

Conducting an article critique for a quantitative research study: perspectives for doctoral students and other novice readers

David E Vance¹
Michele Talley¹
Andres Azuero¹
Patricia F Pearce²
Becky J Christian¹

¹School of Nursing, University of Alabama at Birmingham, Birmingham, AL, USA; ²Loyola University School of Nursing, New Orleans, LA, USA

Abstract: The ability to critically evaluate the merits of a quantitative design research article is a necessary skill for practitioners and researchers of all disciplines, including nursing, in order to judge the integrity and usefulness of the evidence and conclusions made in an article. In general, this skill is automatic for many practitioners and researchers who already possess a good working knowledge of research methodology, including: hypothesis development, sampling techniques, study design, testing procedures and instrumentation, data collection and data management, statistics, and interpretation of findings. For graduate students and junior faculty who have yet to master these skills, completing a formally written article critique can be a useful process to hone such skills. However, a fundamental knowledge of research methods is still needed in order to be successful. Because there are few published examples of critique examples, this article provides the practical points of conducting a formally written quantitative research article critique while providing a brief example to demonstrate the principles and form.

Keywords: quantitative article critique, statistics, methodology, graduate students

Introduction

Evidence-based practice and research are the cornerstones of effective health care and scientific pursuits.^{1,2} Without the objective interpretation of data, breakthroughs and advancements in practice and research would cease. The driving force of evidence-based practice and research in the traditional sense is the ability to measure and quantify a phenomenon, as well as the relationships between phenomena numerically, whether the phenomenon is biological, behavioral, economic, or emotional in scope (Figure 1). This ability allows specific, testable research questions and hypotheses to be proposed so that statistical procedures can be applied to a cacophony of numbers in an attempt to reduce the data into discrete and tractable patterns; this process is often referred to as parsimony and is one of the goals of science.³ From such patterns, objective interpretations can be derived, which over time, in tandem with other interpretations from other researchers investigating the same phenomenon, can drive the application of practices in real life, as well as the development of future research questions that continue to benefit both practice and research. This cascading process can eventually change the phenomenon of interest. For example, the ability to measure CD4+ lymphocytes and human immunodeficiency virus (HIV) plasma load in patients with HIV has allowed teams of dedicated researchers to test various agents to determine whether or not the agents are capable of reducing the amount of virus in the blood and boost immune response. In this way, exhaustive trial and error studies have eventually led to the development of highly active antiretroviral therapy, which has drastically

Correspondence: David E Vance
NB 456, 1701 University Boulevard,
University of Alabama at Birmingham,
Birmingham, AL 35294-1210, USA
Tel +1 205 934 7589
Fax +1 205 996 7183
Email devance@uab.edu

