The POWER Structure and Why an 80% Correct Solution is Sometimes Better Than a 100% Correct Solution

Ian Laga[1]
[1]Pennsylvania State University, 326 Thomas Building, University Park, PA 16802

Abstract
Face-to-face meetings are at the heart of collaboration. The POWER structure (Vance, 2018) is an effective method to make these meetings efficient, informative, and productive. POWER stands for Prepare, Open, Work, End, and Reflect, five key steps that help ensure the domain expert's questions are answers accurately and quickly. On another note, when students are learning statistics, the environment is a sterile system, where problems are typically completely solvable and optimal solutions are well known and achievable. However, when collaborating, data is often messy, timelines are short, and there are different methods to approach the problem. Depending on the situation, different levels of accuracy are expected. While a Bayesian Hierarchical Spatial model may theoretically produce more accurate results, a simpler Generalized Linear Model can offer a sufficiently accurate solution at a considerably smaller cost. Predictive models often have a bias-variance trade-off, while collaborations have a usability-accuracy trade-off.

Key Words: Consulting, collaboration, POWER, communication

1. Introduction

Statisticians frequently collaborate [1] with researchers from other disciplines. This interdisciplinary work requires multiple researchers to come together to answer a common research question. This task is inherently difficult because each party has different objectives, different backgrounds, and different understandings of the research problem. To exacerbate this problem, time is often limited for meetings. In as short as an hour, domain experts must explain their research, their statistical needs, and their timelines, and the statistical collaborator must understand the research, ask clarifying questions, and suggest a path forward.

Statistical collaboration is a skill. This means that individuals new to collaboration may struggle initially, but that every researcher can become an effective collaborator through practice. Young researchers tend to underestimate the difficulty and importance of consulting, only to realize how difficult consulting is in their first meeting. One purpose of this paper is to help these individuals learn techniques that can help them in a collaborate framework before their first meeting.

[1] Throughout this paper, I use the terms consulting and collaboration interchangeably. I also use the client and domain expert interchangeably.

The POWER Structure is one way to organize consulting meetings to facilitate efficient and effective collaboration (Vance and Smith, 2018). The acronym stands for Prepare, Open, Work, End, and Reflect, and represents five ways to divide a consulting meeting. Structuring the meetings helps limit the amount of time wasted during meetings and reduce the need for follow-up meetings. By following the POWER Structure, new collaborators can ensure that the domain expert's needs are met.

The idea that an 80% correct solution is sometimes better than a 100% correct solution was originally an off-hand comment Dr. Vance made in our collaboration class. The original sentiment was not to overcomplicate the statistical analysis and instead offer the right solution for the research problem. I have since developed this idea and hope to illustrate a few of the situations in which collaborators should offer an 80% correct solution.

The reader can think of this paper as two smaller ideas that are designed to achieve the same purpose of successful collaboration. First introduce the reader to the POWER Structure. We summarize the main points introduced in Vance and Smith (2018), while offering a few observations from my time working at the University of Colorado Boulder's Laboratory for Interdisciplinary Statistical Analysis (LISA). For the rest of the paper, we further discuss the details around the idea of why an 80% correct solution is sometimes better than a 100% correct solution.

2. The POWER Structure

2.1 Prepare

The "Prepare" stage includes everything prior to the start of the collaborative meeting. At LISA, we prepared for meeting by trying to understand as much about the domain expert's research as possible before meeting them. This requires participation from both the collaborators and the domain experts. Some domain experts will send over the background files mere minutes before the meeting. This short amount is useful only to understand the research field and what file extension the domain expert uses to save their data. Instead, the collaborators should receive the resources a few days in advance in order to properly study the given material and research additional topics needed for the meeting. For example, if the domain expert supplies an ANOVA table from an extra R package, the collaborators should have time to install the packages themselves, research the implementation of the R function, and fully understand how to interpret that specific ANOVA table. The additional time to read the materials also increases the productivity of both the open and work portions of the POWER structure, as the collaborators will be able to lead the meeting in the proper direction, spend less time on a basic topic introduction, and ask more relevant research questions.

