

## Використання математичних моделей для аналізу результатів психологічного тесту Гілфорда

О. О. Диховичний<sup>1</sup>, Н. В. Круглова<sup>1</sup>, І. О. Москальов<sup>2</sup>

<sup>1</sup>*Кафедра математичного аналізу та теорії ймовірностей,  
КПІ ім. Ігоря Сікорського, Київ, Україна*

<sup>2</sup>*Кафедра суспільних наук,  
Національний університет оборони України імені Івана Черняховського, Київ, Україна*

[a.dyx@ukr.net](mailto:a.dyx@ukr.net), [natahak@ukr.net](mailto:natahak@ukr.net)

### Анотація

У роботі досліджено результати тестування за методикою Гілфорда на професійну придатність 800 військовослужбовців за допомогою відомих моделей IRT. Застосовано модель Раша, а також 1-PL, 2-PL та 3-PL моделі. Для комп'ютерного оброблення застосовано мову R, а саме, пакет `ltm`. Розраховано латентні параметри як респондентів, так і відповідних індикаторів (запитань тесту). Виявлено суттєві порушення адекватності побудованих моделей результатам тестування. Проаналізовано можливі причини таких невідповідностей, головною з яких є порушення об'єктивності умов проведення тестування.

**Ключові слова:** тест Гілфорда; соціальний інтелект; IRT; модель Раша; PL моделі; латентні параметри.

**MSC2010** 62P15

**УДК** 519.248:159.9

# 1 Вступ

Основним завданням професійного консультування є підбір для суб'єкта найоптимальнішого виду діяльності й конкретної професії. Велику увагу при цьому приділяють визначенню індивідуальних психологічних особливостей людини. Профконсультант прогнозує те, як людина себе поведе в тій або іншій ситуації, як його індивідуальні психологічні властивості проявляться у професійній діяльності, чи сприятимуть вони успіху. Для ефективного вирішення цього завдання, профконсультант потребує надійних і зручних у застосуванні вимірювальних процедур. Натепер існує чимало психодіагностичних методик ([Kihlstrom & Cantor, 2000](#)), ([Sternberg & Kaufman, 1998](#)), що дозволяють виявляти стійкі психологічні особливості людини, зокрема особливості її соціального інтелекту.

У США і багатьох європейських країнах широко застосовують методика дослідження соціального інтелекту Гілфорда ([O'Sullivan, Guilford, & de Mille, 1965](#)), який визначив соціальний інтелект як інтегральну інтелектуальну здібність, що визначає успішність спілкування й соціальної адаптації.

Методика дозволяє діагностувати соціальний інтелект і визначає як загальний рівень соціального інтелекту, так і рівень окремих його складових.

Визначальною особливістю методики Гілфорда дослідження соціального інтелекту є її прогностична можливість у професійному відборі.

Методика вимірювання соціального інтелекту впливає із загальної моделі структури інтелекту Гілфорда. Він розуміє соціальний інтелект як систему інтелектуальних здібностей, незалежних від чинника загального інтелекту і пов'язаних, передусім, з пізнанням поведінкової інформації, які визначають у просторі трьох змінних: зміст, операції, результати. Гілфорд виділяє одну операцію — пізнання — і зосереджує свої дослідження на пізнанні поведінки. Ця здатність включає шість чинників:

1. пізнання елементів поведінки — здатність виділяти з контексту вербальну й невербальну експресію поведінки;
2. пізнання класів поведінки — здатність розпізнавати загальні властивості в деякому потоці експресивної або ситуативної інформації про поведінку;
3. пізнання стосунків у поведінці — здатність розуміти стосунки, які існують між одиницями інформації про поведінку;
4. пізнання систем поведінки — здатність розуміти логіку розвитку цілісних ситуацій взаємодії людей, сенс їх поведінки в цих ситуаціях;
5. пізнання перетворень поведінки — здатність розуміти зміну значення схожої поведінки (вербальної або невербальної) у різних ситуаційних контекстах;

6. пізнання результатів поведінки — здатність передбачати наслідки поведінки, виходячи з наявної інформації.

Діагностичну батарею Гілфорда складають чотири субтести, найбільш адекватні (за результатами досліджень) для вимірювання соціального інтелекту, зорієнтовані на визначення наступних показників:

- пізнання результатів поведінки (П1);
- пізнання фактів поведінки (П2);
- пізнання змін поведінки (П3);
- пізнання систем поведінки (П4).

Як відзначено в ([Sternberg & Kaufman, 1998](#)), методика Гілфорда має високі психометричні характеристики (надійність, валідність), а також низку інших переваг. Результати тесту Гілфорда методами математичних моделей IRT ([van der Linden & Hambleton, 1997](#)) раніше не досліджувалось. Застосування IRT до дослідження соціального інтелекту розглядалось у ([Chermahini, Hickendorff, & Hommel, 2012](#)). Нами було поставлено задачу аналізу результатів тестування на профпридатність на підставі тесту Гілфорда 800 військових віком від 30 до 50 років. Схожий аналіз проведено в ([Dykovychnyi, Zlyvkov, Kruglova, Lukomska, & Kotukh, 2018](#)).

## 2 Основна частина

Стандартна процедура оброблення результатів тесту Гілфорда полягає в застосуванні певних ключів, за допомогою яких обраховують цілочисельні бали і респондентів розподіляють на п'ять категорій:

- 1 бал — низькі здібності до пізнання поведінки;
- 2 бали — здібності до пізнання поведінки нижче середнього;
- 3 бали — середні здібності до пізнання поведінки (середньовибіркова норма);
- 4 бали — здібності до пізнання поведінки вище середнього (середньосильні);
- 5 балів — високі здібності до пізнання поведінки.

У субтестах відповіді мають тип множинного вибору, але стандартне оброблення в системі MOODLE, ураховує лише правильність або неправильність відповіді, тобто, за суттю, є дихотомічним. Тому для оброблення результатів ми застосували

такі IRT-моделі: Раша, 1-PL, 2-PL та 3-PL (просте роз'яснення суті моделей, параметрів та змісту характеристичних кривих, а також перевірки на адекватність можна одержати в (Rodionov & Bratischenko, 2014).)

Ми запровадили наступні латентні параметри. Кожному субтесту відповідає один з показників — (П1, П2, П3, П4), відповідні питання субтесту назвімо індикаторами показника, латентні параметри індикатора — це рівні індикатора — визначають відповідно до кожної моделі IRT, зокрема складність, розрізняльна здатність, параметр угадування; для респондентів також визначають латентні параметри згідно з моделлю, тобто рівень показників (П1 —П4).

Латентні параметри було оцінено за допомогою функцій *rasch* (Раш, 1-PL), *ltm* (2-PL), *tpm* (3-PL) пакету *ltm*, одержаного із сайту (*The Comprehensive R Archive Network*, n.d.).