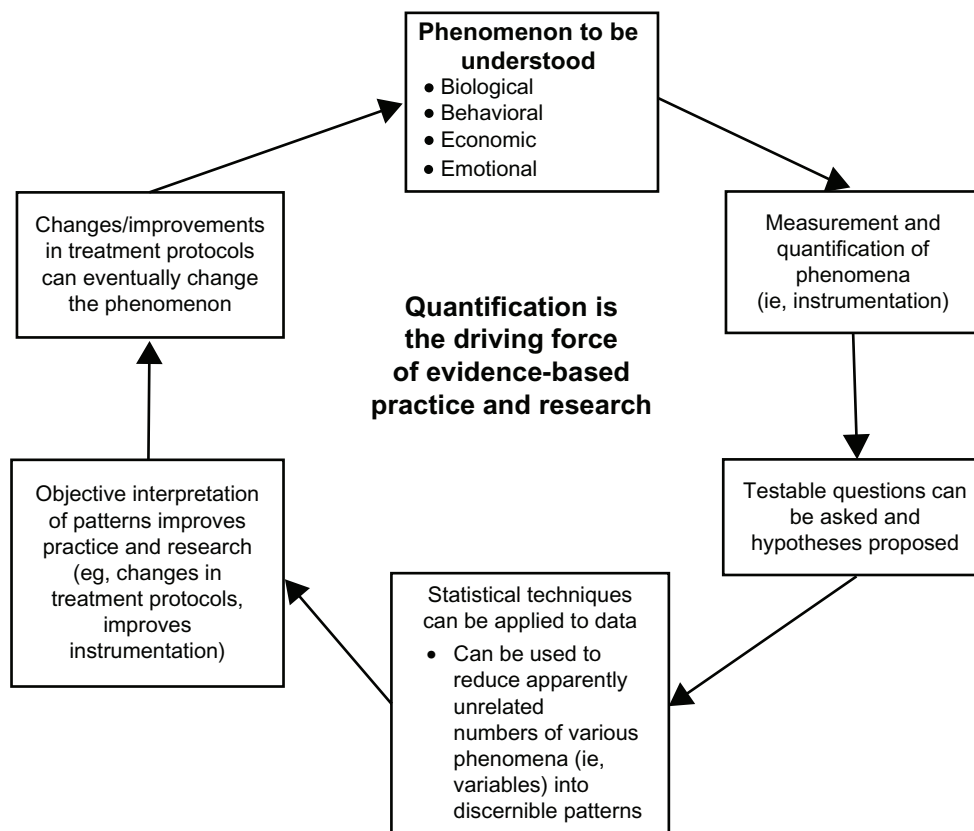


Figure 1 The process of quantitative data-generated evidence.

changed the phenomenon of HIV-related mortality; in fact, the phenomenon of aging with HIV has become an emerging topic of research, as practitioners of several disciplines examine ways to improve health and quality of life in those surviving and growing older from this disease.⁴ Thus, the critical evaluation of the scientific evidence available in the literature, required of researchers and practitioners, has resulted in a continuing, positive cascade of improved treatment practices and research methods that have also enhanced patient care of those infected with HIV.

For many experienced practitioners and researchers, critically evaluating research articles becomes an automatic behavior given their familiarity and involvement with the research process and ease of understanding the published literature. This ease of understanding is a direct result of constant immersion in the literature that arises from frequent reading, which is an essential part of the role of the professional. It is for this reason that published article critiques are rarely found in the scientific literature because most doctorally prepared scientists, both practitioners and researchers, have mastered this skill and do not require someone to interpret the merits of a journal article for them. Moreover, such scientists may be reading a particular article

only to glean one or two points from the article (eg, to see how others recruited from a specific population, because they plan to conduct a similar study) relevant to a specific professional need. In addition, given the cost of labor, resources, and materials required to publish an article,⁵ scientific journals strongly prefer to publish original research rather than a secondary interpretation or critique of a research study. Nevertheless, the skill of critically evaluating published research is essential for the professional.

The purpose of this article is to provide doctoral students, as well as other graduate students and junior faculty, or other novice readers of research, with a balanced perspective on how to critique an article given that this article was written by a psychologist, a biostatistician, and three nurses (with one of them being a current doctoral student herself). More specifically, balancing the methodological and statistical advantages and disadvantages of an article will be reviewed along with considerations of how to evaluate the utility of an article. The exact methodological and statistical skills needed to conduct a quantitative article critique are not explicitly provided in this article (which was far beyond its scope); the article provides: (1) the practical philosophy of conducting an article critique; and (2) an example of how a

written quantitative article critique may be structured as a guide for a future critique. In this way, practical “realistic” points are provided.

Balanced perspective on critiquing an article

In general, a formally written article critique is usually only conducted as a didactic exercise for graduate students developing skills in research methodology and statistics. Conducting a critique serves as a method for students to review the essential scientific elements of a research article against criteria for a well written article, and formulate an informed assessment of the strengths and weaknesses of each component of the research, concluding in an overall evaluation, whether the conclusions and/or recommendations merit consideration for changing practice or incorporating into research. More specifically, this recommendation ranges from “not worthy of consideration” to “very worthy of consideration.” But it is important for graduate students to be mindful that it is usually not a single article that forms the basis of a professional opinion about a phenomenon, but several articles that when taken together drive the field forward. In fact, it may be that some scientifically flawed studies actually energize the field and encourage researchers to examine their phenomenon of interest in new ways. Thus, one’s professional opinion about a phenomenon should be based upon constant reading and interpretation of the scientific literature across disciplines.

