International Journal of Aquatic Research and Education

Volume 2 | Number 4 Article 10

11-1-2008

How to Get Answers to Our Questions: - Finding and Understanding Scientific Literature

Barbara A. Bushman
Missouri State University, barbarabushman@missouristate.edu

Recommended Citation

Bushman, Barbara A. (2008) "How to Get Answers to Our Questions: - Finding and Understanding Scientific Literature," *International Journal of Aquatic Research and Education*: Vol. 2: No. 4, Article 10.
DOI: https://doi.org/10.25035/ijare.02.04.10
Available at: https://scholarworks.bgsu.edu/ijare/vol2/iss4/10

INVITED REVIEWS

International Journal of Aquatic Research and Education, 2008, 2, 367-371

How to Get Answers to Our Questions: Finding and Understanding Scientific Literature

Barbara A. Bushman

It ain't so much the things we don't know that get us in trouble. It's the things we know that ain't so. (Artemus Ward)

I am often struck by how seemingly simple questions on fitness or health are hard to accurately answer with a scientific basis. Finding reliable sources of information and then interpreting the results accurately can be a challenge. Often we rely on news reports, advertisements, or "common knowledge" as the basis for our statements or actions. In this article, I will point out how to find solid sources of information as well as how the design of the study or the statistical analyses used may influence interpretation/application.

Too often we are hit with advertisements and statements that seem to promise the world. Recently, one of my students brought a supplement label to me to review. The content of the label seemed to be full of "good science." It even had a list of references of articles found in the literature. The product promised that it had a unique and special blend of ingredients that would promote muscle growth like no other product could. I took the time to look up the articles listed—most of the articles did not study athletes. Some were done on animals and most had no relationship to performance at all. In this particular case, although the claims were a bit of a stretch scientifically, the product didn't appear to contain any harmful ingredients and was reasonable in the vitamins/mineral included. As consumers become more interested in results, marketing groups realize the power of science. We all want to believe the claims, and thus we buy the products. Hinting at scientific support helps sell products.

Often we see that phrase "statistics show . . . " as a way to convince us to take a particular action or to buy a product. Even our chewing gum and toothpaste are marketed by the claims of "4 out of 5 dentists recommend." So, how do we find good information? There are many resources available through the scientific literature. In the past, accessing these documents was a challenge for the public. Most sources were housed in university or medical libraries. Now with the internet, many more options are available that do not require a link to a college or university. One example of an important scientific resource is PubMed (http://www.ncbi.nlm.nih.gov/pubmed). This is an excellent website that allows anyone with an internet connection to access original research articles.

The author is with Health, Physical Education and Recreation at Missouri State University in Springfield. E-mail: barbarabushman@missouristate.edu

I would like to comment on the type of articles under consideration. Scientific articles undergo a peer review process. This means that the article has undergone a review by other experts in the field before the article being accepted for publication. This process is important in ensuring accurate and appropriate information in print as opposed to merely someone's opinion on a topic. These articles will always have a list of references at the conclusion of the article showing the basis on which the research was conducted. The PubMed website, as opposed to simply using a more commercial search engine, will lead you to published scientific sources. To find on-line access to journals, you can check PubMed Central. This is a free digital archive of biomedical and life sciences journal literature (see www.pubmedcentral.nih.gov/).

Most scientific articles follow a common format—they begin with an abstract, followed by the introduction and background section, methodology, study results, discussion, conclusion, and references. The introduction includes 3–5 paragraphs quickly outlining past research and how the current study relates to previous studies. Typically, the last paragraph of the background includes the purpose statement of the study. The methodology (or "method") section outlines in detail all the tests, protocols, and equipment used in the study. The method section also includes a paragraph on what statistical analyses were conducted. The discussion is the researchers' opportunity to reflect on their results in light of previous research. Whereas the results section is pretty cut and dry, the discussion is an opportunity to provide more expert commentary and interpretation. The conclusion section is the bottom line of the study that should reflect on the purpose of the study. The reference list provides all the background material cited in the article.

Although the previous paragraph describes the typical order of presentation within a research article, my own approach to understanding an article differs. I read the abstract (or summary) first, and then, just like with a good book, I take a peek at the ending by checking the conclusions. Then I want to get a feel for how the researchers feel their work fits with what has been done previous (i.e., the discussion). At that point, if I'm still interested, I head to the method for some detail, then the introduction, and then to the nitty gritty of the statistical results.