Для чотирьох субтестів знайдено всі латентні параметри згідно з обраними IRT-моделями, а також проаналізовано ансамблі характеристичних кривих індикаторів для кожного субтеста (рис. 1 —16).

**Перший субтест.** Для всіх чотирьох моделей (рис. 1–4) ансамблі характеристичних кривих нерівномірні (рівні індикаторів розподілені нерівномірні), але всі ансамблі є подібними. Окрім того, індикатори 13, 14 виділяються із загального ансамблю, вони є надскладними і мають низьку розрізняльну спроможність згідно з 2-PL і 3-PL моделями. Це саме можна сказати і про графіки інформаційних функцій.

**Другий субтест.** Модель Раша (рис. 5) та 1-PL модель (рис. 6) мають подібні ансамблі з очевидною нерівномірністю. У моделі 2-PL (рис. 7) виявляється дуже неоднорідна розрізняльна спроможність. А у 3-PL моделі (рис. 8) дуже високими є ймовірності вгадування для індикаторів 9, 10, 15, це відображено на графіках інформаційних функцій.

**Третій субтест.** Модель Раша (рис. 9) та 1-PL модель (рис. 10) мають подібні ансамблі з невеликою нерівномірністю. У моделі 2-PL (рис. 11) виявляється дуже нерівномірна розрізняльна спроможність, і є індикатори з дуже низькою і, навіть, від'ємною розрізняльною спроможністю (індикатори 9, 11). Таких індикаторів не повинно бути. А у 3-PL моделі (рис. 12), крім того, ще дуже високі ймовірності вгадування для індикаторів 4, 9. Це свідчить про можливість списування. Ці ефекти також відображені на графіках інформаційних функцій.

**Четвертий субтест.** Модель Раша (рис. 13) та 1-PL (рис. 14) модель мають подібні ансамблі з невеликою нерівномірністю. У моделі 2-PL (рис. 15) виявляється приблизно рівномірні розрізняльні спроможності, хоча є індикатори з дуже низькою розрізняльною спроможністю (індикатори 4, 12). А у 3-PL моделі (рис. 16), високі ймовірності вгадування для індикаторів 10, 11 (знову списування).

