Management and Organization Review 7:3 471–479
doi: 10.1111/j.1740-8784.2011.00222.x

Presenting Post Hoc Hypotheses as A Priori: Ethical and Theoretical Issues

Kwok Leung
City University of Hong Kong, Hong Kong

ABSTRACT Presenting post hoc hypotheses based on empirical findings as if they had been developed a priori seems common in management papers. The pure form of this practice is likely to breach research ethics and impede theoretical development by suppressing the falsification process. Two other forms may be more tolerable: deletion of rejected hypotheses and refinement of hypotheses inspired by empirical findings. To address this problem, the field should provide stronger recognition of replication, descriptive research, rejected and post hoc hypotheses, and critical tests of competing hypotheses. These positive changes require the concerted effort of researchers, management associations, and journal editors and reviewers.

KEYWORDS hypothetic–deductive method, positivity bias, post hoc hypothesis

How odd it is that anyone should not see that all observation must be for or against some view if it is to be of any service!

Charles Darwin
Letter to Henry Fawcett

INTRODUCTION

The zeitgeist of top-notch management journals mandates a huge premium on theoretical contributions in determining what papers will see the light within their pages. Take *Academy of Management Journal* as an example, which states in its information for contributors that 'All articles published in the *Academy of Management Journal* must also make strong theoretical contributions' (Academy of Management, 2010). Hambrick (2007: 1349) has provided a lucid account of how the obsession with theory in management research can ironically stifle theory development: 'what we too often see in our journals: a contorted, misshapen, inelegant product, in which an inherently interesting phenomenon has been subjugated by an ill-fitting theoretical framework'. The objective of this paper is to explore a different, but equally consequential result of the over-emphasis on theory in

© 2011 The International Association for Chinese Management Research

https://doi.org/10.1111/j.1740-8784.2011.00222.x Published online by Cambridge University Press

management journals, namely, presenting post hoc hypotheses based on the empirical findings as if they had been developed a priori. There are many reasons why authors engage in this or other practices that may be at odds with desirable research norms (Chen, 2011; Kirkman & Chen, 2011). I focus on the implications of this particular practice from two perspectives: research ethics and theory development, and offer some suggestions to address the predicament associated with this practice.

POST HOC HYPOTHESES PRESENTED AS A PRIORI

Scientific research guided by theories typically follows the hypothetico–deductive method, and requires the formulation of a priori hypotheses and their subsequent empirical evaluation (Ayala, 2009). Confirmed hypotheses will support the theory upon which they are based, whereas unsupported hypotheses will have the potential to lead to theory refinement. In the case of disconfirmation, post hoc theorizing and hypotheses are developed, which will be subjected to further empirical testing (Jaccard & Jacoby, 2010). Presenting post hoc hypotheses based on the findings obtained as a priori differs from post hoc hypotheses in that the post hoc nature of these hypotheses is not acknowledged. Kerr (1998) labelled this practice as 'HARKing' (Hypothesizing After the Results are Known). I prefer the label 'Presenting Post hoc Hypotheses as A priori' (PPHA) to clearly differentiate this practice from the formulation of post hoc hypotheses for subsequent empirical verification.

Kerr (1998) has described several forms of hypothesizing based on one's results, three of which are especially widespread. In the pure form, researchers formulate hypotheses entirely based on the known results to develop the most compelling, most theoretically coherent account of the results. A priori theorizing is ignored, and virtually all theoretical analysis is developed post hoc to justify the hypotheses, which are presented as a priori in the introduction. The second type is known as suppressing loser hypotheses, in which disconfirmed hypotheses are dropped from the introduction. The third type is known as empirical inspiration, in which new hypotheses inspired by the results are added, but the added hypotheses, although post hoc in nature, are presented as a priori in the introduction.

To determine how widespread PPHA is, Kerr and Harris (1998, cited in Kerr, 1998) surveyed a group of researchers from psychology and sociology. This group reported that they personally observed over 30 percent of researchers in their disciplines engaging in the three forms of PPHA. Interestingly, less than 20 percent of this group approved the pure form of PPHA and the suppression of loser hypotheses, but empirical inspiration was seen as almost as preferable as the hypothetico–deductive method.