In interpreting the scientific merits of an article, there are several article critique review guidelines available in the literature from nursing,^{6–8} medicine,^{8–10} and psychology and sociology.¹¹ Each of these review guidelines has merits and is consistent in structure and content by proposing what graduate students should be examining to judge the value of the findings (eg, sample selection, sample size, hypothesis, and appropriateness of statistical tests). Textbooks^{1,2,7} provide similar information, but in a much lengthier and more detailed form. Despite seemingly clear instructions on what to consider when evaluating a published research study report, the instructions are meaningless if the graduate student does not possess a sufficient and working knowledge of foundational elements and basics of research methodology and statistics. Unfortunately, many blindly follow available guidelines, checking “yes” or “no” to a particular item on a critique guideline or checklist, without being able to articulate a clear rationale or understanding of why the study may be well-conducted or compromised scientifically. For example, a critique guideline may suggest that the student

evaluate whether the sample size is sufficient for statistical purposes and for generalizing the findings. Thus, if the sample size is small (eg, $N = 14$), the graduate student may remark that this is too small for the findings to be generalizable to the larger population² and, in general, the determination of insufficient sample would be correct; however, if referencing a qualitative design study, a small sample size is considered appropriate and a sample of 14 cases would be acceptable.^{12–14} Additionally, small samples are also found in descriptive quantitative studies, pilot feasibility studies, or proof of concept research, as well as in cross-sectional design and even in some clinical trial studies involving multiple data collection points or a hard to capture sample,^{2,7} for example those with a rare disease such as progeria, a disease that causes premature aging beginning in childhood. Even a small study with a sample of 14 cases would be of value in learning more about progeria, although it may still be challenging to generalize the findings to the larger population of those with this disease. For many graduate students, the tendency to see research in black and white terms can be troubling, when considering that there is no such thing as a perfect study and that many design decisions have been made by balancing the added value of the element with excessive time and cost. As such, even research with numerous major flaws may have at least some limited value to the scientific literature, even if only to serve as a negative model or suggest how to conduct such research better in future endeavors.

This black and white perspective may also be seen when graduate students critique on power calculations. For example, Polit and Beck² suggest that many nursing studies and most studies, in general, do not report power, or are not based on power calculations. In a perfect world, it would be nice for studies to be based on power calculations and then report on that calculation and observed power. Again, students may get hung up on this suggestion and think that because a researcher did not report power, the researchers must be hiding something and, thus, the findings of the study are suspect or even fraudulent. The reality is that more likely there are several reasons why researchers may not report power. First, if there is more than one statistical analysis, which there often is in studies, this could introduce even more detail to an already complicated article that would reduce the readability of the report; thus, the researchers may have left out power calculations purposely or in some cases may have been urged to do so by the journal editor or reviewers. Second, by providing the sample size, the value of the statistical test, P -value, and effect size, the power can be calculated anyway; so for someone statistically savvy,

power is already embedded in the article, but quite difficult for the novice reader to identify. Third, if the sample size is extremely large, say from a population-based study of two or three thousand participants, sufficient power is automatically assumed even if the effect size of the relationship(s) being examined is small. In fact, in a scenario with dozens or even hundreds of tests, alpha correction techniques are needed (ie, Bonferroni, Benjamini-Hochberg) to set a more stringent significance level and minimize type I error.² Thus, as with this example, it takes a thorough understanding of statistics and a practical understanding of the research process including publication constraints in order to use such article critique guidelines effectively.

Besides these methodological and statistical criteria provided in such article critique guidelines, there are some more cosmetic guidelines that may confuse graduate students as they learn to examine the science reflected in research articles. For example, Polit and Beck² suggest that the title should be less than 15 words and include the independent and dependent variables (or phenomenon) in the study. Ideally, investigators might follow this recommendation, but in reality, there may be too many variables in the research to practically mention them all in a title, and in a title of less than 15 words, it might be impossible. In many studies (eg, Okonkwo et al,¹⁵ Preiss et al,¹⁶ Suzuki-Crumly et al,¹⁷ Vance et al,¹⁸⁻²⁰), there are several independent, dependent, and covariate variables, so that listing all variables in the title would be unwieldy and cumbersome; if the researchers did, the journal editor or reviewers would likely urge the researchers to change the title by selecting only two or three of the key variables to include. Thus, if a graduate student blindly followed Polit and Beck's² suggestion in critiquing a research article, the student might mistakenly conclude that because all of the variables in the study were not included in the title, the study is flawed. Although a title is important in delineating relevant information about the study to the reader, the title has no clear bearing on the scientific merits of a study, although the use of key words or variables in the title enhances the reader's ability to determine the relevance and usefulness of the study.