2.2 Open

The "open" stage allows the client and consultant become acquainted and discuss the plan for the meeting. As noted in Vance and Smith (2018), collaborators will skip the opening phase in the interest of saving time. This may seem like a reasonable concern at first since meetings are often as short as one hour and understanding the domain expert's research seems far more important than talking about the meeting structure and timelines. However, a well-structured opening makes up for any lost time by increasing the efficiency of the work section and reducing the need for follow-up meetings and emails. In collaboration, we want to focus on quality, rather than quantity.

During my tenure at LISA, I found the most effective way to start the meeting is to outline the flow of the meeting for the client. I began all my meetings by outlining the following steps.

1. First, we will talk about how much time we have for this meeting.
2. Second, we want to learn about what you would like to accomplish during this meeting.
3. Next, we will talk about both whether we are willing and able to accomplish what you expect from us. This will include timeline discussions and when we can expect to complete the work.
4. Then, we will discuss your research.
5. About 10 minutes before the end of the meeting, we will have an end discussion. We will review what you expect from us and what we plan to do. If we didn't cover everything today, we can schedule another meeting.
6. Finally, shortly after the meeting, we will email you a summary of the meeting and what the next steps are for both of us.
7. Do you have any questions about the structure of the meeting?

These seven steps enumerate just one example of a potential opening structure. Clearly explaining the structure of the meetings helps the consultant control the pace and organization of the meeting, leading to an effective consultation. Collaborators should find an opening sequence that allows them to make the client feel comfortable and heard, while still leading the meeting.

2.3 Work

The “work” stage is the bulk of the meeting and includes the time spent discussing the actual research question. This section is where most statisticians are comfortable since they typically have experience applying statistics to answer problems. One of the most important components of this work that statisticians may forget is to ask questions. Asking questions leads to a better understanding of the research problem, and thus a better understanding of how to apply the correct statistical methods. Throughout the work stage, the collaborator should restate the research objectives and key components back to the domain expert to confirm they understand the research correctly. As with all stages of the POWER structure, communication is tantamount to a successful collaboration.

2.4 End

The “end” stage is how the collaborator and domain expert close the meeting. Collaborators often underestimate the importance of this end time, leading to a rushed summary of the meeting after the work stage. However, by outlining future steps, the end section ensures that the work section of the meeting is not wasted. In the end stage, the collaborator and the domain expert should review not only the main part of the research problem, but also what the next steps are for all participants and under what timeframes. The collaborator can achieve these goals by drafting an email summarizing the meeting and next steps and sending the email to the domain expert. Summarizing the meeting in text is useful because it creates a written record of tasks for participants to look back to, holds everyone accountable, and allows for easy misunderstandings to be caught early. While the end stage may seem straightforward, it is easy to leave too little time, leading to a less effective collaboration moving forward.

2.5 Reflect

The “reflect” stage is everything that happens after the close of the meeting. We already noted that collaborators love to skip the open stage, but collaborators also avoid reflecting on the meeting. One reason for ignoring the reflect stage is because student collaborators think they need only to hold meetings until the end of the semester, so why should they spend the time improving their meetings when they are getting by just fine.

At LISA, in order to make the reflect stage more tangible, we watched a video of our collaboration meetings. This video feedback is a powerful tool discussed in Vance and Smith (2017). It is easy to convince ourselves that each meeting went better than the previous one, even with no evidence. Seeing the change in a small timeframe allows the collaborator to better reflect and see how any changes they made affects the meeting. Reflecting helps statisticians learn from their mistakes and improve as collaborators.

2.6 Final Thoughts

While the POWER Structure isn't the only way to organize collaborative meetings, it is an easy guide for new collaborators to learn and follow. Ultimately, each collaborator should adapt the POWER Structure to their individual experiences. This may mean drafting the summary email throughout the meeting instead of just at the end or merging the opening and work stage into a new stage. However, the basic idea of the POWER Structure stands: a structured meeting leads to efficient and effective meetings.

3. Why an 80% Correct Solution is Sometimes Better Than a 100% Correct Solution

3.1 Another Pitfall

Researchers have long written about the difficulties present in collaboration. Kimball (1957) suggested that the error of the third kind is answering the wrong problem. This error is common among collaboration projects because often the statistician has not spent as much time understanding the research problem and the data as the domain expert has. This issue is exacerbated when the domain expert withholds key information they deem not important, while the statistician fails to ask the necessary questions to properly understand the problem.

