованості, самооцінки власної професійної придатності на основі аналізу своїх можливостей і ступеня розвитку професійно важливих якостей, активні спроби самовдосконалення, самовиховання, самоосвіти, самоорганізації.

Серед психологічних методик, що їх найчастіше використовують і рекомендують до використання в профорієнтаційному консультуванні Україні відносяться:

- ДДО (диференційно-діагностичний опитувальник Клімова);
- Опитувальник професійної готовності Кабардова (ОПГ);
- "Карта інтересів" Голомштока;
- «Структура інтересів» Хеннінга;
- Адаптована Анкета професійної спрямованості особистості Холланда;
- Конструктивний малюнок із геометричних фігур, 16PF Кетела, Big 5, Каліфорнійський особистісний опитувальник (CPI), MBTI (часто його модифікований варіант - Опитувальник Кейрсі), і навіть MMPI – відомі особистісні тести;
- Тест структури інтелекту Амтхауера, кубики Кооса;
- За вибором психолога в якості допоміжних засобів діагностики використовують також різноманітні завдання на кмітливість, методики визначення структури темпераменту;
- «Магелано університет».

Ясно видно, що профорієнтація особистості ґрунтується загалом на вивченні її інтересів, а також здібностей, особистісних рис і типів особистості. Для порівняння, закордонні видання зазначають, що професійне консультування опирається на теорії вибору професії, і серед психологічних теорій вирізняють: факторні теорії рис, психодинамічні теорії, теорії розвитку, теорії прийняття рішень. Як бачимо з приведеного переліку методик, *теорія прийняття рішень* у професійній орієнтації в

Україні не застосовується взагалі, психодинамічні теорії теж практично не представлені (як правило вони знаходять відображення у проєктивних методиках).

ДДО (диференційно-діагностичний опитувальник Клімова) та Опитувальник професійної готовності Кабардова (ОПГ) є порівнюваними методиками, адже результатом їх застосування є визначення прийнятних для людини сфер професіоналізації: людина-природа, людина-техніка, людина-людина, людина – знакова техніка (знаковий образ), людина – художній образ. Відрізняються опитувальники типом та кількістю запитань, а також процедурою. Варто зауважити, що з усіх варіантів і модифікацій найдовший опитувальник налічує 50 запитань, процедура роботи з ОПГ у порівнянні з ДДО є загалом більш складною і трудомісткою.

При ретельному аналізі сучасних публікацій і довідників з психодіагностики не було знайдено жодних відомостей про адаптацію методик українською, хоча інтернет пропонує багато варіантів українською, що представляють собою не більше ніж дослівний (і часто вельми невдалий) переклад оригіналу. До того ж, якщо теоретична обґрунтованість оригінальних варіантів опитувальників та форма запитань розроблені на високому рівні, то текст опитувальників нагально потребує перегляду, адже запитання з формулюваннями (приводжу з тексту) про *«конверти грампластинок»*, *«вирощування молодняка»*, *«тренування товаришів»*, *«спостереження за життям мікробів»*, *«роботу на клавішних машинах: телетайпі, друкарській машинці, телеграфі, ЕОМ»*, а також описи транспортних і вантажних засобів, способів здійснення креслень та ін. зараз досить застарілі, тим більше для оптанта 12-17 років.

За змістом опитувальники ДДО та ОПГ спрямовані на вимірювання сфери переважаючих *інтересів* особистості. До того ж типу особистісних опитувальників (опитувальників інтересів) відносяться і "Карта інтересів" Голомштока та «Структура інтересів» В. Хеннінга, що вони діагностують

не широку сферу професій, а вузькі спеціальності, які згідно запитань опитувальника виявляються більш чи менш привабливими для респондента. За формою побудови представляють собою своєрідний інвентар інтересів, згрупованих за рубриками, що включають шкільні предмети, а також творчі, трудові рубрики. Знову ж таки, валідизовані версії опитувальників доступні лише російською і знову ж таки проблема старіння формулювань не оминула і ці опитувальники.

Вважається, що, оскільки відповідати на запитання цих двох опитувальників респондент може лише на основі власного позитивного чи негативного досвіду знайомства з пропонованими ситуаціями, заняттями, видами діяльності, то на основі такого опитувальника можна визначити обґрунтованість чи необґрунтованість професійних переваг досліджуваного. Таким чином, в основу опитувальника вже закладено значну обмеженість вибору опитанта наявним досвідом.

Адаптована російською Анкета професійної спрямованості особистості Холланда відрізняється від інших опитувальників інтересів тим, що у ній подаються не твердження про заняття чи приклади виконуваної діяльності, а найменування професій у запитаннях з заданим вибором, і респондент, кожного разу обираючи найбільш і найменш привабливу з поданого ряду професію поступово окреслює найменш популярні варіанти і бачить таку собі «висічену» картинку професійних переваг. Даних стосовно результатів валідизації цієї методики (особливо цікавила прогностична валідність) не знайдено, але з логіки побудови опитувальника зрозуміло, що він ймовірно виявлятиметься тим більш валідним, чим краще проінформований досліджуваний з приводу змісту різних професій, соціальної ролі її представників і таке інше.

Таким чином, зазначені опитувальники інтересів об'єднує те, що вони засновані на самооцінці досліджуваним своїх вподобань і концептуально розроблені на певній класифікації інтересів чи сфер інтересів за напрямком (математика, біологія, людина-художній образ

тощо). Типи запитань (частіше всього, як бачимо, завдання з вимушеним вибором) дозволяють калькулювати переважання одного напрямку інтересу над іншим. Але такі ознаки інтересів особистості, як широта, глибина, стійкість (особливо стійкість) не враховуються.

Дійсно, і у нас, і за кордоном у професійному консультуванні найчастіше використовуються методики діагностики саме професійних інтересів. За межами України найбільш популярним і ретельно валідизованим вважається **Інвентар інтересів Стронга-Кемпбела (SCI)**, **огляд інтересів Кьюдера (KOIS)** і **самоаналіз професійних схильностей (SDS)**. Дані про адаптацію цих опитувальників російською чи українською відсутні.

Результати досліджень свідчать про те, що приблизно 50% із тих, хто використовує опитувальники інтересів, через 10-15 років після проходження тестів займаються тим видами професійної діяльності, які прогнозувалися опитувальником. В цілому, однак, ***характер зв'язку між інтересами, здібностями і виборами професії залишається невстановленим*** [20].

Завдяки значному інтересу дослідників за кордоном до розвитку професійної кар'єри та профконсультування систематичне вимірювання професійних інтересів, загальна оцінка інтересів – все це набуло важливого значення у цій сфері. У число актуальних питань входить адекватність психометричних підходів до вимірювання професійних інтересів, проблеми, пов'язані з сексизмом та/чи расизмом в процедурах вимірювання професійних інтересів, зокрема, як ці упередження відображаються на мові інструментів, які припущення лежать в основі при розробці інструментів, що собою представляють норми, на яких ґрунтуються показники. Так, в інструментах Кьюдера і Стронга-Кемпбела сьогодні існують одностатеві і різностатеві підходи до визначення норм, а сексизм видалено із мови описів.

Отже, в основу розробки покладено різні принципи – теоретичний та емпіричний, про це детальніше йтиметься далі. Ми бачимо, як відрізняються представлені в нас опитувальники від іноземних кількістю запитань, а отже – надійністю і ґрунтовністю аналізу. Тож складається враження, що згадані російськомовні опитувальники інтересів можуть краще працювати на профвідбір, ніж на професійну орієнтацію.

При цьому визнається, що інтереси людини змінюються з часом.

Одна з причин, по якій інтереси недостатньо ефективно прогнозують вибір професії і задоволеність професією у дівчат і юнаків, полягає в тому, що на визначення професійного вибору і досягнення впливають такі змінні, як спеціальні здібності, результати діяльності і можливості сім'ї.

Було виявлено, що у професійному консультуванні можуть стати корисними стандартизовані процедури вимірювання ***здібностей***. В цілому, вивчені позитивні кореляції між здібністю та успішністю в оволодінні професією (0,39), здібностями і виконанням роботи (0,22). Множинні тести чи комбінації тестів (батареї) показують кращу прогностичну валідність у порівнянні з одиничними тестами. Знову ж таки, в контексті неспростовності теорії про наявність генерального фактору інтелекту, здається логічним, що вимірювання спеціальних здібностей або структури здібностей (за Атамхауером) корисно застосовувати у поєднанні з опитувальниками інтересів, адже без прикладного тлумачення результатів вимірювання здібностей останні носитимуть не більш ніж ознайомчий характер.

Використання особистісних опитувальників у профорієнтаційному консультуванні (опитувальників рис чи типів) хоч і є популярним у нас, та потребує наявності чітких вимог до особистості професіонала у тій чи іншій галузі. Звичайно, у багатьох сферах професійної діяльності на основі емпіричних досліджень встановлено, яка саме особистість (її тип чи набір рис) досягає найбільших успіхів у професії, та все ж мені не зустрічалися дослідження, що вказують на «особистісно зумовлені протипоказання» до

того чи іншого виду діяльності, а значить фактор мотивації має потенційно більше значення для здійснення прогностично ефективної профорієнтації, ніж фактор особистості, який може набувати свого критичного значення при профвідборі.

В контексті всього сказаного вище, варто розглянути новітній опитувальник *«Магелано університет»*¹, розроблений спільно з італійським університетом в Падуї (адаптований українською у 2012 р. на вибірці у 500 старшокласників київських шкіл організацією з розробки психологічних тестів ТОВ «ОС Україна»), є батареєю із 5-ти профорієнтаційних тестів. Опитувальник спрямований на аналіз професійних та культурних інтересів, здібностей, учбової мотивації, вибору учбових стратегій і методів, здібностей до вирішення проблем. Бачимо, що поруч з уже «традиційними» змінними інтересів і здібностей, присутні й «нові».

Варто зауважити, що у перелік професійних інтересів розширений, порівняно «Картою інтересів» та «Структурою інтересів» та приведений у відповідність до актуальних сьогоднішніх сфер зайнятості (виокремлено екологічний, освітній і педагогічний, гуманістично-літературний, природнонауковий, лінгвістично-туристичний, техніко-математичний, науково-технічний, художній, психологічний, економіко-торгівельний, економіко-управлінський, юридично-соціальний інтереси та інтерес до охорони здоров'я та соціального забезпечення). Культурні інтереси представлені окремим субтестом.

Модифіковано і типи інтересів (порівняно з ДДО і ОПГ) і виокремлюються: види діяльності, пов'язані із спілкуванням, механічні види діяльності, дослідницька діяльність, діяльність в природних умовах.

У діагностиці здібностей враховується традиційний набір факторів вимірювання інтелекту (загальний фактор, гнучкість мислення, абстрактне

¹ <http://www.osukraine.com>

мислення, вербальні здібності, розуміння тексту, математичні здібності, просторові здібності, технічні здібності).

Новим запровадженням методики є також профіль мотивації і методів навчання (розподіл часу, здатність контролювати тривогу та страх в ситуації іспиту, мотивація до навчання, аналіз і відбір інформації, стратегія навчання, самоконтроль ефективності навчання, вміння конспектувати, виокремлення головних ідей, відношення до навчання, концентрація). Дійсно, такі вміння можуть виявитись критично важливими для певних професій, особливо тих, що пов'язані із переробкою значної кількості інформації і прийняттям рішень.

Здатність вирішення проблем вимірюється через показники самовизначення, здатності передбачати наслідки, здатності приймати критику та зауваження, відношення до проблем, здатність взаємодіяти з іншими.

Автори пропонують таку процедуру дослідження: оптант обирає із переліку можливих професій ті, які його найбільше цікавлять, далі відповідає на запитання комп'ютеризованої методки, відповіді аналізуються і надається оцінка й інтерпретація результатів по кожному з тестів батареї і на основі цього здійснюється аналіз того, який напрямок навчання запропонувати учневі, підтримати його вибір чи порекомендувати інший напрямок подальшого навчання чи роботи над собою.

Варто зауважити, що особистісні характеристики (риси, тип) не вимірюються у батареї методик, але при обговоренні результатів тестування наводиться перелік рис, необхідних для ефективного провадження діяльності, і респондент має сам для себе оцінити, наскільки його особистість відповідає вказаним критеріям. Іншими словами, виходить, що для вдалого рішення на основі методики респондент повинен добре знати себе, адекватно оцінювати свою особистість.

Бачимо, що як за кордоном, так і в Україні, більшість профорієнтаційних методик спрямовані на вимірювання *здібностей та*

інтересів особистості, що відповідає опису відповідного періоду Кьюдером і є загальнопоширеною практикою.

В той час, як у нас і у Європі з'являється нова концепція профорієнтаційного дослідження, що полягає у співставленні вимог привабливої спеціальності із здібностями, інтересами, розвитком певних ділових навичок, рефлексії оптанта, його здатності до навчання і т.д., за кордоном все більшу увагу привертає відносно нова область оцінки - ***визначення стану здатності прийняти рішення***. Три опитувальники, розроблені Самюелем Осиповим, Джоном Холландом і Вінсентом Херреном (шкала прийняття рішень про кар'єру Career Decision Scale, Моя професійна ситуація My vocational Situation, Оцінка прийняття рішень про кар'єру Carrer Decision-Making), вимірюють ***стан здатності особистості прийняти професійне рішення***. Всі три перелічені шкали безпосередньо пов'язані з процедурами втручання професійного консультування. Відомості про адаптацію цих шкал українською чи російською мовами відсутні.

Тож, проаналізувавши концептуальні засади, на яких ґрунтуються доступні в Україні інструменти вивчення професійної спрямованості особистості на стадії вибору професії і порівнявши їх із найбільш поширеними за кордоном методиками, бачимо, що розробка ефективних новітніх методик для здійснення професійної орієнтації особистості є актуальним завданням, нагальною необхідністю. Рішення про те, в якому напрямку – вивчення особистісних мотиваційних і ділових компетенцій чи стану здатності прийняти рішення – розгортати пошуки основи для створення нових методик, яка форма і зміст запитань будуть найбільш корисними для вимірювання і т.д. повинно базуватися на результатах спеціальних досліджень, в чому вбачаємо перспективу подальших наукових розвідок у цьому напрямку.

Концепція факторів інтелекту як основа відповідних тестів

Загальна ідеологія факторного підходу до пояснення інтелекту зводиться до наступних позицій [9]:

- інтелект, як і інша психічна реальність, є латентним, тобто даним дослідникові через множину проявів у поведінці при вирішенні життєвих задач;

- інтелект є латентною властивістю деякої психічної структури («функціональної системи»), яка може бути виміряна на основі його проявів у поведінці. Тобто інтелект є лінійною властивістю (одномірною чи багатомірною);

- множина поведінкових проявів інтелекту (комбінацій «поведінковий акт-задача») завжди більша, ніж множина властивостей. Тобто можна вигадати багато інтелектуальних задач для виявлення лише однієї властивості;

- інтелектуальні задачі об'єктивно розрізняються за рівнем складності;

- рішення задачі може бути правильним або хибним (або може як завгодно наближатися до правильного);

- будь-яку задачу можна вирішити правильно за нескінченний час, тож точність і час вирішення є еквівалентно важливими.

Насідком цих положень є принцип квазівимірювальної процедури: чим складніша задача, тим більш високий рівень розвитку інтелекту потрібен для її правильного вирішення. Отже, людина, що вона має «ідеальний інтелект», може правильно і самостійно вирішити подумки задачу (або множину задач) довільно високої складності за нескінченно малий час і незважаючи на внутрішні чи зовнішні перешкоди.

Уявлення Ч. Спірмена, автора теорії генерального фактору інтелекту в їх класичному вигляді багато в чому застаріли, але, тим не менше, сьогодні успішно розвиваються за кордоном. Ми не ставимо за мету

охарактеризувати на сторінках цього посібника відомі моделі інтелекту. Про концепцію Спірмена згадуємо для того, щоб підкреслити, що інтелект не розпадається на незалежні здібності, а має загальну основу та внутрішню єдність. Властивості інтелекту представляють собою сторони єдиної та нерозривної пізнавальної діяльності, органічно включеної в психічну активність людини.

Отже, жодна із існуючих сьогодні моделей інтелекту не заперечує концепції генерального фактору Спірмена. Будь які спроби роздробити інтелект на множину окремих здібностей зазнали поразки, адже вперто проявляються стійкі кореляції між тестами, що вказують на єдину основу вимірюваного ними явища.

Конструюючи тести інтелекту, варто пам'ятати також і про деякі закономірності його розвитку, встановлені відомими у психодіагностиці дослідженнями.

Так, наприклад, виявлено, що існують статистично значимі відмінності у IQ афроамериканців, світлошкірих, латиноамериканців та американців азійського походження в межах 15 балів.

Генетики визначають роль спадкових чинників, проводячи сімейні дослідження, в яких порівнюють осіб, які знаходяться в різних ступенях генетичних зв'язків. За даними міжнародного дослідження 1997 року щодо кореляції IQ у родичів, виявлено, що найбільш сильний зв'язок між рівнями IQ властивий однояйцевим близнюкам, які росли разом та окремо, різнояйцевим близнюкам які росли разом та окремо, та сиблінгам, які росли разом та батьками і дітьми, які жили разом.

Щодо впливу біологічного середовища, то результати досліджень свідчать, що особливості перебігу вагітності, пологів та неонатальні особливості дитини значимо не пов'язані з її IQ (без урахування випадків патологій плоду). Частково це пояснюється увагою та впливом лікарів, батьків та педагогів до дітей, які мали певні особливості вищезгаданих факторів.

Відмінності в IQ чоловіків і жінок статистично не значимі у дорослому віці, але до 7-ми річного віку дівчатка у інтелектуальному розвитку випереджають хлопчиків. Відомо, що у жінок краще розвинений вербальний інтелект, а у чоловіків зорово-просторовий. Більша частка чоловіків мають вищий та нижчий від середнього показники IQ, тоді як серед жінок переважають середні показники.

Відмінності IQ залежно від кількості дітей у сім'ї та черговості їх народження досліджені Ф. Салоуеєм, який виявив що первістки діти мають перевагу в побудові та підтримці батьківських та сімейних устроїв, тоді як діти, народжені пізніше, більше проявляють бунтарські тенденції та схильні до радикалізму. У дітей з багатодітних родин відмічають більш низькі показники за тестами інтелекту ніж у сім'ях з одним-двома дітьми. К. Зайонц та О. Маркус, дійшли висновків, що рівень IQ дітей (за тестом Равена) знижується по мірі збільшення чисельності сім'ї, а в сім'ях з однаковою чисельністю дітей старші діти мають вищий рівень інтелекту в порівнянні з молодшими. Дослідники пояснювали ці закономірності в теорії «сімейного ресурсу», запропонувавши модель сумісної взаємодії, яка полягає в припущенні, що інтелектуальний рівень дитини визначається інтелектуальним рівнем сім'ї, всіх її членів.

Дослідники описують особливий вплив домашнього середовища на дитину, поділяючи його на два типи: загальний вплив оточуючого середовища (пов'язаний з факторами, які визначають загальну атмосферу в родині: наявність розвиваючих іграшок та книг, батьківське моделювання тощо) та специфічний вплив оточуючого середовища (пов'язаний з факторами, які визначають відмінність сиблінгів: особливе батьківське ставлення, черговість народження, різниця у віці, особливі події, які стосуються одного сиблінга більше ніж інших).

Всі ці відомості варто враховувати, формуючи вибірку для дослідження ефективності тестів інтелекту.

Діагностика спеціальних здібностей особистості

Спеціальна здібність – це психологічні особливості індивіда, які забезпечують можливість виконання певного виду діяльності – музичної, сценічної, літературної, математичної і т.д. Розвиток спеціальних здібностей опирається на відповідні задатки (музичний слух, пам'ять тощо).

Про спеціальні здібності, як правило, говорять у двох аспектах:

- 1) як про здібності у певній сфері психічної діяльності (перцептивні, мнестичні, логічні тощо);
- 2) як про здібності до певного типу діяльності (лінгвістичні, художні, педагогічні).

Здібності другого виду описуються на базі перших і тести для їх діагностування включають *завдання для діагностики здібностей першого типу*.

Відомо, що для розвитку спеціальних здібностей існують певні сензитивні періоди, протягом яких вони прогресують особливо успішно (наприклад, для музичних здібностей це період до 5-ти років, коли активно формується музичний слух та пам'ять). Певний рівень спеціальних здібностей можна сформувати у будь-якої здорової дитини. Тому є реальні підстави для спрямованого розвитку здібностей дітей в умовах спеціально організованого навчання.

Здібності тісно пов'язані з загальною спрямованістю особистості. В. Чудновський зазначає, що співвідношення спрямованості особистості й рівня здібностей неоднозначне: високий рівень здібностей суттєво впливає на стиль поведінки та формування особистості. Ще більше значення має той факт, що розвиток здібностей великою мірою визначається умовами виховання, спрямованістю особистості, яка або сприяє розкриттю здібностей, або, навпаки, призводить до того, що здібності не реалізуються. В основі однакових досягнень за умов виконання певної діяльності можуть

бути різні здібності, водночас одна й та сама здібність може бути умовою успіху різних видів діяльності. Це забезпечує можливості широкої компенсації здібностей.

Здібності можна визначити також як поєднання сприятливих індивідуально-своєрідних особливостей та властивостей психіки, які виявляються у швидкості, результативності та якості виконання відповідної діяльності за мінімальних силових, енергетичних і часових затрат.

Рівень розвитку здібностей залежить від:

- якості знань і вмінь, від міри їх об'єднання в єдине ціле;
- природних задатків людини, якості природних нервових механізмів елементарної психічної діяльності;
- більшої чи меншої "тренуваності" самих мозкових структур, які беруть участь у здійсненні пізнавальних і психомоторних процесів (Б. Додонов).