In general, although there are many good guidelines available to direct critique of a research article, there is no prescribed way to do so. All such guidelines still require that the person conducting a critique has a basic understanding of research methodology and statistics. Furthermore, all studies vary and focus on different populations, employ different sampling designs, use different study designs (eg, experimental, case study, or mixed methods), utilize sundry

behavioral and/or biological instruments, and perform various statistical and/or qualitative analyses specific to the researcher(s)' aims and hypotheses. Thus, a well-reasoned critique, requires familiarity with all these procedures, as well as knowing how they fit together, to form an overall idea of what the study did and how well it was done. These issues are particularly important for the novice to understand, because novice readers of research expect that the research design and other components should be stated clearly in the article. However, researchers often assume that their readers should know which design they are using (eg, experimental, cross-sectional) simply from the description, and researchers may not articulate in their published reports the exact design of the research. That lack of clarity befuddles the novice reader who unfortunately expects to be able to use a checklist for items being critiqued.

In addition, a generous understanding of the business of science and the publication process is helpful in critiquing an article and setting expectations for the reader. For example, some journals actually prefer that if particular instruments are commonly used in the literature, that they should only be mentioned but not described in the article to save on journal space (eg, *Journal of the International Neuropsychological Society*). So it may appear to a novice reviewer that the researchers failed to report, say, the psychometric properties of survey instruments (ie, Cronbach's alpha, test-retest reliability) and that this is a major flaw. Unfortunately, this black and white perspective adopted by many graduate students who use such critiquing guidelines causes these students to end up evaluating the article harshly and attend to the letter of the law and not the spirit or they "throw the baby out with the bathwater" to share two related idioms. In other words, some students focus on a minor detail and generalize those minor infractions to claim that the entire study is flawed and the information generated from it should be disregarded. For example, several research textbooks claim that a multi-item instrument for a study should have a Cronbach's alpha of 0.7 or higher.²¹ Yet, if a study employed a few survey instruments and the Cronbach's alpha for each ranged from 0.61 to 0.68, this may be a minor weakness of the study that should probably be acknowledged, but it does not disqualify the study from having scientific merit. Moreover, some elements of a study are more important and critical to the study design than others, yet many students treat each element as having equal value.

As mentioned, several fine guidelines are already available to help graduate students or novice readers to critique research articles. Unfortunately, there are few

examples in the published literature of how such a critique is structure or conducted. Thus, there is need for an example critique to help guide novices in the science, and especially art, of critiquing. Rather than rehashing what questions to ask in the critique of a research article, the second part of this article includes an actual quantitative article critique which can serve as a practical example of critiquing. Following the example, a few pointers are reviewed to help graduate students balance these guidelines with the practical approach of evaluating the scientific merits.

A brief article critique

Summary

Danno et al²² surveyed 59 homeopathic physicians and their patients ($N = 168$) on whether homeopathy was effective in reducing the frequency and intensity of migraine in children. In addition, these researchers documented the most common homeopathic medicines used to prevent and treat migraine attacks. In general, they claimed that homeopathic medicines may be effective in preventing and reducing the intensity of migraines, which also may result in fewer school days missed. Furthermore, they found that *ignatia amara*, *lycopodium clavatum*, *natrum muriaticum*, and *gelsemium* were the most commonly reported homeopathic medicines used to prevent migraines, and *belladonna*, *ignatia amara*, *iris versicolor*, *kalium phosphoricum*, and *gelsemium* were the most commonly used for treatment of acute migraine attacks.²²

Study rationale

As provided in the brief literature review, migraines occur in 3% to 10% of children and can be so intense that they can disrupt family life and interfere with education due to missed school days. Homeopathy has been used to treat migraine in adult patients aged 16 and older and in children with chronic diseases; however, there appears to be little information on the use of homeopathy in children suffering from migraines. Thus, if this literature review in Danno et al²² was accurate, this study appeared to fill a gap in the literature.

