Is there a Place for Bayesian Statistics in Communication Research?

Jens Vogelgesang* Michael Scharkow**
Freie Universität Berlin Universität der Künste Berlin

March 31, 2009

Accepted for DGPuK Jahrestagung 2009, Bremen

The following manuscript has not been submitted to or presented at any other conference or publication.

*Institut für Publizistik- und Kommunikationswissenschaft, Freie Universität Berlin, gesang@zedat.fu-berlin.de
**Institut für Theorie und Praxis der Kommunikation, Universität der Künste Berlin, scharkow@udk-berlin.de

Abstract

Statistical inference in communication research has been dominated by the frequentist theory of statistics for decades. The alternative Bayesian approach to statistical inference promises to give a more appropriate statistical answer by incorporating a prior state of scientific knowledge. Although the Bayesian approach is rapidly gaining popularity in the American social sciences, communication scholars have not (yet) made much use of it. In this paper, we address the benefits of Bayesian statistics for applied research, focussing on methodological and empirical rather than purely statistical issues. We present findings of a short case study where we examined the handling of prior scientific knowledge by German communication scholars. We conclude with a rather personal view of the future of Bayesian statistics in the discipline of communication research.

1 Introduction

From the 1940s on, the frequentist approach to statistics was the dominant statistical paradigm in the social sciences. In contrast to the Bayesian approach to statistics which incoporates prior knowledge or at least current beliefs, frequentist statistics does not. This is quite remarkable because almost every scholar employing frequentist statistics regularly refers to prior scientifc knowledge. For example, when writing a research paper pre-existing scientific knowledge is usually accounted for by citing seminal theoretical arguments, by adopting previous operationalizations, or by drawing on other empirical findings. Put differently, scholars try to collate all pre-existing information of the subject of inquiry before prospectively collecting new data but suspend any empirical evidence based on past studies when it comes to hypothesis testing using classical frequentist statistics.

When asking for the place of Bayesian statistics in communication research, we presume that many colleagues are frequentists in mind but Bayesians at heart. Of course, becoming a frequentist in mind is rarely a conscious decision. Rather, it just takes place in the foundation phase of statistical education because it is the lecturer who decides which introductory book on statistics is used and whether or not the Bayesian paradigm is covered. In what follows, we first show how misunderstandings of frequentist statistics create Bayesians at heart by reviewing the notion of frequentist probability. Second, we provide a short outline of the basic concepts of Bayesian statistics. Third, we review why the use of prior distributions turns out to be advantageous in the enterprise of scientific research. Fourth, we highlight the key feature of Bayesian statistics, that is the concept of prior distributions. Fifth, we report findings of a short case study where we examined the handling of prior scientific knowledge by German communication scholars. Finally, we briefly present a rather personal view of the future of Bayesian statistics.

2 How to involuntarily become a Bayesian at heart?

We are surprisingly unaware of the history of statistics and probability. This may stem from the fact that introductory textbooks rarely cover the history of the major methodological strains in statistics. Instead, introductory textbooks on statistics usually present the so-called classical approach as if there were never any conflicts about its coherence and applicability. Historically, the roots of the classical approach to statistics can be traced back to the old conflictful days of R.A. Fisher on the one hand, and Jerzy Neymann and Eugen Pearson on the other hand (Savage, 1976; Efron, 1986; Zabell, 1989). Interestingly, it were historians of science who critizied the ritualistic use of the classical approach (Krüger et al., 1987, p. e.g.). Those historians of science concluded that the classical approach to statistics is actually a "hybrid theory of statistical inference" (Gigerenzer et al., 1989) binding together frequentist methodology developed by Neymann and Pearson with likelihood based methodology developed by Fisher. It is the combination of two different methodological strains in statistics what is nowadays referred to as the frequentist paradigm.

A frequentist statistic indicates "how frequently a researcher could expect to obtain a given result if an experiment were repeated and analyzed the same way many times" (Malakoff, 1999, p. 1461). The probability p of an observed result (*data*) from such an infinite series of trials performed under identical conditions given a hypothesis (H) can be expressed as a conditional probability, that is

$$p(data|H). \tag{1}$$

Needless to say, the null hypothesis significance testing (NHST) is the backbone of the frequentist framework. However, NHST is also a sometimes controversial discussed issue in various academic disciplines like sociology and political science, medicine, marketing, psychology, and communication research (Gill, 1999; Altman & Gardner, 2000; Hubbard & Armstrong, 2006; Cumming et al., 2007; Levine et al., 2008a). To cut a long story short, objections against classical NHST include the commonly false probabilistic interpretation of the p-value, the lack of sensitivity to sample size, the ritualistic use of significance levels, the negligence of power and error rates, and the mechanic use of a meaningless null hypothesis. To improve the quality of statistical testing, communication scholars recently published a researchers' guide proposing several state-of-the-art alternatives to NHST like using confidence intervals, effect or equivalence testing instead of unsubstantial null hypothesis testing, or meta-analysis (Levine et al., 2008b). But even when using more appropriate tools from the frequentist toolkit, the frequentist paradigm and it's application can not give any auxiliary answer about a hypothesis. We illustrate this objection by means of the well-known frequentist confidence interval.

