

УДК 378

DOI: 10.15587/2313-8416.2015.57049

ВИКОРИСТАННЯ СТАТИСТИЧНИХ МЕТОДІВ У ПЕДАГОГІЧНИХ ДОСЛІДЖЕННЯХ

© В. М. Шкляр

У статті висвітлено питання актуалізації та розвитку знань у галузі педагогічної інноватики, описано модель статистичного оброблення результатів педагогічного експерименту. Представлено методи виявлення збігів/відмінностей досліджуваних характеристик в експериментальній і контрольній групах. Запропоновано рекомендації щодо формулювання статистичних гіпотез та вибору статистичного критерію. У статті розглянуті можливості застосування програми "Педагогічна статистика" в обробці та аналізі даних емпіричного дослідження

Ключові слова: психолого-педагогічні дослідження, модель педагогічного експерименту, педагогічна статистика, гіпотези дослідження, статистичні методи, статистичні критерії

An issue of actualization and development of knowledge in educational innovatics is highlighted in the article. Model of statistical results processing of pedagogical experiment is described. Methods of detection coincidences/differences of investigated characteristics in the experimental and control groups are presented. Recommendations about formulating statistical hypotheses and statistical selection criteria are offered. Possibilities of usage of «Pedagogical statistics» program for processing and analyzing the empirical data are considered in the article

Keywords: psychological and pedagogical research, pedagogical statistic, pedagogical model of the experiment, research hypotheses, statistical methods, statistical criteria

1. Вступ

Очевидне розширення виховного простору в сучасному суспільстві незмірно підвищує роль загальноосвітніх навчальних закладів у процесі відповідальності за виховання особистості. Лише за умов дійсної реалізації принципу єдності навчання та виховання, а також надання *провідного характеру* навчальній діяльності загальноосвітні навчальні заклади стануть спроможними ефективно здійснювати функцію урахуванням відмінностей вікових можливостей учнів на етапі засвоєння знань. Психологічні дослідження, робота над розробкою та впровадженням нового методу навчання (В. В. Давидов, С. Д. Максименко, В. В. Рєпкін, В. В. Рубцов та ін.) зумовили висновок, що використання активних методів у процесі навчання сприяє фопосередника між іншими інститутами виховання та забезпечити цілісний розвиток особистості. Ш. А. Амонашвілі, В. В. Давидов, Л. В. Занков, В. В. Рєпкін, Д. Б. Ельконін та ін., провели низку психолого-педагогічних досліджень, результати яких засвідчують, що вирішення цього завдання пов'язано саме з докорінними змінами змісту та методів навчання у ЗНЗ. Загальновідомо, що у 60-ті роки було виявлено новий зміст та розроблено нові методи навчання з рмуванням продуктивного (творчого та теоретичного) мислення, а також суттєво впливає на перебіг психічного розвитку учня [1]. У 70–90 рр. ХХ ст. внаслідок упровадження експериментальних систем навчання був практично доведений зв'язок між процесами навчання та виховання, що дозволило дослідити та у повному обсязі вивчити закономірності розвитку особистості залежно від відмінностей у способах навчання (концепція В. В. Давидова та Д. Б. Ельконіна) [2].

2. Аналіз та постановка проблеми

Здійснюючи обробку отриманих в дослідженнях результатів, педагоги, як правило, використовують

найпростіші математичні дії, використовуючи при цьому порядкові шкали. Але такі дії не завжди мають сенс. Наприклад, обсяг знань двох "двієчників" не еквівалентний обсягу знань одного "хорошиста" (хоча $2+2=4$). Або при оцінюванні змін обсягу знань від "трієчника" до "відмінника" він збільшується так само, як при переході від "двієчника" до "хорошиста" (хоча $5-3=4-2$). І це не залежить від того, яка шкала використовується: 5-бальна, 12-бальна чи 100-бальна. Очевидно, що використані для оцінки знань бали дозволяють порівнювати знання учнів, але ці числа безглуздо складати і віднімати, а також множити і ділити.

