

# **A Small Introduction to Critical Thinking for Graduate Students**

by

**William H. Percy, Ph.D., LP, LMFT**

**The Graduate School of America**

**Minneapolis, Minnesota**

Rev: June, 1998

## **Table of Contents**

Introduction 3

Some “levels” of critical thinking 3

Everyday criticism 3

“Professional” criticism 3

“Critical Analysis” 4

Critical Theory in Europe 4

Feminist Criticism 5

Critical Thinking 6

Some basic concepts in “critical thinking” 7

“Argumentation” 7

“Evaluation” 7

Operationalizing “Truth” 8

Form 8

“Trustworthiness” 11

“Relevance” 12

“Significant” or “Meaningful” 13

Summary of the basic concepts discussed thus far 13

Some heuristic devices (“rules”) for evaluating the quality of arguments 14

Evaluating the overall study and the central question 14

Rule 1: Keep it simple. Don’t over do it. 14

Rule 2: If the whole report is well-formed, go on. 14

Rule 3: Write out the central question. Write out the truth–claim it makes. 14

Rule 4: Ensure that the Question is in proper form for the claim made. 14

Rule 5: If the Question is poorly formed (Rule 3), STOP 14

Evaluating the supporting and dissenting evidence 14

Rule 6: Look at the literature review next. 14

Rule 7: Next evaluate the methods. 15

Rule 8: Evaluate the findings and conclusions. 15

Rule 9: Decide whether you wish to consider this study dependable and safe to rely on in your professional and personal life. 16

General Summary 16

References 17

*Anything written by people is written by people*

– Anon.

## Introduction

The epigram says it all: Anything created by people is created by someone real who can make mistakes or be subject to bias. In science and scholarship, errors in methods, in logic, or in analysis can make conclusions undependable. This is why “critical thinking” is important.

My own field, psychology (or more broadly, Human Services) depends on research — scientific scholarship — to show us what facts about human nature and the world around can be depended on when we propose solutions to our clients’ problems. So if the research papers we read are flawed, woe betide us and our clients. Fortunately, we have standard ways to detect lethal flaws in research studies. This brief essay will not describe them in detail, but I will try to review the main issues with which critical thinking concerns itself. I’ll sum them up in four words: form, trustworthiness, relevance, and significance.

Critical thinking concerns itself with evaluating the form, trustworthiness, relevance, and significance of any work of scholarship. Remember these terms! They give a set of tools

for evaluating any scholarly work and for answering this simple question which every scientist-practitioner must ask about the studies she reads: Can I depend on the results of this study in my professional work or my daily life?

After I briefly introduce various “levels” of critical thinking and mention a few basic concepts you’ll need to know, I’ll talk about the four themes (form, trustworthiness, relevance, and significance) in some detail, giving you some “heuristic devices” which I have learned to use to evaluate studies.

### **Some “levels” of critical thinking**

#### **Everyday criticism**

My mother told us not to “criticize” other people. Sometimes we try to improve other people by giving them “constructive criticism.” Monday–morning quarterbacks criticize the coaches or the players from Sunday’s football game. These “everyday” forms of criticism convey a sense that some skeleton is to be dragged from a closet and exposed to the daylight. Often it is helpful, but sometimes it can be mean-spirited, nasty, and wrong. Everyday criticism can be based on biased values, prejudices, or unspoken hidden agendas, so it is not too reliable. Still, note the elements: break some experience down into its elements, then evaluate them.

#### **“Professional” criticism**

Siskel and Ebert are film critics. The newspapers have music critics, art critics, drama and architecture critics, and book critics. In the bookstore, there are shelves devoted to writers of “Literary Criticism.” What these people do is “criticism.” They approach a work (film, art, literature, a building, even a recipe) with a “Let’s take it apart and see how it works” mentality.

This mentality, which is usually called “analytic,” does indeed “take things apart” (which is the Greek root of *analysis*). The critic breaks a work into its parts and then discerns how they are related and how well they accomplish what they are meant to do. So there is some of that everyday meaning here as well: criticism includes some sort of analysis, followed by evaluation and judgment.

Professional criticism is not necessarily negative. Unfortunately, media–based criticism can be sadly mean–spirited and competitive, reflecting more the critic’s ego than a reasoned analysis–and–evaluation. Or it reflects a need to increase readership of the journal. But it can also produce works of profundity, wisdom, and delight. If you want an example of the latter sort of professional criticism, check out Malcolm Cowley’s (1984) stunningly personal critical study of eight American writers, *A Second Flowering*.

#### **“Critical Analysis”**

#### **Critical Theory in Europe**

A third level of “criticism” occurs when philosophers and social theorists use a highly refined discipline for analyzing and evaluating the value of theories, models, and

concepts. Karl Marx was a powerful critical thinker who focused on the analysis and evaluation of capitalism as a social theory (we often forget this, since the Soviets propagandized Marx as merely a revolutionary writer, not a critical philosopher).

After World War II, a group of social philosophers in Europe developed what is called “Critical Theory.” A famous group is the “Frankfort School,” radical young professors in the philosophy, sociology, and economics departments of universities in the city. Their criticism exposed the hidden social arrangements and the elements of theories and institutions that people take for granted, but which may be less “true” than expedient for certain powerful people. The critical theorists would take on such a slogan as “managed care,” and would expose the hidden beliefs about reality that it embodies (for instance, that “health” is a product or commodity that can be bought and sold).

