An Academic Research Publishing Checklist

Brian Whitworth

Institute of Information and Mathematical Sciences,
Massey University, Albany, Auckland, New Zealand
Email: bwhitworth@acm.org

June 26, 2007 Under publication submission – please do not quote without author agreement.

Abstract

While traditionally information systems (IS) PhD students first completed a thesis then published, today they often publish before completing. While at first publishing seems yet another PhD student burden, it can be a useful learning experience, raise motivation, provide helpful feedback, and unite student and advisor. PhD publications also help with grant and job applications. However PhD research publishing is an extra demand, suggesting the need for a learning support tool. The PhD publishing checklist proposed: 1. Chunks knowledge into elements, for easier handling. 2. Grounds the elements, with examples and summary statements, for practicality, and 3. Structures the elements, in academic format, for easier location. Research is currently underway, to discover the value of this checklist.

Introduction

Background

Once PhD students first graduated then published, but IS PhD students today publish an impressive number of papers (Larsen, 1998). One could argue that putting students on the publication treadmill is stressful, but perhaps publishing is part of what a PhD should teach, letting new researchers "cut their teeth", as it were. Publishing benefits include student recognition, motivation, useful committee connections, justifying grant support, and a common advisor/student focus. Third party reviews expose students to comments outside their advisor, and give advisors a useful third-party reference. It is also very unlikely that a student with successful publications will be failed. Assisting students to "publish and prosper" has become an important advisor duty, and students whose advisors don't do this may be disadvantaged when applying for a job. While publishing can seem yet another burden, it can also contribute to what a PhD should be: a higher learning experience.

Goal

PhD publishing now seems established, and so needs to be a PhD planning forethought not an afterthought. This new demand on students and advisors (to publish) suggests the need for a support tool in the process. This paper suggests a first draft checklist for consideration by the community. It is intended to be used as support during student/advisor interaction, not as a stand-alone document. While doing research itself is a major effort, writing it up also deserves a special focus. One reason why some publish frequently is that they submit frequently, as research that is not written up and submitted never gets published (Van Slyke et al, 2003).

Scope

This paper is not about the political, financial and motivational issues of getting a PhD (Grover, 2001), or how to get a PhD (Phillips & Pugh, 1994), nor about how to "work" the academic publishing system (e.g. choosing "least publishable units" or "publication recycling"). Such responses to over-zealous tenure committees are understandable, but seem not to advance science as a whole. As a PhD gives the time and opportunity for genuine research, the focus is on publication excellence not academic advancement.

Approach

Advice to new authors on academic publishing must recognize that research is:

1. ***Complex.*** Science, a way of approaching "Truth" that humanity has developed over hundreds of years, has many aspects. No "formula" or cookie cutter template can define it, and research has no "right" way. Each instance is a complex balance of many requirements, and good research is often a "golden mean" of many things.

2. ***Abstract.*** Given the topic's complexity, knowledge about it is often stated abstractly. Each abstract principle has innumerable concrete examples. To interpret an abstraction in a specific case usually requires an advisor, usually someone who has gone through the process themselves.

3. ***Changing.*** Research is not just mechanical observers impartially gathering data to fill up the picture of knowledge. If science were like this it would already be complete, which it is not. Scientific knowledge not a body of facts but a way to understand more - it is not a static *product* but a dynamic *process.* The scientific

method is not about what "is", but about how to look at what is. Its premise, **that we don't know,** means asking the right questions in the right way to increase understanding.

The *complexity* of research suggested "chunking" the knowledge into digestible elements. Students and advisors can select the elements relevant to their particular situation, and ignore those that do not apply. The *abstractness* of research suggested using the advisor-student interaction, so elements needed summary statements and practical examples. Elements can be read "cold", but are better used as discussion points. The *changing nature* of research suggested structuring the process in an academic layout, to show how the elements are an integrated whole, and to allow for easy location. *Since the PhD research journey is complex, with abstract signposts on an ever changing terrain, students need experienced advisors to give advice, interpret signs, and warn of dangers.* The following aims to divide the knowledge typically exchanged between advisor and student working on an academic publication into relatively discrete elements, to make each element pragmatic, and to structure the elements in a typical order.

Constraints

Some constraints must be noted:

a. *This is a checklist for discussion not a scientific method summary*. The paper aims to aid advisor-student PhD discussion, not to summarize the scientific method, which would need a book (e.g. Rosenthal & Rosnow, 1991). The goal is to "tag" the knowledge element not specify them. If an advisor finds an element wrong or incomplete, they can expand or correct the point, or signpost into a larger text.

b. *Element relevance varies between topics and methods.* Different disciplines will prioritize elements differently, and see some as entirely irrelevant, e.g. qualitative studies don't need statistical analysis. In this "toolbox" approach, that not all tools (elements) are used in every case is not surprising.

c. *Order and terminology will vary between topics and methods,* e.g. element order within sections is often not critical, and different disciplines may use different names for similar things.

A knowledge elements checklist

The following is a work in progress, so critique is invited. Elements specific to a particular field or method are marked with an asterisk (*). Note that the order presented below is not necessarily the order research arises, e.g. it often takes students a very long time to formulate the research question

Overview

1. ***Collegiality***. Research is a social activity as well as an intellectual one, done by people together not alone. The scientific community is a group whose *culture* is the scientific method. This culture combines logic and physical world data, believing that theory and practice work better together than apart. Logic alone seems like a sword, and those who live by the sword may die by it, and the detail of data can overwhelm, like a forest in which one can easily get lost. In this analogy, science uses the sword of logical argument to cut through the forest of physical data and find new pathways to understanding. In this culture members resolve disagreements not by conflict, but by research, using the scientific method. Scientists may disagree about ideas but agree to abide by research results. Critical to this community is respect. One respects research colleagues by reading and quoting their research. To ignore other's work, or to plagiarize it, disrespects the entire scientific community.