Також, не дає підстав для об'єктивних висновків і досить часто використовуваний на практиці освітніми установами "середній бал". З його допомогою оцінюють успішність окремого учня (середній бал з оцінок, виставлених з різних предметів) або порівнюють групи учнів (усереднюються оцінки різних учнів з одного і того ж предмета). Зазвичай під середнім балом мають на увазі середнє арифметичне. Але чи можна складати оцінки, виміряні за допомогою порядкової шкали?

Метою експерименту, в тому числі в педагогічному дослідженні, є емпіричне підтвердження або спростування гіпотези дослідження та/або справедливості теоретичних результатів.

3. Статистичні методи в дослідженнях

Розглянемо наступну *модель педагогічного експерименту*. Оберемо деякий педагогічний *об'єкт*, зміна *стану* якого досліджується в ході експерименту. В якості об'єкта (згідно з працями методологів педагогічного наукознавства окремих індивід, група, колектив є об'єктами педагогічного впливу, але не педагогічного експерименту. Об'єктом педагогічного експерименту є педагогічний процес (навчання, ви-

ховання, розвиток, професійна підготовка та ін.) чи системи (дидактична, методична та ін., ефективність яких перевіряється на певній групі осіб. У психології досліджуються стани і якості цих осіб, але не самі по собі особи) може виступати окремих індивід, група, колектив і т. д., наприклад, безліч учнів, яких навчають за новою методикою. Стан об'єкта вимірюється тими чи іншими критеріями відповідно до його істотних характеристик. Прикладами критеріїв є: успішність, рівень знань, особистісні риси чи характеристики тощо [3].

Таким чином, для того, щоб виділити в явному вигляді результат цілеспрямованого впливу на досліджуваній об'єкт, необхідно взяти аналогічний об'єкт і подивитися, що відбувається з ним за відсутності впливів.

Традиційно ці два об'єкти в експериментальних дослідженнях називають відповідно *експериментальною групою* та *контрольною групою* (рис. 1.) Досліджуються не групи, а якості, які перевіряються на вибірках учасників цих груп відповідно до критеріїв і показників.

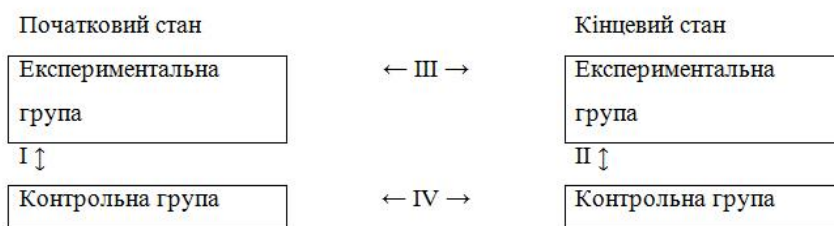


Рис. 1. Структура педагогічного експерименту (але можна у цій структурі додати експериментальний чинник)

Отже, для обґрунтування висновків про вплив досліджуваних в експерименті умов треба коректно і достовірно обґрунтувати збіг чи розходження станів контрольної та експериментальної групи.

Типовим завданням аналізу даних у педагогічних дослідженнях є встановлення збігів або відмінностей досліджуваних характеристик експериментальної і контрольної групи. Для цього формулюються статистичні гіпотези:

- гіпотеза про відсутність відмінностей (так звана *нульова гіпотеза*);
- гіпотеза про значущість відмінностей (так звана *альтернативна гіпотеза*).

Для прийняття рішень про те, яку з гіпотез (нульову або альтернативну) слід прийняти, використовують вирішальні правила – *статистичні критерії*. Тобто, на підставі інформації про результати спостережень (характеристиках членів експериментальної і контрольної групи) обчислюється число, зване *емпіричним значенням* критерію. Це число порівнюється з відомим еталонним числом, званим *критичним значенням* критерію.