Critical theorists analyze-and-evaluate their subject matter as “everyday” and “professional” critics do, but the former go very deep. They inquire into the underlying hidden assumptions and taken-for-granted ideas supporting the theories. Ideas may appear “objective” and “true,” but they break them down and show how they are based on real people’s political, social, and economic needs and desires in given historical situations.

### **Feminist Criticism**

Feminist criticism is critical theory. For instance, take the common observation that a woman on a committee with men often defers to the men. Though some might question her assertiveness, the critical theorist would ask: Don’t socially disadvantaged people act just this way, with no psychological implications at all? (A deferential minority hotel doorman might be quite assertive on his own turf among familiar people.) Thus, the woman’s deference might reflect her “second-class” status.

But critical theorists (feminist or not) would not stop there. They would inquire into the underground assumptions of that particular committee of real people working for real ends in a real corporation.

Nor would that suffice.

Once they had exposed the actual historical, social, political, and economic status of all the people in the meeting, they would expose the participants took for granted. For instance, what assumptions about the world are “taken for granted” by the players here? That money (or status) is more important than integrity? If so, is this true in the wider social-economic system of which this committee is but one part? If so, how does this belief affect how we act with one another in the society? Isn’t her deferent behavior merely an example of how we all, in such a world, defer to our bankers and owners and bosses? Is this the reality we live in? Is it the only possible reality?

Such deep criticism can be unsettling. An early version of my dissertation was rejected, when I thought I had discovered an important truth about reality. My mentor, a Marxist political scientist, showed me how my work was an expression of my unaware middle class bias, and might be irrelevant to the working people I aimed to write about. (By the way, notice that word, “irrelevant.”)

You can distinguish “deep” critical theory from the kind of critical thinking this essay is about by this example: Given a newspaper report of a study, the scientist-practitioner would ask whether the actual study is valid and reliable can be relied on for everyday use. But the critical theorist (“deep critic”) would ask, What assumptions about the world are hidden in this study? Is that “taken-for-granted” world the one we care about? Is it even possible? In fact, how can we actually know anything practical?

“Deep” questioning, of course, means philosophical questioning. What fascinates me about deep criticism is that it is based on the conviction that the world should be improved. It says, We can’t fix the world without understanding of how that world really works, so we must challenge our basic assumptions about how that world works. Without that level of analysis, we will fail to get at the real causes of problems and will merely tinker with the effects. As you can see from this last point, the aim of “deep criticism” is not merely to understand the world.

The point of philosophy, Karl Marx said, is not to understand the world. It is to change the world.

In this essay, I won’t be dealing with the “deep critics” any more, except for an occasional nod in their direction. As you can imagine, critical theory is a field of study in itself, a branch of philosophy and social theory. Here is a list of some people who are important in the field, though, in case you are interested in doing some further study:

· Jurgen Habermas (Frankfort School thinker, showed how knowledge is not objective, but reflects human interests).

M. Merleau-Ponty (1962), a French phenomenologist, social philosopher, and psychologist, wrote:

“True philosophy consists in relearning to look at the world, and in this sense [any research] can give meaning to the world quite as ‘deeply’ as a philosophical treatise. *We take our fate in our hands, we become responsible for our history through reflection, but equally by a decision on which we stake our life* . . . “ (Phenomenology of Perception, p. xx, my emphasis).

Recall Merleau-Ponty’s situation. He and his fellow students suffered through the Nazi occupation. In their youth, they had inherited the idea that European values were humane and that Euro-civilisation was proof against despots. Then Hitler emerged from the bosom of European civilization, and they found barbarism at the heart of Euro-Christianity. The students re-evaluated philosophy, during a period when being critical of the Nazi “sciences” could result in deportation and death. Thus Merleau-Ponty knew that critical thinking means that “we take our fate in our hands.” This is the spirit of the critical thinkers.

- Germaine Greer (American social theorist and feminist philosopher, analyzed capitalism’s — and patriarchy’s — relegation of women to inferior status).
- Michel Foucault (French social philosopher, parent of deconstructivism and post-modernism).

- Jacques Lacan (French psychoanalyst, eloquent critic of Freudian psychoanalysis)
- Michael Polanyi (British biologist and philosopher of science, said science reflects human beings' preoccupations and beliefs, not some sacred "objectivity").
- Simone de Beauvoir (French feminist philosopher, looked at women's isolation in social and professional worlds of critical thinking)

## Critical Thinking

A fourth kind of "criticism" is called for when graduate students are told to "think critically." In a recent useful book on the subject, Julian Meltzoff (1998) writes:

I use the phrase *ability to think critically* as a positive quality, not as a pejorative reference to the characteristics of a meanspirited faultfinder. The phrase refers to the skill that one has in thinking about an issue, analyzing it, looking at it from all sides, and weighing whether there is sufficient evidence of good-enough quality to warrant making a reasoned judgment that is as free of personal bias as possible (p. xi, italics his).