A confidence interval for a population parameter consists of a range of values, restricted by a lower and an upper limit. But even when confidence intervals were computed instead of classical NHST p-values, the true interpretation of the probability statement associated with the frequentist confidence interval is resistently overlooked. Many researchers employing the frequentist toolkit want to and do interpret the classical 95% confidence interval in a Bayesian fashion. Although frequently done, it is incorrect to state, for example, that the 95% confidence interval has a probability of 95% of containing the unknown population parameter θ. Note that within the frequentist framework a population parameter is regarded as a fixed, unknown, and unvarying quantity, whereas within the Bayesian framework a population parameter is regarded as an unobservable random variable which can be described distributionally. If $f(X)$ is normal it follows from repeated sampling theory that

$$\overline{X} \sim N(\mu, \sigma^2/n). \tag{2}$$

Hence, the probability that the random interval

$$(\overline{X} - 1.96 \cdot \sigma/\sqrt{n}, \overline{X} + 1.96 \cdot \sigma/\sqrt{n}) \tag{3}$$

contains the value of μ is 95%. Replacing $\overline{X}$ with the estimate $\overline{x}$ by using the Fisherian "plug-in principle" (Efron, 1998, p. 101) provides the 95% confidence interval for μ. Once having calculated the frequentist confidence interval, the unknown but fixed parameter is either contained in the interval or not. Note, a valid probability statement associated with a 95% confidence interval can only be made with reference to the procedure not the unknown parameter μ itself. Since μ is not a random variable, the 95% frequentist confidence interval for μ means nothing but that the procedure of interval construction is expected to construct intervals that include μ about 95% of the time: "The sampling property of confidence intervals [...] says nothing at all about the probability of θ given the result of the one particular sample of data actually obtained. Rather, it makes probability statements about the interval calculated from hypothetical samples that have not been obtained" (Fisher-Box, 1978, p. 456). In other words, frequentist probability statements are nothing else but "idealized frequencies" (Bernard, 2000, p. 168). To conclude, the frequentist framework is mathematically off target when a researcher is interested in making probability statements about a hypothesis of interest because it only refers to the sample space and not to the parameter space.

Unfortunately, many researchers are prone to assign a probability statement to a frequentist test statistic and, thus, misinterpret frequentist confidence intervals and significance tests. In doing so, those researchers involuntary behave like Bayesians. Although it seems to be a statistical subtlety at first sight, what frequentists in mind but Bayesians at heart are actually interested in is

$$p(H|data) \tag{4}$$

with p being the probability that the null hypothesis H is true given the results (*data*) obtained. As we shall see subsequently, this kind of interpretation will become feasible only if the Bayesian statistical framework is employed.

3 What's Bayesian statistics all about?

In September 2000, The Economist published an article titled "In praise of Bayes" which brilliantly summarized the very basic idea of the Bayesian approach: "The essence of the Bayesian approach is to provide a mathematical rule explaining how you should change your existing beliefs in the light of new evidence. In other words, it allows scientists to combine new data with their existing knowledge or expertise" (Economist, 2000).

The Bayesian framework comprises a prior state of knowledge (i.e. *prior* in the sense of *before* observing the data) and the data likelihood to a *more informed* posterior distribution, that is

$$\textit{Posterior Distribution} \propto \textit{Prior Distribution} \times \textit{Likelihood}, \tag{5}$$

where the symbol "$\propto$" means "is proportional to". More specifically, Bayesian inference always begins with some prior probability statement about an unknown parameter, that is $f(\theta)$, for example. Recall that, in contrast to the frequentist paradigm, all unknown parameters of a Bayesian model are treated as random variables which can be described distributionally. The prior probability statement $f(\theta)$ is nothing else but a summarized expression of the current state of knowledge before gathering or seeing any new data. In general, the distribution of the unknown parameters before any new data are seen is called prior distribution (we thoroughly explain the rationale for selecting prior distributions in section 5). By specifying a prior distribution, only half the Bayesian inference story is told tough. Let, for example, $data = (x_1, \ldots, x_n)$ be a sample from a density f_θ with an unknown parameter θ and with an associated likelihood function

$$L(\theta|data) = \prod_{i=1}^{n} f_\theta(x_i). \quad (6)$$

The likelihood function summarizes the sample information about θ and provides some value of θ that makes the data most *likely* to have occured. The information from the likelihood function is weighted with the prior probability distribution by employing Bayes' Theorem to calculate an updated distribution *posterior* to the former state of knowledge:

$$f(\theta|data) = \frac{f(data|\theta) \times f(\theta)}{f(data)}, \quad (7)$$

where $f(\theta|data)$ denotes the posterior distribution for the parameter θ, $f(data|\theta)$ is a sampling density for the data, $f(\theta)$ is the prior distribution for the parameter, and $f(data)$ is the marginal probability of the data. Whereas frequentists inference about θ follows from inspection of the likelihood only, Bayesian inference contrastly relies on inspecting the posterior distribution using descriptive measures. The shape of the posterior distribution can be described by calculating location parameters such as the posterior mean, that is the expected value of θ under $f(\theta|data)$, or the posterior mode, that is the most likely value under $f(\theta|data)$, or, in addition, by some variability measures. Assuming that a posterior density is approximately normal, derivation of a 95% confidence interval (also called *credible interval* (CI) by Bayesians), for example, is straightforward then

$$posterior\ mean \pm 1.96 \times posterior\ standard\ deviation. \quad (8)$$

To sum up, Bayesian statistics always involves calculating a posterior distribution which is formed by weighting a prior distribution by the likelihood function: "Put generally, the goal of Bayesian statistics is to represent prior uncertainty about model parameters with a probability distribution and to update this prior uncertainty with current data to produce a posterior probability distribution for the parameter that contains less uncertainty" (Lynch, 2007, p. 50).

4 Why should we use Bayesian statistics?

The Bayesian approach to statistical inference is much more compatible to the cumulative nature of science as it is practiced by most quantitative researchers today. A theoretical model of causal mechanisms is developed, data are gathered either by experiment or observation, and the theoretical model is tested against the empirical data. Based on inferences drawn from this analyses, hypotheses are either retained, rejected or modified. Ideally, this process is repeated, using past results as a guideline for new analyses. However, even for newly developed hypotheses, any scholar will have some expectation or belief about the investigated phenomenon – tabula rasa situations are very rare in our discipline. Any statistical framework that utilizes this cumulative process rather than starting from scratch with every further study is therefore more appropriate (cf. Box & Tiao, 1973). Contrary to the frequentist approach, the Bayesian framework explicitly enables, and indeed requires, the researcher to incorporate prior knowledge into the statistical inference: "Bayes' theorem is the *only* consistent way to modify our beliefs about the parameters given the data that actuallly occured" (Bolstad, 2004, p. 7, emphasis added by the authors).

The notion that statistical inference is a way of updating existing knowledge given new data, and that, consequently, data do not speak for themselves might be irritating at first sight. The incorporation of pre-existing information adds a seemingly more subjective flavour to the scientific method than pure likelihood-based inference. Consequently, the subjectivity of the chosen priors is the most prominent objection to the Bayesian approach (Gelman, 2008). The counterargument to this claim is twofold: First, as noted by many authors (Levine et al., 2008a; Gill, 2008) before, frequentist inference rests not only on strong objective assumptions – e.g. asymptotic theory – but also on substantive subjective decisions – e.g. the level of significance. Second, and more importantly, the incorporation of prior knowledge has considerable advantages for the quality of the scientific discourse. Namely, the Bayesian approach forces the adopter to think about, document, and justify the choice of the prior which is, then, subject to scientific debate. If the prior merely reflects the beliefs of a scholar about a social phenomenon, a thorough theoretical explanation will be necessary. A prior based on pre-existing research will provide valuable insights into the author's comprehension of past scholarship. In any case, social sciences will be more transparent and less prone to implicit or undocumented assumptions. We will return to this issue in the next section.

The Bayesian framework does not only have epistemological advantages, but also offers direct and intuitive answers to the question that lies behind most empirical research. That is, what is the probability of H given the data? Only Bayes' theorem allows to assign probability statements to parameters, e.g. to claim that the true value lies within a certain limit. This kind of claim was unintentionally made for decades by many Bayesians at heart.

Since the Bayesian computation of credible intervals (CI) for the posterior distribution takes additional data into account, the estimation is equally or even more precise than with the frequentist procedure. This means that a Bayesian CI is not only intuitively interpretable, but also generally smaller than the frequentist counterpart. This is especially advantageous if an empirical prior is used for a replication study. Since the prior adds additional information, the Bayesian CI will yield more precise estimates most of the time. Moreover, since Bayesian analysis deals with the complete posterior distribution rather than point estimates and standard errors, many alternative ways of summarizing these results are available (Lynch, 2007, p. 71).

Furthermore, Bayesian methods allow to tackle large and complex statistical problems with relative ease, where frequentist methods can only approximate or fail altogether. For example, the incorporation of missing data is straightforward and consistently handled (Jackman, 2000), predictive summaries of parameters of interest are easily obtained (Bolstad, 2004; Gill, 2008), hierarchical and other complex models can be fit even with small samples (cf. Gelman & Hill, 2007). To our view, these are important but rather technical features of the Bayesian framework that a far from typical models of communication research, and, therefore, we do not discuss their use subsequently.