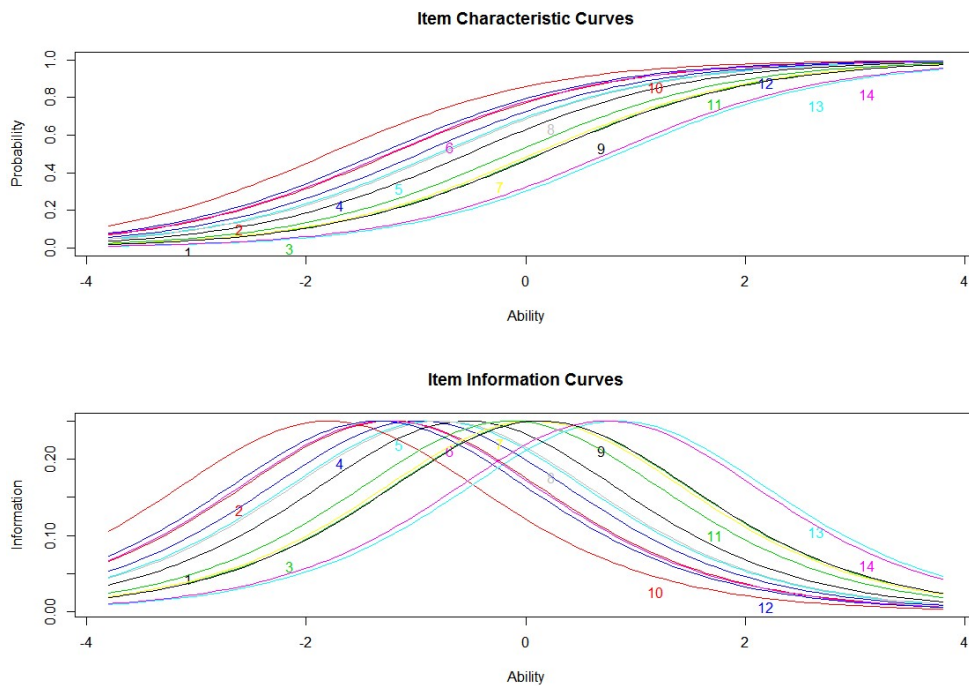


Рис. 1: Перший субтест. Модель Раша

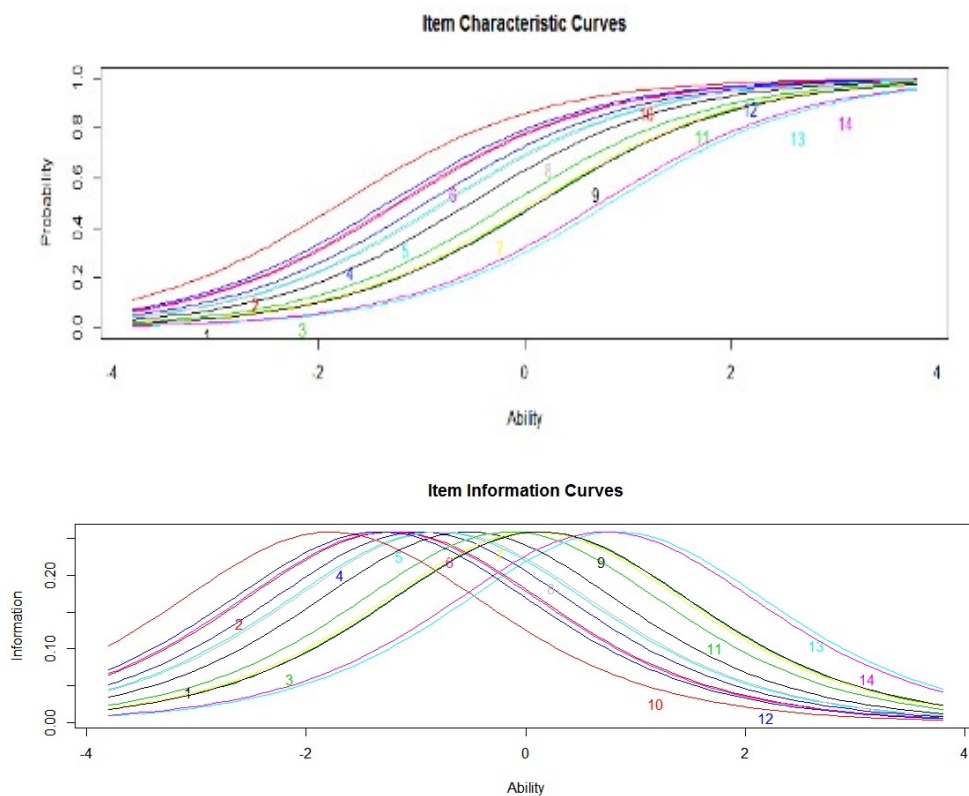


Рис. 2: Перший субтест. 1-PL model

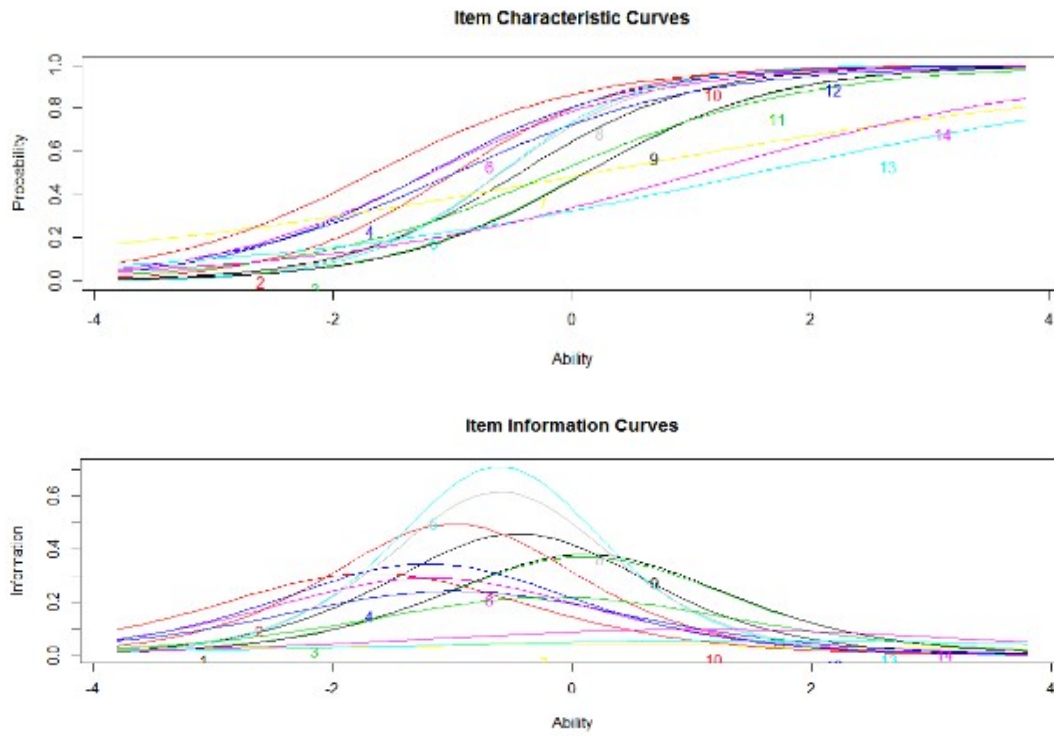


Рис. 3: Перший субтест. 2-PL model

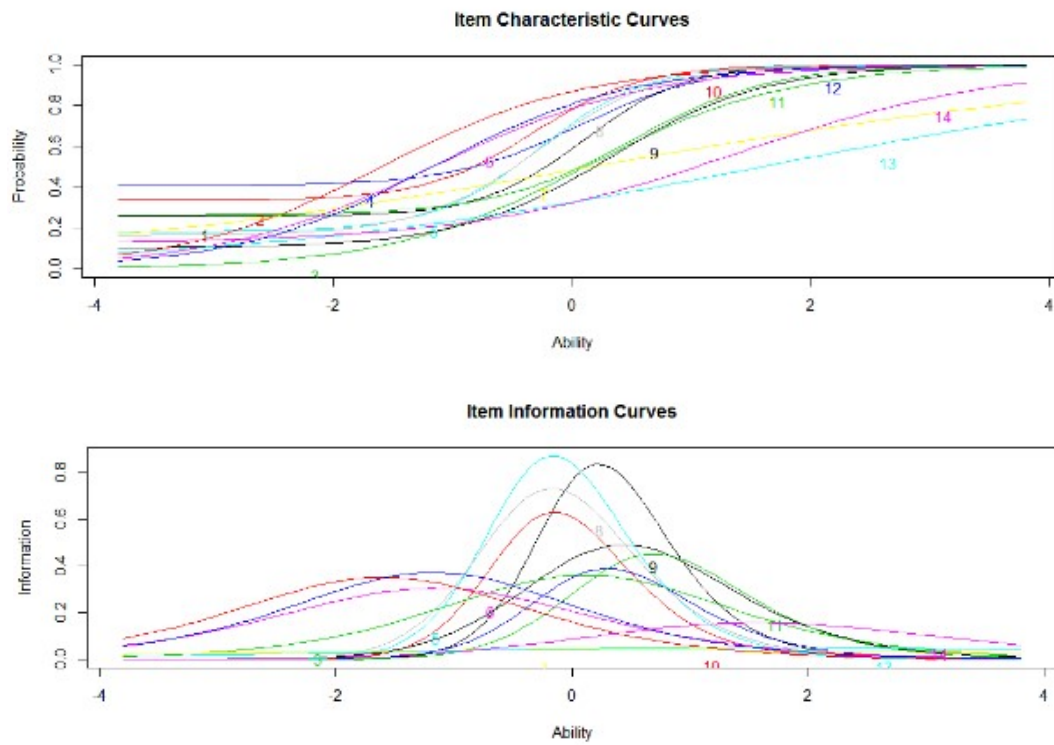