What does he mean? Recall that I said that all the levels involve analysis and evaluation of the subject matter. Let's use Meltzoff's statement for our example. First, notice that he makes a claim: critical thinking is a skill, involving a set of things. Let's break this claim into its components, then evaluate them. The components are:

- *skill at analyzing,*
- *look at it from all sides.*
- *weigh evidence* in order to judge for such things as
- a sufficiency of
- good-enough quality to support our making
- a reasoned judgment that has the character of being
- as free of personal bias as possible.

I would ask additional questions about each element. A few examples:

What does he mean by "analysis"? How does one look at something "from all sides"? What makes up "evidence" and how does evidence get "weighed"? How will I know when my evidence has "good-enough quality"? What, really, is "quality"? (1) In the phrase "reasoned judgment," what does "reasoned" mean? What is "judgment" about? He wants us to root out "personal bias" as much as possible, but how much is possible? How do we identify "personal bias" anyway?

Hey! Why is personal bias bad? Why should we get rid of it? Whose purposes and needs are served by getting rid of it? And who decides how much evidence is sufficient? What gives them the power to decide this and how do they come to their decisions? When they decided these things, what was going on in the world that might have affected their

decisions? Where were they living, and what conditions were they working in that might have affected their decisions?

These are the kinds of questions that you get into, when you start analyzing something critically. This list of questions could go on, but I think you get the drift. In their simplest form, they are the famous journalistic questions: *Who? What? Where? When? Why? How?*

But what would we evaluate each of these components for? Here we encounter form, trustworthiness, relevance, and significance. Before doing this, I want to discuss what I mean by some basic terms you ought to know.

### **Some basic concepts in “critical thinking”**

#### **“Argumentation”**

This is probably the main term to become familiar with. All scholarly and scientific articles (including research papers and dissertations) are put in the form of an argument. Very simply, an “argument” is a claim supported by evidence in its favor. If I say to you, “You need to think critically,” that is a claim. (British philosophers began calling such sentences, indeed all sentences, “truth–claims.”) If I add, “Because there are five studies showing that grad students who do so do better in their careers,” I am adding some evidence (although who knows whether the studies really show that). To paraphrase the famous line from Jerry Maguire, critical thinking says, “Show me the evidence!” This combination of a claim plus evidence in its favor is the basic form of an argument.

The field of rhetoric (now called “Communication Studies” in many schools) studies argumentation in depth. Alan Bensley (1998) has a nice introduction to argumentation in his book *Critical Thinking in Psychology*. It goes into the various types of arguments and how to evaluate them critically.

#### **“Evaluation”**

In critical thinking, we evaluate the quality of arguments. When you stop and think about it, a research paper (say a class paper at TGSA) is really an argument: a claim presented along with supporting evidence. Your instructor will evaluate the quality of your argument. Or if a headline says, “Miracle Drug Cures Impotence!” you might think somewhat critically about the quality of the original research before rushing out to buy the drug. Was the research question well–formed? Were the methods suited to the question? Were the findings trustworthy? Was the study really applicable to the general public in the real world? Is this “cure” important?

So evaluating the “quality of an argument” involves making judgments about how “well–formed” the claim is, how well the evidence actually supports that claim, and how well the results apply to the desired audience. Numerous rules and tools exist for this evaluation, the rules and tools of critical thinking. But you may have noticed so far that I have not defined “quality,” except to analyze it into “well–formed” claims and “well–supportive” evidence. I’d say we are trying to get at the basic evaluative question: Is the study true? What might that mean?

## Operationalizing “Truth”

If I say that critical thinking evaluates the “truth” of an argument (substituting “truth” for “quality”), and if we decide not to “go deep,” (2) “truth” has to mean something we can all accept and use. This is called “operationalizing” a concept, defining it in a way the audience can operate with and agree on. In science, the safest path is usually to stick to what anyone can observe, measure, or perform. Could we agree that a “true” argument has a well-formed question which is properly based in the previous literature, and methods and results which are trustworthy, relevant, and significant?

We might, if we know what “form,” “trustworthiness,” “relevance,” and “significance” mean. OK, let's operationalize each of them.

### Form

Do you recall that in every argument a claim is stated, followed by supporting evidence? (The claim can be supported by presenting dissenting evidence and then refuting it, by the way.) If a paper or study has this structure, it is “well-formed,” at least for our purposes here.

Since each component of the argument can also be evaluated, each should also be well-formed. First, let's deal with proper form of the claim or “central question” (as Bensley, 1998, calls it). Then I will discuss the proper form for evidence. (You'll notice that I use many terms specific to the world of research methodology and design; I'll assume that you will follow up on them if you don't know what they mean.) Bensley (1998, pp. 231 ff) suggests a basic outline for most research (term) papers which includes the components: Introduction of the basic problem or question; development of the argument of the claim embedded in the central question (presentation of the evidence pro and con); and conclusion (including review and evaluation of the main points of the argument). These elements then can be viewed as contributing to good form in a paper. Let me take each in turn.

#### Well-formed topics (theses) and central questions

The central question is usually found in the Introduction to the paper reporting the scholarly study or research project. If it isn't, the report has poor form. It describes the main topic that was studied. As such, it conventionally takes one of these four forms. Each has its own sentence structure, and each has secondary tests for good form.