Before highlighting the issue of prior distributions, we recapitulate the basic arguments in favour of the Bayesian approach to statistical inference:

1. The use and interpretation of p-values or significance testing is often inappropriate or simply wrong, but even if applied correctly, e.g. using confidence intervals, the common approach gives us $p(data|H)$ instead of the more interesting $p(H|data)$: *Bayesian analysis yields the kind of probability statements we are interested in.*
2. Frequentist methods are inefficient because the available information is not incorporated at the analysis stage. Even when meta analyses are conducted beforehand, the actual inference happens on the new data alone: *The Bayesian approach allows to incorporate pre-existing information into the statistical model.*
3. As an additional benefit, Bayesian statistics enforces transparency and clarity by demanding justification for the choice of the statistical model or likelihood *and* the prior: *Using Bayesian methodology increases the quality of the scientific discourse.*

5 Where do prior distributions come from, anyway?

After acknowledging the potential advantages of the Bayesian approach, most scholars usually wonder about which priors to use. In the following, we try to provide a useful taxonomy of priors, as depicted in Figure 1, and a brief discussion of the rationale and consequences of using them.[1]

Not surprisingly, there is considerable dissent on the issue of specifying priors among scholars with regard to objectivity. Although the debate between *subjective* and *objective* Bayesians is not only about prior distributions, most scholarship focusses on this problem.[2] Consequently, the choice of priors is first and foremost between uninformative and informative distributions for a parameter of interest. The former are chosen in such a way that the data, i.e. the likelihood according to Bayes' theorem, speak for themselves. Although there is consensus that no statistical or other scientific method can be truly objective, many scholars argue that using uninformative priors is more justifiable and transparent to both colleagues, students and external reviewers, like regulatory agencies overseeing clinical trials (Berger, 2006). For those, the appearance of objectivity might be more important than the efficiency gained by using informative priors. Statistically, using noninformative priors results in about the same results as conventional frequentist inference but still allows Bayesian *interpretation*, so at least one of the advantages we mentioned above remains.

The simplest way of expressing prior ignorance is often seen in using a uniform distribution for a parameter of interest: Every value is deemed equally probable, possibly within some bounds like $-1 \leq r \leq 1$ for a correlation coefficient. Because this kind of prior is a constant, the posterior distribution is computed only from the likelihood, yielding a pseudo-frequentist result. As easy as the interpretation and justification of this prior may seem, *uniform priors* are not equally well suited for all kinds of parameters, mainly because they are not robust to simple transformations. For example, a uniform prior for a variance σ^2 is different than that for a standard deviation σ, and both yield different posteriors (Lambert et al., 2005).

The construction of uninformative priors that are both robust to transforma-

[1]See Gill (2008) for an excellent overview from both applied and statistical perspectives.

[2]Although even the nomenclature of objective/subjective Bayesian analysis is challenged by many scholars, we keep this distinction in its historical context. See Berger (2006), Goldstein (2006) and the numerous comments in the same issue of *Bayesian Analysis* for a recent discussion of the topic.

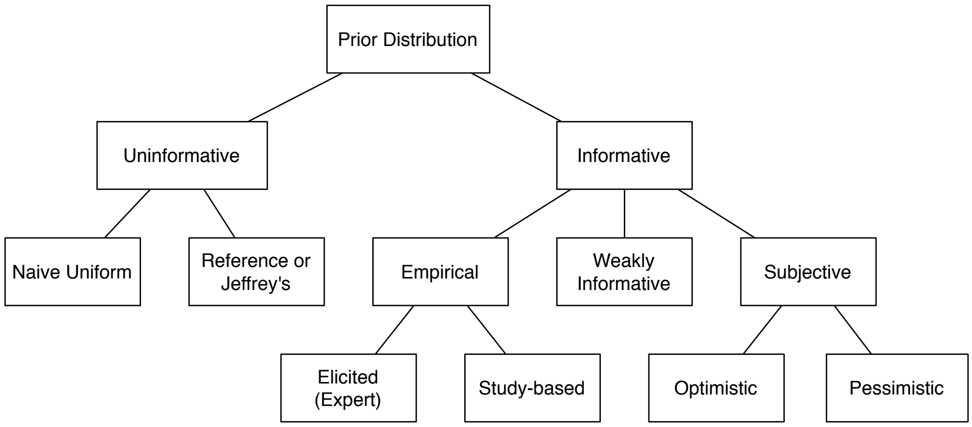

Figure 1: A taxonomy of priors in Bayesian inference

tions and lead to proper posteriors has challenged many Bayesian statisticians from Jeffreys (1961) to Bernardo (1979), and others. Today, most objective Bayesians would probably recommend using these *reference priors* for all-but-trivial models. Important as research in this area may be, we feel that these issues are somewhat statisticians' pets and of less interest for applied researchers.[3]

As argued above, the incorporation of informative prior knowledge is the key to more efficient statistical analyses and a major argument in favour of the Bayesian approach. If one accepts the merits of informative priors in quantitative research, the question of obtaining them arises. The answer is both simple and complicated: Any knowledge or belief, be it from an expert, a pre-existing study, theoretical reasoning or just an educated guess can be used as long as it can be transformed into a probability distribution of some kind. We propose a distinction between empirical priors, weakly informative priors and subjective priors (see Figure 1).