Sample

The sample was composed of both homeopathic physicians and their child patients with their parents/guardians. While this type of convenience sampling can be considered less rigorous than random selection of participants,^{7,23} it was a reasonable and traditional approach given that it may require more resources to recruit participants and ensure their compliance to a treatment protocol such as in a randomized clinical trial. One hundred and sixty-two homeopathic

physicians were contacted who had received training in homeopathy. It was not clear how the researchers identified those physicians who had received homeopathic training. Because the study was sponsored by Boiron Laboratories, Lyon, France and 45% ($n = 27$) of the physicians in the study were from France, it is possible that Boiron Laboratories manufactures homeopathic medicines and provided these physicians with their products, which would explain how these researchers knew how to contact these physicians. Furthermore, since these physicians were from twelve countries spanning three continents, it was not clear if they were all equally proficient in homeopathy given that each country may have different standards on what constitutes a “certified” homeopathic physician. Of the 162 homeopathic physicians contacted, only 36% ($n = 59$) agreed to participate in the study, which was a very low response rate according to recommendations²⁴ and could introduce bias into the study. Also, it was not clear if these physicians received any benefit (ie, payment, discount on medicines) from the sponsor.

These physicians also recruited the first five child patients (5–15 years old) who presented with migraines; the migraine patterns had to meet the International Headache Society criteria for migraine²⁵ which is a strength of the study because a standardized definition of the outcome measure was used. In addition, these patients could not: (1) be receiving long-term medication treatment at the time of presentation, (2) have epilepsy, and (3) fail to take their homeopathic medication regimen for at least 10 days. These criteria were clearly listed and appear appropriate in that they helped standardize the study population.² However, the ages of 5–15 cover several cognitive developmental groups that vary considerably, such that remembering to note the occurrence and intensity of a migraine may be done so inconsistently, which may have obscured the results. Eighty percent of those patients and their families approached for this study had complete data at the end of the study and were eligible to participate; this represents a strong response rate. Danno et al²² did not indicate whether or not the research reported was approved by an ethics board, or participants signed an informed consent form, which may have been due to publication restraints. The report would have been strengthened with inclusion of at least a single comment regarding the ethics board approval process. Although not generally required for a single group study, a diagram of the sample from recruitment to study ending, such as a CONSORT diagram, would have been helpful to understand the sample clearly.²

Although the sample size was not large, at least 30 cases are required to reasonably ensure that sampling distributions

of common estimates such as a sample mean would be approximately normal (ie, central limit theorem).^{2,7,21} However, unlike many observational studies that use convenience samples, prospective studies that attempt to evaluate treatment effectiveness typically require a priori sample size computations. In addition, participants were recruited across 12 countries, ranging from two participants to 53, which made the findings more representative of homeopathic treatments worldwide; yet, this was countered with the fact that findings were hard to apply specifically to any one country.

Study design

The researchers reported that the study's objective was to "evaluate the effectiveness of homeopathic medicines for the prevention and treatment of migraine in children," yet the study design was described as "observational, prospective, open, nonrandomized, noncomparative, multicenter study;"²² therefore, the objective and design were incongruent. Evaluation of treatment effectiveness requires an experimental design, where a treatment is compared to either a placebo/control or a currently accepted treatment. A randomized clinical trial is the commonly accepted study design for obtaining reliable evidence with regard to efficacy or effectiveness evaluations.² Although a nonexperimental design can provide a suggestion of treatment effect, ultimately any claims of treatment effectiveness will have to be sustained by randomized trials.

Further, it was reported that the study was conducted over a 6-month period, but the first 3 months were in fact retrospective in that the physicians assessed at the first visit the patients' migraine symptoms including number of school days missed. This clearly could have introduced recall bias, while the prospective data were gathered through logbooks that the patients and their parents/guardians were provided. The Data Collection section mentioned that this was a "6-month follow-up period," which appears to be a typographical error because it was clear in the rest of the article that a 3-month follow-up period was completed.

Instrumentation

The instrumentation in this study was challenging to understand and was not well articulated. No standardized instruments were used; therefore, no psychometrics such as Cronbach's alpha could be reported. Instead, at the initial visit, the physicians asked the patients and their parents/guardians to recall their migraine symptomatology (ie, duration, intensity, triggering factors, prior treatments) and number

of school days missed for the preceding 3 months. Then after the homeopathic medicine was prescribed, a logbook was provided to the patients and their parents/guardians and they were to record their migraine symptomatology for the following 3 months; then, after the 3-month period, the physicians "completed a record file" using this logbook. It was also not clear whether the patients or the parents/guardians completed the logbook; children completing the logbook may not have been as accurate and diligent in completing the logbook as adults. Furthermore, utilizing the parents' or guardians' report as a proxy for the child's symptoms and medication use may carry its own bias. Whether any training or written instructions were provided for keeping such a logbook, or if any accommodation was made for the older patients versus the younger patients, was not clear. Thus, comparing retrospective data with prospective data, which had been collected in different ways, introduced severe problems, especially because it was not clear by whom or exactly how the prospective data were collected. In addition, it was mentioned that each of the intensity levels (ie, mild, moderate, severe, unbearable) "was described with the symptoms associated with the attack so that a correlation (and a score) could be established between each level of intensity."²² It was not clear how the association was done, nor how a coefficient was derived and then used in the data analysis; these were important details to leave out of the article.