1) The central question can be written as a thesis, that is, an assertion about something. For example, “Depression is caused by a lack of chemicals in the brain” is a thesis. It has two elements of form: First, it should be a declarative sentence which unmistakably makes a truth claim. Second, a thesis must be arguable. This means that evidence supporting it (and dissenting from it) can be brought forward.

2) The central question can be written as a question. When I want to know about a topic and do not know enough to make a claim about it, I may want to learn more first. A typical course research paper is an example. So I pose my topic as a question aimed at learning what is out there. To have good form, first, a question should be a question,

which is, well, a sentence ending in a question mark. Second, a well-formed question is answerable. (In conventional scientific terms, a question like, What is God? is not answerable, because it cannot be operationalized. A question might be unanswerable for other reasons as well, such as the questions: How hot does the interior of the sun feel? If there falls a tree in the forest with no one to hear, will it sound? Or, in my wife's favorite version of that chestnut, If a man, alone in the forest, says something and his wife is not there, is he still wrong?)

The answer to the question will be the conclusion(s) of the study (which themselves are truth claims, which need to be based on evidence).

3) The central question can be written as an hypothesis. When there are well-researched theories about the topic and now I want to extend the theory into some new area, or I want to refine the theory and make it more accurate, I write my research topic as an hypothesis. First, hypo-theses must take the form of an "If . . . then . . ." sentence. Basically, an hypothesis says this: "If this already known theory (or thesis) is true, then this as-yet-unproved aspect will be thus and so."(3)

### **A Poorly-formed Hypothesis**

Here is a well-intentioned hypothesis: *Given that dopamine is associated with pleasure-seeking behaviors, if dopamine levels are too low, certain depressives should experience higher sex drives.* The problem? The then-clause is not testable (how can we observe or test internal subjective experiences like "sex drive"?). Operationalize "higher sex drives" as "more frequent sex-seeking behaviors" and you might succeed. Write hypotheses so that they can be tested and falsified, or your great insight will never see the light of day.

For instance, the catecholamine theory of depression holds that depressed people have an insufficiency of certain neurotransmitters in the brain. An hypothesis (sub-thesis) might be written: "Since other neurotransmitter levels are known to be lower in depressed people, there will also be a shortage of serotonin (an unstudied neurotransmitter) in depressed people." This uses different words ("Since . . . there will be . . ."), but has good "If-then" form.

Secondly, an hypothesis must be testable and it must be falsifiable. (4) "Testable" hypotheses can be tested. Simple enough. Take the hypothesis in the last paragraph. Reliable and valid tests could be devised to measure depression in the participants and serotonin levels in their brains. Then the results could be analyzed to see if indeed there is the hoped for correlation. Notice: the tests could easily show that the hypothesis is false as well: it is falsifiable.

4) Finally, the central question can be written as a prediction. Suppose the above hypothesis proved true and its supporting evidence was replicated a couple of times. How can we use it to treat depression? We need to link the serotonin findings to treatment variables. This calls for an experiment, and an experiment's central question is written as a prediction: *Selective serotonin re-uptake inhibitors (SSRI's) will decrease depressive symptoms in an experimental group, compared with a control group.*

For proper form, the prediction is written as a predictive sentence. And secondly, it must be provable (not just testable) and falsifiable. In other words, predictions are written so they can be studied by means of an experiment design in which variables can be measured against each other and against chance.

### **A Commonly Asked Question**

Does a thesis or a question need to be testable and falsifiable? Nope. Here's the difference. A thesis and a question do not make a prediction. Theses state something, and as long as their argument is solid, they succeed. Questions seek information, and as long as they can be answered, they succeed. Hypotheses and predictions, on the other hand, apply the already known to new areas. It is not simply a matter of arguing a good case or gathering information, but of showing the point. The tests of form here, then, are stricter: testability and falsifiability.

The other component of a good argument (the evidence supporting the claim) also should be well-formed. What is the proper "form" for evidence?

#### Well-formed evidence

Research papers of all kinds have four sections in the "evidence" sections: the literature review, the methods used to gather the evidence, the findings themselves (what was discovered, and the conclusion. You will be delighted to know that each has a proper form.

1) The literature review needs to be well-formed, as if you hadn't guessed. Put simply, the literature review is well-formed when it presents a citation for each idea which is not the author's and is not common knowledge; when it develops a logical argument which supports the hypothesis or prediction (if there is one) or the need for a question; and when each citation is formatted properly (according to APA style, for example).

Conventionally, a research paper presents the lit review before the statement of the hypothesis or prediction (since they should flow logically from the previous research reviewed in the lit review. Of course, a thesis or a question-based paper needs only the statement of the problem or question, since the research review often is the "methods and results" section.

#### 2) The methods.

A simple rule for evaluating the form of the methods is this: The methods need to follow the form of the central question. If it is a thesis, then the method must be a logical argument, and obey the rules of logic. If it is a question, then the methods must be properly designed to elicit information, and they should be appropriate to the question being asked. I saw a study once in which the question was "What do prostitutes say about their own self-esteem?" The method used was to administer a questionnaire about income and consumer spending! (The writer assumed that higher levels of spending relative to income was an indicator of low self-esteem, without showing any previous research which supported that assumption.) Thus, her methods were poorly formed, in that they did not appear to relate to the question.