Empirical priors refer to *previous* observation or other forms of data collection including expert interviews.[4] The latter are often termed *elicited priors*. Eliciting a prior is almost a science because transforming qualitative or vague

[3] See Yang & Berger (1996) and Kass & Wasserman (1996) for a detailed discussion of noninformative priors.

[4] This is not to be confused with the empirical Bayes approach that suggests using the same data for the construction of both the prior *and* the likelihood. Many Bayesians warn against the double use of the data, and we agree.

expert opinions into prior distributions for parameters like regression coefficients or between-group variances is a challenging task for any applied researcher (cf. O'Hagan, 1998; Gill & Walker, 2005; O'Hagan et al., 2006). Another kind of empirical prior comes from the incorporation of previous results, which are widely available in clinical or industrial trials. Those results can either be directly incorporated in strict replication studies or somewhat discounted depending on the similarity of the previous and the current study. While this can be intuitively accomplished by using equivalent sample sizes (Bolstad, 2004) in order to relate prior and likelihood, a more versatile and rigorous approach to this problem are so-called *power priors* (Ibrahim & Chen, 2000).

Subjective priors express personal theories or beliefs about a phenomenon, and as such they are highly debatable. We do not encourage the use of strong subjective priors as they tend to make actual empirical research obsolete. However, it may often be desirable to check the consequences of different priors on the posterior distribution, or compare models with optimistic or pessimistic priors.

While both empirical and subjective priors are domain-specific, a third kind of informative prior distribution is based on statistical convenience and common sense about model parameters. The rationale behind weakly informative priors (Gelman, 2006; Gelman et al., 2009) is that extreme values of parameters such as correlations or regression coefficients are highly unlikely and should therefore be given less prior probability. This generic approach to using default priors leads to efficient and stable estimation without overly affecting the likelihood. We agree that such pragmatic solutions that just work in many applied cases may be essential for the wider adoption of Bayesian analysis.

The taxonomy proposed here and pictured in Figure 1 suggests that uninformative and informative priors are clearly separate, but in applied research any informative prior, subjective or empirical, can be made less informative simply by adjusting the scale parameter (i.e. variance) of the distribution.

Finally, priors can also be classified according to their mathematical properties, namely if they are proper and conjugate. Some, but not all improper prior distributions lead to improper posteriors, i.e. posterior probabilities that do not sum to 1. This is undesirable and should therefore be avoided by using proper priors. Conjugate priors belong to the same distributional family as the likelihood, making it easy to analytically compute posteriors. However, since the advent of numerical methods like Markov Chain Monte Carlo non-conjugacy is no longer a big problem in Bayesian analysis (Lynch, 2007). After all, priors should be chosen for substantial reasons not mathematical convenience, although having both is even better.

Up to now, we totally ignored the question of whether communicaton researchers are able to adopt prior scientific knowledge at all. To shed some light on this transformational process, we carried out a short case study. More specific, we presented some pre-existing empirical results to fellow German communication scholars and examined how they transform the information presented into either a numerical or a graphical representation.

6 Are we ready for Bayesian statistics?

Research questions

No matter how promising the Bayesian framework might sound from a methodological perspective, its widespread adoption depends on the willingness and ability of researchers to apply Bayesian methods in their work. Our question is not whether or not humans are generally capable of Bayesian – or any other probabilistic – thinking (Tversky & Kahneman, 1974; Cosmides & Tooby, 1996; Gigerenzer & Todd, 1999). Instead we focus on the routine use of prior information in empirical research: How do empirically trained scholars read, understand and incorporate preexisting results?

In order to answer this question we conducted a small survey-based field experiment with fellow communication researchers. We had two research questions: *When* is prior knowledge usually given consideration in the research process? And *how* do communication researchers perceive empirical results and transform them into suitable priors?

If one assumes the typical workflow in empirical social science, prior research results may be incorporated at four different stages: During the review of existing scholarship on a specific subject, at the operationalization and measurement step, during the actual statistical analysis, and finally in the discussion of results. Given the difficulty of incorporating prior data in frequentist analyses, we expect that communication researchers will report less frequent use of these data in the analysis step compared with other steps of research.

It is important to note that it is a matter of both willingness and statistical expertise to incorporate prior knowledge in a Bayesian model. However, we are unable to answer the question which of these stages is most important, especially given that most communication researchers are not aware of the alternative statistical approach provided by Bayesian statistics.