У структурі здібностей (загальних та спеціальних) виокремлюють потенційні та актуальні можливості розвитку. Потенційні здібності -це можливості розвитку особистості, які виявляються щоразу, коли перед нею постає необхідність розв'язання нових завдань. Проте розвиток особистості залежить не лише від її психологічних властивостей, а й від тих соціальних умов, у яких ці властивості може або не може бути реалізовано. У такому разі говорять про актуальні здібності. І справді, аж ніяк не кожна людина може реалізувати свої потенційні здібності відповідно до своєї психологічної природи, оскільки для цього може не бути об'єктивних умов і можливостей. Отже, актуальні здібності становлять тільки частину потенційних.

Кожна здібність має свою структуру, в якій розрізняють *провідні й допоміжні* властивості. Зокрема, провідними властивостями в літературних здібностях є особливості творчої уяви та мислення; яскраві наочні образи пам'яті тощо.

Належність особистості до одного з трьох людських типів: "художнього", "мислительного" і "проміжного" (за термінологією І. Павлова) суттєвою мірою визначає особливості її здібностей. Відносна перевага першої сигнальної системи в психічній діяльності людини характеризує художній тип, відносна перевага другої сигнальної системи - мислительний тип, певна їх рівновага - середній тип людей. Ці відмінності пов'язують із функціями лівої (словесно-логічний тип) і правої (образний тип) півкуль головного мозку.

Усередині тих чи інших спеціальних здібностей виявляється загальна обдарованість індивіда, яка співвідноситься з більш загальними умовами провідних форм людської діяльності.

Контрольні запитання

Що таке тест? Як розуміють «тест» у психодіагностиці?

Які існують види тестів за спрямованістю?

Як називаються тести, для яких встановлюються часові обмеження?

Чим відрізняються прості й комплексні тести?

У чому полягають відмінності між стандартизованою і не стандартизованою психодіагностичною методикою?

До яких діагностичних інструментів більш доцільно вживати термін «методика»?

Що таке опитувальник?

Які переваги застосування комп'ютерних та комп'ютеризованих тестів?

Які вимоги перед фахівцями-психологами висувають комп'ютерні та комп'ютеризовані тести?

Чи потрібно емпірично перевіряти ефективність комп'ютеризованих тестів?

Що таке комп'ютерна тривожність?

Скільки завдань або запитань може містити особистісний опитувальник?

Які опитувальники мотивів Вам відомі?

Які особистісні опитувальники розробляв Л. Терстоун?

Які методики можна використовувати для тестування особи з метою профорієнтації?

Як можна зрозуміти, що опитувальник застарів?

Що таке генеральний фактор інтелекту?

Які особливості розвитку інтелекту у різних груп населення варто враховувати при розробці тестів інтелекту?

Що таке спеціальні здібності і які вони бувають?

Як можна діагностувати спеціальні здібності?

Що таке тести досягнень?

Поміркуйте...

До яких заходів можна вдаватися розробнику тестів, щоб запобігти їх безконтрольному розповсюдженню?

Яких етичних норм варто дотримуватися розробнику психологічних тестів?

Що може робити психолог задля сприяння розвитку психологічного тестування в Україні?

ПІДГОТОВКА ДО КОНСТРУЮВАННЯ ТА АДАПТАЦІЇ ПСИХОЛОГІЧНОГО ТЕСТУ

- ✓ Основні етапи конструювання тесту.
- ✓ Адаптація зарубіжних тестів.
- ✓ Характеристики досліджуваного, що можуть вплинути на результати тестування.

Основні етапи конструювання тесту

Алгоритм створення нового психодіагностичного інструментарію виглядає приблизно наступним чином:

- 1) розробка теоретичного конструкту того феномену, на вимірювання якого буде спрямований майбутній тест (можна сказати, що даний конструкт розробляється на трьох рівнях: теоретичному, емпіричному та операціональному);
- 2) специфікація тесту (визначення кількості завдань);
- 3) генерування завдань для тесту й оформлення тесту;
- 4) здійснення пілотажного тестування;
- 5) аналіз завдань;
- 6) визначення надійності тесту;
- 7) визначення дискримінативності тесту;
- 8) доведення валідності тесту;
- 9) збір нормативних показників;
- 10) публікація керівництва по застосуванню тесту;
- 11) перегляд тесту з часом (оскільки нормативи мають тенденцію старіти, особливо це стосується тестів інтелекту).

(послідовність виконання етапів, зазначених за номерами 6-9, може змінюватися залежно від стратегії конкретного дослідження)

П. Кляйн сформулював своєрідну пам'ятку розробнику тесту [19], у якій дещо конкретизує ці основні етапи й рекомендує вдатися до наступних кроків:

- ✓ встановити чіткі обмеження на зміст завдань;
- ✓ сформулювати настільки багато різноманітних завдань, релевантних цьому змісту, наскільки це можливо;
- ✓ випробувати ці завдання на настільки великій вибірці, наскільки це можливо;
- ✓ виконати аналіз завдань тесту, відбираючи ефективні;
- ✓ крос-валідизувати завдання на новій вибірці; переконатися, що завдання адекватно охопили відповідний контекст;
- ✓ обчислити коефіцієнт внутрішньої узгодженості завдань;
- ✓ валідизувати тест;
- ✓ визначити норми для цього тесту;
- ✓ виконати факторний аналіз завдань і порівняти його з процедурою аналізу та відбору завдань;
- ✓ якщо є підстави вважати цей тест ефективним і корисним, представити результати у керівництві до тесту.

Крім того, він наголошує на можливості застосування й інших методів конструювання тестів, зокрема конструювання тесту за допомогою критеріальних груп, факторного аналізу (а не процедури аналізу й відбору завдань), шкалювання завдань тесту за Рашем та ін.

Адаптація зарубіжних тестів

У практиці психодіагностики в Україні активно застосовуються адаптовані варіанти відомих закордонних тестів. Втім, більшість із якісних адаптацій (що про них є відомості щодо надійності, валідності, збору нормативних показників), були реалізовані зусиллями відомих радянських вчених В. Русалова, Л. Собчик, О. Шмельова, В. Зайцева, С. Єнікополова та

багатьох інших досить давно і російською мовою. Тим не менше, в Україні існує прикра тенденція до використання цих адаптацій в карщому випадку мовою оригіналу, в гіршому й просто зухвалому – неапробованому й несанкціонованому перекладі їх українською.

На думку Т. Лукіної, попри наявність у практиці українських психологів великої кількості адаптованих тестів, не можна сказати, що вони є якісними інструментами виміру, які ґрунтуються на теорії. Стандартизація більшості тестів не була здійснена до кінця. Тому тестування, як практика застосування тестів, створює видимість наукового підґрунтя [24].

Згідно думки Л. Бурлачука основні труднощі при адаптації тесту пов'язані з мовними та соціокультурними відмінностями досліджуваних. Вони зумовлені приналежністю людей до різних етносів, національностей, народів, конфесій, країн, географічного положення, тощо. Мовні проблеми здебільшого обумовлені відмінностями часів дієслів, подвійними запереченнями, розмовними ідіомами, родовими закінченнями, лінгвістичними конструкціями.

Варто згадати про нещодавнє дослідження особливостей використання російськомовної версії особистісного опитувальника Г. Айзенка українськими досліджуваними, що його провели студентки Національного університету «Острозька академія» В. Хмель та В. Поліщук. Дослідницями було зафіксовано, що респонденти добре розуміють менше половини (а саме 47,37%) запитань опитувальника. У іншій частині опитувальника у них виникають труднощі із розумінням окремих слів, а іноді й цілих запитань [34]. Недостатньо високий рівень розуміння змісту запитань, встановлений на основі аналізу якості їх перекладу респондентами, труднощі з окремими словами й навіть реченнями, - все це змушує поставити під сумнів надійність використання російськомовної адаптації особистісного опитувальника на українській вибірці. Отже, можна

констатувати, що необхідність у створенні україномовної адаптації існує і це є перспективне завдання досліджень з психодіагностики.

Адаптація будь якого психологічного тесту полягає не тільки у професійному перекладі, а й у перевірці психометричних характеристик репрезентативності, надійності, валідності, дискримінативності. Використання тесту здійснюється на основі різних методик, а процес тестування є змінним і адаптується до нових умов. Запозичення готових зарубіжних методик не є запорукою надійної і валідної психодіагностики. Використання тестів без адаптації до нових соціокультурних умов спотворює результати тестування щодо елементарних психічних властивостей.

Найменш складні в адаптації проєктивні методики, стимули яких не підлягають змінам. Це ж можна сказати і про інші невербальні тести. З найбільшими труднощами адаптуються особистісні опитувальники. Процес їх адаптації часто обмежують підготовкою попереднього варіанту перекладу з оригіналу на другу мову; експертною змістовою оцінкою попереднього перекладу із залученням лінгвістів і професійних психологів, що володіють мовою оригіналу, перевіркою еквівалентності експериментального перекладу шкали і оригіналу, встановлення нової шкали і збір відповідних норм на вітчизняній вибірці.

Завершена адаптація багатьох особистісних опитувальників за обсягом роботи не поступається розробці оригінальної методики. Тож ще до початку адаптації зарубіжного тесту слід бути упевненим в тому, що саме даний тест необхідний і заповнює певні прогалини в діагностичному інструментарії. Тому адаптації тесту повинне передувати вивчення його теоретичних підстав. У багатьох випадках на зміну адаптації зарубіжних тестів повинна прийти розробка оригінальних методик.

Отже, інформація про тест (адаптований чи щойно розроблений), що він пропонується для вирішення певних діагностичних завдань, повинна обов'язково містити:

- ✓ детального опису вмісту діагностичного інструменту;
- ✓ відомості про процедуру проведення і обробки отриманих результатів;
- ✓ опис норм;
- ✓ дані про надійність;
- ✓ дані про валідність.

Характеристики досліджуваного, що можуть вплинути на результати тестування

В ході численних досліджень впливу різноманітних факторів на процес та результативність тестування, виявилися деякі, що дійсно мають вплив. Саме про них йтиметься далі.

Тестова тривожність – один із самих яскравих проявів реакції досліджуваного на ситуацію тестування, що має великий вплив на результати останнього.

С. Саразон виявив, що *показники шкільних досягнень та тестів інтелекту мають значущі негативні кореляції з тривожністю*. Ці дані були неодноразово перевірені і підтверджені іншими дослідниками.

Цікаво, що ***зв'язок тривожності з результативністю тестування носить нелінійний характер*** – невелика тривога позитивно позначається на результативності, сильна ж стурбованість здійснює яскравий пагубний вплив.

Що стосується *природи тестової тривожності*, то тут дослідниками було виділено *два важливих компоненти*, а саме: *емоційність та стурбованість (сумніви в собі)*. Емоційна складова тестової тривожності включає почуття та фізіологічні реакції, такі як напруження та збільшення

частоти серцевих скорочень. Стурбованість (когнітивна складова) включає взаємопов'язані негативні думки, такі як очікування невдачі при виконанні тесту та стурбованість наслідками після очікуваного провалу. Ці думки відволікають досліджуваного від завдань тесту і тим самим порушують його виконання.

Поки що опублікований лише один валідний та надійний *опитувальник тестової тривожності* – TAI (Test Anxiety Inventory), розроблений Спілбергером та його співробітниками [12].

Вплив практичного навчання на виконання тестів. Вихідна гіпотеза відповідних досліджень формулювалася так: *тестовий показник стає невалідним лише в тих випадках, коли конкретний досвід підвищує його, не здійснюючи при цьому помітного впливу на ту сферу поведінки, для вимірювання якого призначений даний тест.* Факт впливу тренувань по виконанню схожих із тестовими завдань на результативність тестування підтвердився численними дослідженнями і виявилось, що завдання, що найбільше піддаються навчінню – це завдання-інсайти та складні завдання з великою кількістю додаткових інструкцій. Але в цьому контексті слід *розмежовувати* поняття «*тренування*» та «*навчання*», виходячи із зазначеної гіпотези (тренування по виконанню конкретних завдань знижують валідність тесту, а навчання цілій галузі знання в тестах досягнень, наприклад, навпаки, підвищують ефективність тесту).

У зв'язку з цією проблемою слід згадати також про *тестову досвідченість*, чи набуття обширної практики виконання тестів. Цей ефект у «бувалих» користувачів виникає внаслідок розвинутої в ході практики виконання різних тестів упевненості в собі і більш позитивного ставлення до тестової ситуації порівняно з тими, хто у використанні тестів не має значного досвіду. В результаті, з роками тренувань показники виконання тестів значно покращуються.

Виходячи з вищесказаного можна зробити висновок, що при розробці тесту необхідно брати до уваги різні тонкощі, що стосуються

подальшого його використання. Він повинен бути придатним для постановки психологічного діагнозу (надавати можливість виміряти різні елементи останнього), бути мінімально чутливим до впливу різних факторів (обстановки, дослідника, досліджуваного тощо) на результати тестування та відповідати іншим стандартам ефективного психологічного тесту.

Контрольні запитання

З чого починається розробка нового психологічного тесту?

Які основні методи конструювання тестів Ви можете виокремити?

Чим адаптація тесту відрізняється від розробки?

Чим адаптація тесту відрізняється від перекладу?

З чим пов'язані основні труднощі адаптації психологічного тесту?

Які тести адаптувати найпростіше?

Що представляє собою завершена адаптація тієї чи іншої методики?

Які характеристики досліджуваного можуть впливати на результати тестування?

Поміркуйте...

Які приклади адаптації методик для специфічних вибірок Ви могли б навести?

У чому полягають проблеми використання відомих і вживаних російськомовних адаптацій різних особистісних опитувальників?

Як можна уникнути впливу тестової досвідченості на результати тестування?

ОСНОВНІ ХАРАКТЕРИСТИКИ ЕФЕКТИВНИХ ПСИХОЛОГІЧНИХ ТЕСТІВ

- ✓ Особливості різних типів шкал вимірювання.
- ✓ Поняття надійності тесту та її види.
- ✓ Валідність як комплексний показник ефективного тесту.
- ✓ Дискримінативність тесту та його окремих частин.
- ✓ Стандартизація та норми.

Психологічний тест можна визнати ефективним, якщо він задовольняє певним вимогам, а саме:

- 1) використання шкали інтервалів;
- 2) перевірено і доведено надійність;
- 3) продемонстровано переконливі докази валідності;
- 4) доведено дискримінативність;
- 5) наявні нормативні дані, або (у випадку індивідуальних тестів) критерії, встановлених експертами.

Усі вимоги, що вони висуваються до психологічних тестів, не випадкові, а пов'язані з покращенням точності вимірювання – обов'язковою умовою науковості. Чим більш ретельно проконтрольовано зазначені параметри і дотримано необхідних процедур їх визначення, тим більша ймовірність того, що психологічний тест буде точним, а похибка вимірювання настільки незначною, що нею можна буде знехтувати.

Особливості різних типів шкал вимірювання

Шкала вимірювання – це форма фіксації сукупності ознак досліджуваного об'єкта з упорядкуванням їх в певну числову систему.

Шкали представляють собою метричні системи, що моделюють досліджуваний феномен шляхом заміни прямих позначень досліджуваних

об'єктів числовими значеннями і відображення пропорцій континуального складу елементів об'єкта у відповідних числах. Кожному елементу сукупності проявів властивостей об'єкта відповідає певний бал чи шкальний індекс, що кількісно встановлює положення досліджуваної одиниці на шкалі, котра охоплює всю сукупність чи її частину, суттєву з точки зору задач дослідження. *Операція упорядкування вихідних емпіричних даних в шкальні носить назву «шкалювання».*

Тобто, вимірювання можна визначити як присвоювання чисел для представлення деяких нечислових властивостей множини об'єктів.

Шкали відрізняються залежно від характеру функції, що лежить в основі їх побудови. Чотири основних типи шкал можна класифікувати за двома основами: початком відліку шкали (наявністю нульової точки) та одиниці вимірювання.

Таблиця 1.

Класифікація шкал вимірювання

	<i>Немає природного початку відліку (умовний нуль)</i>	<i>Наявний природний початок відліку (абсолютний нуль)</i>
<i>Немає одиниці вимірювання (неметрична шкала за С. Стівенсом)</i>	I (номінальна)	II (порядкова)
<i>Наявна одиниця вимірювання (відстань між двома точками; метрична шкала за С. Стівенсом)</i>	III (шкала інтервалів)	IV (шкала відношень)

Номінальна шкала (шкала найменувань) – якісна шкала – встановлює відповідність ознаки тому чи іншому класу. Тобто, об'єкти

класифіковані і ці класи позначені номерами. *Те, що номер одного класу більший чи менший від іншого, ще нічого не говорить про властивості об'єкту, за виключенням того, що вони відрізняються.* Не обов'язково, щоб між виявленими класами існував взаємозв'язок. Сама назва «шкала найменувань» вказує на те, що значення по шкалі грають роль лише назв класів.

Прикладами номінальної шкали можуть бути:

- запитання з дихотомічними відповідями («так»/ «ні»);
- запитання, що стосуються національності, кольору очей, статі, клінічного діагнозу і т.д.;
- завдання закритого типу (наприклад: *«Протилежністю значення «великодушний» є... а) марнотратний; б) малодушний; в) скупий; г) щедрий»*);
- перелік чи набір певних ознак, що будуть групуватися при зборі інформації чи її обробці (*«Ви переважно проводите вільний час... а) з товаришами та приятелями; б) на лоні природи; в) в заняттях спортом; г) в родинному колі»*).

Фактично, використання цієї шкали не дозволяє власне виміряти певну ознаку. Але, не зважаючи на цю обставину, цей вид шкалювання в психодіагностиці має досить велике значення, особливо при фіксації якісної інформації (наприклад, даних проєктивних методик при зборі психологічного анамнезу і т.д.)

Порядкова шкала (ординальна) – кількісна – призначена для розчленування сукупності ознак на елементи, пов'язані відношенням «більше» - «менше», і допускає віднесення змінних до груп, що є впорядкованими (ранжованими) по відношенню одна до одної і представляють собою деяку системну єдність. *Порядкова шкала дає можливість оцінити міру вираженості певної ознаки, кількість якості вимірюваного об'єкту; відповідно, значення шкали можна порівнювати за величиною*

Прикладом такої шкали могли б бути, скажімо, військові ранги, твердість певних матеріалів, в психології - ранжування за індивідуальними рисами особистості, шкала приємності-неприємності, шкала оцінювання якостей уваги: досить стійка – стійка – лабільна – розсіяна тощо.

Тобто, результатом використання шкали є те, що суб'єкти тут ранжовані, наприклад, за вагою, зростом і т.д. Така шкала очевидно груба, оскільки вона не враховує значення різниці між рангами (отже, шкала є неметричною, оскільки відсутні одиниці вимірювання).

Зі шкалами III та IV типу можна зустрітися при психологічному шкалюванні в будь-якій ситуації, коли по відношенню до кожного набору, наприклад, з трьох стимулів (a , b , c), що відрізняються між собою, допустимо запитувати, чи є різниця між a та b більшою чи меншою, ніж різниця між b і c , як при використанні шкал цінностей та переваг.

Шкала інтервалів – кількісна – існує одиниця вимірювання, за допомогою якої ознаки можна не лише впорядкувати, але й приписати їм числа так, щоб рівні різниці чисел, присвоєних об'єктам, відображали рівні різниці в якостях вимірюваної властивості. Нульова точка шкали вибрана довільно і не вказує на відсутність якості.

В таких шкалах різниці між значеннями у всіх точках шкали є рівними. Еквівалентні інтервальні шкали можуть бути лінійно перетворені одна в одну, що дозволяє приводити результати тестування до загальних шкал і таким чином здійснювати порівняння показників. Більше того, багато статистичних процедур передбачають вимірювання за шкалою інтервалів.

Шкалами такого типу є календарний час, шкали температур по Фаренгейту та Цельсію тощо. До розряду шкали інтервалів відносяться шкали IQ-показників, Т-балів, процентів тощо.

Шкала відношень (пропорційна) – кількісна – числові значення присвоюються об'єктам таким чином, щоб між числами та об'єктами простежувалася пропорційність. Тобто, числа, присвоєні предметам мають

всі властивості інтервальної шкали, але, окрім цього, на шкалі існує *абсолютний нуль*. Значення «нуль» свідчить про *відсутність досліджуваної властивості*. Наявність такої точки – проблема для більшості психологічних змінних, хоча існують методи конструювання тестів, що дозволяють використовувати цю можливість.

Приклади шкал відношень: час, температура за Кельвіном (наявний абсолютний нуль). Прикладом використання такої шкали в психологічному вимірюванні може слугувати, наприклад, шкала порогів абсолютної чутливості аналізатора.

П. Кляйн вважає очевидним, що в ідеалі розробник психологічних тестів повинен прагнути використовувати шкали відношень. У випадку, коли це неможливо, бажаними є шкали інтервалів, якщо передбачається, що результати будуть піддаватися статистичному аналізу. Оскільки визначення валідності тесту практично неминуче тягне за собою такий аналіз, висновок очевидний: нічого менш точного, ніж шкали інтервалів, використовувати не можна. Фактично, більшість психологічних тестів наближуються до інтервальних шкал, а обробка тестів з такими шкалами дає позитивні результати [19].

Поняття надійності тесту та її види

***Надійність** – це характеристика методики, що відображає точність психодіагностичних вимірювань, а також стійкість результатів тесту до дії сторонніх випадкових факторів [12].*

Результат психодіагностичного дослідження можуть зазнавати впливу *великої кількості неврахованих факторів* (наприклад, емоційний стан чи втома, якщо вони не входять до кола досліджуваних характеристик, освітленість, температура та інші особливості приміщення, в якому відбувається тестування).

Враховуючи сказане, на мінливість результатів дослідження можуть впливати дві групи причин:

- зміни, що властиві самому вимірюваному явищу (наприклад, коли досліджується психологічний стан, то мінливість отриманих показників не викликає подиву; якщо при дослідженні більш сталих психологічних явищ отримані за тестом показники не є стабільними у часі, досліднику варто з'ясувати, чи могли ці зміни дійсно мати місце);

- факторів нестабільності вимірювальної процедури.