Критичні значення наводяться, як правило, для декількох *рівнів значущості*. Рівнем значущості називається ймовірність помилки, що полягає у відхиленні нульової гіпотези від гіпотези, тобто ймовірність того, що відмінності вважаються істотними,

Але реально вони є випадковими. У педагогічних дослідженнях зазвичай обмежуються значенням 0,05 (допускається не більше 5 % помилки).

Якщо отримане дослідником емпіричне значення критерію виявляється менше або дорівнює критичному, то приймається нульова гіпотеза – вважається, що на заданому рівні значущості характеристики експериментальної і контрольної груп збігаються. В іншому випадку, якщо емпіричне значення критерію виявляється більше за критичне, то нульова гіпотеза відкидається і приймається альтернативна гіпотеза – характеристики експериментальної і контрольної групи вважаються різними з достовірністю відмінностей $1 - a$. Наприклад, якщо $a=0,05$ і прийнята альтернативна гіпотеза, то *вірогідність розходжень* дорівнює 0,95 або 95 %.

Іншими словами, чим менше емпіричне значення критерію (чим лівіше воно знаходиться від критичного значення), тим більше ступінь збігу характеристик порівнюваних об'єктів. І навпаки, чим більше емпіричне значення критерію (чим правіше воно знаходиться від критичного значення), тим сильніше розрізняються характеристики порівнюваних об'єктів.

Для таких, вимірених згідно зі *шкалою відносин*, для перевірки гіпотези про збіг характеристик двох груп доцільним є використання або *критерію Крамера-Уелча*, або *критерію Вілкоксона-Манна-Уїтні*. Критерій Крамера-Уелча призначений для перевірки гіпотези про рівність середніх (математичних очікувань) двох вибірок, критерій Вілкоксона-Манна-Уїтні є більш "тонким" (але і більш трудомістким) – він дозволяє перевіряти гіпотезу про те, що дві вибірки "однакові" (в тому числі, що збігаються їхні *середні, дисперсії і всі інші показники*).

Для даних, вимірених за *порядковою шкалою* доцільно використовувати *критерій "Хі χ^2 -квадрат"*.

Для даних, вимірених за *дихотомічною шкалою* доцільне використання *критерію Фішера* [4].

Алгоритм вибору статистичного критерію

У першому наближенні цей алгоритм надзвичайно простий: якщо дані отримані в результаті вимірів за шкалою відносин, то слід використовувати критерій Вілкоксона-Манна-Уїтні (ВМУ), якщо за порядковою шкалою, то критерій $X_1 \chi^2$ -квадрат.

Якщо число значень, які розрізняються між собою в порівнюваних вибірках велике (більше десяти), то доцільно використовувати критерій ВМУ.

Якщо число розрізнення між собою значень в порівнюваних вибірках мале (менше десяти), то, зробивши групування результатів вимірювань (тобто, перейшовши від шкали відносин до порядкової шкали), можна використовувати критерій χ^2 .

Якщо обсяг вибірок малий (<50), то слід використовувати критерій ВМУ (при малому числі різних значень характеристик в цьому випадку можна використовувати і критерій χ^2).

Якщо обсяг вибірки великий, то, знову ж таки за допомогою групування результатів вимірювань має сенс використовувати критерій χ^2 .

Для порядкової шкали у разі, коли число градацій більше або дорівнює трьом, використовується критерій χ^2 , якщо ж застосовувалася дихотомічна шкала, то можна використовувати або критерій χ^2 , або критерій Фішера.

Обмеження стосовно використання статистичних критеріїв

Критерій ВМУ. Кожна вибірка повинна містити не менше трьох елементів, якщо ж в одній з вибірок всього два елементи, то в інший їх має бути не менше п'яти.

Критерій Крамера-Уелча. Обмежень на використання критерію немає.