Рис. 4: Перший субтест. 3-PL model

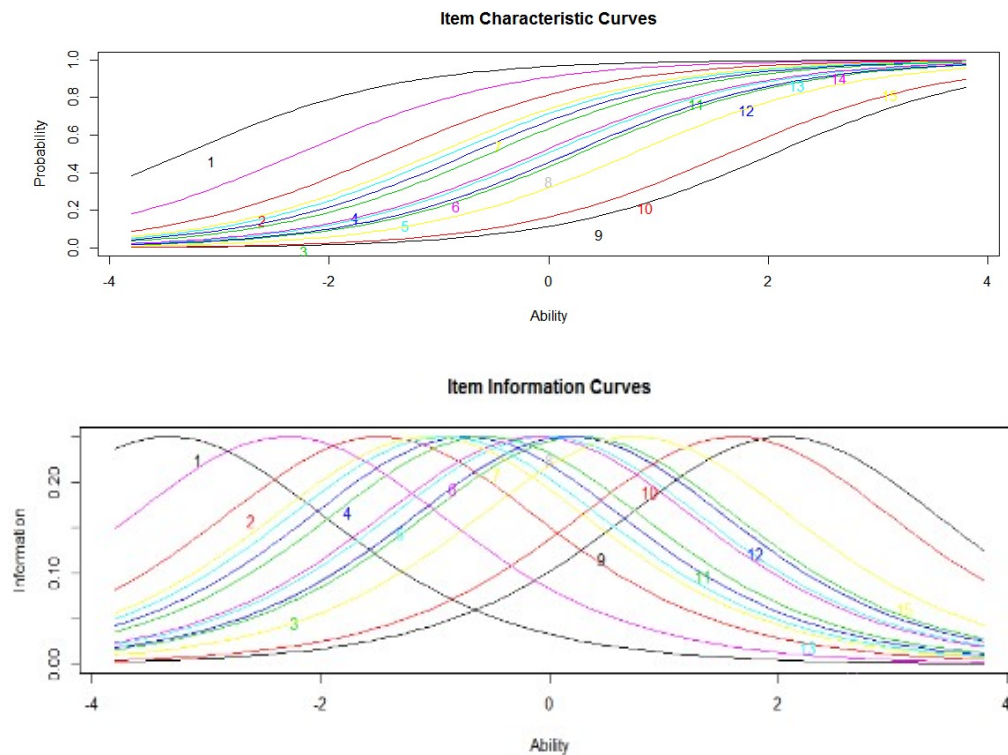


Рис. 5: Другой субтест. Модель Раша

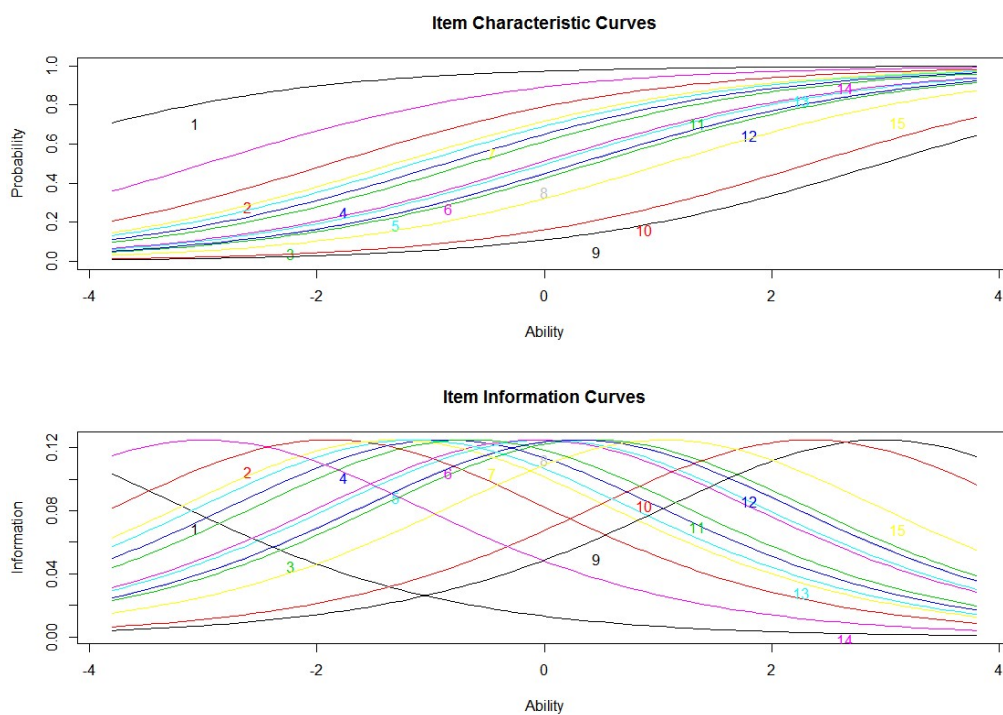


Рис. 6: Другой субтест. 1-PL model

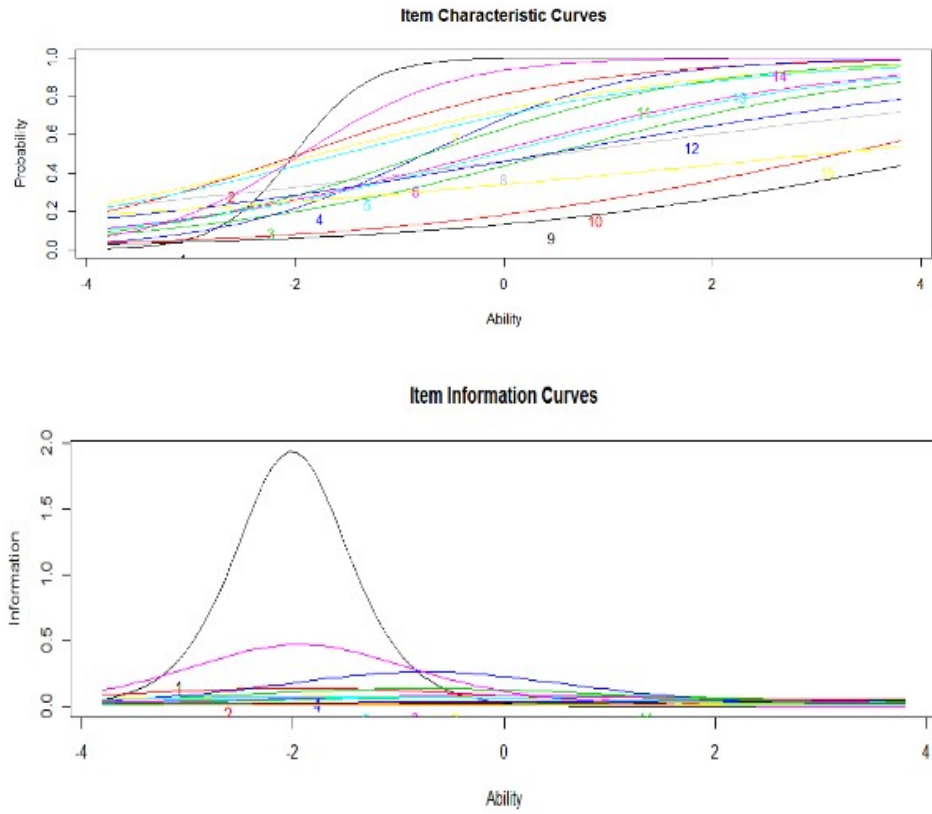


Рис. 7: Другий субтест. 2-PL model

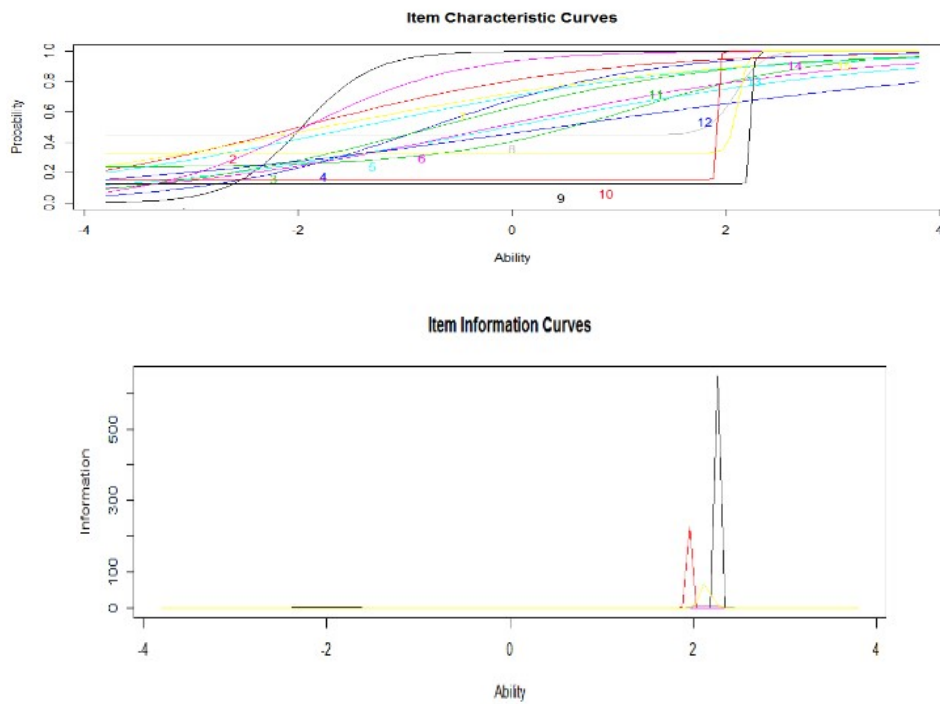