В самому широкому сенсі *надійність тесту* – це характеристика того, якою мірою виявлені у досліджуваних відмінності за тестовими результатами є *відображенням дійсних відмінностей в досліджуваних якостях*, і якою мірою вони можуть бути *приписані випадковим помилкам* [12].

В психометрії термін *«надійність»* має два значення:

1) тест називається надійним, якщо він є *внутрішньо узгодженим*;

2) тест називається надійним, якщо він дає *однакові показники* для кожного досліджуваного *при повторному тестуванні* (за умови, якщо досліджуваний не змінився).

У відповідності із сказаним розрізняють два види надійності: *надійність за внутрішньою узгодженістю* та *ретестову надійність*.

Надійність за внутрішньою узгодженістю. В психометрії вважається: *якщо певна змінна вимірюється частиною тесту, то тоді в інших частинах тесту, якщо вони не узгоджені з першою, ця ж змінна вимірюватися не може* [9]. Така внутрішня узгодженість частин тесту є передумовою його валідності.

Варто зазначити, що Р. Кетел заперечував це формулювання, стверджуючи, що висока внутрішня узгодженість завдань тесту призводить до того, що вони дублюють один одного, не відкривають нічого нового у досліджуваному об'єкті. Більше того, дослідник вважав, що цим узгоджені

завдання можуть знижувати валідність тесту. Втім, на практиці всі тести, валідність яких вдалося переконливо довести, є високою мірою внутрішньо узгодженими. С. Морозов зазначає, що у більшості методик, що використовуються на практиці, рідко вдається досягти коефіцієнта надійності, що перевищує значення 0,7 – 0,8 [25].

Для обчислення внутрішньої узгодженості завдань тесту використовується коефіцієнт α Кронбаха (для неперервних величин) або коефіцієнт К'юдера-Річардсона (для дискретних величин, якщо завдання опитувальника дихотомічні).

Формула для обчислення коефіцієнту α Кронбаха

$$\alpha = \frac{k}{k-1} \left(1 - \frac{\sum \sigma_i^2}{\sigma_y^2} \right), \text{ де}$$

k - кількість завдань;

σ_y^2 - квадрат стандартного відхилення для всього тесту;

$\sum \sigma_i^2$ - сума квадратів стандартних відхилень для завдань.

Формула для обчислення коефіцієнту K-R20 (К'юдера-Річардсона)

$$r_{kr} = \frac{k}{k-1} \left(1 - \frac{\sum PQ}{\sigma_y^2} \right), \text{ де}$$

k - кількість завдань;

P – доля досліджуваних, що дали ключову відповідь для кожного завдання;

$$Q = 1 - P;$$

σ_y^2 - квадрат стандартного відхилення результатів тесту;

$\sum PQ$ є еквівалентом $\sum \sigma_i^2$ формули коефіцієнту α .

Розробника психологічного тесту задовольняє значення показника внутрішньої узгодженості не менше 0,7. Якщо отриманий показник менший, то на це могла вплинути ціла низка причин: суб'єктивність у оцінюванні результатів, відгадування досліджуваними відповіді (або установочна відповідь чи відповідь навмання при відсутності мотивації до

проходження тестування), незрозумілі завдання, недостатня величина тесту, незрозумілі інструкції до тесту.

Ретестова надійність. Коли тест не може дати той самий результат для досліджуваного (за умови, що той не змінився) в різних умовах, - значить, не все гаразд. Спосіб вимірювання ретестової надійності дуже простий: вираховується кореляція показників для вибірки досліджуваних, протестованих двічі з інтервалом приблизно у півроку (достатньо тривалий час, щоб досліджуваний встиг забути, як він відповідав раніше, і водночас достатньо короткий, щоб досліджуваний залишився схожим на себе півроку тому). Найменшим задовільним значенням для ретестової надійності є 0,7 [6]. При меншому значенні коефіцієнту кореляції використання тесту навряд чи буде доцільним, оскільки стандартна похибка отриманих показників буде настільки великою, що їх інтерпретація стане сумнівною.

Ретестова надійність досить проста у обчисленні. Варто остерігатися її штучного завищення шляхом проведення надто близьких у часі опитувань. До того ж, вибірки для перевірки надійності мають бути достатньо великими (200-300 респондентів) і репрезентативними за своїм складом.

Слід згадати також про **надійність паралельних форм** як варіант встановлення стабільності результатів тестування. Для перевірки цього різновиду надійності конструюються еквівалентні (паралельні) набори завдань. Таким чином, досліджувані виконують «абсолютно інший» тест в однакових умовах. Однак, є деякі труднощі у доказі того, що обидві форми є дійсно еквівалентними.

Слід зазначити, що одним із найважливіших **засобів підвищення надійності** психодіагностичних методик є **стандартизація процедури обстеження**. При суворій регламентації процедури (обстановка та умови роботи досліджуваного, характер інструкції, часові обмеження, способи та особливості контакту з досліджуваним, порядок пред'явлення елементів

методики, отримання первинних оцінок тощо) суттєво зменшується дисперсія похибки та підвищується надійність тесту.

Надійність певним чином пов'язана з іншою характеристикою ефективного тесту – валідністю. Про характер цього зв'язку йтиметься пізніше, а поки що слід визначитися із поняттям валідності та її видами.

Валідність як комплексний показник ефективного тесту

Тест називається *валідним*, якщо він вимірює те, для вимірювання чого він призначений.

Валідність – комплексна характеристика методики (тесту), що включає дані про сферу досліджуваних явищ та діагностичної репрезентативності процедури по відношенню до них [12].

У той час, як *надійність* можна охарактеризувати як стійкість процедури по відношенню до об'єктів дослідження, то валідність – це однозначність, стійкість відносно вимірюваних властивостей об'єкта (тобто, предмету вимірювання).

Слід особливо підкреслити твердження про те, що *надійність є необхідною, але недостатньою умовою валідності* [9]. Не можна, однак, це співвідношення характеризувати як прямо пропорційний зв'язок: підвищення надійності зовсім не означає автоматичного підвищення валідності.

На сьогодні розроблено багато способів продемонструвати валідність розробленої чи адаптованої методики. Всі вони складають так звану **конструктну валідність**, яка є тим способом доведення валідності методики, що найбільш повно відповідає визначенню валідності і передбачає врахування різних аспектів валідизації тесту. Висновки про конструктну валідність методики робляться на підставі результатів спеціально організованого дослідження, у ході якого перевіряються

гіпотези, щодо різних «видів» валідності. Приклади таких гіпотез буде подано після опису різних способів доведення валідності тесту.

Очевидна (зовнішня) валідність. Вважається, що тест є очевидно валідним, коли у досліджуваного про нього складається враження, ніби він вимірює саме те, для чого призначений. Звичайно, очевидна валідність не має ніякого відношення до істинної валідності і важлива лише у випадку роботи з дорослими досліджуваними для досягнення співпраці з ними. Л. Бурлачук зазначає, що дорослі не надто охоче включаються в роботу з тестами, яким бракує зовнішньої валідності, вважаючи їх дурними і навіть образливими.

Конкурентна валідність оцінюється за кореляцією результатів досліджуваного тесту з результатами інших тестів, валідність яких уже доведено.

Природньо постає питання про доцільність розробки інструменту у випадку, коли існує інший валідний тест. Так, дійсно, це потрібно лише тоді, коли у новий тест включено якусь значущу складову, яку не враховано у вже розроблених і валідизованих методиках. Крім того, доцільність розробки нового інструменту може бути продиктована необхідністю у коротших, сучасніших, більш простих у використанні й обробці способів вимірювання того чи іншого явища, ніж вже існуючі.

З іншого боку, якщо ефективні тести для вимірювання певної властивості чи особливості індивіда відсутні, тоді зрозуміло, що вивчення конкурентної валідності стає проблематичним. Іноді, коли немає тестів для вимірювання деякої властивості, можна спробувати використати *експертні оцінки*. Однак, і тут можуть існувати певні проблеми [9].

Загалом, конкурентна валідність корисна тоді, коли є незадовільно працюючі тести для вимірювання деяких змінних, а нові тести створюються як спроба підвищити якість вимірювання. В подібних випадках, можна очікувати значущих, але помірних кореляцій. Зрозуміло,

що конкурентна валідність не є достатньо повним, вичерпним аспектом валідності.

Конкурентна валідність також корисна для встановлення факту, чого ж **не вимірює** тест. Тест не повинен мати кореляції з іншими тестами, що вимірюють абсолютно інші змінні. Наприклад, тест для діагностики особистісної тривожності не повинен корелювати на значущому рівні з інтелектом. Якщо ж така кореляція виявилася в результаті відповідного дослідження, це свідчатиме про те, що тест тривожності включає в себе ті характеристики, яких не повинен містити, а отже він не може бути валідним. При цьому цілковито ймовірно, що результати досліджуваних виявляться стабільними у часі (будуть стійким до ретестування, ретестово надійними). Втім, як вже зазначалося вище, надійний тест може виявитися невалідним.

Одним з найбільш переконливих доказів валідності тесту вважається доведення її **прогностичної валідності**. Для її встановлення вивчаються кореляції між показниками тесту і деяким критерієм, що характеризує вимірювану властивість, але через певний час. Наприклад, прогностична валідність тесту на професійну спрямованість могла б бути продемонстрована кореляцією показників досліджуваних за цим тестом з їхніми здобутками у професійній діяльності через декілька років. Навіть з цього прикладу зрозуміло, що *основна складність при такій валідизації тесту полягає у виборі значущого зовнішнього критерію*.

Наприклад, у випадку тестів інтелекту можна було б використовувати майбутні успіхи в навчанні чи рівень доходів у якості такого зовнішнього критерію. Втім, це швидше гіпотеза, яку ще треба довести, адже очевидно існують й інші змінні, окрім інтелекту, котрі пов'язані з цими критеріями, такі як мотивованість, наполегливість, вміння користуватися зовнішніми й внутрішніми ресурсами, а також ряд інших випадкових зовнішніх факторів. Якщо у випадку з тестом інтелекту вивчення прогностичної валідності

може бути досить ефективно сплановане, то у випадку з особистісним опитувальником це є більш проблематично.

Тести для вимірювання нейротизму чи тривожності також можуть вивчатися для встановлення їх прогностичної валідності, оскільки їх показники можуть бути пов'язані зі сферою психічних захворювань та лікуванням, хоча й тут можливі неточності, так як не виключена поява невиявлених випадків чи таких, що не підлягали лікуванню [9].

Таким чином, вивчення прогностичної валідності психологічного тесту є достойним викликом для розробника, адже важко уявити більш переконливий доказ спроможності методики не лише фіксувати вираженість певного явища, а й робити на цій основі певний прогноз.

Інкрементна та диференціальна валідність вказує на роль тесту в покращенні відбору осіб для реальної діяльності і може бути корисною, в основному, у психологічних обстеженнях при профвідборі. *Інкрементна валідність* (від англ. increment – прибуток, приріст) відноситься до випадку, коли один тест з батареї тестів може мати низьку кореляцію з критерієм, але не перекриватися іншими тестами з цієї батареї [12].

Диференціальна валідність добре ілюструється тестами інтересів. Ці тести помірно корелюють з загальною університетською успішністю, але по-різному для різних дисциплін. Таким чином, можна сказати, що вони мають диференціальну валідність для академічної успішності. Тести IQ, з іншого боку, мають більш високі кореляції з прийнятими в університеті вченими ступенями, але не можуть диференціювати за окремими дисциплінами [9].

Зрозуміло, що інкрементну та диференціальну валідність потрібно визначати не для всіх тестів, вони є корисними показниками ефективності для тих тестів, котрі потрібно буде використовувати з метою відбору.

Змістова валідність вивчається у випадку конструювання тестів досягнень. Для її доведення необхідно показати, що тест включає всі аспекти, які він повинен був містити відповідно до певних програм

підготовки чи ін. Прикладом тесту досягнень може бути модульний контроль з будь якої дисципліни у вузі. Щоб переконатися у його змістовій валідності, розробнику необхідно, залучивши експертів з цієї дисципліни, вивчити питання про те, чи до вибірки тестових завдань рівномірно й повно включено всі тематичні блоки, що передбачалися програмою цієї дисципліни. Якщо можна показати, що завдання тесту відображають всі аспекти досліджуваної сфери, то тест є, по суті, валідним за умови, що інструкції викладені зрозуміло. Цей же вид валідності має бути доведеним, наприклад, при конструюванні тестів ЗНО.

Таким чином, вивчивши теоретичний конструкт, з'ясувавши, які види валідності можливо і доцільно вивчити для цього тесту, дослідник формулює гіпотези щодо його конструктної валідності.

Прикладом таких гіпотез можуть бути припущення, сформульовані студентками 3 курсу НаУОА *М. Ткаченко, Т. Дробатою, О. Снігурою* у 2014 році в ході адаптації методики Neo PI-R (Big 5) П. Коста та Р. Мак-Крея:

1. Шкала «екстраверсія» опитувальника буде мати значущу кореляцію з шкалою «екстраверсія/інтроверсія» тесту Г. Айзенка.
2. Шкала «співчутливість» буде помірно корелювати з результатами тесту на емпатійні здібності В. Бойко.
3. Шкала «сумлінність в роботі» буде помірно корелювати з результатами диференційного тесту перфекціонізму А. Золотарьової.
4. Шкала «нейротизм» буде значуще корелювати з відповідною шкалою особистісного опитувальника Г. Айзенка.
5. З опису шкали «відкритість досвіду» витікає, що вона буде мати помірну кореляцію зі шкалою В 16-факторного опитувальника Р. Кетела.
6. Жодна із шкал опитувальника не буде корелювати із методикою САН.

Звичайно, перелік гіпотез щодо конструктивної валідності зазначеної методики може бути розширений, тим не менше, із наведеного прикладу стає зрозуміло, яким чином дослідники планують вивчати валідність психодіагностичного інструменту, які засоби їм знадобляться для цього. *Подальший, більш безпосередній спосіб продемонструвати валідність тесту може полягати у формулюванні останньої гіпотези, а саме: досліджувані, що мають високу вираженість вимірюваних рис особистості, покажуть за тестом Neo PI більш високі показники, ніж ті, у котрих вона низька.*

Дискримінативність тесту та його окремих частин

Дискримінативність – це здатність окремих завдань тесту та тесту в цілому диференціювати досліджуваних відносно максимального та мінімального результатів тесту [19].

Дійсно, досягнення задовільного розподілу показників є однією з цілей розробника тесту. Немає необхідності підкреслювати те, що є очевидним: яка цінність психологічного тесту, за яким всі досліджувані показали однакові результати? За допомогою ретельного конструювання можна забезпечити відповідний рівень дискримінативності, а це те, за допомогою чого тести значно виграють порівняно з іншими формами досліджень.

Дискримінативність тесту вимірюється показником дельта Фергюсона і приймає максимальне значення при рівномірному (прямокутному) розподілі показників ($\delta=1$). Дискримінативність завдання встановлює, чи спроможний конкретний пункт тесту надійно розрізнити досліджуваних, які мають високі та низькі бали за тестом. У випадках, коли значення дискримінативності завдання наближається до нуля, відповідний пункт тесту має бути переглянутий.

Формула обчислення коефіцієнту Фергюсона:

$$\delta = \frac{(n+1)(N^2 - \sum f_i^2)}{nN^2}, \text{ де}$$

N - кількість досліджуваних;

n - кількість завдань;

f_i - частота появи кожного показника.

Значення $\delta=0$ показує, що всі досліджувані отримали однакові показники за тестом, а отже він є абсолютно нечутливим інструментом, тож ефективність його використання сумнівна.

Ще у 1997 році вважалося, що розрахунки, пов'язані з обчисленням індексів дискримінативності, досить трудомісткі, особливо для тестів з великою кількістю завдань [25]. Проте сьогодні доступність програмного забезпечення для здійснення комп'ютерного обчислення усіх математико-статистичних показників значно полегшує здійснення психометричного контролю психодіагностичних тестів.

Стандартизація та норми

Наступна характеристика ефективних тестів, що не є, однак, внутрішньою особливістю, що притаманна самому тесту, - це наявність нормативних даних. *Норми – це, фактично, множина показників із чітко визначених вибірок*, а розробка та процедура отримання цих показників складають процес стандартизації тесту.

Стандартизація – це уніфікація, регламентація, приведення до єдиних нормативів процедури та оцінок тесту [12].

Завдяки стандартизації методики досягається порівнянність результатів у різних досліджуваних, за використання різних методик. В психодіагностиці вирізняють дві форми стандартизації:

1) коли під стандартизацією розуміється *обробка та регламентація процедури проведення*, уніфікація інструкції, бланків обстеження, способів

реєстрації результатів, умов проведення обстеження, характеристика контингентів досліджуваних;

2) коли під стандартизацією розуміється *перетворення* нормальної (чи штучно нормалізованої) шкали оцінок *в нову шкалу, засновану* вже не на кількісних емпіричних значеннях досліджуваного показника, а *на його відносному місці в розподілі результатів у вибірці досліджуваних*.

Наявність норм є надважливою для будь якого тесту, адже саме вони дозволяють кваліфікованому користувачу тесту адекватно інтерпретувати показники, котрі він отримує від досліджуваного. Відповідно *норми мають більше значення для практичного застосування тестів, ніж для дослідницьких цілей*. Для більшості психологічних досліджень, зокрема для тих, що проводяться студентами-психологами під час навчання у вузі, достатньо так званих «сирих», не опрацьованих показників за тестами, а норми додають лише деяку додаткову інформацію. В індивідуальній роботі ситуація зовсім інша – без нормативних показників жодний бал по тесту не має значення, сенсу. Таким чином, за відсутності норм індивідуальний показник по тесту може бути проінтерпретований хоч якимось чином лише за наявності групових даних для порівняння.

Слід зазначити, що *комплектування вибірки досліджуваних для збору нормативних показників за тестом має бути адекватним (вибрка відповідним чином стратифікована), а її обсяг – великим (приблизно 200 чоловік на кожну страту)*. У протилежному випадку *нормування тесту може виявитись швидше шкідливим, ніж просто марним*. Слід також розуміти, що більшість методів досліджень, відмінних від тестів, не можуть бути стандартизовані взагалі, тож ця можливість є важливою та виключною характеристикою психометричних тестів.

Зважаючи на те, що норми збирають для тестів, ефективність яких перевірено (доведено їх надійність, валідність та дискримінативність), питання про особливості цієї процедури не висвітлюватиметься на

сторінках цього посібника, адже реалізація цього етапу розробки тесту не передбачено програмою вивчення дисципліни «Психодіагностика».

Контрольні запитання

Що таке «шкала вимірювання»?

Які типи шкал вимірювання Ви знаєте?

Чому при конструюванні тестів рекомендовано застосовувати шкалу інтервалів?

Що таке надійність тесту?

Які способи вивчення надійності методики Ви знаєте?

Як ре тестова надійність характеризує якість методики?

За допомогою яких математичних коефіцієнтів обчислюється внутрішня узгодженість запитань? Які значення цих коефіцієнтів задовольняють розробника?

Що може бути причиною незадовільної надійності тесту?

Якою повинна бути вибірка для вивчення надійності тесту?

Які засоби підвищення надійності методики Вам відомі?

Що таке валідність тесту?

Як пов'язані між собою валідність та надійність методики?

Які способи доведення валідності тесту Вам відомі?

Що таке прогностична валідність? У чому цінність і проблеми її доведення?

Що представляють собою гіпотези щодо конструктної валідності тесту?

Як характеризує якість тесту його очевидна валідність?

При дослідженні якого виду валідності передбачається робота експертів?

Як підбираються діагностичні методики для дослідження конкурентної валідності розробленого тесту?

Що таке дискримінативність тесту?

Як обчислюється дискримінативність тесту математично?

Що таке норми? Для чого вони потрібні?

Поміркуйте...

Як можна досліджувати прогностичну валідність тестів, призначених для профорієнтації?

Відповідність яким високим стандартам (надійності чи дискримінативності) є важливішою для психологічного тесту?

Як інтерпретувати індивідуальні результати досліджуваного за тестом, якщо він не був стандартизований і відомостей про норми для різних вибірок відсутні?

РОЗРОБКА ЗАВДАНЬ ПСИХОЛОГІЧНОГО ТЕСТУ

- ✓ Розробка теоретичного конструкту та специфікація майбутнього психологічного тесту.
- ✓ Формулювання завдань особистісного опитувальника.
- ✓ Розробка завдань для тестів здібностей.
- ✓ Формулювання завдань для тестів досягнень.
- ✓ Пілотажне дослідження й математичний аналіз завдань тесту.

Розробка теоретичного конструкту та специфікація майбутнього психологічного тесту

Розробці завдань будь якого психодіагностичного інструменту передують ретельна проробка **теоретичного конструкту** того явища, на вимірювання якого буде спрямований цей тест. Саме тут варто детальніше зупинитися на вже згаданих двох можливих принципах конструювання діагностичних методик: критеріально-ключовому та факторно-аналітичному.

Критеріально-ключовий принцип ґрунтується на емпіричному виявленні психологічних ознак, що дозволяють диференціювати релевантні критеріальні групи від контрольних [12]. Результатом реалізації цього принципу конструювання є так звані емпіричні психологічні тести (зокрема Мінесотський багатоаспектний особистісний опитувальник - MMPI). Краще зрозуміти критеріально-ключовий принцип можна на прикладі розробки вірцевого «Бланку інтересів» Е. Стронга. Формуванню остаточного набору завдань цього тесту передувало емпіричне вивчення груп досліджуваних, що мали різні інтереси. З загального масиву сформульованих завдань до остаточної версії опитувальника було відібрано лише ті, які ефективно розділяли ці групи досліджуваних. Така ж стратегія формування завдань може виявитися ефективною і при професійному

доборі: ансамбль завдань, сконструйований на основі результатів емпіричного вивчення відмінностей між групою успішних у певній професії осіб та контрольною групою, має більше шансів бути прогностично валідним, ніж навіть ретельно вивчений теоретично перелік можливих вимог до кандидата. Із сказаного зрозуміло, що в емпіричних тестах основна увага приділяється вивченню дискримінативності.