Критерій χ^2 . Застосовний за умови, що для будь-якого значення в будь-якій з порівнюваних вибірок не менше п'яти її членів повинні мати Це значення і кількість градацій шкали має бути не менше трьох.

Критерій Фішера. Критерій застосовний тільки до даних, виміряних за дихотомічною шкалою (тобто за шкалою, що містить лише дві різні градації).

Обчислення статистичних критеріїв може здійснюватись як "вручну", так і за допомогою програмових засобів. У педагогічних дослідженнях доцільно використовувати спеціальну програму "Педагогічна статистика", яка призначена для аналізу даних, отриманих в результаті педагогічного дослідження [5].

Автоматично розраховуються показники описової статистики:

- обсяг вибірки;
- мінімум;
- максимум;
- інтервал (розмах) ;
- сума;
- середнє;
- медіана;
- дисперсія.

Вибірки порівнюються з використанням таких статистичних критеріїв:

– для даних, виміряних в шкалі *відносин*:

- критерій Вілкоксона-Манна-Уїтні;
- критерій Крамера-Уелча;
- критерій Хі-квадрат;
- для даних, виміряних в порядковій шкалі:

- критерій Хі-квадрат;
- критерій Фішера.

Також доступний режим автовизначення критерію, який працює у відповідності з алгоритмом вибору статистичного критерію (рис. 2) [6].

Введені дані зберігаються у файлі у власному форматі *.psf

Робота з файлами (створення, відкриття, збереження, закриття) може здійснюватися трьома способами:

1. З меню "Файл", вибором відповідного пункту.

2. З панелі інструментів, вибором відповідної кнопки.

3. За допомогою "гарячих клавіш", які вказані праворуч від назви пункту в меню.

Для початкового введення даних потрібно:

1. Створити новий документ, обравши пункт меню "Создать".

2. Обрати шкалу, в якій здійснювалися вимірювання:

3. Обрати метод введення даних ("Індивідуальні дані" або "Сумарні дані").

4. Перейти на вкладку, що відповідає обраному методу, натиснувши на неї.

5. Ввести дані в таблицю.

5. 1. Введення індивідуальних даних. Дані можна вводити як з клавіатури, так і шляхом вставки через буфер обміну.

5. 2. Введення сумарних даних:

Дані можна вводити як з клавіатури, так і шляхом вставки через буфер обміну. Нумерація рядків проводиться автоматично (рис. 3).

Виходячи з введених даних, автоматично розраховуються показники описової статистики. Для перегляду треба перейти на вкладку "Описова статистика" (рис. 4).

Для перегляду результатів порівняння вибірок потрібно перейти на вкладку "Аналіз", на якій відображаються результати *парного* порівняння вибірок (рис. 5):

– синім кольором відображаються результати порівняння співпадаючих вибірок;

– червоним кольором відображаються результати порівняння вибірок, що розрізняються;

– чорним кольором відображаються причини, з яких порівняння відповідних вибірок неможливо.

Будь-яке значення або всю таблицю цілком можна скопіювати в буфер обміну і вставити в будь-яку програму, що підтримує роботу з буфером обміну (на сьогоднішній день це дозволяє виконати практично будь-яка програма, що працює під управлінням Windows).