As I mentioned, my first stop is the abstract. The abstract is intended to be a quick overview of the study. A good abstract contains all the major findings of the study as well as the conclusion. It is very short in length so it is a way to determine if you are really interested in the article. If so, then you can forge ahead and read the rest. Although it is intended to be a complete picture, I would recommend reviewing the entire article before making conclusions because due to the limited length, not all important factors can be included in an abstract.

The design of a research project influences what conclusions can or should be made. Sample selection is one aspect that can drastically change the outcome as well as the application of study results. For example, one area of my research interest is deep water run training. Research participants in this topic area range from previously sedentary to highly trained distance runners. Obviously, it is vital to know the participant pool when examining the results. In addition, details such as the use of a flotation device (belt or vest) are important. Town and Bradley (1991) did not use a flotation device and they reported different findings than others (e.g., Butts, et al., 1991; Frangolias & Rhodes 1995).

Most research in applied areas such as aquatic exercise uses a convenience sample rather than a true random sample. The random sample is the ideal situation in which each member of the population of interest has an equal chance of selection. In 1936, the Literary Digest predicted that Alf Landon would win the presidential election over Franklin Roosevelt. This incorrect prediction, however, was not drawn from a random sample of potential voters. Inadvertently, a bias had been introduced by selecting only individuals who could afford telephones and magazine subscriptions. Similarly, if I asked patrons of an aquatic center about the importance of swimming lessons for children, I likely would arrive at a different conclusion than if I asked a sample of people shopping at a nearby mall. In applied research, we often do use convenience samples. Use of college student volunteers is quite typical in university-based research. Volunteers may be different from the rest of the population, and this is a factor that we must consider when reading a research article and considering its findings.

Another aspect of research design is the use of a placebo, which is an inactive/inert compound with the same appearance as the experimental item. This is very common in pharmaceutical studies where it is important to separate out the effect of a particular drug versus a person's perception of what a particular pill may do. Unlike pharmaceutical studies, with exercise studies this is not possible because it is obvious if a person is not exercising. Control groups are typically a "no intervention" group. Control also can take into account special factors. For example, attention/socialization with exercise may influence mood, so a research design may include a control group that gathers to play board games (so they socialize but are not physically active).

Determining if two groups are truly different requires statistical analyses. If I asked a group of people to flip a coin, I would expect to see a 50–50 split between heads and tails. Often this doesn't come out perfectly equal even though we know the two sides of the coins are the same. This is not uncommon in research either. For example, I had a 6 s difference between pretraining and posttraining 5km times in one study examining the influence of four weeks of deep water run training (Bushman et al., 1997). The question I needed to resolve required a statistical analysis to determine if this was a real statistically significant difference. Sampling error occurs when chance or random effects cause an event to occur in a manner different from what is expected. The way that we determine this is by looking at a preestablished standard value called the alpha level. The error sampling standard should be reported within the method section of a research article. By comparing our calculated statistic to the value expected by chance, we now can determine if the groups are really different. In my example, did the subjects really get slower or was that difference due to sampling error or natural variability (i.e., that a 6 s longer 5K time wasn't different "enough" but was due to "chance")?

As researchers, we often select a standard alpha error rate of 0.05. This means that any difference in two groups that can occur due to chance more than 5% of the time is probably due to sampling error (so one must accept that the performance of the two groups probably is equal). If the difference observed between two groups occurs due to chance less than or equal to 5% of the time, then that would be a statistically unlikely result and probably not due to sampling error or natural variability, but due to something real (i.e., so the performances between the groups *are not* equal). With my previous example with the 5K running times—we did observe what appeared to be a difference between the two groups, but through my statistical analysis, I found that such a small difference was probably due to chance and thus we say the two tests were equal even though looking at the two means we can see the numbers don't match exactly. This acknowledges the chance differences that can and do occur.

To clarify what is typically in scientific articles, the statement " . . . significant difference between the two groups, $p \leq .05$" is saying that the probability that a sampling error or natural variability in scores could cause that difference to occur would be less than or equal to the stated alpha level (preestablished standard) and thus something real (other than sampling error or natural variability) caused it to occur. In contract, if we read " . . . no significant difference occurred between the two methods, $p > .05$," the researchers are saying that the probability that sampling error or natural variability could cause that difference to occur would be greater than the stated alpha level so they conclude that the differences could be attributed to chance occurrences. This is how *p*-values are typically noted within articles. If the probability of the difference due to chance is low, then we can conclude the difference is probably real. Rather than using the term "real," we typically call it a "significant" difference. When the probability of the difference being due to chance is large, then we conclude the difference is "not significant."