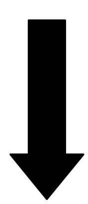
- **Respect the scientific community by reading, referencing and building on other's work**.

2. ***Scientific thinking.*** *Rationalization* uses the intellect to argue backwards, from conclusion to reasons, i.e. the conclusion comes first then the reasons, e.g. after making a mistake the mind quickly finds reasons to "explain" it. These are usually excuses or justifications, thought up "after the fact". In rationalization the arrow of thought goes backwards, and so it adds little new. *Scientific argument* tries to use the intellect to argue forwards, from present "facts" to unknown conclusions, which method Socrates called the *dialectic* (Buchanan, 1948). His conversations began with agreed facts, then bravely "followed the logic" wherever it led. Scientific thinking drives the arrow of thought forward, creates new ideas and predicts the future.

- **Give reasons before conclusions, and derive conclusions from what went before.**

3. ***Academic format.*** While academic publications have many forms, most follow a standard five chapter logic (Perry, 1998):

   a. **Introduce Topic**
   b. **Review Theory**
   c. **Describe Method**
   d. **Analyze Data**



e. Discuss Results

Qualitative methods fit into this structure as well as quantitative ones (Perry & Zuber-Skerritt, 1994). While heading names vary, and different publication types expand, combine, delete or contract sections, the general logic seems common to many forms, e.g. research in progress papers have no results section but otherwise follow the same sequence. This general academic "form" helps researchers think forwards rather than backwards, and weaves a subtle path between the world of abstract theory and the world of concrete events (Figure 1). This path usually begins with a practical problem, then moves to relevant abstract theories, which analysis gives an abstract research question. From that question an appropriate practical method is chosen, to create concrete data which can be analyzed to answer the research question, and impact current theory and the practical problem the research began with. Scientific research is "powered" by the abstract, but "grounded" in the concrete, and understanding this relationship is perhaps one of the most difficult concepts for research students (Swan & Martin, 1994).

- **Connect ideas together to form a logical and consistent scientific argument:**
  - **Use the introduction to define the theory scope.**
  - **Use the theory review to derive a research question**
  - **Use the research question to choose an appropriate method.**
  - **Use the method to create physical based data**
  - **Use the data analysis to form conclusions about the research question**
  - **Use the conclusions to discuss general implications (theoretical and practical).**

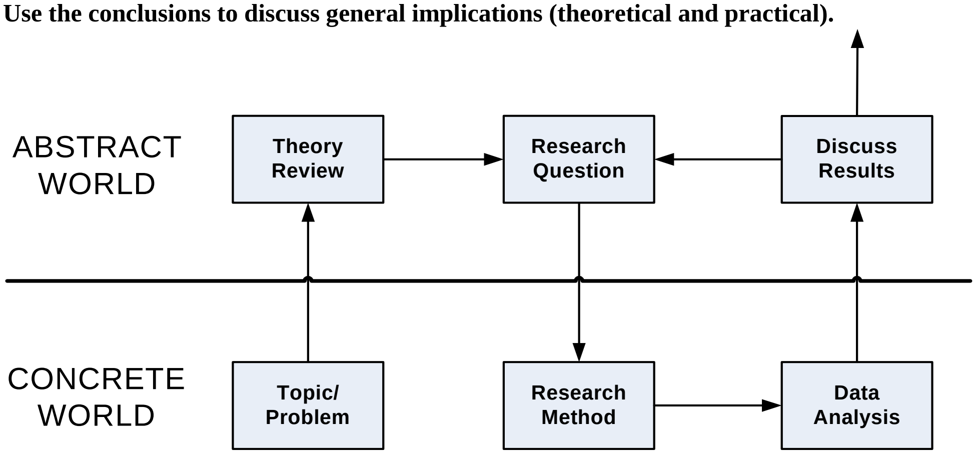


Figure 1. A Research Logic

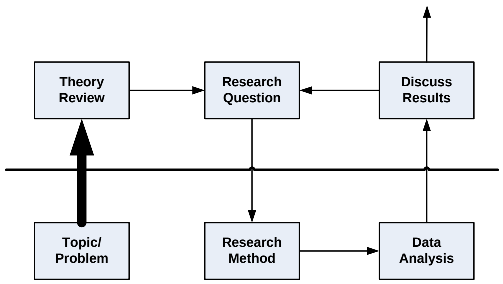
4. **Correct language** While one should not judge a book by its cover, unfortunately many do, assuming that those who cannot spell correctly also cannot think correctly. Spelling or grammar errors, especially early on, suggest a lack of professionalism, causing reviewers to look for and find further faults. Always run a computer spell-check last thing, and have a native language speaker check the document.

- ***Check for and fix language errors, especially early on, e.g. spelling and grammar*.**

5. **Opinion statements**. Indefensible statements cannot be defended against critical review, e.g. "Technology improves efficiency" cannot be defended against the critique that some technology does not. Indefensible statements can reduce reader trust and respect, and to reviewers such statements are like a red flag to a bull. Reduce indefensible statements by providing more support, such as supporting references, or by reducing the statement's strength to only what is necessary for the argument. Reduce the overstatement, e.g. "Technology can improve efficiency", or moderate an absolute, e.g. "Some technology improves efficiency", or hedge the statement, e.g. "Many argue that technology improves efficiency", or make it a question, e.g. "Does technology improve efficiency?" If you state opinions as facts, the reader who disagrees may be the reviewer. Delete or modify indefensible red flags, especially if not crucial to the argument.

