

Critiquing Sampling: A Guided Analysis

Prerequisite Knowledge	<ul style="list-style-type: none"> ● Problems, objectives, hypotheses, and variables in studies ● The steps of research in nonexperimental and experimental designs ● Strengths and threats to validity in various study designs ● Study frameworks
Learner Setting	Online
Strategy Type	Other: Instructional media with corrective feedback
Time	Faculty prep: 30 minutes Delivery with students: 60 minutes Evaluation: 45 minutes
Learning Objectives	<ul style="list-style-type: none"> ● To critically appraise the sampling criteria in studies ● To determine the adequacy of sample size in published studies
Materials/ Resources	Article: Padula, C. A. Yeaw, E., & Mistry, S. (2009). A home-based nurse-coached inspiratory muscle training intervention in heart failure. <i>Applied Nursing Research</i> , 22, 18-25. doi: 10.1016/j.apnr.2007.02.002 Worksheet (see below) Video (see below) Transcript (see below)
Strategy Overview	In this activity, students read an assigned article and submit worksheet related to sampling. They then receive access to a screencast video where the instructor reviews the rationale for the answers. Students then submit a second draft of the worksheet. The order in which these items are required helps students appreciate what they do not understand and experience valuable learning and guidance.
Steps	1. Each student reads assigned article and submits answers to

	<p>worksheet via learning management system (LMS)</p> <ol style="list-style-type: none"> 2. Their submission prompts the LMS to allow student access to a ten-minute guided analysis video which walks student through the rationale for each answer and where to find the answers in the article 3. Student submits a second draft of answers to worksheet 4. Instructor grades answers on second draft and provides feedback
Evaluation	<p>Low stakes – Small credit given for effort on first draft, second draft answers are graded</p>

Critiquing Sampling: Guided Analysis Worksheet

Answer the following questions:

1. What were the inclusion criteria?
2. What were the exclusion criteria?
3. Of the screened subjects, what percent were eligible to participate?
4. What was the refusal rate?
5. What is the independent variable?
6. What are the dependent variables?
7. Did Padula and colleagues indicate that they conducted a power analysis to determine the sample size?
8. Did they report an effect size?
9. In the limitation section, what standard did the researchers use to argue that their sample size was adequate?

Critiquing Sampling: Guided Analysis Video

<https://gvsu.ensemblevideo.com/Watch/Qb8g5AMd>

Critiquing Sampling: Guided analysis transcript

So, you've read – or at least skimmed – this article, and come up with some answers to the questions I've posted in Blackboard. Grab your article, any notes you took, and maybe your textbook for good measure, and let's walk through the answers.

The first two questions ask about the inclusion and exclusion criteria. Of the main, basic sections in a research article (introduction, methods, results, and discussion), this is something you should expect to find in the methods section. In this article, you'll find it on page 21. Under *Recruitment*, you'll see five inclusion criteria listed. But it's a trap! Technically, only the first three of these are inclusion criteria. They are criteria that the participants must have. The last two are the exclusion criteria because they are criteria that the participants must not have – coexisting pulmonary disease and cognitive impairment. The authors say these criteria maintain the quality of the data, so we might assume that they perceive these to be confounding or extraneous variables and would like to control for that. That's great, but on the other hand, if a large proportion of heart failure patients have coexisting pulmonary disease or cognitive impairment, are the authors failing to get a sample that represents the population of heart failure patients? They even mention that nearly half of those screened were ineligible due to COPD! Deciding how broad or narrow to make the inclusion and exclusion criteria is a compromise between excellent control and excellent representativeness. Generally, both will be flawed to at least some extent.

The answers to questions three and four also fall on this page, in Figure 1. Of the screened subjects, what percent were eligible to participate? The answer is 12.5% - the number eligible (36) divided by the number screened (288). Did you notice, higher on the page, the authors stated that 13.8% were eligible? Where does this number come from? It doesn't match the basic math in Figure 1. Also, the number ineligible and the number eligible don't equal the total number screened; there are four people unaccounted for. Let me know if you can figure this out! What a great example of how just because something is published, it doesn't mean that there's nothing to scrutinize. As for the next question about the refusal rate, that's 11.1% -- the number refused (4) out of the number eligible (36). At least this math seems to add up correctly. These

first four questions should get you thinking about representativeness. I've already hinted at that for the first three questions. Think on your own about how the 11% refusal rate reflects the degree of representativeness that this sample might have to the greater population. Why might they have refused?