Data and statistics

Most of the analyses were *t*-tests, Chi-squares, and summaries/tallies. The sample size of 168 was reported; however, as seen in Tables 1 and 2 of the original article,²² varying sample sizes were reported for each of these analyses. It was not clear why the sample sizes changed; no information was provided discussing such aberrations. Polit and Beck² recommend that sample size should be stable throughout a study, or explanations about attrition should be detailed; no explanation was given for the instability, or for changes in analyses based on differences. For example, at the initial visit (ie, inclusion) in Table 1, a sample size of 165 was reported instead of 168. Otherwise, with limited information about how intensity of attacks was measured, it was difficult to determine whether the statistical result of the Chi-square in Table 1²² was meaningful. More importantly, only 38% of patients reported use of homeopathic prescription alone for treatment of acute migraine attacks, which greatly weakens any study conclusions regarding the effectiveness of homeopathic medications. Change in primary outcomes (number, intensity, and duration of migraine attacks)

should have been compared between the three reported acute treatment groups: (1) homeopathic medicine alone, (2) both traditional and homeopathic medicine and, (3) traditional medicine alone – using appropriate statistical methods that try to control for differences in underlying characteristics across these groups (given that a randomized design was not utilized). For these comparisons, the sample size might not be adequate. At a minimum, descriptive statistics (similar to Tables 1 and 2) by acute treatment groups should have been reported.

Internal and external validity

This study reflected many characteristics that threaten both the internal and external validity of the results. As already mentioned, the study's objective and design were incongruent. Lack of a control group threatens internal validity because it was not clear if any homeopathic benefits "observed" accurately reflected their use; perhaps with these participants, just seeing a physician would have been sufficient to encourage better symptom outcomes (as with placebo effect). Further, only 38% of patients reported actually using homeopathic prescriptions alone to manage acute migraine attacks. It is plausible that any benefits observed may have resulted from use of traditional medications by 62% of patients or simply by regression to the mean,²⁴ a statistical phenomenon that can make natural variation in repeated measures look like change over time. In addition, it is not clear how the participating homeopathic physicians were selected, so whether these physicians were affiliated with Boiron Laboratories or not remains uncertain. If they were, then this limits the findings of this study to mainly those who were affiliated with this company. Also, the physicians represented twelve countries from three continents, which threatened the internal validity because homeopathic physicians from different countries may have different training and use different techniques. Likewise, the external validity was threatened because it was not clear to whom the results would apply. Clearly, this study had many weaknesses that limit its value and the applicability of results. Because many of the study weaknesses were acknowledged in the discussion section, it is not clear why the study was not labeled simply as an exploratory pilot study.

Practical pointers

As can be seen from this brief article critique, there were some strengths but several limitations mentioned about the methodology. These strengths and limitations were written

from the perspective of the potential barriers (ie, money, time) the researchers had to address in conducting the study. It is vital to consider these barriers because of their influence on how the data were collected, the quality of these data and, of course, data interpretation. From this example, the following pointers on conducting an article critique should be considered: literature review, balance and practicality, logic, and application.

Literature review

As the reader may have noticed, there was no critique of the literature review; this omission was done purposely. First, a true and complete literature review can be so lengthy as to qualify as an article in itself, yet judging the value of a brief literature review in light of the imposed limitations many journals place on researchers seems rather pointless. Second, unless the reader is quite familiar with the topic of a study, he or she will not have the knowledge base to judge whether all the relevant studies have been cited. Because it may take nearly a year (and in some cases longer) for a research study report to be printed after it has been accepted for publication, the literature review may be outdated as a result, especially if the topic is in a field that progresses quickly.

Finally, some article critique guidelines suggest that only recent articles and reference materials (within the past 5 years) be included; ideally, that makes sense but some areas of science progress faster than others and some classic studies are seminal works and continue to contribute even decades after they have been published (eg, the Stanley Milgram's *Obedience to Authority; An Experimental View*).²⁶ Moreover, evaluating a literature review based solely on the age of the articles cited represents an inadequate manner in which to evaluate an article, especially if the reviewer is unfamiliar with the literature.