- ***Avoid opinion statements, which cannot be defended against critical review*.**

Introduction



The introduction sets the publication's context, answering questions like what is this research about, why was it done and how did it originate? It *positions* the paper relative to others. Many readers use the introduction to decide whether to read on. It may include any or all of the following:

1. ***Title.*** The title should briefly convey what the paper is about usually the topic and contribution, e.g. 'Websites" is too vague, "The usage of Websites" describes the topic, and "A Study of Factors Influencing Website Usage" adds the contribution. Also, an interesting title may attract more readers, e.g. Website Usage: Why people Come? may be even better.

- **A title should describe the topic in an interesting way and invite the reader to read on**.

2. ***Author(s).*** The first author is usually the person who did most of the research and writing. In academic publishing, author order is often based on contribution not seniority, so professors may put their names after their students. Put a student's name first if they did most of the work, even if the idea was the advisors. Students learn a more when they are first authors. All co-authors should understand the paper well enough to adequately present it at a conference. If not, recognize their contribution by acknowledgement rather than authorship.

- **List authors in order of contribution not seniority, where each could present the publication at a conference.**

3. ***Abstract.*** An abstract should convey meaning, e.g. "This paper gathers data relevant to current theories and analyses it" says nothing. Don't "tease" readers, e.g. "The paper draws some critical conclusions." Don't state you made a conclusion, *state the conclusion*. Don't think: "If I give my findings in the abstract, they won't read the paper." If your abstract gives no value, people won't read it either. The abstract usually gives at least: 1. The paper's purpose (e.g. "This paper investigates a new technology acceptance variable."). 2. A findings summary (e.g. "User control increased user acceptance."). 3. Some implications (e.g. "Systems giving user control are more accepted.")

- **An abstract should tell the reader the paper's main purpose, results and implications.**

4. ***Problem.*** The reason for research is usually some practical problem, e.g. cancer research is based on the problem of cancer. Most scientific theories have a practical base because problems activate people to solve them, so highlighting a problem invokes reader interest. If there is no problem, why was this research done? Stating the problem your research addresses answers the reader's question: "Why bother reading this?

- **State a practical problem which the research addresses.**

5. ***Topic.*** The topic is the general subject area. It tells readers if your paper is in their area of interest. Even if the title gives it, state your topic early. Never assume, as authors familiar with their topic can do, that the topic is obvious. It is not. Choose *one* topic, and refer to it with consistent words throughout the paper, e.g. technology acceptance, technology diffusion and technology evaluation are different topics with different target audiences. Don't "morph" the topic into something else as the paper processes, e.g. state the topic is technology evaluation (of existing products), then do technology market research (of potential products).

- **Choose <u>one</u> topic, and refer to it consistently throughout the paper.**

6. ***Purpose.*** The purpose statement adds to the topic (what the paper is about) what the research aims to achieve. Editors expect a sentence like: "This paper aims to …" within a page or two. Finding several different purposes is a bad sign, as a paper should have a one purpose, stated in one sentence. More is not better when it comes to purposes. Reading the purpose addresses the readers question "What is this paper about?"

- **State the research purpose in a single sentence within the first few pages.**

7. ***Publication type.*** Research publications have many forms, e.g. *theory* papers on new models, *reviews* of the current state of the art, *empirical* papers with data, or *proposals* of methods. Science has three sources of value: *data*, *logic* and *past research*. Data is the any information the researcher has collected from the world. Logic is any analysis that combines ideas and information to create new knowledge. Past research is past data and logic, often distilled into theories. All three sources are not necessary, e.g. reviews analyze past research but present no data. Identify your publication type early, so readers know what to expect, e.g. a research proposal should not say: "This paper investigates …" as readers will be disappointed to later find no investigation is presented.

- **State publication type (e.g. review, proposal or empirical study) early so readers know what to expect.**

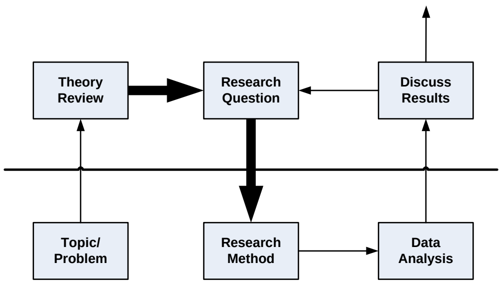
8. ***Background.*** A little background history can "warm-up" readers to a topic, as unarguable historical facts can "set the stage", and explain how the current situation arose. Good papers always "tell a story", and background history helps connect the present to the past.

- **Give a brief background history introducing the research.**

9. ***Target audience.*** The *relevance* of research is indicated by the number of people interested in it, and is distinct from its *rigor* or scientific quality. Defining the relevant target audience, by statements like 'This paper will interest software users." helps those readers relate to it. Stating a target audience addresses the research's relevance, and helps define the way the paper is written, e.g. is it for a technical group or a user group?

- **State the target audience, i.e. who is expected to read this paper.**

Theory



This section reviews relevant theories, explores their agreements and contradictions, and then chooses a theory base to use. It may include any or all of the following:

1. ***Scope.*** Defining topic scope is like drawing a line around a territory, to define what is in and out of the paper. State not only what the research is about, but also what it is *not* about. The scope is often poorly stated by new authors, who like to wander freely in areas irrelevant to their topic.

- **Define the boundaries of the research study and stay within that scope**.

2. ***Literature review.*** Every topic sits within a previous knowledge context, which must be recognized. A "State of the Art" literature review tries first to *honestly* summarize current research, like a reporter reporting facts, including those that disagree with your view. It also aims to be *complete*, to cover those important in the field. Finally it *analyzes* to find themes, agreements, conflicts, contradictions or gaps. Don't just give a sequential summary of what others have done, e.g. A did this, B did that, C did the other, etc. Structure the review by issues, e.g. compare and contrast many authors on one issue, then repeat for other issues.