If the question is put as an hypothesis, the methods need to be tests or instruments which are valid and reliable and can show what the hypothesis asserts. Once I read a study which had this hypothesis: Since it is known that a large number of child abusers were themselves abused as children, if abused children are given training in anger management, they will be less likely to abuse children themselves. Aside from the poor form of the hypothesis (it is not really testable at all, in our lifetimes!), the tests devised included giving the participants MMPI's. Heaven knows what that showed, but it wasn't that these kids would not abuse their own children!

### **But wait . . .**

Isn't common sense a kind of evidence? Take the "prostitutes self-esteem" study: isn't it common sense that higher spending might link to lower self-esteem? The answer: Maybe, but a research project must flow logically from the available prior research. Sorry. That's just a basic rule of research, meant to ensure that we can (in theory at least) depend on each step of the logic chain. First, the researcher should have done a different study to see if lowered self-esteem in prostitutes is actually linked to higher consumer spending — and then this study might make sense!

Finally, if a question is phrased as a prediction, the methods need to be experiments. Experiments can take many forms, of course, which is beyond our scope in this essay, but as long as they are real or quasi-experiments, the methods for a predictive study are "well-formed."

### 3) The Findings

The findings of your study — or any study — need to be presented clearly and logically. That's the real test of the form of the findings. Generally speaking, findings will "flow from" the methods and the question. If those earlier elements are well-formed, the findings section should be well-formed as long as it follows from them.

### 4) The Conclusions:

The main test of form for conclusions is logic: Does this conclusion flow logically from the premises? If the study is a thesis, has the thesis been demonstrated properly? If it is a hypothesis-driven project, do the conclusions support or disconfirm the hypothesis? If the study is in answer to a question, is the question satisfactorily answered and is the answer logical?

Get how important the rules of logic are in this game?

### **"Trustworthiness"**

By trustworthiness, I mean that the study can be depended on. It says what it purports to say and studied what it promised to study. Here are three main tests of trustworthiness: validity, reliability, and objectivity. Ever hear those words before?

1) "Validity" means that something does what it claims to do. You have a valid passport when the passport actually identifies you, the bearer. If I take your passport, I no longer

have a valid passport (because it does not identify me. A research test (say, the MMPI Depression (2) scale) is “valid” if it actually measures depression.(5)

2) “Reliability” simply means that if I repeat the test or study carefully, it gives the same results. A reliable pregnancy test, for instance, will give accurate results each time it is used. Same with your meat thermometer in the Thanksgiving turkey. May the research gods preserve us from unreliable meat thermometers, research instruments, and pregnancy tests (which are all very straightforward research instruments).

3) “Objectivity” does not mean neutrality. It does not mean impersonality. It does not mean getting two sides to every story (as journalists nowadays are apt to say). It can mean impartiality. Objectivity means showing the object clearly and from as many viewpoints as possible. In experimental studies, for instance, objectivity is protected by measuring everything, because there is no argument about how many “223” are, nor how long a meter is. A score on a valid and reliable test should mean about the same for all of us. Likewise, in the report of a field study, there will often be many details describing the setting, the people, the ambiance, interactions, and so on. Increased detail, especially sensory detail, supports objectivity, in that it permits any reader to “re-observe” what was originally observed in the study.

Thus, objectivity is a correlate of reliability.

For example, if you stop in an unfamiliar town and ask directions. Good directions will include landmarks and sensory markers for you (“When you smell the hog farm, you went too far!”), and when you observe them you’ll recognize where you are. Good directions are “objective.” Presumably, the next ten strangers could follow the same directions and arrive at the same destination.

Technically, objectivity is the quality of research which protects and preserves the reader’s ability to “see” the findings as clearly as the researcher “saw” them. That’s why “hard” scientists prefer quantification and statistics -- numbers are the simplest way to “see” what the scientist “saw.” But don’t be fooled by anybody’s claiming that quantified results are the only “truly” objective results. That’s nonsense. The key is whether you can observe what I observed on the basis of my descriptions.

So a trustworthy study has these three virtues: validity (Does the study do what it says it does?), reliability (Does it do it over and over, as often as we try it?), and objectivity (Can anyone look at the study and see the same things?).

What does a relevant study have?

### **“Relevance”**

“Relevance” touches the power of the study and its results to relate to my interests. As I use it, relevance has two components: “generalizability” and “applicability.”

1) The first test of relevance looks at “generalizability.” The best place to test generalizability is by looking at the sample or population studied in the research. If it can

reasonably be generalized to other populations, the study may pass this test. Generalizability usually is a statistical test, but there are some rough-and-ready gauges as well: Is the sample large enough to even think of generalizing? Is the original research question one that could generalize to other populations? If the necessary changes were made, could the conclusions be said to be relate to different populations? If the original research is on, say, animals, were those animals (or the parts of the animals) sufficiently like the groups to which we want to generalize?

2) “Applicability” asks two basic questions: First, is this study really relevant to the field of interest? Does it apply in the field it should apply to? For instance, if an article is published in a psychology journal about the mating habits of fruit flies, how is that applicable —either in broad outline or in one or more of its specific findings — to psychology? It might be, for example, if one finding is that certain neurotransmitters common to fruit flies and humans are overproduced during mating.