But what we can examine is the question of how researchers read and summarize pre-existing empirical results. Any scholar who reviews existing studies will inevitably develop an idea of the direction and strength of relationships between variables of interest. However, most communication researchers rarely state their expectations in terms of effect sizes. Therefore we ask how this implicit knowledge can be elicited using different techniques (cf. O'Hagan, 1998; O'Hagan et al., 2006). Since prior knowledge can be expressed both numerically (as point estimates with uncertainty) and graphically (as probability distributions), we explore whether the mode of elicitation makes a difference.

Method

The case study was carried out as a survey during two meetings of German communication scholars at the annual conferences of the *Media Audiences and Effects* section and a joint conference of the *Journalism Research* and *Methods in Communication Research* sections of the *Deutsche Gesellschaft für Publizistik- und Kommunikationswissenschaft (DGPuK)*. In total, $n = 43$ colleagues participated in the case study. All interviews were conducted using paper-and-pencil questionnaires.

First, we asked how often respondents incorporated prior empirical results at different stages of the research process on a 5-point scale. The elicitation task referred to the graph pictured in Fig. 2. The plot displays the results of different studies in the context of news values research (Shoemaker et al., 1986; Eilders, 2006) and is taken from a meta-analysis conducted by Fretwurst (2008). We chose this topic because it should be familiar to most German communication researchers.

The graph displays point estimates of correlation coefficients between the news factor continuity and the news value and their confidence intervals. After describing the context and the meaning of Figure 2, we asked our colleagues to specify what they learn from these data, i.e. what they think the true correlation is and how certain they are.

In order to answer the question whether numerical and graphical elicitation modes differ, we used a 4-way split ballot design with two numerical and two graphical options.[5] We then asked the respondents how hard they felt the

[5]The wording of the items is presented in the appendix.

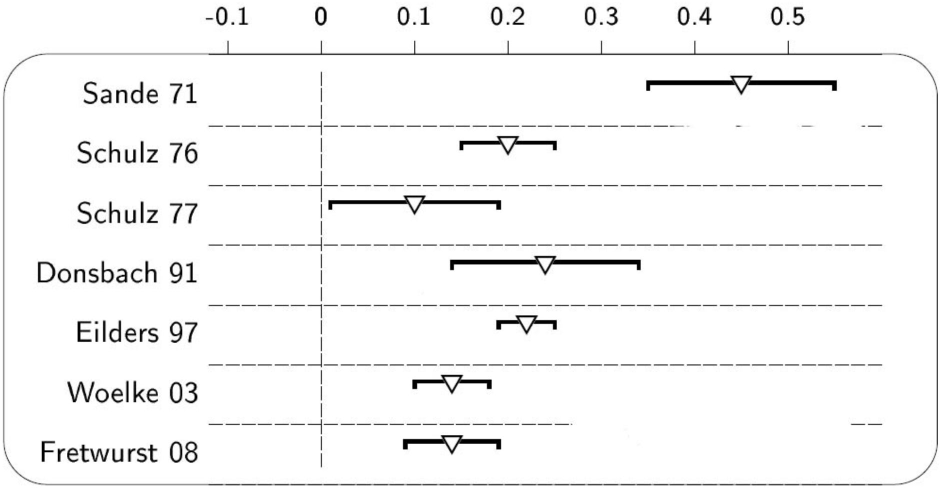

Figure 2: Empirical results presented in the survey

preceding task was on a 5-point scale.

Results

Since we collected responses from $n = 43$ scholars, all inferences should be interpreted with caution. Of course, we used Bayesian methods for the following analysis. Specifically, we used data from a small pretest we ran with seven colleagues before the conferences to form prior distributions. We chose different normal priors which are, however, completely dominated by the survey data.[6]

In order to test the hypothesis that incorporation of prior result happens less frequently during the actual analysis, we computed group means and Bayesian CI. As can be seen in Figure 3, the distribution of answers is quite skewed (remember the scale ranges from 1 to 5), as most respondents claim to incorporate prior results at all stages of the research process. As expected,

[6]See the appendix for the prior distributions used.

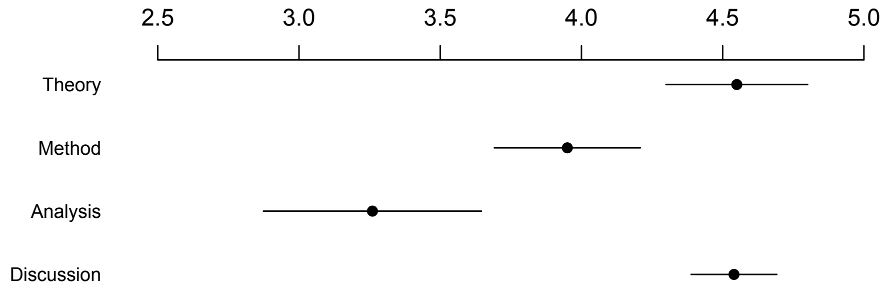

Figure 3: Incorporation of prior results in the research process (means and 95% Bayesian CI)

this incorporation happens significantly less often at the operationalization and especially at the analysis stage. Moreover, while there is almost universal agreement to acknowledge prior results in the literature review and discussion, opinion is divided whether or not to do this at the method or analysis stage.

