

---

## Writing an Empirical Paper

---

For correlational and experimental studies, the research report typically contains the following components:

1. Title
2. Abstract
3. Introduction
4. Hypothesis
5. Method
6. Sample / Participants
7. Stimuli
8. Results
9. Conclusion
10. Discussion
11. References

Now in more detail.

1. **Title.** At least two issues are important in composing a good title. First, what is the size of the community of readers you are targeting? Some titles are obviously intended for a small group of readers: *Coactivator and corepressor regulation of the agonist/antagonist activity of the mixed antiestrogen, 4-hydroxytamoxifen*. Especially in interdisciplinary fields, it is important to have titles that make sense to a large number of readers. Always try to reach out beyond the narrow audience of professionals in your specialty.

The second consideration is whether the title *identifies a topic* or *states a conclusion*. A title such as *The effect of scale degree on melodic accent* identifies a topic. Reading the title does not tell you what the researchers discovered. By contrast, consider the title of a paper by Jane and William Siegel published in 1977: *Categorical perception of tonal intervals: Musicians can't tell sharp from flat*.<sup>[1]</sup> One of the main conclusions of this study is that for small mistuning, musicians can tell that something is out-of-tune, but they can't reliably tell you whether the mistuning is sharp or flat. In general, "conclusion" titles are better than "topic" titles. It is very helpful for readers if the title describes the main finding. Of course, not every study will produce a clear conclusion. Use a "topic" title when the main finding is too complicated to distill into a few words.

Good topical titles often pose a question — the question that motivated the research. Examples include: *What is melodic accent?*<sup>[2]</sup>, *Is music an evolutionary adaptation?*<sup>[3]</sup>, *What is a musical feature?*<sup>[4]</sup> And *Why is sad music pleasurable?*<sup>[5]</sup> A good empirical paper will trace the research path: *from question, to theory, to conjecture, to hypothesis, to protocol*. The title may be a good place to present the motivating question.

2. **Abstract.** Provide a single paragraph that describes the essential elements of your study. Abstracts are typically 100-250 words in length and report in simple terms the hypothesis, a cursory description of the method, and a clear statement of the conclusions. Especially in the arts, scholars are sometimes tempted to organize the abstract like a film trailer — something intended to tantalize a reader into reading the full paper. This is the wrong approach. A bad film trailer is one that gives away the ending of the film. By contrast, a bad research abstract is one that *doesn't* give away the ending. Good abstracts always report the main findings. If the results are negative, say so. If an abstract doesn't report a conclusion then it suggests that the research may not have anything useful to say.
3. **Introduction.** Introduce the problem you are interested in. Introductions should retrace the main intellectual path in research: *Question-Theory-Conjecture-Hypothesis-Protocol*. Once again, if you can, begin with a question: Why do people tap their foot in time to music? Continue by reviewing the main theories people have proposed in the past — that is, provide a literature review. Use the introduction to identify an unresolved issues and set the scene for your study. Your study might provide an additional test of an already existing theory, or test a new theory that you describe.

End the introduction by giving a one- or two-sentence preview of your experiment or correlational study. The final paragraph might begin:

*In brief, a study was carried out to determine ...*

OR:

*In light of this question/debate/issue, three studies are described here. In the first study ...*

Research papers are easier to read when the reader has some sense of where you are going. So end the introduction:

*To anticipate our results, we will see that ...*

Be sure to use the tentative language of empirical research. Say “we will see that the results are *consistent* with the notion that ...” Of course we never say “the results *prove* that ...” or “the results *establish* that ...” Don't say “we will *demonstrate* that ...” or “our results *show* that ...” Simply say “the results are consistent with ...”

4. **Hypothesis.** The introduction should lead you right up to the statement of hypothesis.

*In light of the Smith-Jones theory of music-induced foot-tapping, we might propose the following hypothesis:*

Indent, and state your hypothesis.

*H1. Listeners are more likely to tap their feet as the beat approaches 88 beats per minute.*

If there is more than one hypothesis, label them H1, H2, etc. Then continue by acknowledging that there is no way to directly test the hypothesis:

*As it stands, the main terms of your hypothesis must be operationalized in order to allow us to proceed with a test.*

Identify the main conceptual terms in the hypothesis. For each term describe different ways of operationalizing it; identify the advantages and disadvantages for each of the various operationalizations. For example, some operationalizations may simply be too laborious, and so are impractical. At the end of the hypothesis section, readers should be convinced that you have provided a reasonable operationalization of your hypothesis.