Why do researchers engage in PPHA? Kerr (1998) identified the pervasive positivity bias in academic research as a major cause. For instance, Turner,

© 2011 The International Association for Chinese Management Research

https://doi.org/10.1111/j.1740-8784.2011.00222.x Published online by Cambridge University Press

Matthews, Linardatos, Tell, and Rosenthal (2008) compared the results of trials of antidepressant agents as reported in the literature against those registered with the Food and Drug Administration of the U.S. In the published literature, 94 percent of the trials yielded positive results, whereas only 51 percent of the results registered with the Food and Drug Administration were positive. Because positive results are deemed as generally more valuable, researchers can maximize the amount of positive results by tinkering with the hypotheses based on the empirical findings.

I believe that this positivity bias is also prevalent in management, leading researchers to engage in PPHA. To illustrate this point, I analysed the 54 articles published in *Academy of Management Journal* in 2009. The papers were categorized as either qualitative or quantitative. Among the 54 papers, 45 were quantitative in nature and all involved hypothesis testing, confirming that deductive research is seen as more valuable than research based on the inductive approach. In contrast, none of the nine qualitative studies had explicit hypotheses, confirming the well-known quantitative/hypothesis testing and qualitative/exploration divide.

In two quantitative papers, there were two studies testing the same set of hypotheses with different samples. I treated the two studies in each paper as distinct, resulting in 47 quantitative studies. A total of 251 hypotheses were examined, with seven hypotheses duplicated because two papers tested the same set of hypotheses in two independent studies. Note that subhypotheses in the form of 'Hypothesis 1a' and 'Hypothesis 1b' were counted as distinct hypotheses, because they involved different relationships. One exception was made, which involved three competing subhypotheses, such that if one was supported, the other two would be rejected. The three subhypotheses were counted as one hypothesis, which was classified as having received 'partial support'.

To determine whether a hypothesis was supported or rejected, I relied on the conclusions presented in the papers. In addition, a hypothesis was counted as being 'partially supported' if it was explicitly described as being 'partially supported', or when unhypothesized, mixed results were presented (e.g., 'the hypothesis was supported in Situation X, but not in Situation Y'). As expected, the large majority of the hypotheses (75.3 percent) were fully supported (see Table 1). In 20 out of the 47 studies, all the hypotheses were fully supported, and no study had more than 50 percent of the hypotheses rejected. We may congratulate management researchers

Table 1. Analysis of the papers published in *Academy of Management Journal* in 2009

Type of study	*Number of study*	*Total number of hypotheses*	*Number of hypotheses supported*	*Number of hypotheses partially supported*	*Number of hypotheses not supported*
Quantitative	47	251	189 (75.3%)	21 (8.4%)	41 (16.3%)
Qualitative	9	0	0	0	0

© 2011 The International Association for Chinese Management Research

https://doi.org/10.1111/j.1740-8784.2011.00222.x Published online by Cambridge University Press

on their impressive foresight in their theoretical forecast, but I suspect that some degree of PPHA may have contributed to this high rate. Without additional data, it is hard to determine the prevalence of PPHA in management research, but this crude analysis suggests that it is not a rare phenomenon.

Is PPHA Unethical?

Kerr (1998) reviewed the Ethical Principles and Codes of Conduct of the American Psychological Association (e.g., Canter, Bennett, Jones, & Nagy, 1994) and the list of ethical felonies and misdemeanours of the National Academy of Sciences' Blue-Ribbon Committee (Committee on Science, Engineering, and Public Policy, 1992), and did not find an explicit objection to PPHA. Similarly, the Academy of Management's Code of Ethics is mute about PPHA. The following guideline in the Code seems most relevant to PPHA, but it does not spell out whether the practice of PPHA requires disclosure.

> 4.1.3. AOM members take particular care to present relevant qualifications to their research or to the findings and interpretations of them. AOM members also disclose underlying assumptions, theories, methods, measures, and research designs that are relevant to the findings and interpretations of their work.

Kerr (1998: 209) concluded that presenting post hoc hypotheses as a priori 'violates a fundamental ethical principle of science: the obligation to communicate one's work honestly and completely'. This view is reasonable, but not universally endorsed (e.g., Bem, 1987), and is contradicted by the positive attitude towards empirical inspiration, a mild form of PPHA, in the survey previously described. My own position is more nuanced than a simple right-or-wrong stance. I agree with the view that hypotheses should be set up in a way so that they may be disconfirmed. Because PPHA turns empirical findings into a priori hypotheses and makes their disconfirmation impossible, this practice does involve hidden information. I think the pure form of PPHA is not defensible, because hypotheses in this context are no more than empirical findings disguised as hypotheses.