До проблем, пов'язаних із конструюванням тестів за цим принципом, належать труднощі у відборі критеріальних груп та психологічній інтерпретації показників тестів.

Факторно-аналітичний принцип – принцип конструювання тестів на основі виявлення обмеженого кола факторів, що визначаються розробником (там же). Під фактором розуміють конструкт, що операційно визначається його факторними навантаженнями (виявляються в процесі факторного аналізу). Факторні навантаження – значення кореляції змінних з цим фактором. При розробці тесту згадане коло факторів визначається на основі кореляційного аналізу зв'язку між завданнями, в результаті якого до остаточної версії тесту включають лише ті, які найбільш тісно корелюють з загальним фактором. Саме він потім виступає в якості теоретичного конструкту. Отриманий у результаті тест називають *теоретичним*. Прикладом опитувальників, розроблених на цій основі, є відомі факторні опитувальники «Шістнадцять особистісних факторів» Р. Кетела, особистісний опитувальник Г. Айзенка та ін.

Створений на основі такого емпіричного та «теоретичного» (в лапках, адже ця теорія обов'язково ґрунтується на даних емпіричних досліджень) вивчення досліджуваного феномену конструкт є основою для формування набору тестових завдань.

Для того, щоб визначити оптимальну кількість завдань майбутнього тесту, варто зважити деякі моменти: по-перше, скільки досліджуваних змінних передбачає методика, по-друге, які прояви цих змінних будуть фіксуватися відповідними завданнями. Відомо, що довжина тесту впливає

на його надійність і для однієї шкали методики (у простих тестах – єдиної) кількість запитань має бути не менше 20. Збільшення кількості запитань є одним із способів підвищення надійності методики. Разом з тим, надто довгий тест втомлює досліджуваного. Визначити оптимальну кількість завдань різного змісту допомагає така процедура, як *специфікація тесту*, що, по суті, узагальнює результати розробки теоретичного конструкту.

Л. Бурлачук зазначає, що специфікацію найбільш зручно здійснювати у вигляді таблиці, у якій по горизонталі міститимуться *змістові області*, котрі будуть вимірюватися, а по вертикалі – їх *маніфестації*, або шляхи, якими ці змістові області можуть проявлятися.

Прикладом специфікації тесту у нашому посібнику слугуватиме таблиця специфікації, розроблена студентами під час спроби адаптації ними (процес адаптації триває) відомого індикатора типів Маєрс-Бріггс.

Таблиця 2

**Таблиця специфікації для особистісного опитувальника МВТІ
(розробники: А. Панасюк, А. Толочко, А. Трофимчук, О.**

Росошик)

Змістові області	Маніфестації			Кількість завдань
	Поведінка (61%)	Емоції (20%)	Когніції (19%)	
Екстраверт - інтроверт (EI)	14	5	1	20
Сенсорик - інтуїт (SN)	9	3	8	20
Мислячий – чуттєвий (TF)	11	6	3	20
Сприймаючий – рішучий (PJ)	15	2	3	20
Шкала брехні	12	4	4	20
Кількість завдань	61	20	19	100

З таблиці 2 видно, що на підставі аналізу наявних емпіричних та теоретичних відомостей про можливі прояви психологічного типу особистості дослідниці визначили питому вагу (і, відповідно, кількість запитань) різних змістових областей та їх можливих маніфестацій наступним чином: змістові області визначено як рівно значимі, причому з такою кількістю запитань, яка дозволить обчислювати надійність та інші показники для кожної шкали окремо (а також аналізувати інтеркореляції між шкалами), серед маніфестацій головну роль відведено поведінковим проявам (61%) і майже порівну емоційним (20%) та когнітивним (19%) маніфестаціям виокремлених змістових областей. Оскільки в цілому відсоткова вага всіх комірок по вертикалі й по горизонталі складає 100%, стає зрозуміло, яку кількість завдань варто створити для кожної комірки. Вважається, що решітка розміром від 16 до 25 комірок (4*4, 4*5, 5*4 чи 5*5) є ідеальною для тої довжини тесту, котру цілком ймовірно сконструювати і опрацювати.

Розробка такої таблиці робить абсолютно прозорим і зрозумілим подальший процес генерування завдань опитувальника, адже дослідник точно знає скільки разів запитати про конкретний спосіб прояву конкретного змістового аспекту вимірюваного явища. Варто зауважити, що для пілотажного дослідження методики варто згенерувати як мінімум на 50% завдань більше, ніж передбачається залишити в остаточній редакції, адже велика частина завдань може бути відсіяна у результаті подальшої перевірки їх ефективності.

Формулювання завдань особистісного опитувальника

Питання про ефективні форми завдань особистісних опитувальників відноситься до добре розроблених питань у психодіагностиці. На практиці виявляються найбільш ефективними і часто застосовуються такі типи завдань:

1. Запитання з відповіддю типу «так / ні» (наприклад: «*Чи легко вас роздратувати?*», «*У Вас виникало відчуття, що Ви нещасні без жодної на це причини?*», «*Чи буває у Вас прискорене серцебиття?*» тощо).
2. Задання типу «правда / неправда» (схоже з попереднім завданням альтернативного типу з тією різницею, що пропонується твердження, яке досліджуваній оцінює на предмет того, чи воно правдиве для нього; наприклад: «*Думаю, я маю здібності до наукової діяльності*», «*У мене багато друзів*», «*Я завжди кудись поспішаю*» та ін.).

*Відомо, що на виконання альтернативних завдань особистісного опитувальника впливає так звана **установка досліджуваного на згоду** – його схильність обирати відповідь «так» незалежно від змісту запитання, значно частіше погоджуватись, ніж відповідати запереченням. Цю установку варто враховувати при розробці завдань і формувати якусь їх частину за оберненим принципом, коли за відповідь «ні» буде нараховуватись бал згідно з ключем.*

3. Запитання з відповідями типу «так / дещо середнє (або непевне) / ні». Тут варто зауважити, що, з одного боку, деякі досліджувані нервують, коли не мають можливості обрати середній варіант (у альтернативних завданнях), з іншого боку – надто люблять середній варіант, коли він присутній.

*Схильність досліджуваного обирати невизначений варіант відповіді у трихотомічних завданнях відома як **установка на середню відповідь**, яка приваблює досліджуваного як безпечний компроміс, який, на жаль, знижує валідність методики. Подолати цю установку досліджуваного буває вкрай важко, тож рекомендація може звучати хіба що так: не варто, щоб серединна категорія виражала невпевненість («не знаю», «вагаюсь із відповіддю»), адже тоді вона втрачає будь яку*

інформативність; завдання з середнім варіантом відповіді типу «іноді» все ж може бути опрацьований як деякий рейтинг).

4. Завдання з відповідями «подобається / не подобається» (стимульне слово чи висловлювання). Цей тип завдань вважається в своїй основі проєктивним, адже досліджуваний відповідає виключно на основі свого емоційного відгуку на стимул. Такі завдання представлені зокрема у опитувальнику МВТІ (наприклад, «*Оберіть слово, яке подобається Вам більше: а) бас; б) барабан*»).
5. Завдання з рейтинговими шкалами. Такі завдання зазвичай здаються досліджуваним більш розумними, ніж дихотомічні, а тести з такими завданнями – більш серйозними (наприклад: «*Якби у Вас була можливість викладати, то ви з більшим задоволенням вели б лекції, а не практичні (лабораторні) заняття: а) ні; б) швидше за все, ні; в) можливо; г) швидше за все, так; д) так*»). Правила, що якимось чином регламентують кількість варіантів вибору, в принципі, відсутні. Втім, зазвичай не рекомендується використання в опитувальниках більше семи градацій.

Саме при опрацюванні результатів рейтингових шкал виявлена установка досліджуваних обирати крайні відповіді незалежно від змісту завдання («завжди / ніколи», «абсолютно так / абсолютно ні»). Причому, чим більше градацій шкали, тим яскравіше проявляється ця установка.

6. Завдання з варіантами вибору. Таких варіантів, як правило, 3 або 4, хоча можлива й інша кількість, і вони оцінюються як номінальна шкала. Цей тип завдань часто застосовується і його перевагою є те, що воно дозволяє вкласти в свою лаконічну форму практично будь який зміст (наприклад: «*Мої друзі частіше: а) радяться зі мною; б) дають мені поради; в) роблять і те, і інше*», «*Я б надав перевагу професії, що передбачає взаємодію: а) з людьми; б) з природою; в) з технікою*»).

7. Завдання з вимушеним вибором. Цей тип завдань, зокрема, представлений у відомому і широко застосовуваному у нас тесті ДДО Є. Клімова. У тесті подано наступну інструкцію (щоправда, російською мовою): *«Припустимо, що після відповідного навчання Ви зможете виконати будь яку роботу. Але якби Вам довелося обирати лише з двох можливостей, якій Ви віддали б перевагу?»* - і далі пари варіантів, з яких досліджуваній мусить обрати, навіть якщо ні одні, ні інший варіант йому не близький (наприклад: *«2 а) допомагати хворим; 2 б) складати таблиці, схеми, програми для обчислювальної техніки»*).

При формулюванні будь-якого типу завдань особистісних опитувальників варто уникати можливості проникнення досліджуваного в суть завдання.

При проникненні досліджуваного в суть запитання може проявитися ще одна установка, що значимо впливає на результати багатьох тестів, зокрема MMPI, - установка на соціально схвалювані (соціально бажані) відповіді (тенденція досліджуваного відповідати так, щоб виглядати «соціально позитивним»). Якщо серед варіантів відповіді є соціально бажаний, досліджуваній з високою долею ймовірності обере саме його. Для виявлення установки досліджуваного відповідати соціально позитивно розробляються спеціальні шкали (наприклад, шкала соціального схвалення Д. Марлоу і Д. Крауна). Крім того, відомою технікою відсіювання результатів досліджуваних, які схильні до нещирих відповідей, є розробка й використання шкали брехні із запитаннями типу «Я ніколи не пліткую», «Я завжди приходжу на зустрічі вчасно» і т.д.

Крім того, добре, коли завдання відображають конкретні, а не загальні аспекти досліджуваної області поведінки. Так запитання *«Чи подобається Вам спорт?»* ймовірно виявиться не настільки ефективним,

як запитання «*Ви граєте регулярно у якусь спортивну гру?*» або «*Чи слідкуєте Ви за грою Вашої улюбленої команди?*».

Слід також прослідкувати, щоб запитання не було подвійним, наприклад: «*Вважаю, що неслухняні і невстигаючі діти мають вчитися окремо від інших*». Таке питання варто переформулювати на два однозначних: «*Вважаю, що неслухняні діти мають учитися окремо*», «*Вважаю, що невстигаючі діти мають вчитися окремо*».

Варто по можливості уникати також і слів, що суб'єктивно визначають частоту дій («*інколи*», «*часто*») і заміняти (хоча на практиці таке буває важко забезпечити) їх більш конкретними, об'єктивними варіантами («*тричі на тиждень*», «*кожного дня*»), що звичайно, потребує особливої уваги до операціоналізації основних конструктів.

В особистісних опитувальниках також дуже важливо, щоб досліджувані знали (а вони це дізнаються з інструкції), що правильних і неправильних відповідей немає, а найкраще давати ту, що перша спала на думку. Загальновідомим є факт, що спотворення відповідей потребує часу на роздуми.

Розробка завдань для тестів інтелекту

Перш ніж перейти до завдань тестів здібностей, варто кілька слів зазначити про **рекомендації до інструкцій**. Досліднику варто постаратися, щоб вони були:

- 1) настільки прості й зрозумілі, наскільки це можливо;
- 2) настільки короткі й лаконічні, наскільки це можливо;
- 3) в них (по можливості) мають використовуватися лише короткі прості речення, без уточнюючих підпорядкованих;
- 4) приклади завжди повинні супроводжувати інструкцію.

Типи завдань у тестах здібностей (незалежно від того, чи мова йде про вербальні чи невербальні завдання):

1. Завдання-аналогії. Їх перевагами є те, що можна легко змінювати рівень їх складності, пристосовуючи їх до різних вікових груп та рівнів розвитку. Крім того, практично немає обмежень щодо матеріалу завдань, до яких можна використати принцип аналогій.

2. Завдання на виключення зайвого.

3. Завдання-послідовності. Теж корисна форма завдань, в яку легко включити відношення різних рівнів складності (єдина форма завдань у матрицях Равена).

Загальні рекомендації до конструювання завдань:

- Варто спочатку продумати відношення, з яким буде пов'язане завдання (наприклад, відношення схожості, протилежності).
- Подати приклад цього відношення в завданні (через аналогію: «ліве – праве, біле - ...», чи послідовність, у якій потрібно виключити зайве: «окунь, щука, вугор, рак» або ж додати пропущене «24, 12, 6, ...»).
- Слід обов'язково переконатися, що компоненти завдання випадковим чином не утворюють інші відношення, що призводять до альтернативних рішень.
- За виключенням роботи з освіченими досліджуваними, варто обирати настільки простий матеріал для завдань, наскільки це можливо. Складність має полягати у природі відношень, яку необхідно розкрити, а не у вузькій спеціалізованості використовуваного матеріалу (як у тестах досягнень).

Щоб уникнути моногонії завдань, їх треба урізноманітнювати, варіювати, наприклад, за допомогою таких способів:

✓ Для завдань-послідовностей можна просити досліджуваного знайти деякий середній елемент, а не крайній за розташуванням, або дати завдання пере впорядкувати елементи у відповідності з виявленою закономірною послідовністю.

✓ Для завдань на виключення зайвого приводяться елементи з загальними характеристиками, а досліджуваним пропонується знайти ті, у яких цих характеристик немає.

Ці форми завдань не єдині, які використовуються у тестах інтелекту. Тим не менше, вони детально вивчені і тому найбільше підходять для конструювання ефективних завдань.

Формулювання завдань для тестів досягнень

При формулюванні завдань до тестів досягнень дослідник має вирішити проблему *форми та змісту* запитань. На відміну від тестів загальних здібностей, у тестах спеціальних здібностей та досягнень критичною складовою є зміст запитань. Типи завдань розроблено з метою забезпечення об'єктивності оцінювання. Відповідно, розробники тесту мають спочатку знайти деякий ідеальний зміст запитання, а потім – формулювати завдання, яке б найкращим чином представляло зміст. (не всі завдання однаково добре підходять для діагностики того чи іншого змістового компонента здібностей).

Тести досягнень оцінюються експертами на предмет відповідності змісту завдання поставленим задачам. Ці тести представляють собою набори завдань у сфері освіти, спрямовані на вимірювання результатів навчання, але це, фактично, не психометрія, як чітко вказує Л. Бурлачук.

Типи завдань для тестів досягнень:

1. Завдання з декількома варіантами вибору використовуються найчастіше. Вони складаються з двох частин: основна частина, що містить твердження або запитання; варіанти вибору чи деяка кількість можливих відповідей, з яких досліджуваний має обрати правильну (як правило, таких варіантів є 4 або 5).

Для формулювання ефективних завдань цього типу важливими є *простота* (не бажано, щоб на результати тестування впливали рівень словникового запасу досліджуваного чи його загальних здібностей); всі *дистрактори* мають бути такими, щоб кожен з них міг привернути увагу досліджуваного (один із способів отримання ефективних дистракторів – використання незакінченого переліку варіантів вибору і подальше використання неправильних відповідей, записаних самими досліджуваними), проте, слід обов'язково переконатися, що обрані дистрактори *не заплутують найкращих досліджуваних*.

2. Альтернативні завдання (твердження або запитання, що передбачають відповіді «так», «ні», наприклад: «*Чи погоджуєтесь Ви з тим, що ненадійний тест валідним бути не може?*») не надто корисні для тестів досягнення через високу ймовірність відгадування відповіді. Тим не менше, вони зустрічаються у тестах. доцільним їх використання може бути тоді, коли вони спрямовані на визначення *розуміння* досліджуваним деякого матеріалу (наприклад, *подано графік*, а потім кілька описів (правильних чи неправильних) матеріалу, який міститься у графіку, а досліджуваний повинен погодитися у випадку правильного твердження чи не погодитися у випадку хибного).
3. Завдання на відновлення відповідності теж успішно застосовуються у тестах досягнень. Прикладом цього завдання може бути таке:

Відновіть, будь ласка, відповідність між автором і його працею:

- | | |
|-------------|---|
| 1. К. Юнг | а) «Наука жити» |
| 2. З. Фрейд | б) «Основні психологічні теорії в психоаналізі» |
| 3. А. Адлер | в) «Психологічні типи» |
| 4. Ф. Перлз | |

г) «Всередині і поза помийним відром»

У цьому типі завдань дистрактори (неправильні варіанти відповіді) відсутні і є ймовірність, що певні елементи завдання можуть бути визначені за остаточним принципом. Тим не менше, досліджуваним вони подобаються і часто використовуються при перевірці досягнень.

Пілотажне дослідження та математичний аналіз завдань тесту

Після того, як банк запитань для пілотажного дослідження сформовано, вони підлягають емпіричній перевірці на вибірці, яка має схожі характеристики з тими досліджуваними, для кого призначений тест. Обсяг вибірки для пілотажного дослідження може бути не таким значним, як, наприклад, для вивчення надійності, але, все ж, пілотажну версію опитувальника варто прагнути випробувати на якомога більшій кількості респондентів. А. Анастасі пропонує, щоб досліджуваних при пілотажі було вдвічі більше, ніж завдань у тесті.

Для більшості тестів при формуванні вибірки пілотажного дослідження слід враховувати такі характеристики досліджуваних, як *стать* (можливі відмінності між відповідями чоловіків та жінок корисно передбачати ще перед пілотажем, адже накопичено достатньо відомостей про те, у яких тестах та при якому змісті завдань статеві розбіжності у відповідях найчастіше мають місце [5, 6, 9, 19], *вік* (особливо впливає на результати тестів здібностей та досягнень) досліджуваних, а часом й інші змінні (національність, рівень освіти, приналежність до певного соціального класу), що можуть бути важливими з огляду на особливості вимірюваного явища.

Якщо досліднику вдалося сформуванати достатньо велику і адекватну за складом вибірку пілотажного дослідження, наступним кроком після його реалізації буде математичне опрацювання результатів з метою вибору завдань, що будуть здатні забезпечити надійність і валідність розробленого психодіагностичного інструменту. Відповідно до цієї мети, дослідника цікавитиме:

- 1) доля досліджуваних, що дали ключові (для особистісних тестів) або правильні (для тестів здібностей) відповіді (цей показник покаже, які запитання погано диференціюють досліджуваних, а отже не мають цінності);
- 2) кореляція завдання з загальним показником за тестом чи шкалою (для забезпечення однорідності, внутрішньої узгодженості тесту).

За використання пакетів математико-статистичної обробки інформації ці обчислення робляться швидко й легко. Найбільш трудомістким процесом, що потребує пильності дослідника, є формування матриці даних, куди буде перенесено закодовані відповіді усіх досліджуваних на кожне запитання тесту. При розробці комп'ютеризованого тесту за допомогою відповідного програмного забезпечення формування матриці автоматизується, що робить процес аналізу даних ще менш рутинним і затяжним. Зразок матриці даних представлено у додатках до посібника.

Розрахунок долі досліджуваних, що дали ключову (або правильну) відповідь (індексу ефективності завдання) здійснюється шляхом поділу кількості досліджуваних, що відповіли згідно з ключем (або правильно, коли мова йде про тест здібностей) на їх загальну кількість:

$$i = n_{\text{пр}} / N, \text{ де}$$

i – індекс ефективності завдання;

$n_{\text{пр}}$ – кількість досліджуваних, які відповідали згідно з ключем або дали правильну відповідь;

N – загальна кількість досліджуваних (обсяг вибірки пілотажного дослідження).

Значення індексу ефективності завдання $i < 0,25$ свідчить про те, що надто мало досліджуваних відповідає згідно з ключем (або правильно), значення $i > 0,85$ теж є незадовільним, адже говорить про те, що надто багато досліджуваних відповідає згідно з ключем (правильно). У обох випадках чутливість завдання до відмінностей між досліджуваними слабка, тож завдання під сумнівом.

Отже, індекс ефективності завдання лежить у межах:

$$0,25 < i < 0,75$$

i в середньому для всього тесту становить 0,5. Втім, необхідно також переконатися, що значення $i = 0,5$ не є результатом вибору середніх відповідей у континуумі оцінювання (особливо, коли у тесті використовуються описані вище трихотомічні завдання), адже це теж погано характеризує чутливість завдання і обов'язково позначиться на інших характеристиках ефективності методики. 0,5 – це бажаний показник середньої ефективності завдань тесту, які представляють собою *варіацію різних оцінок*.

Коли мова йде про тести здібностей чи досягнень, то рішення про складність завдань, що увійдуть до остаточної версії опитувальника, приймається з урахуванням мети подальшого тестування. Якщо, наприклад, створюється методика для відбору чи відсіювання досліджуваних (наприклад, визначення питання про те, хто отримуватиме стипендію, або виявлення осіб, які потребують корекційно-педагогічного втручання), то значення оптимального показника складності може переглядатися.

Ще важливішою ознакою якісного завдання є його зв'язок з загальним показником досліджуваного по тесту. Для його встановлення

обчислюється **дискримінативність завдання**, що представляє собою коефіцієнт кореляції кожного завдання з загальним балом усього тесту (або відповідної шкали, якщо тест комплексний). Саме обчислення не представляє собою чогось складного, для дискретних величин (дихотомічних завдань) використовується *коефіцієнт точково-бісеріальної кореляції*, для неперервних величин – *коефіцієнт добутку моментів Пірсона*.