“Applicability” also questions if the study or its findings actually apply to the clientele of interest. Suppose I am a psychotherapist reading a study showing St. John’s wort to be successful in treating depression. The sample studied was 25 college-age students with mild depression as measured (operationalized) by scores of less-than-25 on the Beck Depression Inventory. I doubt this meets the first test, generalizability. But what about applicability?

Well, the field of interest, psychology and the psychotherapy of depression, could find this study quite applicable. Maybe it needs a bigger sample and certainly some replication and extension, but it certainly applies. But my “clientele of interest” is a 53 year-old twice-divorced mother of four drug-addicted kids, currently living with her abusive pimp and feels trapped in that life because she needs money for her youngest daughter’s abortion. Therefore, I must question the applicability of the findings of the study to my client (since she is so dissimilar to the experimental group, which may lack generalizability anyway), though perhaps not to my field.

So I must question the relevance of any study: Is the sample generalizable? Is the question applicable either to my field or to my clientele? Are the findings and conclusions applicable to either one?

As you can see, this is a restricted, operational meaning to the word “relevant.”

Next, we check the significance or meaningfulness of the study and its parts.

### **“Significant” or “Meaningful”**

Here’s a simple way to ask about significance: Is this study non-trivial?(6) If it is trivial, is is not significant. For instance, if I prove beyond any doubt that humans who don’t brush their teeth have bad breath, my findings are trivial. Another kind of triviality occurs if I do a study proving something again which has already been well-established (re-proving Einstein’s proof of the general relativity theory would be trivial, for example). Don’t waste time either doing or reading trivial studies.

Significance addresses how meaningful the study actually is in the real-world, whereas relevance checks the applicability of the study or its findings to various audiences. Here we ask, If the findings are true, do they matter to anyone? Are they important to someone, someplace? And of course, the “someones” are the audience: the field of interest and the clientele of interest.

Rephrase the question to be: Are these findings meaningful to anyone in the field or to my particular clientele? If they prove true, are they important enough to follow up?

Here we can think about the famous “statistical significance,” which is a different thing entirely. That simply means that the findings are very likely not to be due to chance. If you give St. John’s wort to the experimental group and all of them get better, while none of the control group, who got a sugar pill, improve — you probably have a statistically significant finding. It’s probably not an accident (assuming nothing contaminated your study, like the city pumping Prozac into the water supply during your project).

Since so many of the experimental folks improved, I would say most depressed people and most psychotherapists and psychologists would probably say that if these findings prove true, its important. Thus, the study may be both statistically significant, and have high significance in the sense of meaningfulness.

Now, if you do a study in which there is no statistical significance to your findings, you may still have meaningful findings which you should publish. A TGSA learner recently did a dissertation (Martin, 1998) in which he taught a group of sex offenders to practice mindfulness meditation, hoping that they would improve their self-esteem. But the measure of self-esteem showed no statistically significant improvement. However, he interviewed the participants after the study, finding strong feelings of enhanced self-esteem in most. His original self-esteem measure apparently did not get at these feelings (i.e., was invalid for these expressions of self-esteem), but the self-reports are important data. Further studies are needed, using different ways to measure self-esteem, to establish what mindfulness meditation actually accomplished. This study, although statistically insignificant, could be quite meaningful to sexual offenders and the people who work with them. Thus, I would call it significant, though not statistically so.

### ***Summary of the basic concepts discussed thus far***

I have suggested that critical thinking is a set of rules and tools for evaluating the quality of arguments. I described an argument as a truth-claim plus its supporting evidence. After discussing the idea of operationalizing definitions, I proposed four related concepts to operationalize our definition of “truth” in studies and scholarly reports: form, trustworthiness, relevance, and significance.

I applied the notion of an argument being “well-formed,” and applied it to the central question, the literature review, the methods, the findings, and the conclusions. All should follow the proper form.

Under trustworthiness, three terms played important roles: validity, reliability, and objectivity.

Under relevance, we looked at the generalizability of the findings and the applicability of the study or its findings to the field and to the needs of the clientele to whom it would be applied.

Under significance, we first looked at whether the findings are statistically significant (to be sure they are better than chance happenings), then whether they are important or meaningful, either to the field at large or to the clientele of interest. Another term for insignificance: non-triviality.

### **Some heuristic devices (“rules”) for evaluating the quality of arguments**

In this section, I’ll share some simple “rules” for critical thinking which you might find handy. Some I have borrowed from various authors, especially Bensley (1998), others I have made up. (Does that make them fiction?)

#### ***Evaluating the overall study and the central question***

##### **Rule 1: Keep it simple. Don’t over do it.**

Just run through the drill as you read an article and don’t try to get it perfect. Nothing in science is. Perfect, that is.

##### **Rule 2: If the whole report is well-formed, go on.**

(If not, I may bag it right here and find a better study to read, unless the potential significance and/or relevance of the study is great enough to warrant taking a chance.)

##### **Rule 3: Write out the central question. Write out the truth–claim it makes.**

##### **Rule 4: Ensure that the Question is in proper form for the claim made.**

A. If the claim is a straightforward assertion of some proposition, is the question written as a thesis? Is it a declarative sentence with an unmistakable claim in it? Is that claim arguable?