- **Analyze the literature by issues (agreements/contradictions), not as a sequential list of other's research.**

3. ***Topic construct/variable.*** One result of a good literature review analysis is the abstract *topic construct*(s). Quantitatively measured constructs are also called *variable*(s), e.g. the construct "Usage", measured by the number of web hits or time on site, is a variable. Don't confuse the topic construct/variable with its causes or its effects, e.g. a thesis entitled "Web Site Effects" begs the question "Effects on what?" This thesis wandered across many topics like customer relationship, user acceptance, web usage, and company marketing, any one of which was a study in itself. It gave weak results because the literature review failed to well define the topic construct.

- **Clearly define the main topic construct/variable, which the research is primarily about, and distinguish it from other construct/variables.**

4. **Causal Effects.** A literature review should identify constructs or variables that affect the topic. In a causal relation, those affected by others are *dependent* (e.g. web usage), and those that cause change are *independent* (e.g. web site ease of use). A variable can be both dependent and independent, e.g. ease of use affects web site usage which in turn affects web site sales. *Moderating variables* alter a causal effect, e.g. user age may moderate how ease of use affects usage, as young people learn faster. *Intervening variables* mediate causal effects, e.g. user attitude may intervene between ease of use and usage. In describing theoretical relations, or "effects", distinguish what is being caused from what is causing.

- **Distinguish what is causing from what is being caused in casual relations between constructs.**

5. ***Theory base.*** The value of theories that connect abstract constructs is that they can describe innumerable specifics, e.g. the law of gravity applies generally to any falling object. The disadvantage of theory is that, being abstract, it can also become disconnected from concrete reality, e.g. superstitions. Every publication needs a *theoretical framework*, as research without theory has no context. Collecting data without theory often wastes time and effort – like a lost person who wastes time wandering about. Use the literature review to choose the model you will use, e.g. an existing theory, a theory combination, a theory modified by you, or a new theory of your own. This conceptual framework directs your research, and ideally should be given in a diagram.

- **State the theory base (or conceptual framework) this research will use, ideally with a diagram.**

6. ***Research question.*** The research question (RQ) is the core of your research. It comes before the method, as one chooses from many methods the best way to answer research question. A good research question should: 1. *constitute a single question* (state in a single sentence). Different questions need different studies to answer them, though one RQ may break down into sub-questions, a., b., c. etc. 2. *put a real choice*, e.g. "Does fear make one afraid?" is a circular question, as fear by definition is being afraid, so this is like asking "Is 1 = 1?" Real questions have more than one possible answer, and so are "falsifiable" by data. Science tests its theories by asking risky questions. 3. *be as focused as possible*, so that to answer it is feasible, e.g. "What affects life?" is a bad RQ as everything affects life. The more focused an RQ is, the more likely the research is to finish.

- **A research question is a single sentence (clear), with many possible answers (falsifiable), focused enough to allow feasible data collection (testable).**

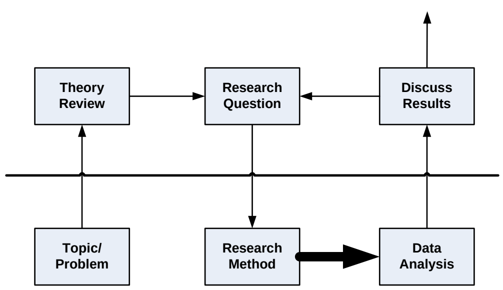
7. ***Research type.*** Empirical research (qualitative or quantitative) can discover, describe, connect, or causally link constructs. The research types are exploratory research, descriptive research, correlational research and experimental research: *1. Exploratory research*, like grounded theory, uses data to "uncover" initially unknown or poorly defined variable(s), so is useful for new topics*. 2. Descriptive research* gathers data to better define variable(s), e.g. questionnaire factors, and is useful for evolving areas. *3. Correlational research* shows connections between variables, but not what causes what, e.g. longitudinal studies of web site usage. Use it where one cannot realistically manipulate variables, *4. Explanatory research* establishes cause and effect by manipulating a *treatment* (independent variable) and measuring its *effect* on a dependent variable in "experimental" conditions. While research *control* usually increases in the order given, *realism* often decreases in the same order, e.g. case studies are usually more realistic than experiments. Exploratory research is the most widely applicable, and explanatory studies much less so. Different research types are not "higher" or "lower", merely more or less appropriate to different situations. To use exploratory research for a well established topic is as misplaced as to do an experiment in a new and poorly understood one.

- **Choose a suitable research type (exploratory, descriptive, correlational or explanatory) that is also feasible to do.**

8. ***Hypotheses*.*** Explanatory research restates the RQ as hypotheses, i.e. abstract statements whose opposite results can falsify, e.g. the RQ: "Do subjects prefer polite software?" gives the hypothesis "Subjects will use polite more than impolite software", and its "null" hypothesis "Subjects will respond equally to polite and impolite software". The null hypothesis assumes a random data distribution, so if it is statistically unlikely, at say less than 1 in 100 chance, we can reject it. Then, its opposite, the hypothesis, is "supported", else it is "not supported". Science does not "prove" its theories, but continuously opens them to falsification by the real world.

- **A hypothesis is a statement whose opposite the data results can falsify.**

Method



The research method states how the research question was put so that physical world data can answer it, and can include any or all of the following:

1. ***Methodology theory.*** If the method is uncommon or new, briefly present its theory, e.g. factor analysis theory, interpretive method theory or conjoint analysis theory. The method theory is not your research theory, so keep it brief – readers familiar with factor analysis don't want pages of factor analysis textbook theory. While your *research theory* needs extended coverage, your *method theory* does not. Refer readers elsewhere for details.