Коефіцієнт точково-бісеріальної кореляції визначається за формулою:

$$r_{pbis} = \frac{\bar{x}_n - \bar{x}}{\sigma_x} \sqrt{\frac{N_n}{N - N_n}}, \text{ де}$$

r_{pbis} – коефіцієнт точково-бісеріальної кореляції (індекс дискримінативності завдання);

\bar{x} – середнє арифметичне всіх індивідуальних оцінок за тестом;

\bar{x}_n – середнє арифметичне оцінок за тестом у тих досліджуваних, які відповіли згідно з ключем (або виконали дане завдання правильно, якщо мова йде про тест здібностей чи досягнень);

σ_x – середньоквадратичне відхилення індивідуальних оцінок за тестом для всієї вибірки;

N_n – кількість досліджуваних, що дали ключову відповідь (виконали завдання правильно);

N – загальна кількість досліджуваних.

П. Кляйн пропонує також дещо інший алгоритм обчислення цього індексу [19], який, втім, дає той же результат. Можливо, деяким дослідникам його варіант може видатися зручнішим.

Формула коефіцієнту кореляції Пірсона виглядає наступним чином:

$$r = \frac{n \sum XY - (\sum X)(\sum Y)}{\sqrt{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}}, \text{ де}$$

r - коефіцієнт кореляції;

X – результат за кожним завданням;

Y - бал (результат) за тестом в цілому (або за відповідною шкалою для комплексного тесту);

n - кількість попарних добутків;

\sum - сума.

Чим ближче отримане значення до одиниці, тим тісніший зв'язок між результатом виконання досліджуваними конкретного завдання і їх результатом за тестом в цілому. Зрозуміло, що індекс обчислюється для кожного завдання. Як правило, до остаточної версії опитувальника входять ті запитання, для яких $r > 0,7$.

У літературі також описано процедуру аналізу завдань із використанням таблиць Фена (або методом порівняння контрастних груп) [19]. Втім, це було доцільно лише за необхідності швидко (і з значною втратою точності) і часто вручну здійснити відповідні обчислення. Сьогодні, за наявності необмежених можливостей комп'ютерного аналізу даних, застосування цього методу втратило актуальність.

Отже, після математичного аналізу ефективності завдань дослідник в змозі відібрати ті пункти, що складуть остаточну версію психологічного тесту. Нагадаємо, що основними критеріями відбору завдань серед рівно-ефективних за математичними показниками буде величина тесту (від якої значною мірою залежить надійність методики) та зміст завдань (варто уважно переконатися, що всі заплановані змістові області теоретичного конструкту адекватно відображено у відібраних завданнях). Якщо перед дослідником постає питання про те, який математичний індекс важливіший, – індекс ефективності завдання чи індекс його дискримінативності – то воно, як правило, вирішується на користь останнього показника, адже саме кореляція завдання з загальним тестовим

показником забезпечує у подальшому надійність за внутрішньою узгодженістю.

Якщо досліднику не вдалося забезпечити набір ефективних завдань запланованого обсягу, йому доведеться критично проаналізувати свої помилки, згенерувати нові завдання, провести нове пілотажне дослідження, знову обчислити згадані показники і т.д.

Коли кількість та зміст ефективних завдань буде відповідати запланованим, наступним етапом роботи дослідника стане формування вибірки для вивчення надійності методики та реалізація відповідного дослідження. Якщо раніше широко застосовувалися допоміжні прийоми швидкого (але менш точного) попереднього обрахунку надійності за внутрішньою узгодженістю (наприклад, обчислення надійності при розбитті тесту на частини) для того, щоб зорієнтуватися, чи є підстави сподіватися з таким набором запитань на достатній показник узгодженості, то сучасне програмне забезпечення дозволяє одразу швидко отримувати коефіцієнти внутрішньої узгодженості запитань з високою точністю. Якщо дослідник виявив, що значення отриманого коефіцієнту надійності перевищує 0,7, то випробування завдань може бути завершеним і дослідник приступає до вивчення валідності й дискримінативності свого тесту. Якщо ж цей показник нижчий, ніж 0,7, розробнику варто знову повернутися до формулювання кращих пунктів, адже тест не може бути кращим, ніж його завдання [19].

Контрольні запитання

У чому полягають відмінності між критеріально-ключовим та факторно-аналітичним принципами конструювання тесту?

Як здійснюється розробка теоретичного конструкту?

Що представляє собою процедура специфікації тесту? Для чого вона потрібна?

Які типи завдань особистісного опитувальника Ви знаєте?

Які установки на відповідь можуть виникати у досліджуваного і як їх уникнути?

Що таке соціально бажані відповіді?

Для чого у деяких тестах потрібна шкала брехні? У яких саме тестах її доцільно використовувати?

Які типи завдань використовуються для розробки тестів інтелекту?

Як формулюються завдання для тестів досягнень?

Якою має бути вибірка пілотажного дослідження завдань тесту?

На основі чого можна робити висновок про ефективність згенерованих завдань?

Які математичні індекси використовують для перевірки ефективності завдань?

На основі який критеріїв формується остаточний набір завдань тесту?

Чи потрібно проводити новий пілотаж тесту, якщо у ньому змінилася інструкція?

Яке значення індексу дискримінативності завдання задовольнить розробника тесту? Про що воно свідчить?

Чому не рекомендують включати у тест завдання, на які, наприклад, більше ніж 75% досліджуваних відповідає згідно з ключем?

Які кроки слідують за перевіркою ефективності запитань?

Чи може статися так, що на пізніших етапах роботи виникне необхідність переформулювати завдання? Якщо так, то з чим це може бути пов'язано?

Поміркуйте...

Що є запорукою об'єктивності особистісних опитувальників?

Якими можуть бути дії дослідника, коли виявилось, що серед сформульованих ним до пілотажу запитань ефективних недостатньо?

Який критерій (індекс ефективності чи індекс дискримінативності завдання) є важливішим для забезпечення надійності тесту?

КОРОТКИЙ ТЕРМІНОЛОГІЧНИЙ СЛОВНИК²

Аналіз завдань – процедура оцінки ефективності завдань у тесті; може включати, зокрема, визначення кореляції завдання з загальним показником за тестом та індексу ефективності завдання.

Валідність – комплексна характеристика тесту, що містить відомості про область досліджуваних явищ і репрезентативність досліджуваної по відношенню до неї діагностичної процедури.

Вибіркова сукупність – репрезентативна частина генеральної сукупності, певна множина: 1) завдань для вимірювання певної якості, із якої формуються вибірки завдань тесту; 2) досліджуваних, із якої формуються репрезентативні вибірки досліджуваних.

Генеральна сукупність – гіпотетична множина елементів, об'єднаних спільною характеристикою: 1) всіх можливих завдань для виявлення вимірюваної якості; 2) всього контингенту досліджуваних.

Генеральний фактор – фактор, що відповідає всім елементам певної множини змінних, тобто має навантаження по всіх змінних.

Груповий тест – методика, придатна для одночасного тестування групи досліджуваних.

Дискримінативність – здатність окремих завдань тесту та тесту в цілому диференціювати досліджуваних відносно максимального і мінімального результатів тесту.

Дистрактори – (у завданнях тестів здібностей і досягнень) – варіанти відповідей, близькі до правильного.

Диференційна валідність – один із компонентів валідності, що відображає здатність методики диференціювати досліджуваних за окремими сферами прояву досліджуваних властивостей.

² Словник укладено переважно за матеріалами довідкового посібника з конструювання тестів Пола Кляйна, повний бібліографічний опис якого подану у переліку використаної літератури.

Змістова валідність – один із типів валідності методики, що характеризує міру репрезентативності змісту завдань тесту вимірюваної сфери психічних властивостей.

Індивідуальний тест – методика, що призначена виключно для індивідуального тестування.

Ключова відповідь – відповідь на завдання (запитання чи твердження) особистісного опитувальника, що вказує на наявність вимірюваної особливості особистості.

Конкурентна валідність – один із способів доведення валідності тесту, який полягає у кореляційному аналізі його результатів з даними, отриманими за допомогою інших тестів, призначених для вимірювання тієї ж змінної.

Коефіцієнт кореляції – показник ступеню взаємозалежності двох змінних.

Конструктна валідність – один з основних типів валідності, що відображає міру репрезентації досліджуваного психологічного конструкту у результатах тесту; чим більше результати відповідають теоретичній гіпотезі про природу вимірюваної змінної, тим вища конструктна валідність тесту.

Кореляційний аналіз – комплекс методів статистичного дослідження взаємозв'язку між змінними, що пов'язані кореляційними відношеннями.

Критеріально-орієнтований тест – тип тестів, що призначені для визначення рівня індивідуальних досягнень відносно деякого критерію (знань, умінь, навичок), а не відносно групових норм.

Надійність – характеристика методики, що відображає точність психодіагностичних вимірювань, а також стійкість результатів тесту до дії сторонніх випадкових факторів.

Надійність за внутрішньою узгодженістю – характеристика тесту, що вказує на ступінь однорідності складу завдань з точки зору вимірюваної якості.

Надійність паралельних форм – характеристика надійності психодіагностичної методики, що визначається за допомогою взаємозамінних форм тесту, які пред'являються одній і тій самій групі досліджуваних.

Норми – показники для різних груп досліджуваних, отримані при стандартизації тесту.

Об'єктивний тест – тест, результати якого можуть бути оцінені об'єктивно (тобто для обчислення показників по тесту та їх інтерпретації не використовуються суб'єктивні судження) і про мету й призначення якого досліджуваний здогадатися не може.

Однорідність тесту – характеристика тесту, що відображає міру відповідності всіх завдань вимірюваній якості чи рисі; визначається як середня кореляція між завданнями тесту.

Особистісні опитувальники – психодіагностичні методики, спрямовані на вимірювання різних особливостей особистості, що їх завдання представлені у вигляді запитань або тверджень.

Очевидна валідність – уявлення про тест, сферу його застосування, результативність і прогностичну цінність, яке виникає у досліджуваного або іншої особи, що не володіє відомостями про характер використання та мету методики.

Показник по тесту – кількісна характеристика вираженості вимірюваної психологічної якості, що отримується за допомогою тесту.

Прогностична валідність – інформація про тест, що характеризує точність і обґрунтованість суджень про вимірювану психологічну якість за її результатом через певний час після вимірювання; чим краще тест може прогнозувати критерій, тим вища його прогностична валідність.

Ретестова надійність – характеристика надійності психодіагностичної методики, отримана шляхом повторного дослідження вибірки за допомогою одного й того ж тесту через деякий час.

Сирий показник – первинний тестовий показник, отриманий досліджуванним за певним тестом, що не підлягав жодній обробці.

Стандартизація тесту – процедура отримання показників по тесту для репрезентативних груп досліджуваних, завдяки якій досягається співставляюваність отриманих результатів у різних досліджуваних.

Стандартизований показник – показник, що відображає міру вираженості вимірюваної психологічної якості у досліджуваного відносно деякої нормативної групи.

Стандартна похибка вимірювання – статистична величина, що відображає точність окремих вимірювань.

Стандартне відхилення характеристика мінливості, що визначається як квадратний корінь з дисперсії.

Репрезентативність – здатність вибіркової сукупності представляти характеристики генеральної сукупності; означає, що з певною похибкою можна вважати, що представлений у вибірковій сукупності розподіл досліджуваних показників відповідає їх реальному розподілу.

Тест - ансамбль стандартизованих, стимулюючих певну форму активності, часто обмежених у часі виконання завдань, результати яких піддаються кількісній (і якісній) оцінці і дозволяють встановити індивідуально-психологічні особливості людини.

Тест атитюдів – група опитувальників, призначена для вимірювання відносного орієнтування особи в одновимірному континуумі (шкалі) атитюду.

Тест досягнень – тест, спрямований на оцінку досягнутого рівня розвитку вмінь, навичок та знань.

Тест інтелекту – сукупність методик, утворених в рамках об'єктивного діагностичного підходу, призначених для вимірювання рівня інтелектуального розвитку.

Тест мотивів – група методик, призначених для діагностики потребнісно-мотиваційної сфери особистості, що вони дозволяють

визначити, на що спрямована активність особи і як здійснюється регуляція динаміки її поведінки.

Тест результативності – тест, що він не має обмежень за часом виконання. Завдання в такому тесті підібрані таким чином, що лише невелика кількість досліджуваних може виконати їх всі правильно, як би довго вони над ними не працювали.

Тест спеціальних здібностей – група тестів, що призначені для вимірювання рівня розвитку окремих аспектів інтелекту і психомоторних функцій, які переважно забезпечують ефективність у конкретних, досить вузьких сферах діяльності.

Тести швидкості – тип тестів, у яких основним показником продуктивності роботи досліджуваних є час виконання (обсяг) завдань тесту.

Установка на відповідь – стереотипи відповідей на завдання тесту, що призводять до порушення валідності; найбільш поширені – установка на згоду та на соціально схвалювані відповіді; найбільше значення мають для особистісних опитувальників.

Фактор – вимірювання, конструктор або вектор, котрий може математично відобразити коваріацію між змінними (факторний аналіз).

Факторний аналіз – розділ багатовимірного статистичного аналізу, що об'єднує математико-статистичні методи зниження розмірності досліджуваної багатовимірної ознаки. У дослідженнях Р. Кеттела і його послідовників факторний аналіз виступає не лише в якості статистичного методу, а й як інструмент пізнання особистості.

Факторні навантаження – значення кореляції факторів із змінними; фактор може бути визначений за факторними навантаженнями.

Факторно-аналітичний тест – тест, що він сконструйований за допомогою факторного аналізу і вимірює певний фактор (або фактори).

Шкала вимірювання – це форма фіксації сукупності ознак досліджуваного об'єкту з упорядкуванням їх в певну числову систему.

Шкала брехні – шкала для тестів особистості (опитувальників) що складається з запитань щодо незначних проступків, властивих більшості людей; використовується для виявлення осіб, які схильні давати соціально схвалювані відповіді.

Шкала Терстоуна – тип шкали для тестів атитюдів, за якою відповіді досліджуваного оцінюються відносно 11-більшої шкали з рівномірно розподіленими на ній завданнями (твердженнями відносно цього атитюду), що попередньо були оцінені експертами.

Шкалювання - операція упорядкування вихідних емпіричних даних у шкальні.

СПИСОК ВИКОРИСТАНОЇ ТА РЕКОМЕНДОВАНОЇ

ЛІТЕРАТУРИ:

1. Айзенк Г. Ю. Интеллект: новый взгляд / Г. Ю. Айзенк // Вопросы психологии. – 1995. - №1. – С.111-131.
2. Айзенк Г. Ю. Структура личности / Г. Ю. Айзенк. – СПб. : Ювента, 1999.
3. Ананьев Б. Г. Комплексное изучение человека и психологическая диагностика / Б. Г. Ананьев // Вопросы психологии. – 1968. - №6. – С.21 – 33.
4. Анастаси А. Дифференциальная психология. Индивидуальные и групповые различия / А.Анастаси. – М. : Апрель-Пресс, 2001.
5. Анастаси А. Психологическое тестирование / А.Анастаси. – М. : Педагогика, 1982. – Кн.1; Кн. 2.
6. Анастаси А. Психологическое тестирование / А.Анастаси, С.Урбина – СПб.: Питер, 2001.
7. Бурлачук Л. Ф. Герман Роршах и проэктивный поход к исследованию личности в советской психологии / Л. Ф. Бурлачук // Проблемы философии. – 1987. – Вып. 74.
8. Бурлачук Л. Ф. Психодиагностика личности / Л. Ф. Бурлачук. – Киев: Здоров'я, 1989.
9. Бурлачук Л. Ф. Психодиагностика: Учебник для вузов / Л. Ф. Бурлачук. – СПб. : Питер, 2006.
10. Бурлачук Л. Ф. Акцентуации личности: что диагностируем? / Л. Ф.Бурлачук, В. Н. Духнович // Вопросы психологии. – 1998. - №2. – С. 136 – 143.
11. Бурлачук Л. Ф. Адаптация опросника для диагностики пяти факторов личности / Л. Ф.Бурлачук, Д. К. Королев // Вопросы психологии. – 2000. - №1. – 126 – 134.
12. Бурлачук Л. Ф. Словарь-справочник по психодиагностике / Л. Ф.Бурлачук, С. М.Морозов. – СПб. : Питер, 2001.

13. Войтко В. И. О некоторых основных понятиях психодиагностики / В. И.Войтко, Ю. З. Гильбух // Вопросы психологии. – 1976. - № 4. – С.18.
14. Гайда В. К. Психологическое тестирование / В. К.Гайда, В. П.Захаров. – Л.,1982.
15. Гуревич К. М. Дифференциальная психология и психодиагностика. Избранные труды / К. М.Гуревич. – СПб. : Питер, 2008.
16. Дружинин В. Н. Психодиагностика общих способностей / В. Н.Дружинин. – М.: Академия, 1996.
17. Дружинин В. Н. Психология общих способностей / В. Н.Дружинин. – СПб. : Питер, 1999.
18. Елисеев О. П. Конструктивная типология и психодиагностика личности / О. П. Елисеев. – Псков, 1994.
19. Кляйн П. Справочное руководство по конструированию тестов / П.Кляйн. – К.,1994.
20. Корсини Р. Психологическая энциклопедия / Р. Корсини, А. Ауэрбах. 2-е изд.– СПб : Питер, 2006. – 1876 с.
21. Кулагин Б. В. Основы профессиональной психодиагностики / Б. В.Кулагин. – Л. : Медицина, 1984.
22. Леонтьев Д. А. Тематический апперцептивный тест / Д. А.Леонтьев. – М. : Смысл, 1999.
23. Либин А. В. Дифференциальная психология: на пересечении европейских, российских и американских традиций / А. В.Либин. – М. : Смысл, 1999.
24. Лукіна Т. Педагогічна діагностика: завдання, методи, інструменти. Модуль. Навчально-методичні матеріали / Т.Лукіна. - К. : Проект "Рівний доступ до якісної освіти", 2007.
25. Морозов С. М. Психометричний контроль та конструювання психодіагностичних тестів / С. М.Морозов. – Київ : Редакційно-видавничий центр «Київський університет», 1997.

26. Мюррей Г. Тематический апперцептивный тест: Руководство / Г.Мюррей / [Электронный ресурс]. – Режим доступа: <http://flogiston.ru/library/tat>.
27. Норакидзе В. Г. Методы исследования характера личности / В. Г.Норакидзе. – Тбилисси : Мецниереба, 1975.
28. Носс И. Н. Психодиагностика / И. Н.Носс. – М.: КСП+, 1999.
29. Собчик Л. Н. Введение в психологию индивидуальности / Л. Н.Собчик. – М.: Институт прикладной психологи, 1998.
30. Общая психодиагностика / Под. ред. А. А. Бодалева, В. В. Столина. – М. : МГУ, 1987.
31. Пиаже Ж. Психология интеллекта / Ж.Пиаже // Избранные психологические труды / Пер.с англ. – М. : Просвещение, 1969.
32. Психологическая диагностика: Учебное пособие / Под ред. К. М. Гуревича, Е. М. Борисовой. – М. : УРАО, 2000.
33. Терлецька Л. Основи психодіагностики: Навчальний посібник / Л.Терлецька. – К. : Главник, 2006. - (серія «Психологічний інструментарій»).
34. Хмель В. Проблеми застосування російськомовного особистісного опитувальника Г. Айзенка (ЕРІ) / В.Хмель, В.Поліщук // Актуальні питання соціально-політичних наук: політологія, психологія, документознавство та інформаційна діяльність. Матеріали ІV міждисциплінарної науково-практичної конференції студентів та молодих вчених. – Острог : Національний університет "Острозька академія", 2015. -С. 278 – 284.
35. Шванцара Й. Диагностика психического развития / Й.Шванцара. – Прага: Авиценум, 1978.
36. Ярошевский М. Г. История психологии / М. Г.Ярошевский. – М. : Мысль, 1976.

**План лекційних та практичних занять з курсу
«Психодіагностика»**

ЛЕКЦІЙНІ ЗАНЯТТЯ

Лекція 1, 2: Тест як основний інструмент психодіагностики (4 год.)

1. Поняття та види тестів у психодіагностиці.
2. Комп'ютеризовані та комп'ютерні тести в психодіагностиці.
3. Адаптація зарубіжних тестів.
4. Особливості постановки психологічного діагнозу.
5. Досліджуваний в психодіагностиці: процесуальні особливості тестування.
6. Характеристика процесу створення нового психодіагностичного інструментарію.

Лекція 3, 4, 5: Основні характеристики ефективних психологічних тестів (6 год.)

1. Особливості різних типів шкал вимірювання.
2. Поняття надійності тесту та її види.
3. Валідність як один з головних показників ефективного тесту.
4. Дискримінативність тесту та його окремих частин.
5. Класична теорія похибок вимірювання.
6. Стандартизація та норми.

Лекція 6, 7: Інтелект людини як предмет психологічного дослідження (4 год.)

1. Поняття інтелекту та його структура.
2. Концепція генерального фактору: вчора та сьогодні.
3. Основні детермінанти інтелектуального статусу людини.
4. Тести для вимірювання інтелекту, їх види.

Лекція 8: Розробка завдань для тестів інтелекту (2 год.)

1. Типи завдань у тестах інтелекту.

2. Поняття емоційного інтелекту людини, тести для його вимірювання.
3. Правила ефективного формулювання завдань для тестів інтелекту.

Лекція 9: Особливості тестів спеціальних здібностей та досягнень (2 год.)

1. Поняття спеціальних здібностей психології.
2. Зміст та типи завдань тестів досягнень.
3. Впорядкування завдань для пілотажу тесту.

Лекція 10-11: Особливості застосування та конструювання особистісних опитувальників (4 год.)

1. Види опитувальників, форми запитань та представлення результатів.
2. Теорія особистості в розробці тестів.
3. Сучасний стан особистісних опитувальників.

Лекція 12: Формулювання завдань особистісних опитувальників (2 год.)

1. Проблеми конструювання особистісних опитувальників.
2. Види завдань особистісного опитувальника.
3. Правила формулювання завдань.