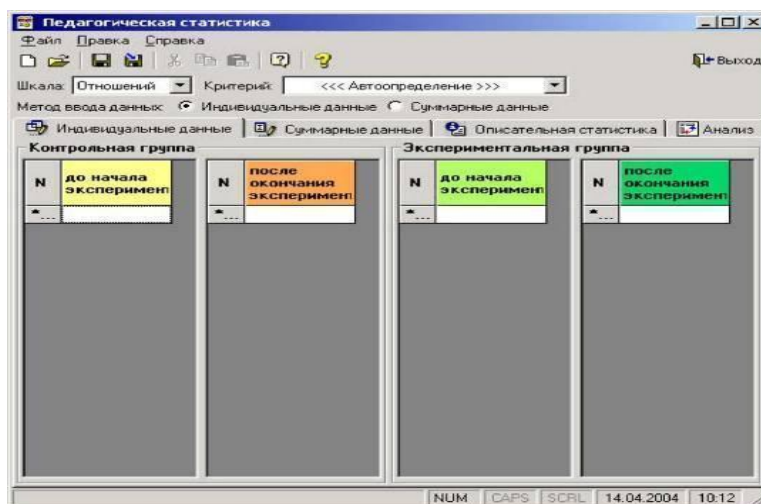


Рис. 2. Інтерфейс програми Педагогічна статистика

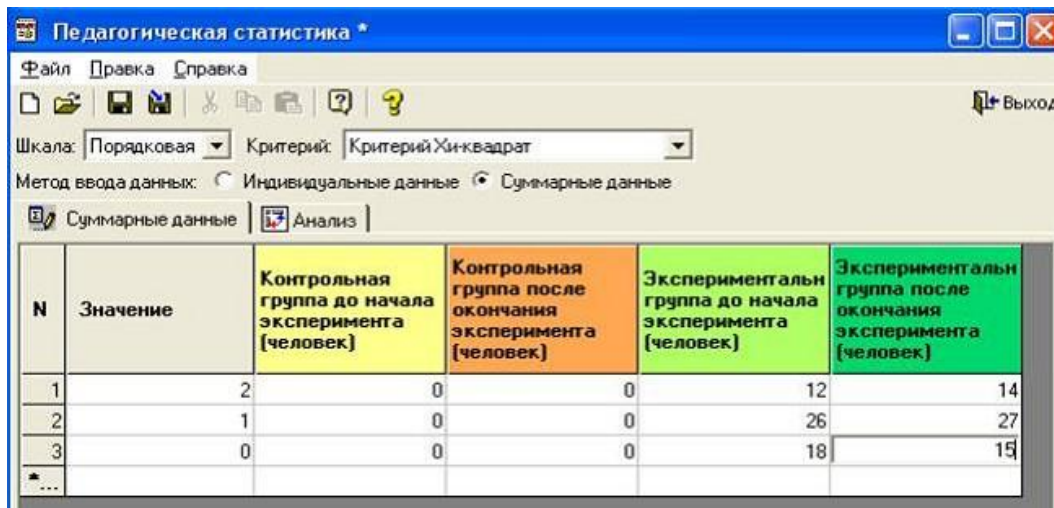


Рис. 3. Комп'ютерна програма "Педагогічна статистика" вкладка ведення сумарних даних