B. If the claim is that some basic information can or must be put together about the topic, is the topic written as a question? Is it in question form and can it be answered?

C. If the claim is that, based on existent theory, certain sub-theories should prove true, is it written as an hypothesis? Does it follow the “If– then” form; is it logically derived from the theories; can it be both tested and falsified?

D. If the claim is that certain variables will interact somehow, is it written as a prediction? Is it in a predictive sentence (with the future tense of the verb) and can it be proved or disproved?

##### **Rule 5: If the Question is poorly formed (Rule 3), STOP**

(Unless the truth claim is potentially significant and relevant enough to either my field or my clientele to warrant risking wasting my time.)

### *Evaluating the supporting and dissenting evidence*

#### **Rule 6: Look at the literature review next.**

A. Has the lit review good form? Does it cite all the previous research related to the central question, pro and con? Is there a logical order showing the basis for the hypothesis or prediction? Is each idea from someone else or not common knowledge cited properly?

B. Is the lit review trustworthy? Does it look into the literature thoroughly, showing studies that both support and disconfirm the central claim? Does it acknowledge weaknesses in the argument?

C. Is the lit review relevant to the field and to my clientele?

D. Does the lit review offer reasons to think this study, if it proves out, will be meaningful or important either to my field or to my clientele?

E. Do the hypotheses or predictions flow naturally and logically from the literature cited? Is the case strong enough that they are warranted and well-founded in previous research?

#### **Rule 7: Next evaluate the methods.**

A. Form: Do the methods fit the research question? Will they generate the kind of data that is being looked for?

B. Trustworthiness: Are the methods valid? Have they been proven to tap into what is being studied?

C. Trustworthiness: Are the methods reliable? Have they been previously shown to work time after time, no matter who uses them?

D. Trustworthiness: Are the methods such that the presence of the researcher did not contaminate the findings? Can they be described clearly enough that any reader will have little trouble understanding what was done and how each method was carried out? Do the methods give results which any reader can understand in the same way (with appropriate education or training)?

E. Relevance: Does the sample permit generalization? To whom? Is that “whom” a group that belongs in either the field of interest or the clientele of interest?

F. Relevance: Do the methods relate properly to the question? If it is a thesis, are the methods some kind of logical demonstration from evidence? If it is a question, do the methods involve literature searching, analysis of texts, gathering of answers in some form? If a hypothesis, is there at least one straightforward test of the hypothesis? Can that test prove the hypothesis false? If it's a prediction, can the tests prove or disprove the predicted correlation?

G. Significance: do the methods chosen generate the sort of information which could be meaningful to the audience — either to the field or to the clientele of interest?

H. Trustworthiness: Does the writer show awareness of limitations or weaknesses in the methods or design of the study? Are these analyzed honestly? Were there attempts to shore up the weaknesses or otherwise account for them? Did these save the study from lethal flaws? If the answer here is NO, stop, unless the potential meaningfulness outweighs the risks of depending on this study.

**Rule 8: Evaluate the findings and conclusions.**

A. Apply the same tests to the findings, as appropriate: form, trustworthiness, relevance, and significance.

B. Form: Look at the logic of the discussion of results. Is a case made that supports the claim inherent in the question? Is the question even answered? Does the writer seem to be aware of this and discuss it?

C. Form: Is the logic of the discussion strong enough to make the case that the findings indeed support the original truth claim? In the case of the thesis, do they successfully demonstrate the point? If the study is answering a question, does it succeed in doing so thoroughly? In short, are you persuaded by the discussion and the conclusions sections that the author's original claim is true?

D. Trustworthiness: Does the writer present contrary findings? Are they discussed logically, weighed against the supportive evidence? Are weaknesses and limitations in the study and in the conclusions discussed frankly, and remedies offered if available?

E. Trustworthiness: Do the findings appear to be valid, reliable, and objective? If not, does the author discuss this and offer any ameliorating arguments or perspectives?

F. Trustworthiness: Could other extraneous influences account for the findings? Does the writer appear aware of these? Are redeeming circumstances offered to offset the unexpected contamination by outside factors? Does these work? Is the study fatally flawed?

G. Relevance: Can the findings and results generalize relatively well? Are they applicable to the audiences: the field of interest or the clientele of interest?

H. Significance: Do the findings mean anything important to the relevant audiences? If not, are they trivial? If they were "tweaked" or elaborated, might they become significant? Would they be worth the work or the wait?

**Rule 9: Decide whether you wish to consider this study dependable and safe to rely on in your professional and personal life.**

A. There is no sure score any article should have. I generally discard articles or studies which don't meet basic criteria for form, trustworthiness, relevance, and significance. But even if a research report is poorly put together, it may have important and powerful results, so I tend to offer much leeway in my reading.

B. Where are you going to rely on these studies? If you want to use them in court, be sure that nearly everything came up "yes." If you use the results in clinical work with clients,

be sure that the vast majority of them came up “yes.” The farther away you get from “handling” other people’s well being, the more loose you can be with your own criteria.