Лекція 13-14: Особливості застосування та розробки проєктивних методик у психодіагностиці (4 год.)

1. Проєкція: від феномену до принципу дослідження.
2. Особливості розробки проєктивних методик.
3. Види проєктивних методик дослідження особистості.
4. Особливості оцінки проєктивних методик.

Лекція 15: Конструювання тестів інших типів (тести мотивів, станів, атитюдів; об'єктивні тести) (2 год.)

1. Конструювання тестів мотивів: опитувальники на виявлення динаміки, інтересів, тест триваючої реакції.
2. Шкали настроїв та станів.

3. Особливості вимірювання атитюдів.

Лекція 16, 17: Особливості відбору та оцінювання завдань тесту (4 год.)

1. Проблеми аналізу завдань тесту.
2. Важливі змінні вибірки досліджуваних при пілотажі опитувальника.
3. Аналіз завдань: процедура та математичні показники.
4. Відбір завдань після їх аналізу.
5. Дії дослідника при невдалій спробі ефективний перелік завдань.

Лекція 18: Визначення надійності як показника якості психологічного тесту (2 год.)

1. Надійність за внутрішньою узгодженістю та її значення при конструюванні тестів.
2. Джерела незадовільної надійності.
3. Вибірка досліджуваних для вивчення надійності методики.
4. Обчислення коефіцієнтів надійності.

Лекція 19: Особливості вивчення й доведення валідності тесту (2 год.)

1. Способи доведення валідності методики.
2. Конкурентна валідність. Правила відбору конкурентних методик.
3. Формулювання гіпотез щодо конструктної валідності особистісного тесту.
4. Математичні методи у перевірці валідності методики.

Лекція 20: Стандартизація тестів (2 год.)

1. Способи отримання репрезентативної нормативної вибірки.
2. Нормування показників. Процентілі, z-показники.
3. Стени та станаїни, приклади їх застосування при стандартизації особистісних опитувальників.

ПРАКТИЧНІ ЗАНТТЯ

Практичне заняття № 1

Тема: Етичні та соціальні аспекти тестування (2 год.)

1. Етичні правила психодіагностичного обстеження та психологічної оцінки. Особистість діагноста.
2. Проблема кваліфікованої інтерпретації та повідомлення результатів тесту.
3. Професійна відповідальність видавців тестів.
4. Тестування особливих популяцій.

Практичне заняття № 2

Тема: Основні характеристики ефективних психологічних тестів (2 год.)

1. Особливості різних типів шкал вимірювання.
2. Поняття надійності тесту та її види.
3. Валідність як один з головних показників ефективного тесту.
4. Дискримінативність тесту та його окремих частин.
5. Класична теорія похибок вимірювання.
6. Стандартизація та норми.

Практичне заняття № 3

Тема: Специфікація тесту як результат вивчення вимірюваного конструкту (2 год.)

Студенти розпочинають роботу над створенням власного тесту.

Передбачається групова (по 3-4 людини в групі) робота. На цьому практичному занятті студентам даються наступні **завдання**:

- сформулювати мету майбутнього тесту;
- розробити специфікацію майбутнього тесту.

Попереднє «домашнє» самостійне завдання: теоретично дослідити та охарактеризувати той психологічний феномен, що буде вимірюватися майбутнім тестом.

Практичне заняття № 4

Тема: Модульний контроль з характеристик ефективних психологічних тестів (2 год.)

Практичне заняття № 5

Тема: Інтелект людини як предмет психологічного дослідження (2 год.)

1. Психологічні проблеми тестування здібностей.
2. Концепція генерального фактору: вчора та сьогодні.
3. Основні детермінанти інтелектуального статусу людини.

Завдання для самостійного опрацювання: Психологічні проблеми тестування здібностей. Тести для специфічних популяцій.

Практичне заняття № 6

Тема: Розробка завдань для тестів здібностей та досягнень (2 год.)

1. Поняття спеціальних здібностей та досягнень в психології.
2. Зміст завдань тестів спеціальних здібностей та досягнень.
3. Типи завдань в тестах спеціальних здібностей та досягнень.
4. Впорядкування завдань для пілотажу тесту.

Практичне заняття № 7

Тема: Модульний контроль із створення надійних тестів інтелекту та здібностей (2 год.)

Практичне заняття № 8

Тема: Проблема достовірності особистісних опитувальників (2 год.)

1. Фактори, що детермінують відповіді на запитання: фальсифікація та установки на відповіді.
2. Види установок на відповідь та способи усунення їх впливу на результати.

3. Способи підвищення надійності особистісного опитувальника за допомогою формулювання запитань.

Практичне заняття № 9

Тема: Формулювання запитань особистісних опитувальників (2 год.)

1. Правила формулювання завдань особистісних опитувальників.
2. Типи завдань особистісного опитувальника.
3. Особливості змісту запитань.

Практичне заняття № 10

Тема: Особливості застосування та розробки проєктивних методик у психодіагностиці (2 год.)

1. Проекція як психологічний феномен.
2. Проєктивна гіпотеза та проєктивні методики, їх види.
3. Роль стимулу в проєктивних методиках.
4. Проблема об'єктивності проєктивних методик.

Практичне заняття №11

Тема: Проблема розуміння запитань та мінливості відповідей особистісного опитувальника (2 год.)

1. Оцінка розуміння респондентом запитання.
2. Результати дослідження зв'язку між розумінням запитань та мінливістю відповідей.
3. Психометричний парадокс.

Практичне заняття №12

Тема: Модульний контроль із створення особистісних опитувальників (2 год.)

Практичне заняття № 13

Тема: Проведення пілотажу самостійно розробленого тесту

Завдання: Здійснити пілотаж попередньо розробленого тесту на запланованій вибірці.

Практичне заняття № 14

Тема: Особливості відбору та оцінювання завдань тесту (2 год.)

Завдання: Оцінити завдання тесту за відповідними критеріями та здійснити подальший відбір завдань або їх переформулювання.

Практичне заняття № 15

Тема: Визначення надійності як показника якості психологічного тесту (2 год.)

Завдання: Обчислити показники надійності створеного психологічного інструменту.

Практичне заняття №16

Тема: Особливості визначення валідності тесту (2 год.)

Завдання: Здійснити кореляційний аналіз і проаналізувати конкурентну валідність створеного психологічного інструменту.

Практичне заняття №17

Тема: Стандартизація тестів (2 год.)

1. Способи отримання репрезентативної нормативної вибірки.
2. Нормування показників. Процентілі, z-показники.
3. Стени та станайни, приклади їх застосування при стандартизації особистісних опитувальників.

Завдання: Підготувати план стандартизації створеної методики, що він включає опис вибірки стандартизації, процедури стандартизації, методів нормування показників та розрахунок необхідного часу й коштів для здійснення стандартизації.

**Рекомендації щодо структури звітної роботи з конструювання
тесту**

(рекомендації стосуються структури роботи та орієнтовного змісту її складових частин; формальні вимоги щодо оформлення повністю відповідають загальноприйнятим нормам, визначеним для наукових робіт; вимоги до змісту не строгі: можна доповнювати, розширювати тією інформацією, яку ви вважаєте необхідною, але звужувати заданий перелік змістовних аспектів роботи не рекомендується)

1. **Титульна сторінка** (інформація про виконавців робіт (якщо кожен виконавець відповідав за свій етап роботи, це має бути відповідно вказано), назву тесту і т.д., що оформляється згідно загальних вимог до написання наукових робіт).

2. **Вступ** (мета створення методики, очікуваний результат: психічні (психологічні) особливості людини, що стануть відомими після використання тесту; актуальність вибору саме цієї методики, основні методологічні засади створення тесту, характеристики популяції, для якої призначений тест, припущення щодо конструктивної валідності методики).

3. **Опис теоретичного конструкту** (вимірюване явище як психологічний феномен, його структура, операціоналізація основних складових досліджуваного феномену (т-то обґрунтування їх можливих емпіричних проявів), основні параметри (змінні) подальшого вимірювання).

4. **Специфікація тесту** (може бути подана у вигляді таблиці з зазначенням змістовних областей досліджуваного конструкту та їх маніфестацій (проявів), а також з обґрунтуванням відповідної кількості запитань).

5. **Набори запитань для пілотажу тесту** (у випадку наявності кількох параметрів вимірювання, запитання подаються тут окремо для кожної змінної).

6. **Опис вибірки для пілотажу тесту** (основні параметри вибірки, їх значення для подальших процедур аналізу завдань, визначення валідності, надійності тощо, особливості формування вибіркової сукупності).

7. **Версія опитувальника для проведення пілотажного дослідження тесту** з відповідними інструкціями та ключами.

8. **Оцінювання та відбір завдань** (процесуальні особливості та результати обчислення індексів ефективності та дискримінативності завдань, їх значення для подальшого відбору/відсіювання пунктів опитувальника; обґрунтування вибору запитань для остаточної версії опитувальника; аналіз неефективних завдань: ймовірні причини їх низької ефективності та можливі шляхи вдосконалення їх у майбутньому)

9. **Остаточна версія опитувальника** з відповідною інструкцією, ключами, бланком для відповідей, інтерпретацією результатів.

10. **Визначення надійності тесту** (обґрунтування використання тих чи інших коефіцієнтів для її обчислення; результати за різними видами надійності, їх значення для визначення ефективності створеного тесту).

11. **Визначення дискримінативності тесту** та значення цієї характеристики для прийняття рішення щодо ефективності методики.

12. **Обґрунтування валідності тесту** (обґрунтування обраних для аналізу видів валідності; «конструктні» гіпотези та результати їх перевірки; висновки щодо валідності створеного тесту).

13. **Висновки** (основні результати процесу створення нового психодіагностичного інструментарію: аналіз досягнутого та нереалізованого, труднощі, що виникали в роботі над створенням тесту, перспективи подальшої роботи над цією методикою).

14. **Список використаної літератури.**

15. **Додатки.**

**Зразок звіту про апробацію психодіагностичної розробки
студентів На УОА**

АДАПТАЦІ ОСОБИСТІСНОГО ОПИТУВАЛЬНИКА Г. АЙЗЕНКА (ЕРІ)

*Підготували студентки групи ПС-31 Хмель Валентина, Поліщук Вікторія,
Лехкобит Марія, Труш Аліна (Острог-2014)*

Вступ

Психодіагностичне вивчення психологічних явищ нейротизму, екстраверсії-інтроверсії, психотизму, типів темпераменту є актуальною проблемою для сучасної психології. Особистісний опитувальник Айзенка ґрунтується на основі диспозиційного напрямку у вивченні особистості, який базується на двох загальних ідеях. Перша полягає в тому, що люди володіють широким набором схильностей реагувати певним чином у різних ситуаціях, тобто набором рис особистості. Це означає, що люди демонструють певну сталість у своїх вчинках, думках і емоціях, незалежно від перебігу часу, подій і життєвого досвіду. Друга основна ідея вказує на, що немає цілком схожих один на одного людей. Особистість володіє певним набором характерних рис, що відрізняють її від інших індивідів. Знаючи якими рисами наділена особистість, можна робити прогнози про її поведінку в типових ситуаціях, реакції на подразники зовнішнього і внутрішнього світу. Також за результатами тесту можна визначити емоційні реакції людини, тобто наскільки в них врівноважені чи не врівноважені процеси збудження та гальмування. Врівноважена людина здатна контролювати свої емоційні прояви, володіти ними. Саме це і визначає рівень емоційної стабільності людини. Виходячи з цього, люди поділяються на емоційно стабільних і емоційно нестабільних, які відрізняються рівнем контролю над своїми емоційними проявами.

Особистісний опитувальник Айзенка розрахований на осіб з 16 років. Окремо розроблені версії опитувальника для дітей та підлітків. Немає ніяких обмежень щодо статі, раси, соціального статусу, культурних чи релігійних поглядів. Проте на україномовних досліджуваних застосовується російськомовна версія опитувальника, що може викривлювати результати по тесту, створювати незручності у розумінні питань та надавати неправдиву інформацію про респондентів.

У більшість україномовних досліджуваних можуть виникнути труднощі у нерозумінні як деяких слів так і цілого питання. Це пов'язано з тим, що освітня система в Україні не передбачає обов'язкове вивчення російської мови та не має статусу державної мови в нашій країні. Тому для кращого розуміння питань україномовними досліджуваними та отримання достовірних результатів, необхідно адаптувати російськомовну версію Особистісного опитувальника Айзенка на українську мову.

Щодо конкурентної валідності ми можемо сформулювати такі гіпотези:

1. Шкала «нейротизму» особистісний опитувальник Г.Айзенка (EPI) повинна достатньо корелювати з шкалами проявів тривоги Ж.Тейлора (MAS).

2. Шкала «екстраверсії» особистісного опитувальника Г.Айзенка (EPI) повинна не корелювати з шкалами проявів тривоги Ж.Тейлора (MAS).

3. Кореляція шкали локусу контролю Дж. Б. Роттера не повинна корелювати з результатами особистісного опитувальника Г.Айзенка (EPI).

Опис теоретичного конструкту у зразку до посібника не включено

Специфікація тесту

При створенні змістовних областей конструкту та їх маніфестації (прояви) було взято за основу теорію рис особистості за Айзенком. Він виділяє 32 основних рис особистості, які відповідають чотирьом типам темпераменту. Відповідно до цієї теорії є чотири типи особистості: стабільний екстраверт, стабільний інтроверт, нестабільний екстраверт і нестабільний інтроверт. Кожен із 4 типів особистості включає 8 видів особистісних рис, які відповідають певній особистості.

При виділенні маніфестацій ми не враховували шкалу брехні, тому на прояви розподілялися 48 питань по двох ортогональних шкалах: екстраверсія та нейротизм. При аналізі вимірювального конструкту питання опитувальника співставлялися із переліком рис особистості за Айзенком, оскільки кожне питання передбачає вимірювання ступінь прояву певної риси.

Ми виділили такі маніфестації: чуттєвість, спокійність, комунікативність, оптимізм, імпульсивність, схильність до роздумів, тривожність.

Прояви

Маніфестації	Кількість завдань	
	Екстраверсія-інтроверсія (50%)	Нейротизм (50%)
Чуттєвість (10 %)	2 (37, 50)	3 (26, 45, 49)
Спокійність (6%)	3 (1, 5, 41)	3
Комунікативність (25%)	9 (15, 17, 20, 29, 32, 44, 46, 51, 53)	3 (2, 11, 19)

Оптимізм (життєрадісність) (8%)	4 (3, 25, 27, 56)	4
Імпульсивність (збудливість) 21%)	6 (8, 10, 13, 22, 34, 39)	4 (16, 33, 10 40 47)
Схильність до роздумів (13 %)		6 (4, 14, 21, 28, 31, 52)
Тривожність (17%)		8 (7, 9, 23, 8 35, 38, 43, 55, 57)
Кількість завдань	24	24 48

Опис вибірки для пілотажу тесту

Обсяг нашої вибірки для пілотажного дослідження складав 40 респондентів. Вибірка була вибрана стихійно, основну частку якої склали студенти Національного університету «Острозька академія». Пілотажне дослідження було проведено на вибірці осіб від 17 до 43 років. Середній вік досліджуваних складав ($M = 21,9$). Опитувальник розрахований на осіб від 16 років і більше, тому вибірку склали респонденти із різних вікових груп.

Переклад методики

Російська версія	Адаптована українським перекладом
1. Часто ли вы испытываете тягу к новым впечатлениям, к тому, чтобы отвлечься, испытывать сильные ощущения?	95.Чи виникає у Вас потяг до нових вражень, бажання відволіктися, отримати нові сильні відчуття? 5.Чи потрібно Вам, щоб відволіктися, пережити сильні враження?
2. Часто ли вы чувствуете, что нуждаетесь в друзьях, которые могут вас понять, ободрить или посочувствовать?	1.Чи відчуваєте Ви потребу в спілкуванні з друзями, які можуть Вас зрозуміти, підбадьорити та поспівчувати? 78.Чи часто Ви відчуваєте бажання в спілкуванні та підтримці?
3. Считаете ли вы себя беззаботным человеком?	38.Ви вважаєте себе безтурботною людиною? 54.Ви безтурботні?
4. Очень ли трудно вам отказываться от своих намерений?	33.Вам складно відмовитися від своїх намірів? 24. Вам важко відмовитися від своїх намірів?
5. Обдумываете ли вы свои дела не спеша и предпочитаете ли подождать, прежде чем действовать?	77.Чи обдумуєте Ви свої справи не поспішаючи, надаєте перевагу тому, щоб почекати перш ніж діяти?

	84.Перед тим як щось робити чи ретельно Ви обмірковуєте можливі наслідки, кінцевий результат?
6. Всегда ли вы сдерживаете свои обещания, даже если это вам невыгодно?	104.Чи завжди Ви дотримуетесь своїх обіцянок, навіть якщо це Вам не вигідно?
7. Часто ли у вас бывают спады и подъемы настроения?	93.Чи часто у Вас бувають переміни настрою? 3.Чи часто у Вас бувають спади та підйоми настрою?
8. Быстро ли вы обычно действуете и говорите, не тратите ли много времени на обдумывание?	103. Ви імпульсивні та швидкі у своїх діях та рухах? 47.Зазвичай Ви дієте та говорите швидко, не витрачаючи багато часу на роздуми?
9. Возникло ли у вас когда-нибудь чувство, что вы несчастны, хотя никакой серьезной причины для этого не было?	79.У Вас виникало почуття, що ви нещасний без ніякої на те причини? 14.Чи буває у Вас відчуття, що Вас переслідують невдачі?
10. Верно ли, что на спор вы способны решиться на все?	36.Ви можете зробити будь-що побившись об заклад? 96.Чи притаманно Вам діяти спонтанно?
11. Смущаетесь ли вы, когда хотите познакомиться с человеком противоположного пола, который вам симпатичен?	81.Ви соромитесь, коли хочете познайомитися із симпатичною людиною протилежної статі? 53.Ви сором'язлива людина?
12. Бывает ли когда-нибудь, что, разозлившись, вы выходите из себя?	27.Розізлившись Ви втрачаєте самоконтроль?
13. Часто ли бывает, что вы действуете необдуманно, под влиянием момента?	16.В неочікуваній ситуації Ви можете діяти необдуманно? 98.Буває, що Ви дієте необдуманно під впливом ситуації?
14. Часто ли вас беспокоят мысли о том, что вам не следовало чего-либо делать или говорить?	68.Чи часто Ви жалієте про свої вчинки? 83.Буває так, що ви жалкуєте про сказане чи зроблене?
15. Предпочитаете ли вы чтение книг встречам с людьми?	40.Ви надаєте перевагу читанню книг, ніж зустрічами з людьми? 41.Ви надаєте перевагу самотійним прогулянкам, ніж в компанії?
16. Верно ли, что вас легко задеть?	30.Вас легко образити? 70. Ви людина, яка сприймає все близько до серця?
17. Любите ли вы часто бывать в компании?	69.Чи подобається бути Вам в компанії друзів? 67.Чи комфортно Вам перебувати серед людей?
18. Бывают ли иногда у вас такие мысли, которыми вам не хотелось бы делиться с другими людьми?	57.Чи виникають у Вас інколи думки, якими ви не хочете ділитися з іншими людьми?

19. Верно ли, что иногда вы настолько полны энергии, что все горит в руках, а иногда чувствуете сильную вялость?	64.Чи характерні Вам зміни активності від енергійності до слабкості? 34.Чи правда, що інколи вас настільки переповнює енергія, що у Вас все горить в руках, а інколи відчуваєте сильну слабкість?
20. Стараетесь ли вы ограничить круг своих знакомых небольшим числом самых близких друзей?	35.У Вас обмежене коло друзів? 17.Ви намагаєтесь обмежити коло своїх знайомих невеликою кількістю близьких людей?
21. Много ли вы мечтаете?	43.Ви багато мрієте? 101.Чи часто Ви роздумуєте про своє життя?
22. Когда на вас кричат, отвечаете ли вы тем же?	51.Коли на Вас підвищують голос, ви робите те ж саме у відповідь? 20. Вам властиво діяти за принципом «як ти мені, так я тобі»?
23. Считаете ли вы все свои привычки хорошими?	21.Чи вважаєте Ви свої звички хорошими? 68.Ви вважаєте свої звички корисними?
24. Часто ли у вас появляется чувство, что вы в чем-то виноваты?	15.Чи часто у вас з'являється почуття провини?
25. Способны ли вы иногда дать волю своим чувствам и беззаботно развлечься с веселой компанией?	58.Чи можете Ви дати волю почуттям, щоб безтурботно відпочити у веселій компанії? 73.Чи можете ви забути про свої турботи і відірватися на повну?
26. Можно ли сказать, что нервы у вас часто бывают натянуты до предела?	85.Чи часто Ви буваєте на межі нервового зриву? 97.Чи часто Ви відчуваєте напругу?
27. Слывете ли вы за человека живого и веселого?	80.Ви схожі на жваву і веселу людину? 25. Вас вважають веселою та жвавою людиною?
28. После того как дело сделано, часто ли вы мысленно возвращаетесь к нему и думаете, что могли бы сделать лучше?	88.Після того як справа зроблена, Ви думаете про те, що могли зробити це краще? 42.Чи обдумаєте ви справи, які вже давно були зроблені?
29. Чувствуете ли вы себя беспокойно, находясь в большой компании?	90.Ви почуваєте себе неспокійно, перебуваючи у великому колі знайомих? 6.Вам буває некомфортно у великій компанії?
30. Бывает ли, что вы передаете слухи?	19.Чи буває, що ви пліткуєте?
31. Бывает ли, что вам не спится из-за того, что в голову лезут разные мысли?	92.Чи буває у вас безсоння через різні думки? 13.Ви часто не можете заснути, через те, що Вас щось турбує?
32. Что вы предпочитаете, если хотите узнать что-либо: найти в книге или спросить у друзей?	10.Відповідь на питання, яке вас цікавить, Ви краще пошукаєте в книзі, ніж спитаєте в друзів? 87.Відповідаючи на питання, Ви надасте перевагу самостійно знайти відповідь, ніж перепитувати у знайомих?
33. Бывают ли у вас сильные сердцебиения?	100.Чи властива Вам надмірна пітливість?