Notice that intelligent people can disagree about what makes a good operationalization. Especially rigorous research will identify two or more ways of operationalization a hypothesis. One operationalization might be considered rather *narrow* (looking at a single composer, or single style, or looking at notated scores). Another operationalization might be considered rather *broad* (sampling with very loose criteria, having judges listen to recordings, etc.). The most impressive studies will seek *converging evidence* by presenting two or more tests of the hypothesis, where each test operationalizes the terms differently.

5. **Method.** The “Method” section usually has several subsections. The section may begin with one or two sentences describing the method in general:

*In brief, the study involved floor-level video capture of foot activity for seated patrons in a discotheque. The tempo of the music was manipulated and the videos coded in order to relate foot-tapping activity to musical tempo.*

In some cases, there is no need for such a general summary since the overall approach is clear from the discussion of the operationalized hypothesis. In these cases, the “Method” title is followed immediately by a subtitle, such as “Sample.”

6. **Sample / Participants.** Describe the method by which you sampled materials or recruited participants. Indicate whether you are using a convenience sample, using quota sampling, stratified sampling, systematic sampling, etc. For example, what procedure did you use when selecting the scores for analysis? How did you recruit participants? Identify any “exclusionary criteria” — that is, conditions for excluding particular people, performances, recordings, scores, etc.

*For the purposes of this study, we excluded musicians who reported possessing absolute pitch.*

OR

*In analyzing the videos, we excluded dancers who were holding objects like drinks or purses, as well as people who were deemed to be merely walking across the dance floor.*

For research participants, provide some basic demographic information (average age,

number of participants of each sex, musical training, etc.).

*Twenty-eight musicians participated in the study, 16 females and 12 males, with an average age of 21.6 years (range 18-27). Most (23) were instrumentalists; five were voice majors.*

7. **Stimuli.** Describe any stimuli used. If the stimuli were created by the researcher, what criterion were used in their construction? How long were the sounds? What timbres were used? Etc. Where possible, describe the stimuli in sufficient detail that another researcher might be able to duplicate them. Consider posting the stimuli on the web so that readers can hear them for themselves.
8. **Procedure.** Describe what happened. What were the precise instructions given to participants? How long did the experiment last? How many stimuli did they hear? How were the data collected? Etc. Report the instructions verbatim. For example: *Participants received the following verbal instructions:*

*INSTRUCTIONS: "In this experiment you will view 20 photographs of people in everyday situations. Imagine that you are present in the scene. For each photograph, identify whether you think there is music playing in the background. If you think there is music, click on the "music" button. If you think there is no music present, click on the "no music" button. When you are finished, click on the NEXT button to view the next photo.*

*Do you have any questions?"*

9. **Results.** Conceptually, four things need to occur in a Results section. First, you should establish that the data are not purely noise. Second, you should providing a broad description of the data. Third, you should provide a formal statistical test of the hypothesis or hypotheses. Finally, you should test any supplementary or *post hoc* hypotheses.

In working with data, you want to know that the data are not simply random noise. This is especially pertinent when you ask people to judge things. Test for both *intra-subjective reliability* and *inter-subjective reliability*. INTRA-subjective reliability means that a participant judges a stimulus roughly the same way each time. Poor intra-subjective reliability commonly occurs when a participant is not paying attention to the task, or the task is too difficult, or the participant has no task-pertinent skill. In order to test for intra-subjective reliability you need to include some duplicate stimuli. For example, if you have 100 stimuli, you should probably repeat 20 or so. This will allow you to determine whether the participant responded in a similar way when the same stimuli appeared. Measure the correlation between repeated trials in order to determine whether there is intra-subjective reliability. This procedure is referred to as *test-retest reliability*.

*In order to determine intra-subjective reliability, the correlation between responses for duplicate stimuli was calculated. The mean correlation for all 30 participants was +0.38.*

You may want to consider excluding data from unreliable participants. Before you test

your hypothesis, establish a possible exclusionary criterion.

*Prior to hypothesis testing it was resolved that data would be discarded for any subject whose test-retest correlation was less than +0.1. According to this criterion, data for two participants was eliminated.*

OR

*Although the mean test-retest correlation was +0.72, the correlation for two participants were considered outliers at +0.31 and +0.22. Before testing the hypothesis, it was decided to exclude the data from these two less consistent participants. Having excluded this data, the mean test-retest correlation was +0.82.*

Are all of the participants doing similar things? Measure the correlation between the same trials across multiple participants. If the correlations are significantly positive, then it suggests that all of the participants are behaving in similar ways. If there is no significant positive correlation between subjects, then it suggests that (i) participants may be engaged in different tasks; for example, the participants may have interpreted the instructions in different ways; or (ii) the participants have no pertinent skill related to the task; for example, if asked to judge how “snerky” a musical passage is, participants may be at a loss; or (iii) there is some other problem. If there is no significant positive correlation between subjects, then there is no good reason to average all the data together. If there is a significant positive correlation between subjects, then there is reason to suppose that they were engaged in similar behaviors and so it is reasonable to aggregate the data.