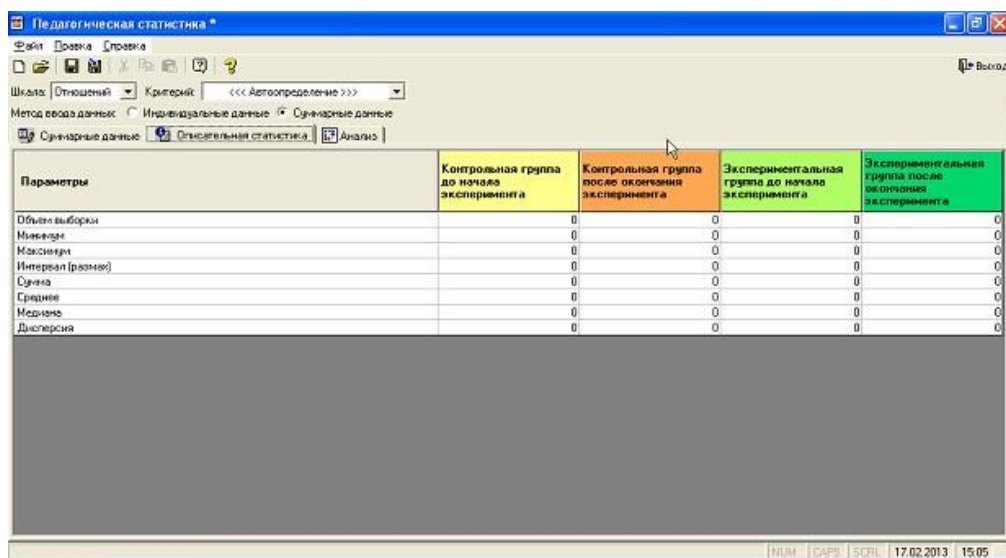


Рис. 4. Комп'ютерна програма "Педагогічна статистика" обчислює величини описової статистики для даних за шкалою відносин а

	Контрольная группа до начала эксперимента	Контрольная группа после окончания эксперимента	Экспериментальная группа до начала эксперимента	Экспериментальная группа после окончания эксперимента
Контрольная группа до начала эксперимента		Эмпирическое значение критерия Вилкоксона-Манна-Уитни 0,3031, критическое 1,96. Характеристики сравнимых выборок совпадают на уровне значимости 0,05.	Эмпирическое значение критерия Вилкоксона-Манна-Уитни 0,0338, критическое 1,96. Характеристики сравнимых выборок совпадают на уровне значимости 0,05.	Эмпирическое значение критерия Крамера-Валла 2,1568, критическое 1,96. Достоверность различий характеристик сравнимых выборок составляет 95%.
Контрольная группа после окончания эксперимента	Эмпирическое значение критерия Вилкоксона-Манна-Уитни 0,3031, критическое 1,96. Характеристики сравнимых выборок совпадают на уровне значимости 0,05.		Эмпирическое значение критерия Вилкоксона-Манна-Уитни 0,3803, критическое 1,96. Характеристики сравнимых выборок совпадают на уровне значимости 0,05.	Эмпирическое значение критерия Вилкоксона-Манна-Уитни 2,1974, критическое 1,96. Достоверность различий характеристик сравнимых выборок составляет 95%.
Экспериментальная группа до начала эксперимента	Эмпирическое значение критерия Вилкоксона-Манна-Уитни 0,0338, критическое 1,96. Характеристики сравнимых выборок совпадают на уровне значимости 0,05.	Эмпирическое значение критерия Вилкоксона-Манна-Уитни 0,3803, критическое 1,96. Характеристики сравнимых выборок совпадают на уровне значимости 0,05.		Эмпирическое значение критерия Вилкоксона-Манна-Уитни 2,0179, критическое 1,96. Достоверность различий характеристик сравнимых выборок составляет 95%.
Экспериментальная группа после окончания эксперимента	Эмпирическое значение критерия Крамера-Валла 2,1868, критическое 1,96. Достоверность различий характеристик сравнимых выборок составляет 95%.	Эмпирическое значение критерия Вилкоксона-Манна-Уитни 2,1974, критическое 1,96. Достоверность различий характеристик сравнимых выборок составляет 95%.	Эмпирическое значение критерия Вилкоксона-Манна-Уитни 2,0179, критическое 1,96. Достоверность различий характеристик сравнимых выборок составляет 95%.	

Рис. 5. Результати попарного рівняння у вікні комп'ютерної програми "Педагогічна статистика"

4. Результати дослідження

Наведені вище статистичні викладки не охоплюють всіх наявних можливостей обробки експериментальних даних із застосуванням комп'ютерних програм. Вони наведені як приклад своєрідної технології їх аналізу за допомогою методів математичної статистики, а також ілюстрації можливостей програми "Педагогічна статистика" у структуруванні та візуальному представленні даних. Крім того, очевидно є значна економія часу дослідника, що користується комп'ютерними програмами (факторний аналіз, наприклад, взагалі неможливо провести вручну через громіздкість обчислень). Таким чином, комп'ютерні технології дозволяють значно інтенсифікувати обробку та аналіз даних прикладних досліджень і відкривають широкі можливості їх використання в різних галузях науки.