Chatfield (1991) also offers a guide to avoiding common issues that statisticians face when working with real data. Outside of the sterility of classroom datasets and simulations, real data often contains typos, strange codings, and unusual relationships, among other complications. While no researcher could enumerate all the possible pitfalls, Chatfield does a wonderful job of guiding researchers to avoid the most common of these issues.

We expand on the works above in this rest of this section by suggesting another pitfall of statistical consulting: creating the perfect answer for the data rather than for the research problem. Too often statistician analyze the data without understanding the context surrounding the research question. It is attractive to view the perfect answer in terms of objective measures, like AIC, predictive power, or even diagnostic visualization. However, the consultant should strive for the solution which the client can use. In the following subsections, we explore the most common situations in which collaborators should consider an 80% correct solution instead of the 100% correct solution.

3.2 Lack of Time

Time is a finite resource in the collaboration setting, both for the domain expert and the collaborator. When time is limited, researchers must often sacrifice model performance in place of practicality. These situations when time is a limiting factor are becoming more common as models become more complicated. For example, despite the rapid increase in computational ability and statistical advancements, fitting Bayesian models on large data sets can often be prohibitive. Consider one collaborative story, inspired by an experience I had at LISA when a client approached me with a strict time limit, forcing me to choose between the more accurate model and the faster model.

Example: Two statistician consultants are meeting with a master's student from the ecology department. The student is interested in mapping the locations of ant colonies in a nearby wooded area. The project is for her master's thesis, which is due in two weeks. The client explains that over the past year, she recorded locations of ant colonies in a small subsection of the total wooded area. The sampled locations were very dense, resulting in almost 1000 observations. In addition to the locations, researchers also measured the pH of the soil and the percentage of ten common elements found in soil, like potassium and iron. The locations, she explained, were measured via incredibly accurate GPS devices. However, the devices used to measure pH and element quantities were often inaccurate and had high margins of error relative to the quantities being measured. The client stated she wanted to use a Bayesian approach to logistic regression to model the presence of ant colonies, using the pH and mineral content as predictive covariates. Furthermore, she wanted to account for the high measurement error of the covariates. She also explained that the model should account for the spatial dependency of the ant colonies, since colonies tended to exist clustered together.

How should the statisticians approach this collaboration? They agreed with the client that her proposed model was the best approach for handling the data in terms of predictive accuracy. However, the client had only two weeks to finish everything for her master's thesis, including incorporating the statistical analysis into her manuscript. After discussing the situation, the statisticians offered the following two options:

- Fit a Bayesian model which accounts for the measurement error of the covariates and the spatial dependency of the ant colonies. A single model takes around 30 hours to converge.
- Fit a frequentist model which ignores both the measurement error and the spatial dependency. A single model takes 30 seconds to fit.

Given infinite time, the Bayesian model is the “100% correct solution.” However, in this collaborative situation, the better approach is to choose the fast, frequentist model. Ideally, the statisticians and the master's student could take the time to fit the Bayesian models and select the best performing model. However, with only two weeks to finish a master's thesis, the speed of the Bayesian model is too prohibitive. A single model takes 30 hours to converge, and the statisticians should perform appropriate model selection, which requires refitting each model. While the frequentist model will likely suffer in predictive performance compared to the Bayesian model, in this situation, it is the 80% correct solution that is better than the 100% correct solution. The frequentist approach not only allows the statisticians to perform the appropriate model selection, but also to create helpful visualizations and leave the client with enough time to include the results in her manuscript.

There is also a different kind of “lack of time” constraint. While the above example concerns the time it takes to construct a model, we can also consider the time it takes to perform prediction. To clarify, sometimes researchers have enough time to perform model selection and analyze the data, but not enough time to make predictions of future events. This situation is best illustrated by a short theoretical example concerning hurricanes.

Example: A researcher is developing models for predicting hurricanes. These models often rely on complicated deterministic equations, like in Davis (2008). The researcher approaches a statistical collaborator and asks for help fitting a Bayesian ensemble model to forecast the path of a hurricane. The goal of the model is to facilitate a warning system to aid in early-evacuation procedures. The statistician codes the model but finds that the supercomputer he was using takes three days to make a next-day forecast. The statistician returns to the researcher with a simpler model that takes only 3 hours to fit a next-day forecast, but consistently performs worse in predicting the path of past hurricanes compared to the more complicated model.