Balance and practicality

Generating an article critique is different for every study; and providing that balance in terms of strengths and weakness is an art that develops with time and experience. All studies have strengths and limitations that must be weighed according to their importance. For example, having a larger sample size (a study strength) may be more important than having two popularly used and well-respected instruments that happen to have a mediocre Cronbach's alpha (a small study limitation). Some of the strengths and limitations are tempered with practical explanations on why the researchers may or may have not done something (eg, difficulty of giving an actual vision or

hearing test to nursing home residents with various degrees of dementia). It is important not to blindly dismiss the scientific value of a study based on some single limitation that may be unavoidable across studies.

Logic

Obviously, the various components of the study (ie, study rationale, sample, design, etc) should logically fit together. This implicit assumption means that how to articulate exactly how to examine the cohesiveness of the study components is not always clear. As already mentioned, an effective critique requires a working knowledge of research methods and statistics, as well as being immersed in reading the scientific literature to understand both topic and research processes. And it is from this immersion that one will understand how the components of research articles are assembled to form an overall picture of what the researchers did and why.

Application

Usually researchers reporting in published literature make a few general comments and recommendations based upon the interpretation of their findings; these vary from research to clinical or policy recommendations, and some applications of the findings may not be obvious to the researchers who lack focus in those areas (eg, a clinical focus), or the comments and recommendations may not be written with a discipline-specific orientation. Likewise, researchers who make recommendations in a nondiscipline-specific journal will not usually make recommendations for a specific alternate discipline such as nursing or psychology; therefore, it is up to the reader of the article to determine the usefulness and the application of the information generated from the study for their discipline and their purposes. Thus, if the article being critiqued is published in a predominantly psychology journal, the researchers would not be expected to address nursing implications. However, if the article is published in a nursing journal, the expectation would be to include nursing implications. Furthermore, researchers cannot meet the expectations of every reader. Thus, it is critical to understand the research reported, and to draw conclusions applicable in the reader's interest area and discipline.

Conclusion

Quantitative article critique guidelines can be helpful tools for graduate students learning this valuable skill, but only if they have a working knowledge and familiarity with research methods and statistics. It also helps to have an understanding of the publication process because certain details about the

study (ie, lack of detail about instrumentation, abbreviated literature review) may be severely shortened at the behest of the journal's editor or publication guidelines (ie, page limitations). Thus, for a novice, it may appear that the study is flawed by not having enough information reported. Furthermore, evaluating the scientific merits of a quantitative research article is more art than science in that everyone approaches this process from their own disciplinary and practice perspective. Because every reader, whether practitioner or researcher, reads such articles with different agendas, competencies, and needs, it is difficult to compare their evaluations of scientific merit because the purpose of their critique differs. To that end, the ultimate point of conducting such research article critiques is to evaluate and judge the study's internal validity in order to decide its external validity. Only then can one determine to what degree the interpretation of the findings can be accepted and applied. Moreover, if we can improve the ability of the novice reader, including graduate students, to critically evaluate the scientific evidence, then we can enhance their professional development by improving research skills in order to improve evidence-based practice and contribution to knowledge development overall.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Melnyk BM, Fineout-Overholt E. *Evidence-Based Practice in Nursing and Healthcare: A Guide to Best Practice*, 2nd ed. Philadelphia, PA: Wolters Kluwer/Lippincott Williams & Wilkins; 2011.
2. Polit DF, Beck CT. *Nursing Research: Generating and Assessing Evidence for Nursing Practice*, 9th ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2012.
3. Voelker DH, Orton PZ. *Cliff's Quick Review of Statistics*. Lincoln: Cliff Notes; 1993.
4. Vance DE, Bayless H, Kempf MC, Keltner N, Fazeli PL. Aging, HIV and wellness: augmenting the components of successful aging. *Aging Health*. 2011;7(3):435–446.
5. Ranasinghe P, Perera YS, Abeygunasekara AM. The process and costs of publishing medical journals in Sri Lanka: an economic evaluation. *BMJ Open*. 2011;1(1):e000057.
6. Coughlan M, Cronin P, Ryan F. Step-by-step guide to critiquing research. Part 1: quantitative research. *Br J Nurs*. 2007;16(11):658–663.
7. Grove SK, Burns N, Gray J. *The Practice of Nursing Research: Appraisal, Synthesis, and Generation of Evidence*, 7th ed. St Louis: Elsevier/Saunders; 2013.
8. Ryan-Wenger NM. Guidelines for critique of a research report. *Heart Lung*. 1992;21(4):394–401.
9. Giacomini MK. The rocky road: qualitative research as evidence. *ACP J Club*. 2001;134(1):A11–A13.
10. Giacomini MK, Cook DJ. Users' guides to the medical literature: XXIII. Qualitative research in health care B. What are the results and how do they help me care for my patients? Evidence-Based Medicine Working Group. *JAMA*. 2000;284(4):478–482.
11. Babbie ER. *The Practice of Social Research*, 13th ed. Belmont, CA: Wadsworth Cengage Learning; 2013.