- **Only present methodological theory for a new method, and even then only briefly.**

2. ***Qualitative vs. Quantitative***. The basic principles of science (validity, reliability, replicability, generalizability, etc) apply both to qualitative and quantitative data. Data types are not in competition, though the jargon differs, e.g. qualitative "constructs" vs quantitative "variables" are just different methodology names for the same thing. Qualitative studies can better convey meaning but may struggle for precision, while more precise quantitative studies can struggle for meaning. Either way, justify the main data type gathered. A *mixed-method approach* can gather complementary qualitative/quantitative data, e.g. interview subjects after an experiment.

- **Use qualitative and quantitative data (mixed-method approach) as appropriate**

3. ***Pilot study.*** Research is complex, and one mistake can nullify the project. A pilot study tries out the method and research tools on a few cases to uncover problems, e.g. poor questions. Briefly report pilot test data, and state how the method was modified as a result, e.g. how research tools like questionnaires were improved.

- **Report how pilot testing improved the research method or tools.**

4. ***Research Design.*** The research design is *the method logic that ensures that the data collected correctly answers the research question*, e.g. case study design, repeated measures design, longitudinal design, etc. Both qualitative and quantitative methods try to ask questions in an *unbiased* way, so the answer is from the world, not researcher bias. Good research design ensures the right answer to the research question is obtained. Different designs have different requirements, e.g. in repeated measures design, where subjects try first one software product then another, product order is randomized, so it does not affect the results. Many experimental designs

require *subjects to be randomly assigned to treatments*, e.g. subjects evaluating two web sites cannot choose the one they view, lest differences found be due to this *spurious* reason (their preference). Qualitative research is equally vulnerable to spurious effects, e.g. the effect of the researcher watching.

- **Describe the method logic that ensures the data really answers the research question**

5. ***Control group*.*** The control group are subjects who do everything the same but do not receive the treatment, e.g. in medical research they receive a placebo sugar pill not the new drug. This controls for spurious influences if subjects are randomly allocated, i.e. don't let subjects choose to be in the control group. A control group is important where subjects know they are being observed, e.g. In the "Hawthorne effect" experimenters found that painting a room a brighter color improved work rate, but later found painting it dark had the same effect! What improved the work rate was not the painting but *being observed by researchers*. A control group avoids such problems.

- **Randomly allocate subjects to a control group to reduce spurious effects**

6. ***Measurement.*** State how you actually measured/investigated each variable/construct, as data can be gathered about one construct in many ways. Justify your measurement tools or method, whether taking qualitative or quantitative measures. Qualitative research may reference methods like action research, case study, grounded theory or ethnography.

- **State how you measured what you are trying to measure, and why you chose this way to gather data.**

7. ***Reliability.*** Good data should be as reliable as possible, i.e. change little when measured at different times or with different people. Reliable measures are stable, with low errors caused by the measuring itself. Qualitative studies use standard methods, keep detailed journals and analyze self-consistency to get reliability. Quantitative studies use 1. *Test-retest* reliability checks how a measurement changes if repeated, e.g. give a test then give it again a day later, or 2. *Split-half* internal-consistency reliability checks if the measure is consistent within itself, e.g. compare the first half with the second half of a test. Reliability coefficients like Cronbach's alpha are accepted if 0.85 or higher. One way to ensure a measure is reliable is to use one that someone else has tested.

- **Argue why the data is reliable, e.g. by precedence, test-retest or split-half checks**

8. ***Validity.*** Good data should be as valid as possible, i.e. *measure what it is supposed to measure*, e.g. foot size is a reliable but invalid measure of intelligence. The measurement should be agreed to represent the abstract construct. Validity can be established in three ways: *1. Content validity*, where the construct content is self evident in the measure, e.g. asking "Did you enjoy the web site?" as a measure of web site enjoyment. *2. Criterion validity*, where one measure is validated against another already accepted as valid, e.g. validating a web clicks "interest" measure against purchase data. *3. Construct validity*, that the measure is a single construct that "hangs together" (convergent validity) and separates itself from other measures (discriminant validity), e.g. a variable like "Circumstances of the Economy" lacks construct validity, if it is not a unitary variable. Statistics like factor analysis can define the factors of a multi-dimensional construct. One way to ensure a measure is valid is to use a tool someone else has validated.

- **Argue why the data is valid, with content, criterion or construct validity**

9. ***Unit of research.*** The unit of research, or *case*, is the data collection unit, i.e. a case is one data collecting act, e.g. in an online group voting study the case could be the <u>vote</u> (choice), the <u>individual</u> (satisfaction), or the <u>group</u> (agreement), or all three. In a raw data table, where the variables are columns, each line or row is a single case. Often the data case is one subject, but it can be a transaction or a group. The case affects the number of measurements (N) of the analysis, e.g. 90 subjects could give 900 votes (N=900), 90 satisfaction ratings (N=90), or 18 five person group agreement ratings (N=18), depending on the study's case.

- **Specify the unit of research of your study, i.e. one data gathering "case"**

10. ***Procedure.*** Describe the specific data collection procedure step by step, which can be face-to-face, by telephone, by self-administered questionnaire or website, etc. e.g. "After entering the room subjects were read the study instructions (Annex A), then signed the consent form (Annex B)…" Provide copies of any scripts, e.g. the statement read to introduce subjects to the study. Also describe any procedure deviations, e.g. subjects who left the room or talked to a friend.

- **Describe in sequence the data collection procedure, with any tools used, e.g. introduction script**

11. ***Task*.*** Describe what subjects had to do, e.g. give the task instruction script given to the subjects. Also describe any training given (to reduce initial skill variations).