C. Remember, though, that most of the professional ethical codes demand of us that we not speak in our professional capacity unless what we have to say is reasonably well grounded in empirical research. In other words, if you are relying on the findings of a given study, keep it to yourself unless you have a reasonably strong set of “yes” answers to these many questions. Got it?

## General Summary

When we think critically about research, we are breaking cultural, scholarly, or scientific reports into their elements and evaluating each of those components. In research, the claim takes various forms, most often stated as a research question or an hypothesis/prediction. Methods are chosen to collect answers to the question, and these data are then analyzed in some fashion and conclusions are drawn which either support or disconfirm the original claim. In other cultural and scholarly works, similar components are used.

Critical thinking evaluates each of these components to discern whether they do their job. That is, whether they logically and sensibly support or disconfirm the claim asserted at the opening of the report. Specifically, we critique four dimensions on all the components: form, trustworthiness, relevance, and significance.

In general, the trustworthiness of a study is shown by its validity, reliability, and objectivity, although other terms with broader applicability can be used in qualitative research models. Relevance is demonstrated by the generalizability of the results from the sample and by the study’s applicability to the field of study and to the professional clientele for whom its findings will be used. Significance means the degree to which the findings and conclusions are or will be meaningful and important in developing the field of study or assisting the clientele in the real world: will the research be helpful in ways that improve on the current situation? Another measure of significance is non-triviality. Or as Marx might put it, the point of research is ultimately not merely to understand things, but to change things.

This is a truth claim about research for which I have not provided any evidence in this small introduction. I leave that to another day.

My aim has been to say two basic things:

First, that all scholarly projects are presented in the form of an argument, having at least a claim about something which is then supported by evidence.

Second, that each component of the research–argument — the central question or hypothesis, the methods, the findings, and the logic of the conclusions — can be criticized along four key dimensions: form, trustworthiness, relevance, and significance.

If you have taken these two points, you have mastered the heart of this essay.

## References

- American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders*, 4th Edition. Washington, DC: Author.
- Bensley, D.A. (1998). *Critical Thinking in Psychology*. Pacific Grove, CA: Brooks/Cole Publishing Co.
- Cowley, M. (1984). *A Second Flowering: Works and Days of the Lost Generation*. New York: Penguin Books (Viking).
- Fennick, J.H. (1997). *Studies Show: A Popular Guide to Understanding Scientific Studies*. Amherst, NY: Prometheus Press.
- Feyerabend, P. (Verso Edition, 1978). *Against Method: Outline of an Anarchistic Theory of Knowledge*. London: Verso.
- Habermas, J. (trans. 1971). *Knowledge and Human Interests*. (trans. S. Shapiro). Boston: Beacon Press.
- Kuhn, T.S. (1972). *The Structure of Scientific Revolutions*, Second Edition. Chicago: University of Chicago Press.
- Meltzoff, J. (1997). *Critical Thinking about Research*. Washington, DC: American Psychological Press.
- Merleau-Ponty, M. (1962). *Phenomenology of Perception*. London: Routledge & Kegan Paul.
- Martin, G. (May, 1998). *The Practice of Mindfulness Meditation and Its Effects on the Social Self-Esteem of Sexual Offenders*. Dissertation approved, May, 1998, The Graduate School of America. Publication date unavailable.
- Polanyi, M. (1962). *Personal Knowledge: Toward to Post-critical Philosophy*. Chicago: University of Chicago Press.
- Robson, C. (1993). *Real-World Research*. Oxford (UK) and Cambridge (US): Blackwell.
- Ross, C.A. & Pam, A. (1995). *Pseudoscience in Biological Psychiatry*. New York: John Wiley and Sons.
- (1) Didn't Phaedrus, in Robert Pirsig's classic, *Zen and the Art of Motorcycle Maintenance*, struggle his entire life to define that term, and give up in the end? Isn't it the hardest word in English to define, except "God"?
- (2) Recall how the British philosophers called sentences "truth claims." What does "truth" mean? Do you notice right away the tendency to go "deep" (that is, get philosophical)? If we're going to talk about "truth," shouldn't we discuss "absolute" vs. "relative" truth? Are there more than one kind of truth? Does truth mean something else

than “conforming with the norm”? You can think of many such questions. But here we won’t “go deep.”

(3) Thus the term, “hypothesis.” The Greek preposition “hypo-” means “under” in the English sense of “subsidiary.” So the “hypo-thesis” is the sub-thesis, a subsidiary which follows from the thesis or theory in question.

(4) No use getting into the technical stuff here. K. Popper (1962) first proposed these as evaluative tests of a good hypothesis. You should know that other philosophers of science such as Paul Feyerabend (1972), disagree, and say the only good evaluation of an hypothesis is whether it lets us find interesting things. Put another way, Feyerabend argues that falsification is not a test of an hypothesis, but only whether the facts falsify it. But here, we will stick with the conventional rules and tools.

(5) Here we hit “operationalizing” again. “Depression” has to be defined in a measurable way for the MMPI-2 (2) scale to measure it, eh? So we define depression operationalizably, workably, usably. For an example, look at the DSM–IV definition of depression (American Psychiatric Assoc., 1994, p. 327).

(6) In statistics, non-trivial has a technical meaning. But I am using it here more broadly, although statistical significance is included in my usage.