	долонь? 31. Чи буває у Вас прискорене серцебиття?
34. Нравится ли вам работа, требующая сосредоточения?	18. Чи подобається Вам робота, яка потребує тривалого зосередження? 82. Ви надаєте перевагу дрібній клопіткій роботі, яка вимагає надмірної концентрації уваги?
35. Бывают ли у вас приступы дрожи?	8. Чи буває так, що у Вас тремтить все тіло? 94. Чи знайоме Вам відчуття кому в горлі?
36. Всегда ли вы говорите только правду?	44. Ви завжди говорите тільки правду?
37. Бывает ли вам неприятно находиться в компании, где все подшучивают друг над другом?	9. Вам неприємно знаходитися в кімнаті, де кепкують один над одним? 4. Ви відчуваєте дискомфорт, коли бачите, що над кимось насміхаються?
38. Раздражительны ли вы?	55. Чи легко вас роздратовати? 72. Коли Ви роздратовані Вам хочеться кричати?
39. Нравится ли вам работа, требующая быстрого действия?	52. Чи подобається Вам робота, яка потребує швидких дій? 86. Вам подобається виконувати роботу, яка вимагає швидкої реакції?
40. Верно ли, что вам часто не дают покоя мысли о разных неприятностях и «ужасах», которые могли бы произойти, хотя все кончилось благополучно?	48. Чи правда, що коли все закінчується благополучно, Вас не покидають думки про різні неприємності, які могли трапитися? 76. Після розмови з людиною, ви задумаетесь над тим, що могли поводити себе по-іншому?
41. Верно ли, что вы неторопливы в движениях и несколько медлительны?	74. Чи правда, що Ваші рухи є дещо повільними? 45. Чи важко Вам робити декілька справ одночасно?
42. Опаздывали ли вы когда-нибудь на работу или на встречу с кем-либо?	50. Чи запізнювалися ви колись на зустріч чи на роботу?
43. Часто ли вам снятся кошмары?	7. Чи часто Вам сняться кошмари? 26. Чи буває так, що після перегляду фільму жахів Ви не можете заснути?
44. Верно ли, что вы так любите поговорить, что не упускаете любого удобного случая побеседовать с новым человеком?	12. Чи любите ви розповідати про історії зі свого життя? 99. Ви настільки любите поговорити, що не втрачаєте нагоди поспілкуватися з новою людиною?
45. Беспокоят ли вас какие-нибудь боли?	60. Час від часу Вас турбують різні болі? 2. Чи відчуваєте Ви дискомфорт у животі, коли Ви знервовані?
46. Огорчились бы вы, если бы долго не могли видетсья со своими друзьями?	29. Коли Ви їдете на відпочинок, то сумуете з своїми рідними та друзями? 63. Ви б засмутилися, якби довго не могли бачитися зі своїми друзями?

47. Можете ли вы назвать себя нервным человеком?	37.Чи вважаєте ви себе нервовою людиною? 61.Чи нервуєте Ви, коли хтось бере Ваші речі без дозволу?
48. Есть ли среди ваших знакомых такие, которые вам явно не нравятся?	65.Чи є серед ваших знайомих люди, які Вам не подобаються?
49. Могли бы вы сказать, что вы уверенный в себе человек?	22.Ви впевнені у своїх у своїх діях? 23.Ви впевнена в собі людина?
50. Легко ли вас задевает критика ваших недостатков или вашей работы?	28.Чи буває так, що коли вказують на Ваші недоліки, то у Вас пропадає бажання щось робити далі? 89.Чи легко Вас образить, якщо вказати на Ваші недоліки?
51. Трудно ли вам получить настоящее удовольствие от мероприятий, в которых участвует много народа?	49.Чи отримуєте ви задоволення працюючи над груповим завданням? 71. Чи важко Вам отримати задоволення від заходів на яких є багато людей?
52. Беспокоит ли вас чувство, что вы чем-то хуже других?	62.Чи турбує вас думка, що ви є гірші чимось за інших? 59.Ви часто себе порівнюєте з іншими людьми?
53. Сумели бы вы внести оживление в скучную компанию?	91.Чи здатні Ви організувати веселу вечірку? 102.Чи здатні Ви внести пожвавлення в сумну компанію?
54. Бывает ли, что вы говорите о вещах, в которых совсем не разбираетесь?	46.Чи буває, що Ви говорите про речі в яких не тямите?
55. Беспокоитесь ли вы о своем здоровье?	39.Чи намагаєтесь Ви вести здоровий спосіб життя? 105.Ви турбуєтесь про своє здоров'я?
56. Любите ли вы подшутить над другими?	75.Ви часто кепкуєте над іншими? 11.Чи подобається Вам веселити своїх друзів?
57. Страдаете ли вы бессонницей?	32.Ви страждаєте безсонням? 56.Чи можете ви розслабитися після напруженого дня?

Оцінка експертів

Перед тим як відібрати питання для пілотажного дослідження ми запропонували для 80 респондентів перекласти по 14 питання методики з російської мови на українську, для того щоб побачити як вони розуміють текст опитувальника. У вибірку увійшли респонденти віком від 18-43 років. Потім до перевірки якості перекладу було залучено 4 експерти (студенти котрі навчаються на спеціальності «українська філологія» 4 курсу). Вони оцінювали якість перекладу за шкалою від 0-5. В подальшому питання які були найбільш якісно перекладені увійшли до версії опитувальника для пілотажного дослідження (Додаток Б).

Набори запитань для пілотажу

Екстраверсія-інтроверсія

1. Ви відчуваєте дискомфорт, коли бачите, що над кимось насміхаються?
2. Чи потрібно Вам, щоб відволіктися, пережити сильні враження?
3. Вам неприємно знаходитися в кімнаті, де кепкують один над одним?
4. Відповідь на питання, яке вас цікавить, Ви краще пошукаєте в книзі, ніж спитаєте в друзів?
5. Чи подобається Вам веселити своїх друзів?
6. Чи любите ви розповідати про історії зі свого життя?
7. Буває, що Ви дієте необдуманно під впливом ситуації?
8. Ви намагаєтеся обмежити коло своїх знайомих невеликою кількістю близьких людей?
9. Чи подобається Вам робота, яка потребує тривалого зосередження?
10. Вам властиво діяти за принципом «як ти мені, так я тобі»?
11. Ви впевнені у своїх у своїх діях?
12. Ви впевнена в собі людина?
13. Вас вважають веселою та жвавою людиною?
14. Коли Ви їдете на відпочинок, то сумуєте за своїми рідними та друзями?
15. У Вас обмежене коло друзів?
16. Ви можете зробити будь-що побившись об заклад?
17. Ви вважаєте себе безтурботною людиною?
18. Ви надаєте перевагу читанню книг, ніж зустрічами з людьми?
19. Ви надаєте перевагу самотійним прогулянкам, ніж в компанії?
20. Чи важко Вам робити декілька справ одночасно?
21. Ви імпульсивні та швидкі у своїх діях та рухах?
22. Чи отримуєте ви задоволення працюючи над груповим завданням?
23. Коли на Вас підвищують голос, ви робите те ж саме у відповідь?
24. Чи подобається Вам робота, яка потребує швидких дій?
25. Ви безтурботні?
26. Чи можете Ви дати волю почуттям, щоб безтурботно відпочити у веселій компанії?
27. Ви б засмутилися, якби довго не могли б бачитися зі своїми друзями?
28. Чи комфортно Вам перебувати серед людей?
29. Чи подобається бути Вам в компанії друзів?
30. Чи важко Вам отримати задоволення від заходів на яких є багато людей?
31. Чи можете ви забути про свої турботи і відірватися на повну?
32. Чи правда, що Ваші рухи є дещо повільними?
33. Ви часто кепкуєте над іншими?
34. Перед тим як щось робити чи ретельно Ви обмірковуєте можливі наслідки, кінцевий результат?
35. Ви схожі на жваву і веселу людину?
36. Ви надаєте перевагу дрібній клопіткій роботі, яка вимагає надмірної концентрації уваги?

37. Чи обдумуєте Ви свої справи не поспішаючи, надаєте перевагу тому, щоб почекати перш ніж діяти?
38. Вам подобається виконувати роботу, яка вимагає швидкої реакції?
39. Відповідаючи на питання, Ви надаєте перевагу самостійно знайти відповідь, ніж перепитувати у знайомих?
40. Ви відчуваєте себе неспокійно, перебуваючи у великому колі знайомих?
41. Чи здатні Ви організувати веселу вечірку?
42. Чи виникає у Вас потяг до нових вражень, бажання відволіктися, отримати нові сильні відчуття?
43. Чи притаманно Вам діяти спонтанно?
44. В неочікуваній ситуації Ви можете діяти необдуманно?
45. Ви настільки любите поговорити, що не втрачаєте нагоди поспілкуватися з новою людиною?
46. Чи здатні Ви внести пожвавлення в галасливу компанію?
47. Зазвичай Ви дієте та говорите швидко, не витрачаючи багато часу на роздуми?
48. Чи відчуваєте Ви потребу в спілкуванні з друзями, які можуть Вас зрозуміти, підбадьорити та поспівчувати?

Нейротизм

1. Ви турбуєтеся про своє здоров'я?
2. Чи властива Вам надмірна пітливість долонь?
3. Чи часто Ви роздумуєте про своє життя?
4. Чи часто Ви відчуваєте напругу?
5. Чи буває у вас безсоння через різні думки?
6. Чи часто у Вас бувають перемини настрою?
7. Чи знайоме Вам відчуття кому в горлі?
8. Після того як справа зроблена, Ви думаєте про те, що могли зробити це краще?
9. Чи легко Вас образить, якщо вказати на Ваші недоліки?
10. Чи часто Ви буваєте на межі нервового зриву?
11. Буває так, що ви жалкуєте про сказане чи зроблене?
12. Ви соромитеся, коли хочете познайомитися із симпатичною людиною протилежної статі?
13. Чи часто Ви відчуваєте бажання в спілкуванні та підтримці?
14. У Вас виникало почуття, що ви нещасний без ніякої на те причини?
15. Після розмови з людиною, ви задумаетесь над тим, що могли поводити себе по-іншому?
16. Коли Ви роздратовані Вам хочеться кричати?
17. Ви людина, яка сприймає все близько до серця?
18. Чи часто Ви жалієте про свої вчинки?
19. Ви вважаєте свої звички корисними?
20. Чи характерні Вам зміни активності від енергійності до слабкості?
21. Ви часто себе порівнюєте з іншими людьми?
22. Час від часу Вас турбують різні болі?

23. Чи нервуєте Ви, коли хтось бере Ваші речі без дозволу?
24. Чи турбує вас думка, що ви є гірші чимось за інших?
25. Чи легко вас роздратувати?
26. Чи можете ви розслабитися після напруженого дня?
27. Ви сором'язлива людина?
28. Чи правда, що коли все закінчується благополучно, Вас не покидають думки про різні неприємності, які могли трапитися?
29. Чи обдумаете ви справи, які вже давно були зроблені?
30. Ви багато мрієте?
31. Чи намагаєтесь Ви вести здоровий спосіб життя?
32. Чи вважаєте ви себе нервовою людиною?
33. Вас легко образити?
34. Чи буває у Вас прискорене серцебиття?
35. Ви страждаєте безсонням?
36. Вам складно відмовитися від своїх намірів?
37. Чи правда, що інколи вас настільки переповнює енергія, що у Вас все горить в руках, а інколи відчуваєте сильну слабкість?
38. Чи буває так, що після перегляду фільму жахів Ви не можете заснути?
39. Чи буває так, що коли вказують на Ваші недоліки, то у Вас пропадає бажання щось робити далі?
40. Вам важко відмовитися від своїх намірів?
41. Чи вважаєте Ви свої звички хорошими?
42. Ви часто не можете заснути, через те, що Вас щось турбує?
43. Чи буває у Вас відчуття, що Вас переслідують невдачі?
44. Вам буває некомфортно у великій компанії?
45. Чи часто Вам сняться кошмари?
46. Чи буває так, що у Вас тремтить все тіло?
47. Чи відчуваєте Ви дискомфорт у животі, коли ви знервовані?
48. Чи часто у Вас бувають спади та підйоми настрою?

Шкала брехні

1. Ви виникають у Вас інколи думки, якими ви не хочете ділитися з іншими людьми?
2. Чи часто у вас з'являється почуття провини?
3. Чи буває, що ви пліткуєте?
4. Розізлившись Ви втрачаєте самоконтроль?
5. Ви завжди говорите тільки правду?
6. Чи буває, що Ви говорите про речі в яких не тямите?
7. Чи запізнювалися ви колись на зустріч чи на роботу?
8. Чи є серед ваших знайомих люди, які Вам не подобаються?
9. Чи завжди Ви дотримуетесь своїх обіцянок, навіть якщо це Вам не вигідно?

Інструкція. Вам пропонується 105 питань. Питання спрямовані на визначення Вашої поведінки в повсякденному житті. Спробуйте уявити типові ситуації і дайте першу

відповідь, яка прийде вам в голову. Відповідайте швидко і точно. Пам'ятайте, що немає «хороших» або «поганих» відповідей. Якщо ви згодні з твердженням, поставте поруч з його номером знак (+) так, якщо ні, то знак (-). Анонімність гарантована. Бажаємо успіхів!

Оцінювання та відбір завдань

Аналіз завдань за результатом пілотажного дослідження є досить важливим моментом, оскільки відбір найефективніших завдань дасть можливість створити більш досконалий опитувальник.

Для оцінки запитань нами були використані індекс ефективності завдань, який дозволяє визначити долю досліджуваних, що відповіли згідно з ключем та коефіцієнт дискримінативності завдань, а саме точечно-бисеріальний коефіцієнт кореляції для дихотомічних шкал, що показує кореляцію кожного завдання із загальним показником.

Оскільки наш тест складається із двох областей (нейротизм, екстра-, інтроверсія) то відповідно і завдання оцінювалися окремо по кожній із них

В області нейротизм було відсіяно 24 запитань - завдання, індекс ефективності яких був меншим від 0,25 (мала кількість досліджуваних відповіли згідно з ключем) і більшим 0,75 (завдання вважається не ефективним, адже велика кількість досліджуваних відповіли згідно з ключем), а також коефіцієнт дискримінації був нижче за 0,2 – недостатня кореляція питання з загальним показником по тесту.

Більшість з відкинутих нами запитань мали високий показник ефективності, але низький коефіцієнт дискримінативності: №11, 12, 40, 69, 82. Так наприклад №69 «Чи подобається Вам бути в компанії друзів?» мав індекс ефективності 0,90, а коефіцієнт дискримінативності лише 0,08 – відповідно можна пояснити такий результат загальністю запитання, відсутністю ситуативності.

Проте питання №4, 5, 6, 35, 38 показали низький показник ефективності та низький коефіцієнт дискримінативності. Так наприклад питання №4 « Ви відчуваєте дискомфорт, коли бачите, що над кимось насміхаються?» мав такі показники: індекс ефективності 0,20, а коефіцієнт дискримінативності лише 0,01, які в свою чергу зумовлені саме ситуативністю питання.

В області екстра-, інтроверсія було відсіяно 30 запитань, проте серед них до остаточного тесту були відібрані запитання №42, 94, 101, індекс ефективності яких був вище 0,75 проте коефіцієнт дискримінативності був значимий. Наприклад №42 $I_{ef}=0,78$, коефіцієнт дискримінативності=0,53.

Більшість з відкинутих нами запитань мали високий показник ефективності, але низький коефіцієнт дискримінативності, наприклад питання №1 «Чи відчуваєте Ви потребу в спілкуванні з друзями, які можуть Вас зрозуміти, підбадьорити та поспівчувати?» №31 «Чи буває у Вас прискорене серцебиття?», №61 «Чинервуєте Ви, коли хтось бере Ваші речі без дозволу?»

На нашу думку такі результати зумовлені загальністю питання та очевидністю стверджувальної відповіді.

При цьому лише одне питання №7 “ Чи часто Вам сняться кошмари?”, показало низький показник ефективності 0,18 та низький коефіцієнт дискримінативності -0,19, оскільки воно відноситься до нейротизму.

**Далі подається набір запитань, який може увійти до остаточної версії
опитувальника, який у цьому зразку ми теж не подаємо**

Визначення надійності тесту

Для визначення надійності нашого тесту ми застосовували коефіцієнт Кьюдера-Річардсона для дихотомічних шкал.

Результатом проведених нами обрахунків став наступний показник надійності 0,82, що свідчить про високу внутрішню узгодженість частин тесту, тобто всі запитання тесту спрямовані на вимірювання однієї ознаки – нейротизму/екстра -, інтроверсії.

Звичайно, стверджувати про надійність тесту лише за показниками внутрішньої узгодженості не є репрезентативним для надійності тесту в цілому, адже не менш важливим показниками надійності є ретестова надійність та надійність паралельних форм. Проте проведення ретестової надійності потребує наявності проміжку часу між тестуваннями 1-3 місяці, а визначення надійності паралельних форм потребує значної кількості часу на створення еквівалентних запитань/завдань тесту що власне унеможливило визначення нами цих виду надійності в межах даного дослідження.

Проте, спираючись на отримані нами показники надійності тесту за внутрішньою узгодженістю, ми можемо таки стверджувати про високу надійність нашого тесту.

Визначення дискримінативності тесту

Для визначення дискримінативності нашого тесту ми скористалися показником дискримінативності дельта - Фергюсона, який визначає відношення між показником дискримінативності, отриманим для даного тесту, і максимальним значенням дискримінативності, яке може забезпечити такий тест.

Результатами наших обрахунків стали наступні:

- для нейротизму = 0,97

- для екстра -, інтроверсії = 0,93

Спираючись на подані результати, можна стверджувати високий рівень дискримінативності завдань тесту та тесту в цілому.

Співставляючи отримані нами результати по рівню дискримінативності тесту з його надійністю за внутрішньою узгодженістю ми бачимо дещо нижчі показники за внутрішньою узгодженістю тесту в порівнянні з показниками дискримінативності, що можна пояснити представленням в завданнях тесту багатьох шкал, що характеризують різні психологічні особливості.

Обґрунтування валідності тесту

Для визначення валідності нашого тесту ми використовували сукупність аналізу інформації про різні види валідності.

У нашому дослідженні ми висували такі гіпотез:

1. Шкала «нейротизму» особистісний опитувальник Г.Айзенка (EPI) повинна достатньо корелювати з шкалами проявів тривоги Ж.Тейлора (MAS).
2. Шкала «екстраверсії» особистісного опитувальника Г.Айзенка (EPI) повинна не корелювати з шкалами проявів тривоги Ж.Тейлора (MAS).
3. Кореляція шкали локусу контролю Дж. Б. Роттера не повинна корелювати з результатами особистісного опитувальника Г.Айзенка (EPI).
4. Досліджувані, які мають високий рівень емоційної нестійкості, показують високі результати за шкалою «нейротизму».

Зважаючи на те, що ми лише припускали в гіпотезах про кореляції тестів, але не здійснювали їх, тому ми не можемо зробити остаточний висновок про конкурентну валідність нашого тесту. Ми можемо лише припускати можливість, що адаптована нами версія особистісного опитувальника Г. Айзенка (EPI) для україномовних досліджуваних буде корелювати із зазначеними вище тестами.

6. Конструктна валідність відображає міру репрезентації досліджуваного психологічного тесту у результатах тесту. У зв'язку з обмеженим часом, відсутності проведення гіпотез та недостатньою кількістю респондентів ми не можемо зробити обґрунтовані висновки про конструктну валідність нашого тесту.

Висновки

Отже, нами було адаптовано особистісний опитувальник Г. Айзенка для україномовних респондентів, який визначає рівень прояву нейротизму та екстра-, інтроверсії. Відповідно результати по шкалі екстраверсії визначають орієнтацію особистості на зовнішній чи внутрішній світ, типові прояви поведінки людини у повсякденному житті. Результати по шкалі нейротизму вказують на здатність особистості адаптуватися до змін середовища, емоційну стабільність чи нестабільність.

Результати дискримінативності тесту (вираховувалися окремо для кожної шкали) показали, що він є чутливим до індивідуальн-психологічних особливостей кожного, а високі показники надійності за внутрішньою узгодженістю свідчать про узгодженість завдань тесту з загальним показником. Проте відсутність хороших показників по конструктній валідності не дає змогу з впевненістю вказати, що наш тест є ефективним. Окрім цього тест має ще декілька недоліків. Не було визначено ретестову надійність, оскільки вона передбачає повторне проходження тесту через певний проміжок часу. Тому в майбутньому варто провести повторне дослідження, результати якого сприяють зробленню більш точних висновків, щодо ефективності тесту.

Також не була визначена надійність паралельних форм. Причиною цього також є брак часу, хоча оригінал особистісного опитувальника Г. Айзенка передбачає наявність форми тесту А і Б.

При повторному проведенні тесту, слід формувати вибірку враховуючи вік респондентів, вибирати однакову кількість досліджуваних із різних вікових категорій.

Зразок елементів звіту про апробацію
Опитувальника особливостей харчової поведінки у жінок

Розробник опитувальника: Нижник А. Є,

У проведенні дослідження та підготовці звіту взяли участь
студенти рупи ПС-31:
 Дерейко О., Кравчук Д., Маянцева Т., Михальчук З., Піцикевич Є., Руденок Д.,
 Савчук І., Сінчук Я., Стецюк І., Шермер Я., Шешеня Т., Шостак І.
Отрог - 2013

Гіпотези щодо конкурентної валідності:

1. Тест на виявлення особливостей харчової поведінки повинен корелювати зі «Шкала оценки уровня удовлетворения собственным телом»;
2. Тест на виявлення особливостей харчової поведінки повинен корелювати із «Опросник пищевых предпочтений-26»;
3. Тест на виявлення особливостей харчової поведінки повинен корелювати із «MPS – многомерная шкала перфекционизма П.Хьюитта и Т.Флетта»;
4. Тест на виявлення особливостей харчової поведінки повинен корелювати із «Торонквя шкала алекситимии»;
5. Тест на виявлення особливостей харчової поведінки повинен корелювати зі шкалами №1-соціальна неспроможність, №III – ставлення до людей, №4 – регуляція потягів «Теста-опросник самооотношения В. В. Столина, С. Р. Пантелеева».