In this situation, it is even more clear that the 80% correct solution meets the client's needs more than the 100% correct solution, since the more accurate model won't produce a prediction in time to warn the population of an approaching hurricane. This type of time constraint is very common in applied settings. Sir Robert Alexander Watson-Watt, a Scottish physicist, advocated for simpler, quicker solutions during World War II. Watson-Watt developed a system of a system of radars that operated at non-optimal frequencies in order to warn the population about enemy aircrafts. Concerning his choice of an “80%” correct solution, Watson-Watt argued, “Give them the third best to go on with; the second best comes too late, the best never comes” (“Sir Robert Alexander Watson-Watt”). This quote echoes the idea above, where a timely solution is often preferable to the solution that provides a better answer but is useful because of lack of time.

3.3 Prioritizing Tasks

The second situation in which worse may be better is when there are more tasks than just modeling. We've already discussed this situation briefly with the master's student's need to incorporate the analysis results in her manuscript. In that example, there were two tasks: analyze the data and write the manuscript. In practice, analyzing the data is almost never the only task. In addition to manuscripts, researchers may also need to consider making a poster, presenting the results, and publishing a software package. In order to handle the variety of tasks, we will rely on the Pareto principle.

Joseph Juran, an engineer, introduced the Pareto principle in his “Juran's Quality Handbook”, a guide for ensuring operational excellence. In his handbook, Juran wrote, “This [Pareto] principle states than in any population that contributes to a common effect, a relative few of the contributors — the vital few — account for the bulk of the effect” (Juran and Godfrey, 1999, p. 5.20) Specifically, Juran noted that 80% of a product comes from 20% percent of the effort, while the last 20% of the effect comes from 80% of the effort. This idea is analogous to the economic principle of diminishing returns. Conveniently, Juran's Pareto principle aligns nicely with our concept of an 80% correct solution.

In short, when a project has multiple tasks, the overall quality of the product is best when the time is spent working on 80% solutions rather than 100% solutions. Let us assign some numbers to a simplified collaborative setting. First assume that a researcher has five tasks to complete, like the five example tasks considered above. Furthermore, assume these tasks all take the same amount of effort to complete. The researcher initially spent two hours creating an 80% correct statistical model. According to the Pareto principle, in order to reach a 100% solution, the researcher would have to spend an additional eight hours developing the model, spending 10 hours total on the model while having no manuscript, no poster, no presentation, and no software package. On the other hand, the researcher could have spent two hours on each task and had 80% of a perfect product for all five tasks. Ideally, the researcher could spend the necessary time to complete every task, but time is always a finite resource. In this situation, the researcher is better off with an 80% correct solution in every task rather than just 100% correct model. When there are multiple objectives, collaborators can improve the overall quality of the project by understanding when an 80% correct solution is appropriate.

3.4 Lack of Usability

Another compelling reason for using an 80% correct solution is when the best solution is too complex to be useful. In these situations, time is no longer a constraint, so even the most complex models can be fit. However, complex models are often more difficult to interpret. Often even relatively simple models are often difficult to interpret. Consider linear regression. Almost every regression class includes lessons on how to interpret interaction effects because even one seemingly simple interaction makes it difficult to communicate the effect that the covariates have on the response. The first example in this section concerns the use of these interaction effects.

Example: A statistics collaborator is working on a project with a university professor. The professor wants to know whether review sessions improve students' exam grades. Over the past ten years, the professor has collected data on review session attendance, exam scores, GPA, gender, and year. The collaborator compares all possible regression models, checks modeling assumptions, and concludes that by all relevant metrics, the model that includes a three-way interaction between session attendance, GPA, and gender and a time-series relationship on year is the best model for estimating exam scores. Proud of his work, the collaborator presents the model to the professor, who is disappointed. He explains that the results are useless to him. The collaborator tries to explain that given a student's GPA and gender, the professor could determine what effect a review session might have on the student's exam score. However, the professor’s research question is only whether he should continue to offer review sessions or not. The professor cannot restrict a student's participation based on GPA or gender. The collaborator returns after a week with a simpler model where review session attendance has no interaction terms. The model suggests that the review sessions have a negative effect on exam scores. The simple model has a worse fit than the complex model with respect to metrics like AIC and leave-one-out cross-validation, but still does a reasonable job of explaining the data and does not violate any assumptions.