Рис. 8: Другий субтест. 3-PL model



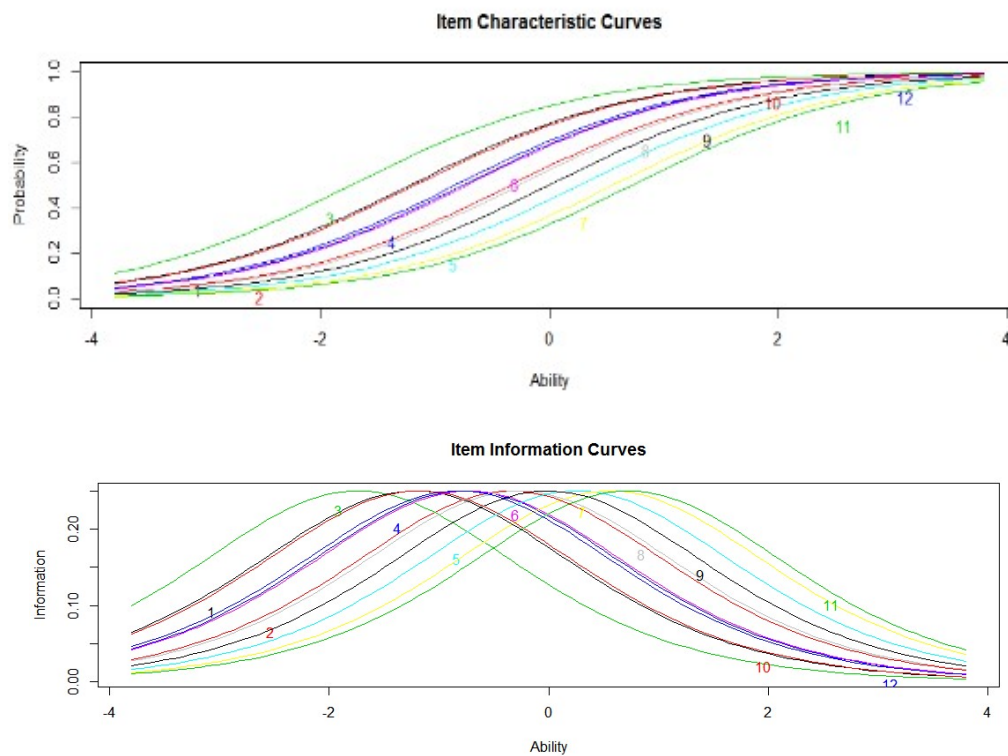


Рис. 9: Третій субтест. Модель Раша

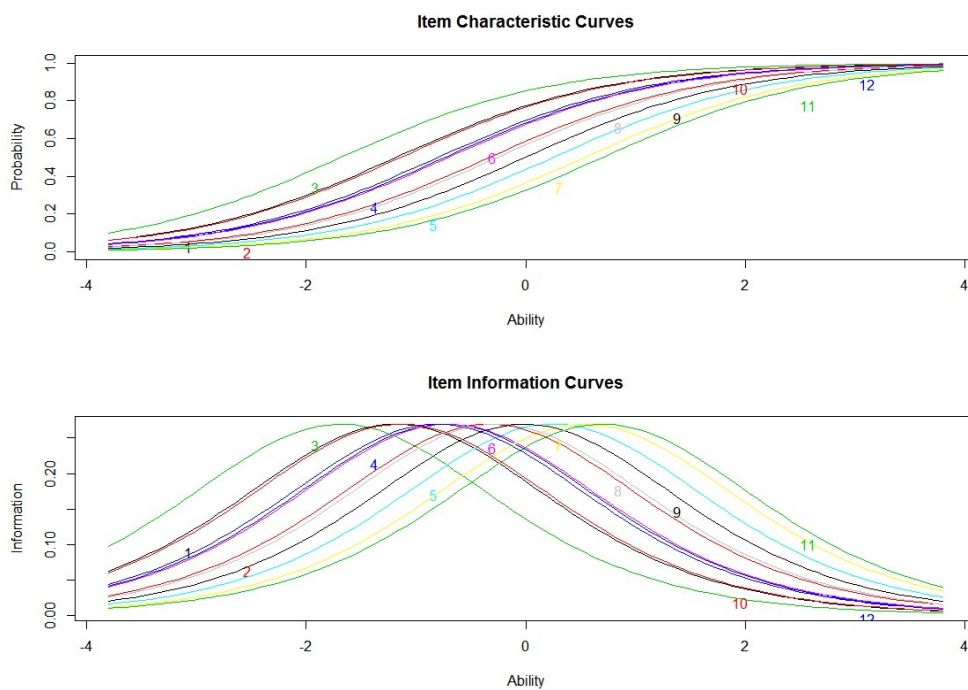


Рис. 10: Третій субтест. 1-PL model

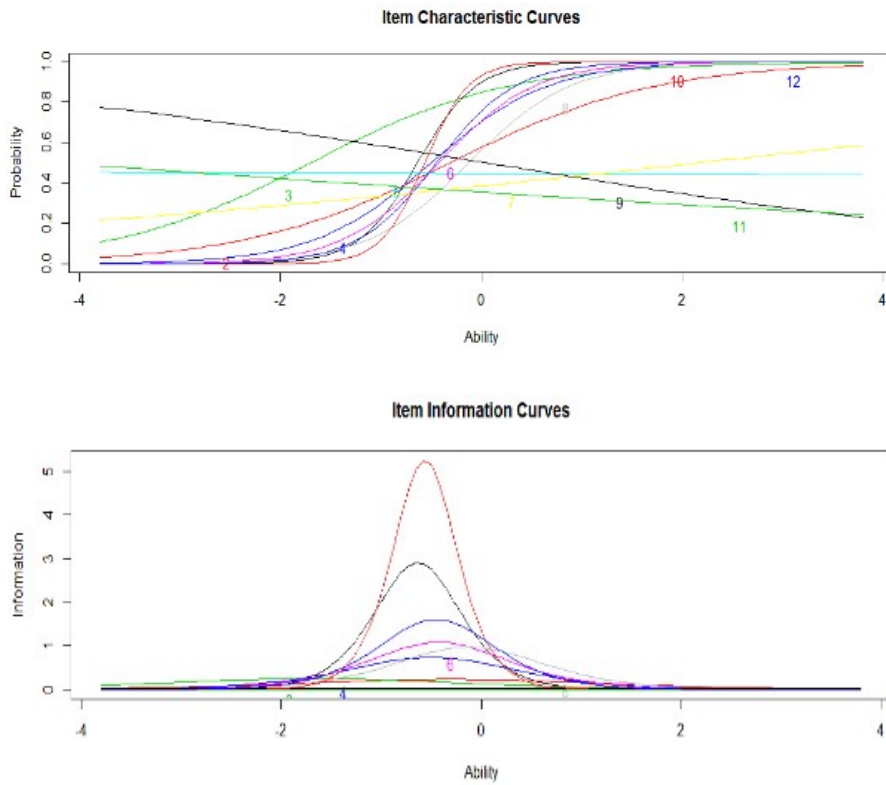


Рис. 11: Третій субтест. 2-PL model