Специфікація тесту				
<u>МАНИФЕСТАЦІЇ</u>	<u>ЗМІСТОВІ ОБЛАСТІ</u>			
	Нервова анорексія	Нервова булімія	Переїдання	Загальні показники
сприйняття власного тіла (прагнення бути худим, незадоволення фігурою, інтероцепція);	18	18	17	53

соціальний фактор (соціальна неспроможність, ставлення до людей);	12	10	12	34
вольовий фактор (перфекціонізм, регуляція потягів)	12	12	12	36
нервові симптоми (симптоми нервової анорексії, булімії та переїдання)	20	20	21	61

Оцінювання та відбір завдань

Аналіз завдань за результатом пілотажного дослідження є досить важливим моментом, оскільки відбір найефективніших завдань дасть можливість створити більш досконалий опитувальник.

Для оцінки запитань нами були використані індекс ефективності завдань, який дозволяє визначити долю досліджуваних, що відповіли згідно з ключем. Також для кращої оцінки завдань ми використали коефіцієнт дискримінативності завдань, що показує кореляцію кожного завдання із загальним показником.

Оскільки наш тест складається із трьох областей то, відповідно, і завдання оцінювалися окремо по кожній із них.

В області переїдання було відсіяно 26 запитань - завдання, індекс ефективності яких був меншим від 0,25(мала кількість досліджуваних відповіли згідно з ключем) і більшим 0,75(завдання вважається не ефективним, адже велика кількість досліджуваних відповіли згідно з ключем), а також коефіцієнт дискримінативності був нижче за 0,2 – недостатня кореляція питання з загальним показником по тесту. Деякі питання мали високий коефіцієнт ефективності, але низький коефіцієнт дискримінативності - №96, №140, №144. Наприклад, питання №122 – (0, 84/0,07): «Ви здатні відкласти свої важливі справи, щоб виконати маловажливе і нетермінове доручення тому, що Ви обіцяли другу?» Отримані результати можна пояснити тим, що формулювання даного запитання передбачає наявність установки на соціальну схвалюваність. Проте більшість з відкинутих нами запитань мали низькі коефіцієнти як ефективності, так і дискримінативності - №18, №94, №123, №142 тощо . Наприклад, питання №49 –

(0,01/0,12): «Вам важко взуватися?». Отримані результати можуть свідчити про неоднозначність запитання і про те, що не було враховано додаткові умови.

В області булімії було відсіяно 24 запитання, серед яких були ті, індекс ефективності яких був високим, а індекс дискримінативності, навпаки, низьким - № 63, №138, №173. Серед них було питання №88 – (0,84/0,05): «Ви контролюєте себе в закладах, де можете поїсти?». У цьому питанні не вказано конкретних уточнюючих умов, що, імовірно, вплинуло на отримання саме таких даних. Варто зазначити, що більшість запитань було відкинуто через низькі показники коефіцієнтів ефективності та дискримінативності - №16, №35, №67. Зокрема, питання №137 (0,04/0,12): «Ви потрапляли в лікарню під час дієти?»; №177 (0,01/0,11): «Чи використовуєте Ви засоби для схуднення мінімум 2 рази за тиждень?» Зміст, закладений у сутність запитань, є вузькоспеціалізованим і не конкретизованим.

В області анорексії нами було відсіяно 27 запитань, які не підійшли по коефіцієнту дискримінативності, оскільки їх показники були нижче 0, 2 та по коефіцієнту ефективності нижче 0, 25 та вище 0, 75. Деякі з них мали високий коефіцієнт ефективності та низький коефіцієнт дискримінативності - №81, №128, №162. Наприклад, питання №126 – (0,87/-0,005): «Коли Ви п'єте гарячий чай, Ви відчуваєте як тепло проходить по вашому тілу?», - описує фізіологічно обумовлений процес, який не розкриває лише суть досліджуваного нами явища. №77 – (0,84/0,03): «Яскраві емоції викликають у Вас мурашки по всьому тілу?» - дане питання було вилучено за рахунок того, що такої схильності не було виявлено. Проте більшість запитань було відкинуто через низькі показники коефіцієнтів ефективності та дискримінативності - №54, №76, №78, №117 тощо. Наприклад, №32 – (0,07/0,14): «Ви підраховуєте кількість з'їдених вами калорій?», - передбачає наявність специфічно запланованої установки, є занадто вузьким, розраховане на малу категорію людей і можливою є установка на соціальну схвалюваність. Ці запитання можна було б перевірити на більшій кількості досліджуваних.

Визначення надійності тесту

Для перевірки нашого тесту на надійність ми вираховували надійність частин тесту шляхом поділу на парні та непарні номери запитань. Надійність за внутрішньою узгодженістю забезпечує узгодженість запитань тесту між собою. Теоретичним обґрунтуванням є припущення того, що у разі нормального або близького до нього розподілу оцінок по тесту в цілому виконання будь-якого випадкового набору завдань з частин тесту дає аналогічний розподіл (за умови, що частини однорідні за характером завдань, а також щодо результату тесту загалом). Це дає можливість встановити міру

еквівалентності, адекватності підбору запитань (А.Анастасі); дає інформацію про внутрішню однорідність тесту, його гомогенність (М.К.Гуревич).

Визначення такої надійності не потребує повторного дослідження і забезпечує відсутність впливу запам'ятовування відповідей (як ре тестова надійність), не передбачає складання додаткових еквівалентних форм запитань (як надійність паралельних форм), тому є хорошим інструментом для швидкого отримання результатів.

Результатами проведених нами обрахунків стали наступні показники надійності:

Коефіцієнт Пірсона - 0,80

Коефіцієнт Спірмена-Брауна - 0,89

Дані показники свідчать про високу внутрішню узгодженість частин тесту, тобто всі запитання тесту спрямовані на вимірювання однієї ознаки – схильності до виникнення розладів харчової поведінки.

Звичайно, стверджувати про надійність тесту лише за показниками його внутрішньої узгодженості не є репрезентативним для надійності тесту в цілому, адже не менш важливим показником надійності є *ретестова надійність*, яка, власне, і підтверджує, що при повторному дослідженні отримано ті ж самі результати у тих самих досліджуваних, а отже можливий вплив зовнішніх факторів при конструюванні тесту враховано і зведено до мінімуму. Проте проведення ретестової надійності потребує наявності проміжку часу між тестуваннями 1-3 місяці, що власне унеможливило визначення нами цього виду надійності в межах даного дослідження.

Ще одним вагомим показником у визначенні надійності є *надійність паралельних форм*, яка, за умови високого показника кореляції між отриманими результатами, виступає гарантією того, що при розробці тесту було знайдено чітку діагностичну ознаку, а отже доказом надійності психологічного змісту тесту. Проте визначення надійності паралельних форм потребує значної кількості часу на створення еквівалентних запитань/завдань тесту, та (особливо!) доведення цієї еквівалентності; з огляду на обмеженість строків нашого дослідження цей вид надійності нами визначено не було.

Отже, спираючись на отримані нами показники надійності частин тесту, ми можемо стверджувати про високу надійність нашого тесту.

Визначення дискримінативності тесту

Тест вважається максимально дискримінативним, коли кожен з досліджуваних відповів по різному – тоді показник дискримінативності = 1. Визначення

дискримінативності полягає у визначенні загальної кількості можливих відношень між парами досліджуваних, які порівнюються між собою і їх результати виявляються або рівними, або різними.

Метою тестування є зазвичай вимірювання ознаки, яка виражена у кожного досліджуваного в різній мірі, тому показником цього може служити досягнення нормального розподілу результатів. Для визначення наявності такого розподілу ми скористалися показником Колмогорова-Смірнова, який $= 1,236$ ($p=0,094$).

Тож можна стверджувати високий рівень дискримінативності завдань тесту та тесту в цілому.

Оскільки наш тест призначений для визначення схильності до виникнення розладів харчової поведінки, яка передбачає високий рівень індивідуальності, варіативності харчової поведінки кожної з досліджуваних, то отримання високих показників дискримінативності було для нас важливою і бажаною характеристикою тесту, як такого, що чутливий до індивідуальних особливостей кожного.

Співставляючи отримані нами результати по рівню дискримінативності тесту з його надійністю за внутрішньою узгодженістю ми бачимо дещо нижчі показники за внутрішньою узгодженістю тесту в порівнянні з показниками дискримінативності, що можна пояснити представленням в завданнях тесту багатьох шкал, що характеризують різні психологічні особливості, тому загалом можуть не бути тісно пов'язані між собою, але все ж корелювати з загальним результатом по тесту.

Саме такого результату ми передбачали досягти і отримали його.

Визначення валідності тесту

Валідність – це здатність тесту вимірювати те, для чого він призначений.

Слід зазначити, що валідність тесту не вимірюється, про неї можна лише судити за сукупністю інформації про різні види валідності, яка інтегрується в результаті в остаточний висновок про «адекватність», «задовільність», «недостатність» валідності тощо. До таких різновидів валідності відносяться:

Очевидна валідність. В нашому випадку, з огляду на вік вибірки очевидна валідність з одного боку мала б викликати серйозне ставлення досліджуваних до виконання завдань тесту, але при цьому не викривати повністю призначення даного психодіагностичного інструменту з метою запобігти неправдивим соціально бажаним відповідям.

Ми перевіряли очевидну валідність запитуючи респондентів, що на їх думку, вимірює цей тест. Більшість відповідей стосувалися теми їжі і зовнішності, звучали типу: «в мене немає проблем з повнотою», «поїсти я люблю» то ж очевидна валідність в

цьому плані була забезпечена. Що стосується серйозного ставлення до виконання завдань, то можна вважати нам це також вдалося з огляду на коментарії : «ого, у вас тут такі *серйозні* запитання!». Крім того ми починали наше спілкування з досліджуваними зі слів: «Наша кафедра психолого-педагогічних дисциплін проводить дослідження жінок вашого віку (кожен з нас, дослідників, вказував свій віковий проміжок) з метою визначення особливостей харчової поведінки. Чи погодитесь Ви(!) бути учасником дослідження?» що посилювало зацікавленість респондентів вже не лише як жінок, а й представниць певного вікового періоду.

Тож ми вважаємо що досягнули бажаної очевидної валідності: з одного боку досліджувані здогадалися лише про область запитань, але не звернули уваги на визначення розладів в харчовій поведінці (на що, власне, і був спрямований наш тест); з іншого боку ставилися до цього дослідження досить серйозно.

Прогностична валідність. В нашому випадку таким показником розладів харчової поведінки могла б стати вага тіла жінок, яка б, прогностично, суттєво збільшилась у тих досліджуваних, які мали високі показники по переїданню, і знизилась у жінок з ознаками схильності до анорексії. Ще одним критерієм міг би стати стан ШКТ (шлунково-кишкового тракту), ендокринної, нервової, серцево-судинної системи тощо. Проте дослідити такі зміни проблематично вже навіть на стадії визначення реального стану організму досліджуваних через витратність таких досліджень в часовому і грошовому відношенні. Тож прогностична валідність в нашому випадку залишається як керування на перспективу.

Конкурентна валідність оцінюється за рівнем кореляції результатів розробленого тесту з іншими, валідність яких вже доведено.

Конкурентну валідність ми визначали шляхом кореляції показників, які отримали досліджувані по певній шкалі з показниками вже раніше стандартизованих тестів, спрямованих на визначення цих показників.

Шкала «Прагнення бути худим»

На нашу думку, респонденти, яким властиві харчові розлади проявляють тенденцію до похудання. Тому для конкурентної валідності для даної шкали нами була обрана методика ОПП-26 («Опросни пищевих предпочетений»), оскільки ОПП дозволяє виявити такі проблеми з харчовими розладами, як: прагнення бути худим, страх набрати зайву вагу, обмеження у споживанні їжі.

Відповідно нами була висунута наступна гіпотеза: результати по шкалі «Прагнення бути худим», повинні корелювати із показниками ОПП-26.

Ми отримали наступні результати $0,473^{**}$, $p=0,000$ (прямий, помірний, статистично значимий зв'язок). Таким чином, бачимо, що показники мають сильну кореляцію між собою.

Шкала «Невдоволеність фігурою»

Після проведення теоретичного аналізу ми встановили, що люди, які страждають кожним з цих порушень харчової поведінки незадоволені власною фігурою. Тож у вигляді конкурентної методики для даної шкали ми використали «Шкалу оцінки задоволеності собственным телом». Ми передбачали, що досліджувані, які мають схильність до розладів харчової поведінки отримують також і високі бали по цій конкурентній методиці.

Відповідно нами була висунута наступна гіпотеза: результати по шкалі «Невдоволеності фігурою» повинні помірно корелювати із показниками «Шкалы оцінки задоволеності собственным телом».

Ми отримали наступні результати: $0,184$ ($p=0,152$), (прямий, слабкий зв'язок).

Тож можна зробити висновок, що дійсно досліджуваним з високим рівнем незадоволеності фігурою характерний високий показник по конкурентному тесту, що означає невдоволеність власним тілом. Даний зв'язок є, але він слабкий. Перевірити наявність такої закономірності можна було б повторним тестуванням на більшій за обсягом вибірці.

Шкала «Інтроцепція»

Дана шкала визначала те, наскільки досліджувані є чутливими до свого тіла. Ми визначили, що усі люди, в яких є порушення харчової поведінки погано відчують власне тіло та ідентифікують те, що в ньому відбувається. Конкурентною методикою для даної шкали ми використали «Торонтську шкалу алекитимії», оскільки вона досліджує здатність людей ідентифікувати те, що з ними відбувається, наприклад емоції.

Нами була висунута гіпотеза: результати по шкалі «Інтроцепція» повинні помірно корелювати із показниками «Торонтської шкали алекитимії».

Ми отримали наступні результати: $0,246$ ($p=0,054$), (прямий, помірний зв'язок).

Тож можна зробити висновок, що дійсно досліджуваним з високим рівнем по шкалі «Інтроцепція» характерний високий показник по конкурентному тесту, що означає невідчуття власного тіла.

Шкала «Соціальна неспроможність»

Дана шкала визначала те, наскільки досліджувані є чутливими до свого тіла. Ми визначили, що усі люди, в яких є порушення харчової поведінки погано відчують

власне тіло та ідентифікують те, що в ньому відбувається. Конкурентною методикою для даної шкали ми використали «Торонтську шкалу алекитимії», оскільки вона досліджує здатність людей ідентифікувати те, що з ними відбувається, наприклад емоції.

Нами була висунута гіпотеза: результати по шкалі «Інтроцепція» повинні помірно корелювати із показниками «Торонтської шкали алекитимії».

Ми отримали наступні результати: 0,246 (прямий, слабкий зв'язок, статистично значимий зв'язок).

Тож можна зробити висновок, що дійсно досліджуваним з високим рівнем по шкалі «Інтроцепція» характерний високий показник по конкурентному тесту, що означає невідчуття власного тіла.

Конкурентну валідність ми визначали шляхом кореляції показників, які отримали досліджувані по певній шкалі з показниками вже раніше стандартизованих тестів, спрямованих на визначення цих показників.

Шкала «Ставлення до людей»

На нашу думку, яскраво характеризує респондентів, яким властиві харчові розлади анорексія, булімія та переїдання. Тому для конкурентної валідності для даної шкали нами був обраний опитувальник «Тест-опросник самоотношения В.В. Столина, С. Р. Пантелеева», зокрема шкала – очікуване відношення від людей.

Відповідно нами була висунута наступна гіпотеза: результати по шкалі «Ставлення до людей», повинні мати обернену кореляцію із показниками шкали очікуваного відношення від людей.

Ми отримали наступні результати $-0,271^*$ ($p=0,033$) (обернений, слабкий, статистично значимий зв'язок). Таким чином, наші припущення підтвердилася, а обернений зв'язок продемонстрував відмінності між нашим тестом (у нас ставлення до людей), та опитувальник «Тестом-опросником самоотношения В.В. Столина, С. Р. Пантелеева» (очікуване ставлення від людей).

Шкала «Регуляція потягів»

Дана шкала визначала те, наскільки досліджувані здатні контролювати свої бажання в плані харчування. Ми визначили, що усі люди, в яких є порушення харчової поведінки погано себе контролюють. Конкурентною методикою для даної шкали ми використали Шкалу 4 «Саморуководство, самопоследовательность» Теста-опросника самоотношения (В.В. Столин, С.Р. Пантелеев), оскільки вона досліджує здатність людей керувати своїми діями.

Нами була висунута гіпотеза: результати по шкалі «Регуляція потягів» повинні обернено помірно корелювати із показниками Шкали 4 «Саморуководство, самопоследовательность» Теста-опросника самоотношения (В.В. Столин, С.Р. Пантелеев),

Ми отримали наступні результати: $-0,133$ ($p=0,301$), (обернений, слабкий зв'язок).

Тож можна зробити висновок, що дійсно досліджуваним з високим рівнем по шкалі «Регуляція потягів» характерний низький показник по конкурентному тесту, що означає неможливість людей з порушеннями харчової поведінки чітко керувати своїми діями. Даний зв'язок є, але він слабкий. Перевірити наявність такої закономірності можна було б повторним тестуванням на більшій за обсягом вибірці.

Шкала «Перфекціонізм»

На нашу думку, респонденти, яким властиві харчові розлади анорексія, булімія повинні мати високі показники поданій шкалі, а переїдання – низькі. Тому для конкурентної валідності для даної шкали нами була обрана методика MPS («Многомерная шкала перфекционизма П. Хьюитта и Г. Флетта»), оскільки MPS дозволяє виявити різні прояви перефекціонізму (на себе, на інших тощо).

Відповідно нами була висунута наступна гіпотеза: результати по шкалі «Перфекціонізм», повинні корелювати із показниками MPS.

Ми отримали наступні результати для анорексії та булімії $-0,229$ $p=0,074$ (обернений, слабкий, статистично незначимий зв'язок), для переїдання $-0,031$ $p=0,812$ (обернений, дуже слабкий, статистично незначимий зв'язок). Таким чином, кореляція слабка, проте для анорексії та булімії вона все-таки сильніша, ніж для переїдання.

Шкала «Нервові симптоми»

Дана шкала визначала саме наявність в досліджуваних певних характеристик, які визначають саме ці порушення харчової поведінки. Проаналізувавши теоретичні джерела ми визначили особливості, характерні для кожного порушення. Конкурентною методикою для даної шкали ми використали «Опросник пищевых предпочтений», який дозволяє виявити вірогідність наявності відхилень в харчовій поведінці

Нами була висунута гіпотеза: результати по шкалі «Нервові симптоми» повинні корелювати із показниками «Опросника пищевых предпочтений».

Ми отримали наступні результати: $0,543^{**}$ ($p=0,000$), (обернений, середній зв'язок).

Тож можна зробити висновок, що дійсно досліджуваним з високим рівнем по шкалі «нервові симптоми» характерний високий показник по конкурентному тесту, що

означає, що в досліджуваних виявлено таку вірогідність наявності харчових відхилень, яку й показує наш тест.

б. Конструктна валідність поєднує в собі всі інші підходи до визначення валідності і дозволяє визначити, чи дійсно існує певна властивість, яку вимірюють, і наскільки надійно даний тест вимірює індивідуальні відмінності певної даної властивості.

Стосовно існування особливостей харчової поведінки можна лише зазначити що поняття нервової анорексії, булімії і переїдання офіційно відносяться до розладів харчової поведінки і є актуальним питанням вивчення в психології на сьогодні (детальніше це описано в теоретичному конструкті, питання №3).

Чи чутливий тест до індивідуальних відмінностей у рівні вираженості особливостей харчової поведінки? – про це свідчать показники дискримінативності тесту, що є високими, і при цьому співвідносяться з показниками надійності (зокрема надійності частин тесту) які є, як і передбачалося, дещо нижчими від показників дискримінативності, але достатніми щоб стверджувати про надійність нашого тесту.

З огляду на теоретичну обґрунтованість явища, яке ми вивчаємо та емпірично доведено здатність тесту надійно вимірювати індивідуальні відмінності у особливостях харчової поведінки різних індивідів, ми можемо стверджувати про валідність даного тесту.

Висновки

Отже, нами був створений тест на дослідження особливостей харчової поведінки, який дає змогу респонденту визначити схильність до певного розладу харчової поведінки: нервової анорексії, нервової булімії та переїдання, а також відповідний рівень схильності: дуже низький, низький, середній, високий та дуже високий.

При математичній обробці тест показав хороші показники по дискримінативності завдань тесту та тесту в цілому, що було нам дуже важливо, адже це свідчить про те, що наш тест чутливий до індивідуальних особливостей кожного. Також ми отримали високі показники по надійності тесту за внутрішньою узгодженістю, задовільні оцінки по конкурентній валідності тесту та досягнули бажаної очевидної валідності. Це дає змогу говорити про певну ефективність нашого тесту.

Проте, існують певні недоліки. Не було обраховано ретестової надійності тесту, що є важливим показником, який показує вплив зовнішніх факторів на результати тесту, проте проведення ретестової надійності потребує наявності проміжку часу між тестуваннями 1-3 місяці, що власне унеможливило визначення нами цього виду надійності в межах даного дослідження.

Також не було визначено надійність паралельних форм, яка є доказом надійності психологічного змісту тесту. Проте її визначення потребує значної кількості часу на створення еквівалентних запитань тесту, та (особливо!) доведення цієї еквівалентності; з огляду на обмеженість строків нашого дослідження цей вид надійності нами визначено не було.