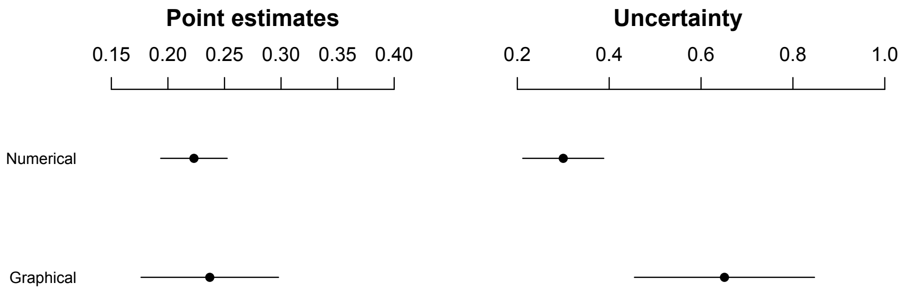

Figure 4: Elicited point estimates and widths of confidence intervals (uncertainty) for the correlation coefficient (means and 95% Bayesian CI)

Our second question was how our fellow researchers handle previous empirical results. Generally, most respondents acknowledge that they found the given task hard to solve ($\bar{x} = 4.24$, CI= [3.85 : 4.63]). Apparently, communication scholars are not used to summarizing prior knowledge quantitatively. More interesting is the fact that the amount of uncertainty given in their answers is significantly higher when elicited graphically than numerically. Apparently, asking for precise lower and upper bounds leads to more certainty in the answers than asking for a histogram or probability curve. As can be seen in Figure 4, the elicited point estimates are not biased by the elicitation method, but the width of the confidence interval for the prior is. In short: people tend to give more informed priors when asked in familiar terms. Finally, the elicited prior distributions from our example study are certainly *very* informative, as they cover at most about a third of the theoretical range for a correlation coefficient.

Summarizing our empirical findings, we can conclude that our fellow communication researchers do recognize that incorporation of prior knowledge does happen less frequently during statistical analyses. A finding that may well reflect the difficulty of doing so with conventional frequentist methods. Moreover, many of our colleagues had quite a hard time summarizing prior empirical data and expressing their perception with estimates of uncertainty. The mode of elicitation makes a significant difference with respect to the priors obtained: Graphical descriptions yield more diffuse priors than conventional confidence intervals which appear to be almost too certain. This problem of expert's overconfidence has been investigated before (O'Hagan, 1998) and is an important, but often neglected problem within the Bayesian framework.

7 Is Bayesian statistics ready for us?

The aim of this paper was to introduce the Bayesian approach to statistical inference for communication research. We hope that interested scholars will find the arguments for this frameworks compelling enough for applied research. Yet, there is still a lot to do both from statisticians and communication scholars in order to better understand the implications and consequences of Bayesian thinking. Our discipline might not be quite ready yet, as is the Bayesian analysis toolkit.

Looking back on statistics in the 20th century, Bradley Efron (1998, p. 96) subsumed that statistical thinking and methodology have "grown from a small obscure field into a big obscure field". The rediscovery of the Bayesian paradigm and its practicability due to the advent of computer intensive methods made the big obscure field even bigger. Bayesian statistics would be seen less obscure when communication scholars become more and more familiar with Bayesian thinking.

References

Altman, D. & Gardner, M. (2000). *Statistics with confidence: confidence intervals and statistical guidelines*. BMJ Books.

Berger, J. (2006). The case for objective Bayesian analysis. *Bayesian Analysis*, 1(3), 385–402.

Bernard, J. (2000). Bayesian inference for categorized data. In J. B. H. Rouanet, M.-C. Bert, B. Lecoutre, M.-P. Lecoutre, & B. L. Roux (Eds.), *New ways in statistical methodology from significance tests to Bayesian inference* (pp. 159–226). Bern: Peter Lang.

Bolstad, W. (2004). *Introduction to Bayesian Statistics*. Hoboken: Wiley.

Box, G. E. P. & Tiao, G. C. (1973). *Bayesian inference in statistical analysis*. Reading: Addison-Wesley.

Cosmides, L. & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition*, 58(1), 1–73.

Cumming, G., Fidler, F., Leonard, M., Kalinowski, P., Christiansen, A., Kleinig, A., Lo, J., McMenamin, N., & Wilson, S. (2007). Statistical Reform in Psychology: Is Anything Changing? *Psychological Science*, 18(3), 230–232.

Economist (2000). In Praise of Bayes. published, Sep 30, 2000.

Efron, B. (1986). Why Isn't Everyone a Bayesian? *The American Statistician*, 40(1), 1–5.

Efron, B. (1998). RA Fisher in the 21st century. *Statististical Science*, 13, 95–122.

Eilders, C. (2006). News factors and news decisions. Theoretical and methodological advances in Germany. *Communications*, 31(1), 5–24.