- **Describe the task and any instructions involved.**

12. ***Question design*.*** Good questions reduce error variance, e.g. subjects who just "tick the boxes". Good questions are: *1. Understandable*, e.g. simple words and grammar, *2. Unambiguous*, e.g. don't use double-barreled questions, *3. Unbiased*, e.g. a question phrased "Don't you agree that …?" is biased, A question can be stated in both negative and positive ways to avoid a "Yes saying" subject bias. *4. Answerable*, e.g. an ordinary person may not know "How good is the national budget?", and *5. Not offensive*, e.g. some may not answer rude questions.

- **Design questions that are understandable, unambiguous, unbiased, answerable and not offensive.**

13. ***Response scale design*.*** Multi-choice question response scales should be 1. *Easy to use*, *2. Unbiased* (do not imply a preferred answer), *3. Exhaustive* (cover all answer choices) and 4. *Sensitive* (produce varied responses). Studies suggest that while more options give more information, above about 5-7 options, scale sensitivity and reliability increase little.

- **Design response scales that are easy to use, unbiased, exhaustive and sensitive.**

14. ***Sample selection.*** While a study collects data from a *sample*, conclusions invariably apply to a *population*, e.g. a study of college students using browsers may conclude how "people" use browsers. The sample is the specific students who gave the data, and the population is "people" in general, to whom conclusions are addressed. The researcher must argue this *generalization*, from sample to population. One way is to show the sample is: 1. Big *enough*, 2. *Represents* the population, and 3. *Unbiased*, e.g. 3000 subjects were chosen at random with the same demographics as the population. Another way is to reduce the generalization, e.g. a qualitative study may *suggest* possible implications with general applicability.

- **State why sample data may have population implications, e.g. the sample is unbiased, big enough, and represents the population**

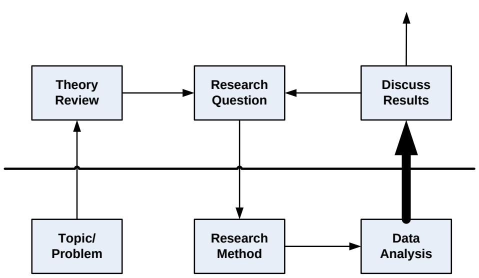
15. ***Sample size*.*** Some experimental designs are summarized as independent variable grids, e.g. how perceived usability affects web site usage for males and females could be *2x3 study* of Gender (Male/Female) by Perceived Usability (High, Medium, Low) upon web site usage. In such designs each cell needs about 15 subjects, e.g. a 2x3 study has 6 cells and so needs at least 6x15 = 90+ cases, while a 5x5x4 study might need 1,500 subjects. The methodology limits the number of variables investigated at once.

- **Check the sample is big enough for the question(s) asked of the data.**

16. ***Replicability.*** Replicability means the research can be repeated by another researcher. Method descriptions like sampling method, missing values, etc. lets others replicate the research. Give copies of research tools, like questionnaire, task script, or subject instructions, in an Annex.

- **Describe your research method well enough so another researcher can repeat it.**

Results



The results section analyses the raw data to give conclusions, and can include any or all of the following:

1. ***Error type.*** The worst result in research is not negative data (that contradicts theory), but data that shows nothing (nil results). Good research can give negative results, but nil results suggest error "noise" drowned out the effect, e.g. an unfocused RQ, a method that confounds constructs, or wrong analysis. In **s**cience a Type I error gives a false result, and a Type II error misses a true result. Insufficient research *rigor* gives Type I errors (of commission), but insufficient research *sensitivity* gives type II errors (of omission). Research must be *sensitive* enough to avoid false results, but still *powerful* enough to gain true results. Improve research power by methods that enhance responses (e.g. motivating subjects with a competition), methods that reduce subject error (like subject training or clearer questions), and more powerful statistical methods. Unfortunately reducing one error type tends to increase the other, so making errors of commission zero can increase errors of omission to 100%.

- **Use research rigor to avoid Type I false claim errors, and research sensitivity to avoid Type II null result errors**

2. ***Data Conversion.*** Describe how the study's raw data was converted into descriptive data, e.g. transcribing tapes or translating computer data. In particular, what was left out? Attempts to gather data can give three outcomes: 1. *Valid result*, 2. *Nil response* (NR), or 3. *Not applicable (*NA), e.g. asking someone "Why did you buy your mobile phone?" may give a reason, they may walk away (NR), or may say "I don't have a mobile phone" (NA). Both NR and NA are *missing values.* The number of missing values, and how they were dealt with, should be stated.

- **State how raw data was converted to descriptive data, and how missing values were handled.**

3. ***Demographics.*** To show the sample is representative, give sample demographics, e.g. males vs. females etc. to compare the sample to the population of interest. Even if your sample was one person, was that person typical?

- **Present sample demographics and compare to population demographics.**

4. *Information type.* How information is handled depends on its type. Qualitative information, like a book, requires human interpretation. Quantitative information being numbers uses statistics. Which statistic to use varies with information type: **1.** *Interval*, a continuous number series like 1,2,3,… **2.** *Ordinal*, a series of ranks, like 1st, 2nd, 3rd, etc, **3.** *Categorical*, a set of categories, like PC, Laptop and Handheld. Information type affects information power and sensitivity, e.g. measuring computer experience in years (interval data) gives more information than self-ratings of Expert, Competent or Novice (categorical), which involves fewer choices. Yet the latter may be more sensitive to reality (as many years may not make one an expert).