The simpler model is the better model for answering the professor’s research question, even though the model is not as accurate. This example demonstrates that sometimes even “simple” models are too complicated to be useful. However, the reader might argue that the consequences of choosing a bad model are small for the professor. While this is true for the above example, we can still extend the idea to collaborations where the consequences of performing unsound statistics is large. Now consider an example from the pharmaceutical industry.

Example: A team of physicians is interested in modeling the dose-response relationship of a new drug. The endpoint of this study is binary, measuring whether the patients were painfree at the end of the sample. They have data from a previous drug study, and they expect the new drug to behave similarly. For the modeling approach they wish to use, the physicians must first specify different dose-response curves, but on the transformed scale. The physicians reach out to collaborate with a statistician, who starts work by considering the data from the previous study. After a week of looking at the previous data, the statistician presents five possible models for the dose-response relationship, all specified on the complementary log-log scale, since this link for the binary regression offered the best fit according to all popular metrics. The physicians are confused and ask what this complementary log-log scale represents. The statistician shows them the formula and explains that the complementary log-log is useful for this data because the probabilities show severe asymmetry and this link offers the best fit. Understandably, the physicians and cannot determine whether the proposed curves accurately reflect the suspected dose-response relationship. Here, the statisticians failed to understand the objective of the collaboration and offered the physicians an answer than was useless.

In this example, the “80% correct solution” is one that uses a logit-scale, for example, that is more easily understood by the physicians, even though the complementary log-log scale seems to offer a better fit for the historical data. In the pharmaceutical setting, the interpretation is often just as important as the model performance. So while the logit-scale is not the best model in terms of objective metrics, it is the best model for answering the research question.

It is important to note that collaborators should not always offer simpler solutions whenever the domain expert asks for one. In these cases, the domain expert and the collaborators must understand the trade-off between performance and usability in order to best answer their research question. These two examples demonstrate that for collaborative projects, optimizing model metrics like AIC are not the only objective. Instead, collaborators must understand when an 80% correct solution does a better job of answering the research question.

3.4 Final Thoughts

The idea that an 80% correct solution is sometimes better than a 100% correct solution revolves around meeting the clients’ needs. Clients want to answer a research question. A solution’s ability to answer the research question is an important metric, although more difficult to quantify than AIC and mean squared error, for example. The 80% correct solutions discussed in this section do a better job at answering the research questions than the 100% correct solutions. It is up to the collaborator and the client to find a solution that is usable in application while still being statistically sound.

4. Conclusion

As research becomes more collaborate, statisticians have a greater opportunity to offer statistical guidance in a variety of disciplines. Organizing meetings, working in a team, and applying often complex statistics all make collaboration difficult. The POWER Structure and the idea that an 80% correct solution is sometimes better than a 100% correct solution are tools that statisticians can use to improve collaborations. While these are not the only tools that collaborators can use, we hope that they make collaborations a little easier for everyone.

Acknowledgements

Thank you to Eric Vance for not only organizing this JSM topic-contributed session, but also introducing me to the world of statistical collaboration, teaching me the POWER structure, and inspiring the topic of this paper.

References

Chatfield, C. (1991). Avoiding statistical pitfalls. *Statistical Science*, 240-252.

Davis, C., W. Wang, S. S. Chen, Y. Chen, K. Corbosiero, M. DeMaria, J. Duhia, G. Holland, J. Klemp, J. Michalakes, et al. (2008). Prediction of landfalling hurricanes with the advanced hurricane wrf model. *Monthly weather review 136*(6), 1990-2005.

Juran, J., & Godfrey, A. B. (1999). Quality handbook. *Republished McGraw-Hill*, 173-178.

Kimball, A. (1957). Errors of the third kind in statistical consulting. *Journal of the American Statistical Association 52*(278), 133-142.

"Sir Robert Alexander Watson-Watt." *Encyclopaedia Brittanica*, edited by the editors of Encyclopaedia Brittanica, 2019. Encyclopaedia Brittanica.

Vance, E. A. and H. Smith (2017). Using video to improve your practice of statistics. Joint Statistical Meetings, Baltimore, MD.

Vance, E. A. and H. Smith (2018). Essential collaboration skills: The ASCCR frame for collaboration. *arXiv preprint arXiv:1811.03578*