5. Висновки

Теоретичний аналіз отриманих результатів дав можливість виявити виникнення зміни егоцентричної позиції учня на об'єктивну (від суб'єктивного ставлення змінювати дану пізнавальну перспективу оточуючого світу до етичного, пізнавального), але лише за умов докорінних змістовних змін та методів активного навчання.

Таким чином не постало завдання перепідготовки вчителів нового типу, які мають глибокі знання у галузі психології навчання та розвитку особистості учня. У наш час руйнується консерватизм у педагогіці, це зумовлює інтенсивний пошук нових педагогічних технологій, оновлення змісту навчальних програм, форм та методів досліджень, залучення широкого спектру інновацій в освітній процес.

Отже, школа сьогодні – це територія прогресу. А розвиток школи – це процес засвоєння інновацій. Тобто розвиток можливий лише за рахунок нового: технологій, підходів, методів, змісту роботи, організаційних структур і т. д.

Література

1. Ващенко, Г. Г. Загальні методи навчання. Підручник для педагогів [Текст] / Г. Г. Ващенко. – К.: Всеукраїнське Педагогічне Товариство ім. Г. Ващенка, 1997. – 410 с.
2. Методи педагогічних досліджень. Бібліотека онлайн [Електронний ресурс]. – Київ, МОН, 2007. – Режим доступу: <http://readbookz.com/book/172/5401.html>
3. Ангеловски, К. Учителя и инновации. Книга для учителя [Текст] / К. Ангеловски. – М.: Просвещение, 1991. – 159 с.
4. Гласс, Дж. Статистические методы в педагогике и психологии [Текст] / Дж. Гласс, Дж. Стэнли. – М.: Прогресс, 1976. – 496 с.
5. Горбунова, Н. В. Внутрешкольное управление: теория и опыт педагогических и управленческих инноваций [Текст] / Н. В. Горбунова. – М.: Педагогическое общество России, 1995. – 47 с.
6. Чепан, М. Л. Проблемы обеспечения валидности эксперимента в условиях виртуальной среды [Текст] / М. Л. Чепан, Н. М. Бугаева. – Экспериментальный метод в структуре психологического знания. – М.: "Институт психологии РАН", 2012. – С. 82–86.

References

1. Vashchenko, G. G. (Ed.) (1997). *Zagal'ni metodi navchannya. Pidruchnik dlya pedagogiv* [Common methods. Textbook for teachers]. Kyiv: All-Ukrainian Pedagogical Society of G. Vashchenko, 410.
2. *Methods of educational research. Library online*. Kyiv, Ministry of Education. Available at: <http://readbookz.com/book/172/5401.html>
3. Angelovski, K. (Ed.) (1991). *Uchitelya i innovatsii. Kniga dlya uchitelya* [Teachers and innovation. Book for teachers]. Moscow: Education, 159.
4. Glass, J., Stanley, J. (1976). *Statistical Methods in Pedagogy and Psychology*. Moscow, Progress, 496.
5. Gorbunova, N. V. (Ed.) (1995). *Vnutrishkol'noye upravleniye: teoriya i opyt pedagogicheskikh i upravlencheskikh innovatsiy* [Intraschool management theory and experience of pedagogical and management innovations]. Moscow: Russian Pedagogical Society, 47.
6. Chepan, M. L., Bugaev, N. M. (2012). *Problems of validity of the experiment in a virtual environment/experimental method in the structure of psychological knowledge*. Moscow: "Russian Academy of Sciences Institute of Psychology", 82–86.

*Рекомендовано до публікації д-р пед. наук, професор Отич О. М.
Дата надходження рукопису 18.11.2015*

Шкляр Вікторія Миколаївна, аспірант, відділ моніторингу обдарованості дітей та молоді, Інститут Обдарованої Дитини НАПН України, вул. Артема, 52-д, м. Київ, Україна, 04053
E-mail: info.shklyar@gmail.com