It is interesting to note that the number of participants in a study can influence this interpretation about whether observed differences were real or chance occurrences. The power of a statistical test is the probability that the test can detect a difference when a real difference exists between groups. Higher power is available with more subjects. In applied research this is often a limitation when dealing with the practicality of training human subjects.

Examination of tables, figures and graphs within a research article also is important. For example, changes in body mass index (BMI) between an exercise group and a control (inactive) group are shown in Figure 1. Taking a quick look might lead me to believe there was a large difference between the groups based on

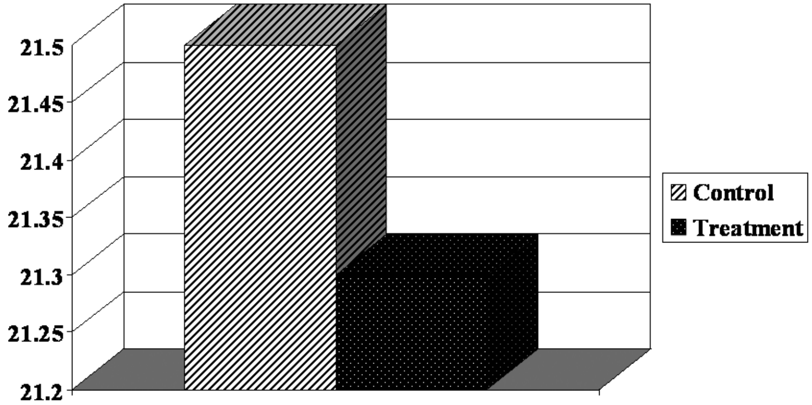

Figure 1 — Difference in BMI between treatment and control groups.

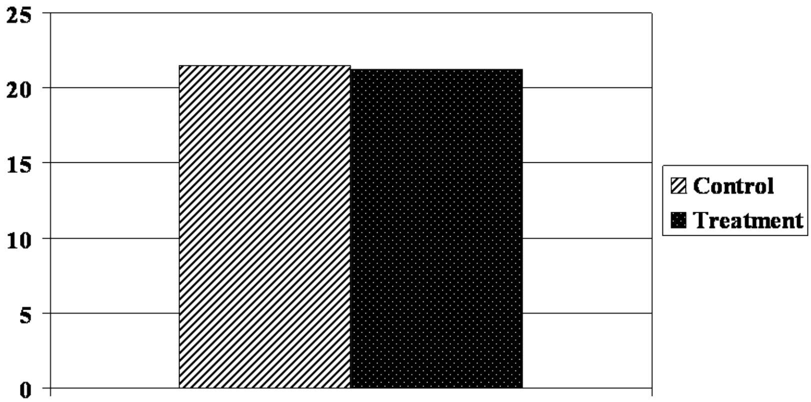

Figure 2 — Difference in BMI between treatment and control groups.

the height of the bars. Now look at Figure 2, and it appears there is no difference between the control and treatment groups. Look more carefully at the vertical, or y-axis, of the two figures. The data are exactly the same but in Figure 1, the y-axis scale has been limited and makes a small, nonimportant difference look huge. Readers should be careful when reading and reviewing studies to make sure your first impressions are accurate.

In conclusion, it is vital to find reputable, peer-reviewed sources of information. Although the internet is a vast resource, you should focus on scientifically-based sources. Use resources like PubMed to find peer reviewed scientific articles. Check out the original (primary) articles for yourself to determine if the participant selection or study design fits with your particular questions. Look at the results section to see what differences are being reported. Carefully examine the graphs and figures. And finally, be willing to question what you read.

References

Bushman, B.A., Flynn, M.G., Andres, F.F., Lambert, C.P., Taylor, M.S., & Braun, W.A. (1997). Effect of four weeks of deep water run training on running performance. *Medicine and Science in Sports and Exercise, 29*(5), 694–699.

Butts, N.K., Tucker, M., & Greening, C. (1991). Physiologic responses to maximal treadmill and deep water running in men and women. *American Journal of Sports Medicine, 19*(6), 612–614.

Frangolias, D.D., & Rhodes, E.C. (1995). Maximal and ventilatory threshold responses to treadmill and water immersion running. *Medicine and Science in Sports and Exercise, 27*(7), 1007–1013.

Town, G.P., & Bradley, S.S. (1991). Maximal metabolic responses of deep and shallow water running in trained runners. *Medicine and Science in Sports and Exercise, 23*(2), 238–241.