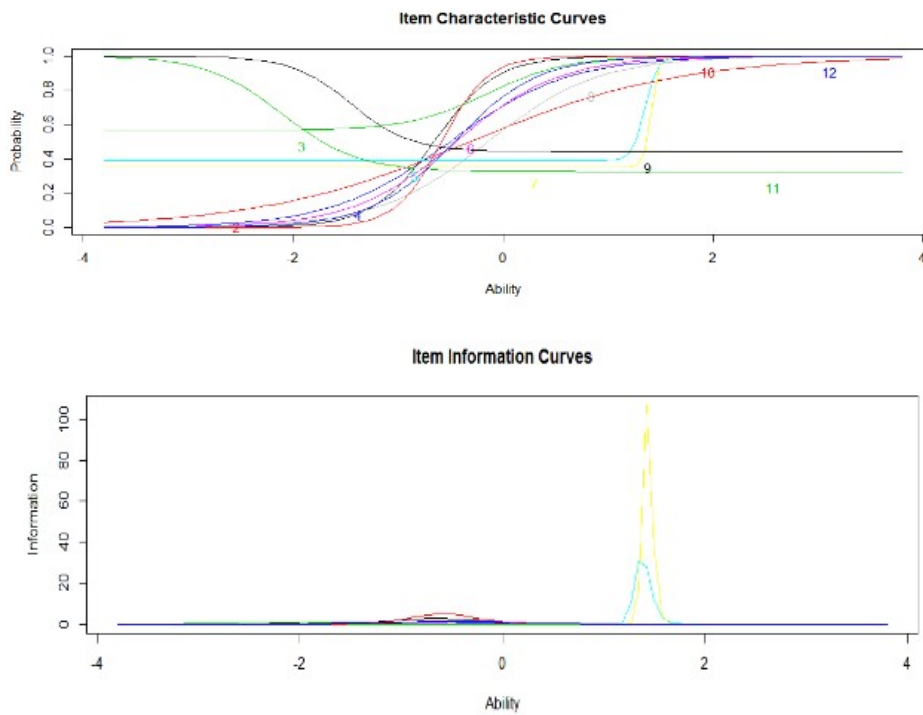


Рис. 12: Третій субтест. 3-PL model

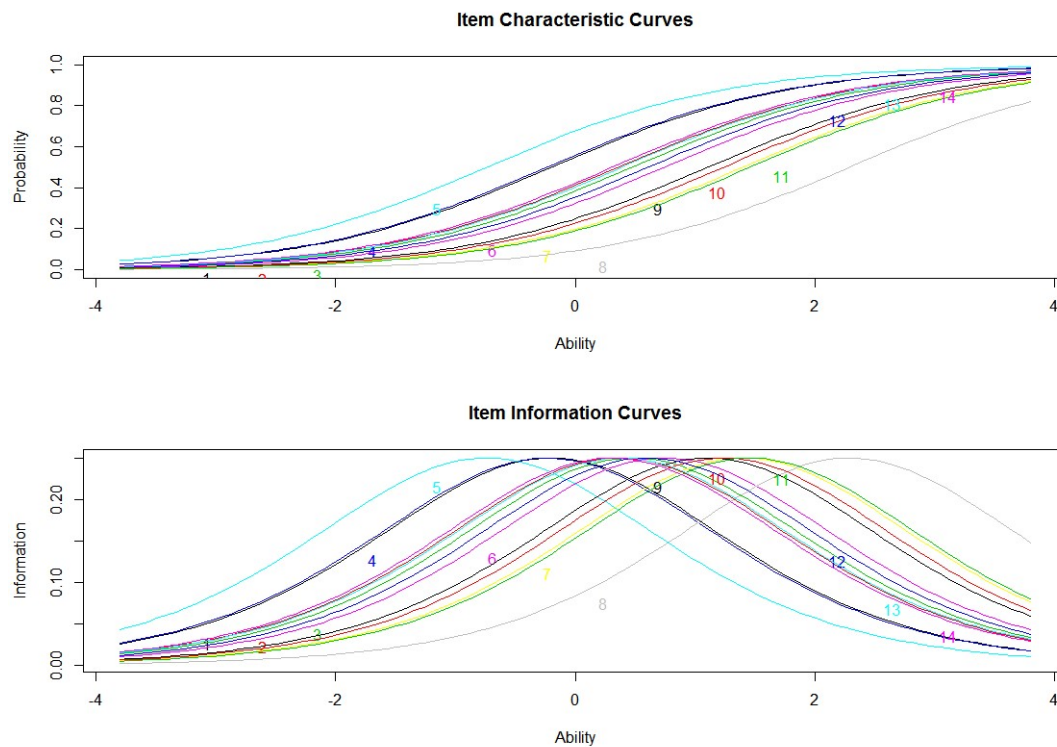


Рис. 13: Четвертый субтест. Модель Раша

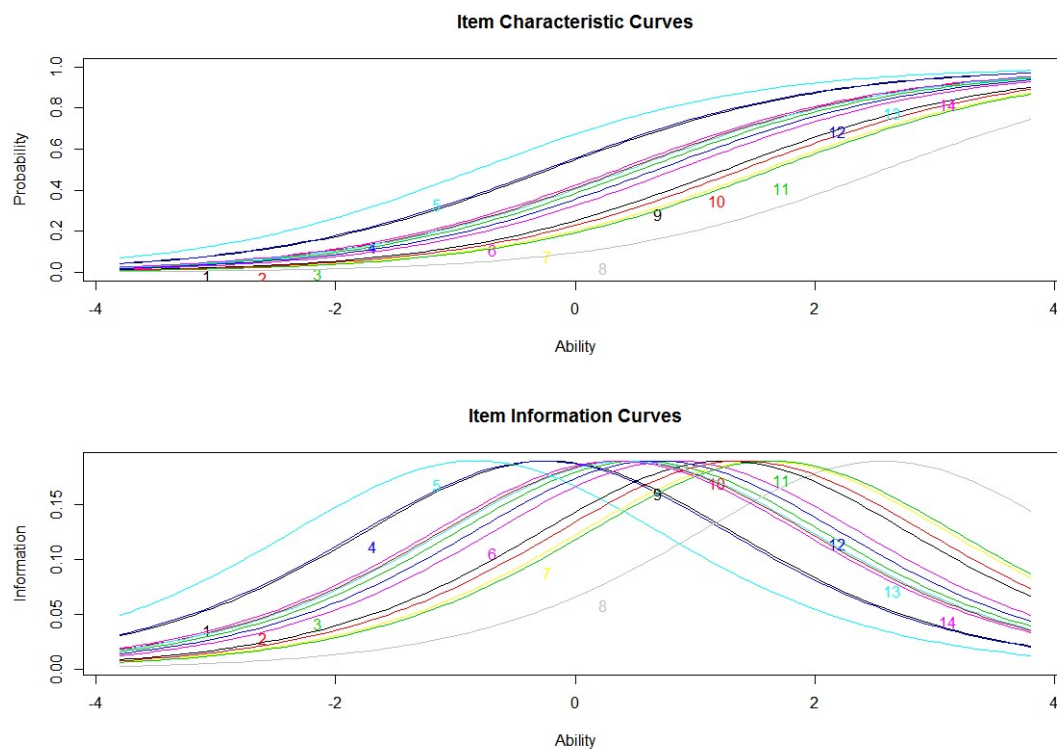


Рис. 14: Четвертый субтест. 1-PL model

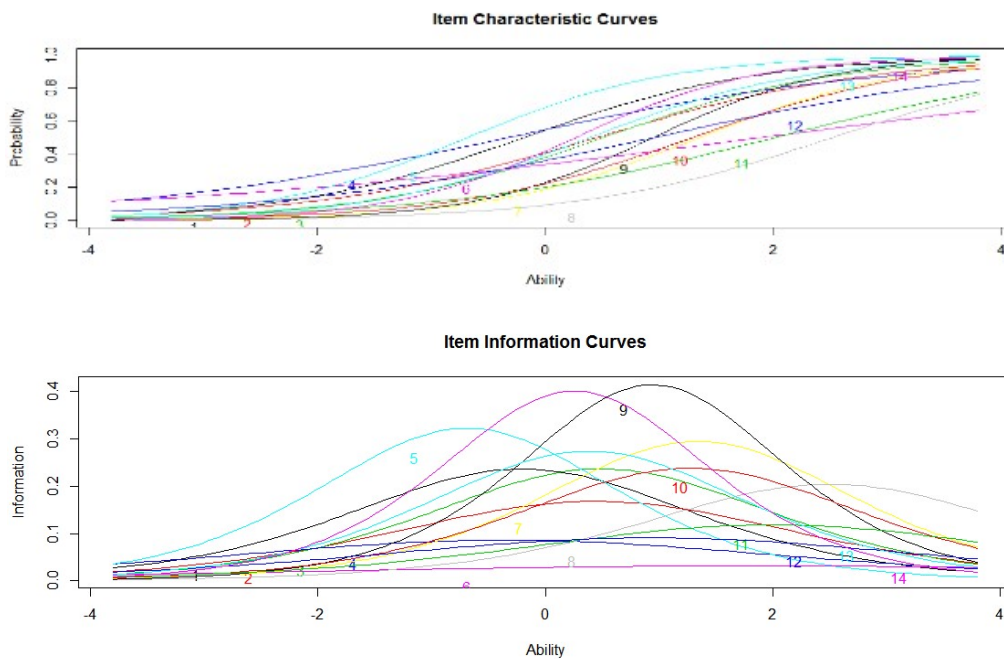