Suppressing loser hypotheses is a more difficult case to judge. This practice seems analogous to giving up the publication of data sets that provide no useful results. I suppose no one would accuse researchers of being unethical if they decide not to publish what they regard as un-interpretable data. One major function of suppressing loser hypotheses is to enhance the focus and coherence of the arguments and theorizing presented. As long as the dropped loser hypotheses and the associated results do not contradict the arguments and theorizing promulgated in a paper, there may not be any ethical issues involved. I have been frequently asked by editors and reviewers to drop results that they found tangential or uninteresting. However, I often ask them to allow me to add a short description of the results that

© 2011 The International Association for Chinese Management Research

https://doi.org/10.1111/j.1740-8784.2011.00222.x Published online by Cambridge University Press

I find important. I place these descriptions either in a footnote or in the discussion section of the text to alert readers for possible future research.

It is important to note that suppressing loser hypotheses is not the same as suppressing results that cast doubt on one's argument and theorizing. In the latter case research ethics are clearly breached. I think the crucial issue is whether dropping the loser hypotheses and the associated results will create the impression that the conclusions of a study are more valid than when a complete account is provided. If the deletion does not make the conclusions appear more valid, but only improves the coherence and focus of a study, this practice would not be a breach of research ethics.

The case of empirical inspiration is equally hard to judge. A scrutiny of the findings often inspires researchers to modify their earlier hypotheses and add new ones. In fact, researchers are often asked by reviewers and editors to do so, partly because they have the benefit of the hindsight provided by the findings. The crucial issue is whether researchers are truly inspired by the findings, or just reformulating the hypotheses to increase the number of confirmed hypotheses. A good case can be made for researchers who are truly inspired by the empirical findings and make a good-faith effort to reformulate the theoretical underpinnings of their research. The reformulation does not necessarily increase the number of confirmed hypotheses, as the primary goal should be to provide an insightful, coherent theoretical basis for the research, but not to artificially inflate the positive findings.

PPHA and Theory Development

A major harm of PPHA is its prevention of the falsification strategy for theoretical breakthroughs (Kerr, 1998). Popper (1959) argued that falsification is crucial for scientific progress, and that an unfalsifiable theory is a useless theory. The confirmatory bias in general, and PPHA in particular, eliminate the opportunity to falsify a theory. Uchino, Thoman, and Byerly (2010) analysed the papers published in *Journal of Personality and Social Psychology* during 1982 and 2005. They found that the vast majority of the research was conducted with a confirmation or verification perspective, which, in their opinion, impedes theoretical development and innovation. My analysis of the papers published in *Academy of Management Journal* in 2009 suggests that the confirmatory strategy seems even more rampant in management research.

Kerr (1998) has described other negative effects of PPHA, some of which are also concerned with the obstruction of theory development. Specifically, the focus on generating hypotheses to fit a set of specific findings inhibits the development of general theories that can account for diverse findings in different contexts. This practice leads researchers to focus on generating hypotheses that best fit their findings in their research context. It also discourages them from looking beyond their findings to develop theoretical accounts of seemingly inconsistent findings

© 2011 The International Association for Chinese Management Research

https://doi.org/10.1111/j.1740-8784.2011.00222.x Published online by Cambridge University Press

across contexts. Another problem with PPHA is that it promotes the development of loose, undisconfirmable theories that can fit a wide range of findings, and hinders the development of precise theories. While a general theory that can account for findings from diverse contexts is appealing, it is useless if it cannot be disconfirmed or falsified. Finally, PPHA suppresses the consideration of rival explanations and the formulation of alternative theoretical models that are more powerful in accounting for diverse empirical evidence.

Managing PPHA

There is no simple solution to the potential problems brought about by PPHA. Whether PPHA is categorically unethical is debatable and controversial, not to mention that PPHA is hard to detect and prove. I do not see rigorous policing as effective in curbing the negative consequences of PPHA. What would be more effective is a reorientation of the standards used for evaluating scholarly publications, and I discuss several major strategies below:

Replication. Replications are required if unexpected findings are reported, and this practice has been around for centuries in science (Tsang & Kwan, 1999). If someone announces a cure for AIDS, I am sure that multiple laboratories across the globe will rush to replicate the effectiveness of the cure. Replications reduce the chance element involved in PPHA and promote genuine a priori hypotheses (Kerr, 1998). In top psychology journals, it is common to find multistudy papers that include replications. However, replications are not emphasized in management, and it is rare to find replications in a single paper. The scarcity of replication may be partly attributable to the emphasis on theory in management, and partly to the difficulty of obtaining data in the field, which is the primary source of data for management research.