- **Analyze information according to type (qualitative, interval, ordinal, categorical)**

5. *Descriptive summary.* Give a descriptive summary of the results, whether as selected quotations or summary numbers. Thousands of data points can be summarized by frequency (N), mean and variance. Round off the data appropriately to the error, e.g. 3.74 not 3.742356, even if the statistical analysis gives that number. An effect's direction can be deduced from descriptive data, e.g. that subjects prefer more usable computer systems.

- **Give a descriptive summary of the information (e.g. themes, or N, mean and variance) and comment on it.**

6. ***Analytic statistics*.*** Analytic statistics show an effect's *strength* or *significance*. For two-variables, strength is how much the variables correlate, and significance is the degree they correlate by chance. Two variables may correlate weakly but highly significantly, e.g. a weak correlation over a million subjects may be very significant, i.e. a weak $r = 0.50$ (strength) but a strong $p < 0.01$ (significance). Likewise a strong correlation for a few subjects may be insignificant. The *strength* statistic used depends on information type: 1. for both variables interval use Pearson's r, 2. for independent variable categorical but dependent variable interval, use Eta, 3. for both are ordinal use Gamma, 4. for both categorical use Phi or Lambda. The *significance* statistic also depends on type: 1. for both variables interval/ordinal use a t-test, 2. for independent variable categorical and dependent variable interval use F-test, 3. for both categorical use Chi-squared.

- **Select analytic statistics for strength (e.g. Pearson's r, Eta, Gamma, Phi, Lamda) and significance (e.g. t-test, F-test, Chi-squared) based on data type.**

7. *Assumptions.* Any analysis has assumptions, so they must be stated, e.g. the F-test statistic assumes normally distributed responses, similar measure variances (homogeneity), and independent measure errors. Any assumptions must be tested or justified.

- **State and justify analysis assumptions.**

8. *Findings.* In the results analysis, argue each finding individually, one at a time, directly from what was found. Do not simply state findings without a preceding argument.

- **Argue each finding individually, one at a time, from the results.**

9. *Tables*.* Tables which summarize results need a title at the top, e.g. "Usefulness by Usability" (topic variable by causal variable). The heading should describe what the rows and columns represent. Tables presented must be referenced in the main text, e.g. "See Table 1."

- **Tables need a title at the top, row/column headings, and a reference in the text.**

10. *Frequency tables** Frequency tables (or cross-tab tables) show *categorical* variables, usually a topic column variable by a causal row variable, e.g. Table 1 shows how Usability (causal) affects Usefulness (topic). The row percentages add to 100%, and show how the topic variable column percentages change from one row to the next.

Table 1. Usefulness by Usability

	Usefulness		
Usability	*Low*	*Medium*	*High*
Low	47%	25%	28%
Medium	26%	51%	23%
High	22%	26%	52%

- **Frequency tables usually show how the topic variable column percentages change for each row**.

6. *Mean tables** Mean tables show an *interval* topic (dependent) variable, broken down by row and column variables. Each cell is a mean of several subjects/cases (not a frequency), e.g. Table 2 shows how Acceptance varies with Usefulness and Usability. Note, the row percentages need not now add to 100%

Table 2. Percentage Acceptance by Usefulness and Usability

	Usefulness		
Usability	*Low*	*Medium*	*High*
Low	47%	55%	68%
Medium	53%	67%	75%
High	58%	70%	86%

- **Mean tables usually show a topic variable cell mean varying by row/column causal variables**.

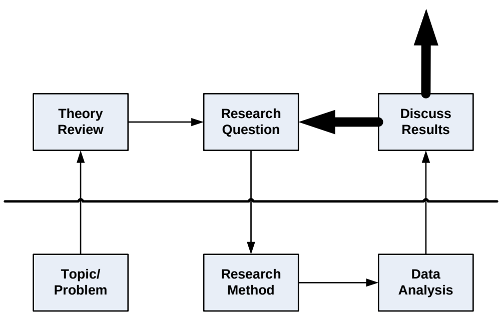
11. ***Graphs*.*** Like tables, graphs need a title (at the bottom) and must be referenced in the main text, e.g. "See Figure 1." For a line graph, the causal variable is the horizontal X-axis, and the topic variable the vertical Y-axis. Each axis must be labeled. Graph type varies with information type, e.g. an interval causal variable (X-axis) suggests a line graph, whereas a nominal causal variable (with an interval topic variable) suggests a bar graph.

- **Use graph types (line, bar or pie) depending on information type, and label each graph axis clearly**.

12. ***Summarize conclusions.*** Summarize the findings as a numbered or bulleted list of single statements. Though the findings have been concluded individually, help the reader see the results in one place. An experimental study will list the hypotheses supported or not supported. Below this, clearly answer the research question raised earlier, and provide a foundation for the next section's discussion.

- **Summarize the main conclusions of your study in a numbered or bulleted list, and answer the research question.**

Discussion



The discussion section considers how the conclusions affect the big picture, and can include any or all of the following:

1. ***Research Contribution.*** A paper should state its research contribution, i.e. what it adds that is new. Make clear what come from others and what is your own. Contributions include: 1. *adding* to knowledge currently lacking, 2. *resolving* a theory conflict, 3. *specifying* more detail on a known theory, or 4. *summarizing* current work. Don't leave readers confused regarding the publication's contribution, i.e. answer the question "W*hat value does this paper add?"*

- **Distinguish the value this research adds, apart from what others have done.**

2. ***Limitations.*** Limitations are the weaknesses of the research, e.g. small sample. They are the assumptions of your scientific argument that may not be true. Just as every argument has assumptions, and a "foolproof" argument is simply one whose assumptions are undeclared, so every research has limitations. In science, it is better to declare limitations than to try to hide them, so then they can be taken into account. Knowledgeable readers will see the limitations anyway, and wonder if you did not mention them from ignorance.