The remaining questions get at evaluating the sample size. Let's start with variables (questions five and six). These also fall in the methods section. The independent variable is very clearly specified on page 20. The four dependent variables are in section 4.6.1 but also right up front in the abstract, listed as outcome measures. Is that discrepancy in terminology confusing? Dependent variables and outcome measures are the same thing because they're the variables that depend on the effect of the independent variable and are measured as the outcome of the implementation of that independent variable. So if these questions are supposedly about sample size, what do variables have to do with sample size? There's a section in your textbook starting on page 266 that covers five factors that influence the adequacy of sample size in quantitative studies; these factors are all important in determining adequate sample size because they impact the statistical ability to accurately detect relationships among variables or differences in outcomes among experimental and control groups. One of these factors is the number of variables. Statistically speaking, the greater the number of dependent variables being tested, the greater the sample size needs to be. This study has four dependent variables; let's keep that in mind as we move on.

Was there a power analysis reported? Did the researchers make this calculation ahead of time to figure out what their sample size should be? I have nothing to show you; no, they did not. Why is this important? Well, among all the study results, there *were* some results that were not significant, such as the scores on HRQOL. The non-significant results raise a red flag. Is there really no difference, or was the sample size just too small to detect a difference? Did the authors at least perform a power analysis after their results were obtained and include it in the discussion to evaluate this concern? No. Shame, shame.

Let's briefly examine the next question and look for the effect size. It's at the top of page 23, and it's a medium effect size of 0.48. Hold that thought as well.

In the final question, I asked you how the authors argued that their sample size was adequate. This is on page 24. They do acknowledge that it was somewhat small (at 31 participants), but they compare it to similar studies published over the prior 12 years and point out that it was one of the largest samples. Is this good justification? Are you satisfied with the answer? If we compare what we've found in this article to the five factors I mentioned in your book, 1) the four dependent variables indicate that the size should be larger. Larger than what remains to be seen, but bear with me. 2) The medium effect size calls for a medium sized sample...okay... 3) The authors claim at different times that this is a quasi-experimental study and a RCT, either way, the sample can be smaller than a descriptive or correlational study. I wonder what the previous studies they cited were. 4) With regard to measurement sensitivity (this is on page 268 in your text book), we haven't covered this yet, but the authors cite some great information on the sensitivity of some of their measures; on others they're more vague. So that's kind of a mixed bag when it comes to drawing a conclusion about the adequacy of their sample size. 5) For data analysis, they used ANOVA. We haven't covered this either, but your book chapter does briefly mention that this test requires larger sample sizes when the groups are uneven. The groups were a little uneven, 15 and 17. Is this even enough? How do you decide? You know what I think? It sure would have been nice if they had just done a power analysis and put the whole issue to rest. Their sample size could be just fine, but their argument that it's bigger than other studies...not really good enough for me.

Now, while I've got you here, let's take a minute to flashback and integrate a couple of the other topics we've discussed this semester. It's bonus learning!

The first thing I'll point out is the interwoven cluster of problem, purpose, aims, objective, and hypothesis. The first paragraph is full of this stuff! The authors set the stage for the problem (HF affects a lot of people and the symptoms are disabling, and they hint that there may be a yet unknown answer in inspiratory muscle training); they give a condensed purpose about determining the effect of this muscle training; and state their hypothesis (which you may note is complex, causal, and directional). You

may notice that some of this information is also present right in the abstract. On the next page, concluding their review of the literature (which is quite substantial), they contribute more to the problem and the purpose (muscle training *is* effective but we don't know the right doses or how it actually impacts dyspnea and quality of life, previous studies have not been based on theory, and this study serves to fill a research gap). On page 20, you find their overall goal, (which you can consider their purpose), their aims, and their research question nicely laid out.

Let's touch on theory too. This study is amazingly integrated with Self-Efficacy Theory. This is clearly not something that the authors tried to make fit their research after the fact so it would look good for publication. The constructs spelled out in the theory were incorporated in the interventions, one of the aims and one of the dependent variables deal directly with self efficacy, their data collection points were informed by the theory, and they relate their discussion of clinical application of these findings back to the theory as well. This is an excellent example of integration of theory and research.

I hope this guided critique helped you in critically appraising the sampling criteria in studies. Let me know if you have any questions!