The American Psychological Association (2010, p.34) recommends that all empirical papers report effect sizes. In addition, in order to facilitate possible future meta-analyses, research reports should provide complete descriptive statistics (number of subjects, means, standard deviations), and report statistical values to three significant digits, including non-significant results.

10. **Conclusion.** The conclusion is essentially an expanded abstract. Like the abstract, the conclusion will reiterate the question or hypothesis, recap the method, and state the results. Where the abstract may state only the main finding, the conclusion restates *all* of the findings. The conclusion should also assemble and restate all of the caveats and assumptions. Typically, the conclusion is roughly three times the length of the abstract.

A good conclusion should be written as a “stand-alone” section. Having read the title and abstract, an interested reader should be able to skip directly to the conclusion, and find a good summary of the research. In stating the conclusion, always use that circumspect language of empirical research. No “the results *prove ... establish ..., demonstrate ..., or show ...*”. Instead — “the results are consistent with the theory that ...” or “the results are consistent with the hypothesis that ...”

11. **Discussion.** Only a minority of research articles include a Discussion section. Unlike the Conclusion, the Discussion section provides a more open forum to consider the repercussions of your work. The Discussion section allows you to speculate about what

you think may be going on. You may also propose *post hoc* interpretations of your data. For example, post-experiment interviews might have alerted you to other possible phenomena:

*In post-experiment interviews, several participants indicated that they had been singing-along with the stimuli. A post-hoc test showed that these participants performed better than the other participants. Moreover, when all of the trained vocalists were included, these participants collectively performed well above the chance level. It may be that “singing” — either overtly or clandestinely makes it easier for people to perform the task.*

You may also make suggestions for future research.

12. **References.** Only include references to works you cite in the article. Be sure that you have read everything you cite. Don't ignore the early and historical literature.

## Helping Readers

In general, an empirical research paper should be written to provide readers with an efficient way of accessing information. Everyone is busy and no one has time to read everything. Don't expect scholars to read your entire paper. Some readers will be casually interested in your work. Others will be looking for results related to a particular interest they have. Yet other readers will be very interested in the subject and will read your report carefully — looking to ensure that the method is solid, that the research was done carefully, etc.

Different levels of interest are echoed in which parts of a research paper people read. Suppose that 100 people encounter your paper and read the title. Of that 100, perhaps 50 will flip through the article looking at the figures or illustrations. Perhaps 35 will read the abstract. Of those 35, 10 may also skip ahead and read the conclusions. Perhaps 10 people will begin reading the entire paper, but only 3 will finish it. A good research paper will cater to *all* of these people. Ideally, a good title will convey the essence of your work. This is not always possible since results are often messy. Nevertheless, readers appreciate titles that are informative. This is the reason why the titles for research papers are often so long. Similarly, the *Abstract* should be crafted so that readers have some idea of what question/hypothesis you addressed, the approach your study took, and what your main results were. The *Conclusion* should act like an expanded Abstract.

**Tip:** Newspaper, magazine, and book editors have long known about the importance of good photos and illustrations. Readers are attracted to the pictures, and that is just as true of research articles. Spend time creating interesting and instructive figures. However, the most under-valued aspect of a good research article is not the *figures* but the *figure captions*. Wherever possible, write a figure caption that tells the whole story. Imagine a reader who has read the title and perhaps skimmed the abstract. They have read nothing else in your paper, but now they are looking at a figure you have provided. Try to make your figure caption summarize the main result so that a

casual reader will understand what you have achieved.

## References:

American Psychological Association (2010). *Publication manual of the American Psychological Association*, Sixth edition. Washington, DC: American Psychological Association.

David Huron (2001). Is music an evolutionary adaptation? *Annals of the New York Academy of Sciences*, Vol. 930, pp. 43-61.

David Huron (2001). What is a musical feature? Forte's analysis of Brahms's Opus 51, No. 1, Revisited. *Music Theory Online*, Vol. 7, No. 4.

David Huron (2011). Why is sad music pleasurable? A possible role for prolactin. *Musicae Scientiae*, Vol. 15, No. 2, pp. 146-158.

David Huron and Matthew Royal, (1996). What is melodic accent? Converging evidence from musical practice. *Music Perception*, Vol. 13, No. 4, pp. 489-516.

Jane Siegel and William Siegel (1977). Categorical perception of tonal intervals: Musicians can't tell sharp from flat. *Perception & Psychophysics*, Vol 21, No. 5, pp. 399-407.