- **Declare rather than hide research limitations, i.e. honestly criticize your own research.**

3. ***Implications*** Note the implications of the conclusions for both theory and practice. Implications include challenging theories, supporting theories, changing current views, suggesting new points of view, changing current practices, and enabling new applications or practices. How does this research fit with current trends?

- **Discuss probable and possible implications of the findings for both theory and practice.**

4. ***Future research.*** Help the research community by suggesting how it could build upon your work. Imagine future research possibilities and envision possible future outcomes. Consider what is possible as well as likely, as in this part of the publication one can freely imagine, since many things that once seemed very unlikely have still actually occurred, e.g. the Internet.

- **Suggest future research possibilities to help other researchers.**

5. ***Acknowledgements.*** It is considerate to acknowledge anyone who helped you, whether an assistant who helped gather data, or a colleague who gratuitously commented. Also acknowledge any grant support.

- **Acknowledge those who have helped you.**

6. ***References.*** Some reviewers judge the quality of a publication from the quality of the references, looking for things like: Are journal article page numbers given? Are author names spelled correctly? Are major authors in the field referenced? Are the references up to date? Are there journal references, not just books, web sites or magazines?

- **Ensure the references reflect the quality, breadth and up to date nature of your publication.**

Final Points

Ask the following questions of your publication (relevant sections are in brackets):

1. **Is it relevant?** – Is it useful to someone? (*Introduction, Theory*)
2. **Is it rigorous?** – Is it scientifically correct? (*Method, Results*)
3. **Is it generalizable? –** Is it widely applicable? (*Theory, Discussion*)
4. **Is it logical? –** Is it logically consistent and logically sequential? (*All*)
5. **Is it well written**? **–** Is it readable and interesting? (*All*)

The diagram below shows how academic publications begin broad, focus at the research question, then expand into the discussion. The publication linchpin is the research question, perhaps why finalizing it can take longer than expected, especially for more original research.

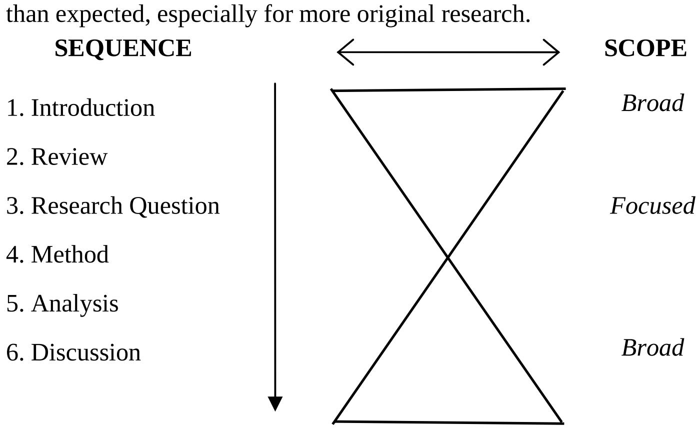


Figure 2. Publication scope by sequence

Using the checklist

The PhD publishing checklist is designed to be used in discussion with students, rather than as a stand-alone document. Preliminary results suggest that most students find it to be a helpful overview. Advisors can use it to review a student's paper with them, or to focus discussion on a particular element. It can help when a student is selectively ignoring an issue inconvenient to or difficult for their research. Using the checklist early on helps students to be forewarning of issues to come, and makes them better prepared to handle them.

Discussion and further work

In seeking common knowledge elements across different disciplines, the paper assumes such commonality exists. The alternative is that different fields have mainly different methods, with only a small "scientific core". This assumption, of a common scientific base, cannot be justified at this point, but many Universities regard a PhD as a trans-disciplinary degree, and it is hoped further research will suggest it is indeed so.

To develop this approach, the elements will be presented to students and advisors to discover which are of higher or lower priority. Some elements may appeal more to qualitative researchers than to quantitative researchers, allowing a profile for each group. This would allow different checklists, tailored to different types of research, with irrelevant elements removed. This first draft of the Academic Research Publishing Checklist is presented to the community in the hope that it will be useful to PhD and Masters students and their advisors.

Acknowledgements

Many thanks to Professor Tony Norris and Professor Ken Hawick for helpful comments and advice on early drafts, and to my colleagues for invaluable feedback.

References

Buchanan, S. (Ed.). (1948). *The Portable Plato*. New York: The Viking Press.

Grover, V. (2001). 10 mistakes doctoral students make in managing their program. *Decision Line* (May), 11-13.

Larsen, K. R. (1998) The Damocles Sword of Academic Publishing: Sharper Sword or Duller Sword in the MIS Field? *Crossroads*, 5(2), 27-31

Perry, C., & Zuber-Skerritt, O. (1994). Doctorates by action research for senior practicing managers. *Management Learning, 25*(2), 341-364.

Perry, C. (1998). A structured approach to presenting phd theses: notes for candidates and their supervisors. *Australasian Marketing Journal , (6 1), 63-86, 6*(1), 63-86.

Phillips, E. M., & Pugh, D. M. (1994). *How to get a PhD* (2nd ed.). Buckingham: Open University Press.

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of Behavioral Research; Methods and Data Analysis* (Second ed.). Boston: McGraw-Hill

Swan, J. E., & Martin, W. S. (1994). The theory-setting-testable hypothesis model: A framework to assist doctoral students in linking theory and empirical research. *Marketing Education Review*, *4*(Fall), 2-15.
VanSlyke, C., Bostrom, R., Courtney, J., McLean, E., Snyder, C., & Watson, R. (2003). Experts' Advice to Information Systems Doctoral Students. *Communications of the AIS, 12*, 469-478.