Рис. 15: Четвертый субтест. 2-PL model

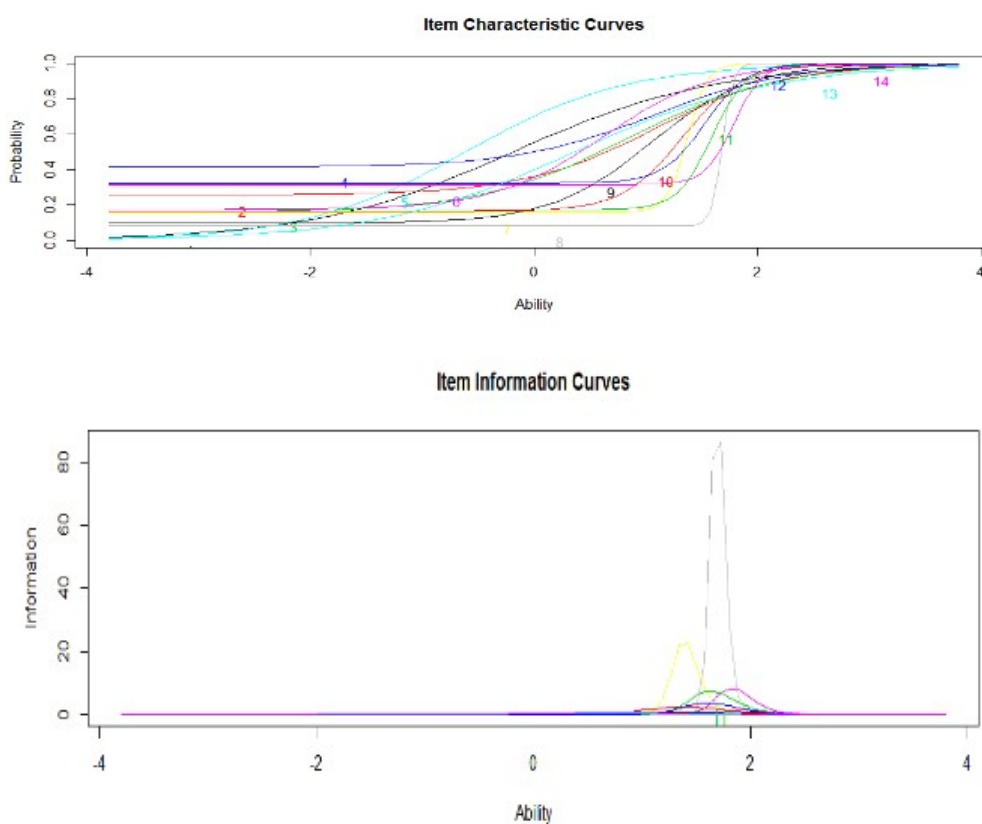


Рис. 16: Четвертый субтест. 3-PL model

Побудовані моделі перевірено на адекватність за допомогою функції `fit.item` пакету `ltm`. Визначено ті індикатори показників, які порушують адекватність (табл. 1).