12. Morse J. Strategies for sampling. In: Morse JM, editor. *Qualitative Nursing Research: A Contemporary Dialogue*, 2nd ed. Newbury Park, CA: Sage Publications; 1991:127–145.
13. Sandelowski M. Sample size in qualitative research. *Res Nurs Health*. 1995;18(2):179–183.
14. Denzin NK, Lincoln YS. *The Sage Handbook of Qualitative Research*, 4th ed. Thousand Oaks, CA: Sage; 2011.
15. Okonkwo O, Vance D, Antia L, et al. Service utilization and cognitive complaints in adults with HIV: results from a statewide survey. *J HIV AIDS Soc Serv*. 2008;7(2):175–194.
16. Preiss D, Thomas LE, Sun JL, et al. Predictors of cardiovascular events in a contemporary population with impaired glucose tolerance: an observational analysis of the Nateglinide and Valsartan in impaired glucose tolerance outcomes research (NAVIGATOR) trial. *BMJ Open*. 2012;2(6).
17. Suzuki-Crumly J, Ackerman ML, Vance DE, et al. The role of religiosity in mediating biophysical outcomes in homosexuals with HIV: a structural equation modeling comparison study. *Journal of Spirituality in Mental Health*. 2010;12:209–223.
18. Vance DE, Burgio LD, Roth DL, Stevens AB, Fairchild JK, Yurick A. Predictors of agitation in nursing home residents. *J Gerontol B Psychol Sci Soc Sci*. 2003;58(2):P129–P137.
19. Vance DE, Roenker DL, Cissell GM, Edwards JD, Wadley VG, Ball KK. Predictors of driving exposure and avoidance in a field study of older drivers from the state of Maryland. *Accid Anal Prev*. 2006; 38(4):823–831.
20. Vance DE, Okonkwo O, Antia L, et al. Factors of cognitive complaints in adults with HIV: a structural equation model analysis. *Occupational Therapy in Mental Health*. 2009;25:4–25.
21. Vogt WP. *Dictionary of Statistics and Methodology: A Nontechnical Guide for the Social Sciences*, 3rd ed. Thousand Oaks, CA: Sage Publications; 2005.
22. Danno K, Colas A, Masson JL, Bordet MF. Homeopathic treatment of migraine in children: results of a prospective, multicenter, observational study. *J Altern Complement Med*. 2013;19(2):119–123.
23. Polit DF, Beck CT. *Essentials of Nursing Research: Appraising Evidence for Nursing Practice*, 8th ed. Philadelphia, PA: Wolters Kluwer/Lippincott/Williams & Wilkins Health; 2013.
24. Barnett AG, van der Pols JC, Dobson AJ. Regression to the mean: what it is and how to deal with it. *Int J Epidemiol*. 2005;34(1):215–220.
25. Headache classification subcommittee of the International Headache Society. The international classification of headache disorders, 2nd edition. *Cephalalgia* 2004;24(suppl 1):9–160.
26. Milgram S. *Obedience to Authority; An Experimental View*, 1st ed. New York: Harper & Row; 1974.

Nursing: Research and Reviews

Publish your work in this journal

Nursing: Research and Reviews is an international, peer-reviewed, open access journal publishing original research, reports, reviews and commentaries on all aspects of nursing and patient care. These include patient education and counselling, ethics, management and organizational issues, diagnostics and prescribing, economics and

Submit your manuscript here: <http://www.dovepress.com/nursing-research-and-reviews-journal>

resource management, health outcomes, and improving patient safety in all settings. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Dovepress