Fisher-Box, J. (1978). RA Fisher: the Life of a Scientist.

Fretwurst, B. (2008). *Nachrichten im Interesse der Rezipienten*. UVK.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, 1(3), 515–534.

Gelman, A. (2008). Objections to Bayesian statistics. *Bayesian Analysis*, 3(3), 445–450.

Gelman, A., Jakulin, A., Pittau, M., & Su, Y. (2009). A default prior distribution for logistic and other regression models. *Annals of Applied Statistics*, 2(4), 1360–1383. mimeo.

Gigerenzer, G., Swijinck, Z., Porter, T., Daston, L., Beatty, J., & Kruger, L. (1989). The Empire of Chance.

Gigerenzer, G. & Todd, P. (1999). *Simple heuristics that make us smart*. Oxford: Oxford University Press.

Gill, J. (1999). The Insignificance of Null Hypothesis Significance Testing. *Political Research Quarterly*, 52(3), 647–674.

Gill, J. (2008). *Bayesian methods: a social and behavioral sciences approach*. Boca Raton: Chapman & Hall, 2nd edition.

Gill, J. & Walker, L. (2005). Elicited Priors for Bayesian Model Specifications in Political Science Research. *Journal of Politics*, 67(3), 841–872.

Goldstein, M. (2006). Subjective Bayesian Analysis: Principles and Practice. *Bayesian Analysis*, 1(3), 403–420.

Hubbard, R. & Armstrong, J. (2006). Why We Don't Really Know What Statistical Significance Means: A Major Educational Failure. *Journal of Marketing Education*, 28, 114–120.

Ibrahim, J. & Chen, M. (2000). Power Prior Distributions for Regression Models. *Statistical Science*, 15(1), 46–60.

Jackman, S. (2000). Estimation and Inference Are Missing Data Problems: Unifying Social Science Statistics via Bayesian Simulation. *Political Analysis*, 8(4), 307.

Kass, R. & Wasserman, L. (1996). The selection of prior distributions by formal rules. *Journal of the American Statistical Association*, 91(435), 1343–1370.

Krüger, L., Gigerenzer, G., & Morgan, M. (1987). The probabilistic revolution. Vol. II: Ideas in the sciences.

Lambert, P. C., Sutton, A. J., Burton, P. R., Abrams, K. R., & Jones, D. R. (2005). How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. *Statistics in Medicine*, 24, 2401–2428.

Levine, T., Weber, R., Hullett, C., Park, H., & Lindsey, L. (2008a). A Critical Assessment of Null Hypothesis Significance Testing in Quantitative Communication Research. *Human Communication Research*, 34(2), 171–187.

Levine, T., Weber, R., Park, H., & Hullett, C. (2008b). A Communication Researchers' Guide to Null Hypothesis Significance Testing and Alternatives. *Human Communication Research*, 34(2), 188–209.

Lynch, S. (2007). *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*. New York: Springer.

Malakoff, D. (1999). STATISTICS: Bayes Offers a'New'Way to Make Sense of Numbers. *Science*, 286(5444), 1460.

O'Hagan, A. (1998). Eliciting Expert Beliefs in Substantial Practical Applications. *Statistician*, 47(1), 21–35.

O'Hagan, A., Buck, C., Daneshkhah, A., & Eiser, J. (2006). *Uncertain Judgements: Eliciting Experts' Probabilities*. Wiley.

Savage, L. J. (1976). On Rereading R. A. Fisher. *The Annals of Statistics*, 4(3), 441–500.

Shoemaker, P., Chang, T., & Brendlinger, N. (1986). Deviance as a predictor of newsworthiness: Coverage of international events in the US media. *Communication Yearbook*, 10, 348–365.

Tversky, A. & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185(4157), 1124–1131.

Yang, R. & Berger, J. (1996). A Catalog of Noninformative Priors. mimeo.

Zabell, S. (1989). R. A. Fisher on the History of Inverse Probability. *Statistical Science*, 4(3), 247–256.

Appendix: Survey questions for the prior elicitation task

What do you think is the true correlation between the news factor *continuity* and the journalists' attention?

1. Please specify a point estimate and the bounds of a 95 percent confidence interval for the correlation.
2. Please specify a value for the true correlation. How big is the coefficient at least and at most?
3. Please draw a pobability distribution for the true correlation coefficient.
4. Imagine 100 replication studies will yield new correlation coefficients. Please draw a histogram for the frequency distribution of those coefficients.

Priors used for the analyses

Importance of prior results during the analysis (see Fig 3)

Theory $N(4.428, .786)$

Method $N(4.00, 1.414)$

Analysis $N(3.143, 1.574)$

Discussion $N(3.714, .951)$

Elicited prior distributions (see Fig 4)

Numerical Mean $N(.184, .023)$

Numerical Interval $N(.362, .301)$

Graphical Mean $N(.225, .035)$

Graphical Interval $N(.500, .141)$