Табл. 1:

	Моделі			
	Раш	1-Pl	2-Pl	3-Pl
Субтест 1	1, 2, 3, 5, 8, 9, 12	1, 2, 5, 8, 9, 12, 13	1, 8	8
Субтест 2	1, 4, 9, 14	1, 4, 9, 11, 14	3, 4, 6, 9, 10, 11, 14	3, 4, 5, 6, 11, 12, 14
Субтест 3	7	7	2, 4, 5, 8, 11	2, 4, 5, 6, 8, 12
Субтест 4	1, 5, 9, 14	1, 5, 9, 10, 13, 14	1, 2, 3, 5, 10, 13, 14	1, 5, 14

Як видно з таблиці значна кількість індикаторів неадекватна запропонованим моделям.

Як правило, особистісні параметри повинні бути розподілені за нормальним законом. Проведено перевірку параметрів респондентів на нормальність розподілу за критерієм Колмогорова – Смірнова (функція `ks.test`). У табл. 2 наведено значення «*p-value*». Як видно в більшості моделей гіпотеза про нормальність розподілу відхиляється (значення  $p - value < 0,05$ ).

Табл. 2:

	Моделі			
	Раш	1-Pl	2-Pl	3-Pl
Субтест 1	2.608e-10	1.166e-14	0.2859	0.4162
Субтест 2	5.333e-06	3.264e-05	0.007212	0.01806
Субтест 3	0.000142	1.433e-05	0.006279	0.005868
Субтест 4	8.416e-06	9.445e-10	0.1409	0.2486

Оцінено відповідні латентні параметри респондентів.

### 3 Висновки

1. Оцінено латентні параметри респондентів — рівні показників (П1 — П4). Це дозволило провести точнішу диференціацію респондентів, ніж стандартну дискретну за первинними балами, які дають лише п'ять можливих категорій.
2. Ансамблі характеристичних кривих не є рівномірними. Виділяються індикатори з надто високими рівнями та індикатори з від'ємною розрізняльною здатністю.

3. Для 3-Р1 моделі для всіх субтестів, яка враховує можливість угадування, ряд індикаторів має високе значення параметру вгадування.
4. Перевірки адекватності рівнів індикаторів обраним моделям показали високий відсоток неадекватних значень індикаторів.
5. Перевірка параметрів респондентів продемонструвала високий рівень відхилення від нормального розподілу, що є свідченням необ'єктивності тестування.
6. Аналіз часу, витраченого на відповідь (він був надто коротким), а також ідентичність за всіма субтестами та абсолютна правильність відповідей певних респондентів доводить можливість взаємного списування або використання відомих ключів.
7. Аналізуючи очевидні недоліки, можна спробувати пояснити їх наступними причинами: а) тест є застарілим, деякі рисунки не зовсім зрозумілі сучасній людині; б) тест не створювався для перевірки профпридатності військових; в) хибним, на наш погляд, є застосування дихотомічної моделі для завдань з явним множинним вибором. Надалі під час використання даного тесту в системі Moodle доречно додати опцію збереження матриць відповідей, у яких вказано номери варіантів відповідей. Це забезпечить можливість використання моделей із множинним вибором; г) очевидно, що проведення подібних тестів дистанційно, за наявності майже в кожного респондента доступу до інтернету, є також хибним, оскільки правильні відповіді можуть бути запозичені з мережі.
8. На наш погляд, застосування IRT-моделі для аналізу результатів тесту Гілфорда є цілком доречним, оскільки воно дозволило виявити недоліки тестування.

## References

- Chermahini, S. A., Hickendorff, M., & Hommel, B. (2012). Development and validity of a Dutch version of the Remote Associates Task: An item-response theory approach. *Thinking Skills and Creativity*, 7(3), 177–186.  
<https://doi.org/10.1016/j.tsc.2012.02.003>
- The Comprehensive R Archive Network*. (n.d.).  
<https://cran.cnr.berkeley.edu/>
- Dykhovychnyi, O., Zlyvkov, V., Kruglova, N., Lukomska, S., & Kotukh, O. (2018). Using the multidimensional models to the teacher authenticity scale adaptation. *Actual Problems of Psychology*, 14(1), 137–146.

- Kihlstrom, J. F., & Cantor, N. (2000). Social intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 359–379). Cambridge University Press.  
<https://doi.org/10.1017/CB09780511807947.017>
- O’Sullivan, M., Guilford, J. P., & de Mille, R. (1965). *The measurement of social intelligence* (Reports from the Psychological laboratory No. BR-5-0 198). Los Angeles: University of Southern California.  
<https://files.eric.ed.gov/fulltext/ED010278.pdf>
- Rodionov, A. V., & Bratischenko, V. V. (2014). Application of IRT models to analysis of results competency-based learning [in Russian]. *Sovremennyye problemy nauki i obrazovaniya*, 2014(4), 1–11.  
<https://www.science-education.ru/ru/article/view?id=13858>
- Sternberg, R. J., & Kaufman, J. C. (1998). Human abilities. *Annual review of psychology*, 49(1), 479–502.  
<https://doi.org/10.1146/annurev.psych.49.1.479>
- van der Linden, W. J., & Hambleton, R. K. (Eds.). (1997). *Handbook of modern item response theory*. New York: Springer.

---

О. О. Диховичний<sup>1</sup>, Н. В. Круглова<sup>1</sup>, І. О. Москальов<sup>2</sup> (2018). Використання математичних моделей для аналізу результатів психологічного тесту Гілфорда. *Mathematics in Modern Technical University*, 2018(1), 75–89.

Submitted: 2018-10-01

Accepted: 2018-11-14

O. O. Dykhovychnyi, N. V. Kruglova, I. O. Moskalov (2018). Using of mathematical models for analyzing of the results in psychological Guilford test. *Mathematics in Modern Technical University*, 2018(1), 75–89.

**Abstract.** Results of the survey, conducted using Joy Guilford methodology, of 800 servicemen regarding their professional competence was analyzed with IRT models. In particular, Rasch model and 1-PL, 2-PL and 3-PL models were used. All computations were performed with R statistical software along with ltm package. Latent parameters for both respondents and corresponding indicators (survey questions) were obtained. It was found that there exists significant discrepancy between the models and results of the survey. Analysis of this discrepancy has shown that the major cause of their existence are inadequate conditions under which the survey was conducted.

**Keywords:** Guilford test; social intelligence; IRT; Rasch model; PL models; latent parameters.