I do not advocate the wholesale requirement for replication in empirical papers, but replications are highly useful for surprising findings that contradict well-established theories or common belief. An example is given by the surprising association between better physical well-being and economic recession, which replicates across different indicators of physical well-being (Ruhm, 2000) and across countries (Gerdtham & Ruhm, 2006). Complex findings based on small effect sizes, such as three-way interactions, also require replications. For a good example, see the three-way interaction between procedural fairness, outcome favourability, and self-construal that replicates across different studies (Brockner, Chen, Mannix, Leung, & Skarlicki, 2000).

Recognition of descriptive research. Hambrick (2007) has made a strong plea for recognizing descriptive research and provided several impressive examples from other business disciplines. Locke (2007) has illustrated the usefulness of inductive

© 2011 The International Association for Chinese Management Research

https://doi.org/10.1111/j.1740-8784.2011.00222.x Published online by Cambridge University Press

research for theory building. In fact, Nobel prizes have been awarded to researchers who made ground-breaking discoveries in science, and it is not crucial whether they have also come up with a compelling theoretical account of their findings. *Science* and *Nature*, perhaps the most sought after outlets in science, focus on empirical discoveries. In psychology, there is a recent trend to establish top journals that showcase empirical discoveries, and the prime example is *Psychological Science*.

I think management is in dire need of publication space in top journals that caters to ground-breaking empirical discoveries. Interestingly, the *Journal of Applied Psychology*, a major journal for organizational behaviour researchers, makes provisions for this type of research. According to the description of the journal, it does publish 'Descriptive research on applied psychological phenomena lacking basic knowledge in the literature that will provide a foundation for building new knowledge and theory' (American Psychological Association, 2010). The availability of prestigious outlets for descriptive research will reduce the need for PPHA.

Recognizing the value of rejected and post hoc hypotheses. Despite good-faith effort, researchers often fail to confirm all their hypotheses. In fact, this failure may be success in disguise, as Popper (1959) viewed falsification as the primary driver of theoretical progress. It is important for the field to recognize and accept well-reasoned hypotheses that turn out to be unsupported, and restrain from criticizing research with unconfirmed hypotheses as lacking a strong theoretical foundation. Reviewers and editors should not compel authors without strong and evidence-based reasons to delete unsupported hypotheses or modify their hypotheses to accommodate the empirical findings.

Encouragement of critical tests of competing hypotheses. To correct the positivity bias in reporting empirical results, Uchino et al. (2010) advocated for more frequent use of critical tests to evaluate competing hypotheses. This suggestion is valid for management research, because this research strategy is rare. Among the papers published in the *Academy of Management Journal* in 2009, only one presented competing hypotheses. An excellent example of this approach is provided by the debate on the competing mechanisms underlying the effect of participation on goal commitment and performance, which was settled by four critical experiments (Latham, Erez, & Locke, 1988). For a recent example of a critical test of competing hypotheses concerning the relationship between repeated partnership in interorganizational alliance and firm performance, see Goerzen (2007). The promotion of critical tests of competing hypotheses will lead to a lesser need for PPHA, because disconfirmed hypotheses in this context provide useful information about the validity of competing theories.

© 2011 The International Association for Chinese Management Research

https://doi.org/10.1111/j.1740-8784.2011.00222.x Published online by Cambridge University Press

CONCLUSION

PPHA can be problematic, leading to the inflation of positive results and suppression of theoretical innovation. Whether or not PPHA should be categorically discouraged is controversial, but I think most researchers would agree that this issue needs to be addressed. I believe that the successful management of PPHA requires a concerted effort at different levels. Researchers should be vigilant about whether they have unjustifiably enhanced the proportion of positive results and the validity of their theoretical account by engaging in PPHA. Researchers should abstain from PPHA that dresses up the validity of their research despite the publish-or-perish pressure. Management associations need to develop clearer ethical guidelines to regulate PPHA practices. Journal editors and reviewers should accord more recognition to replication, descriptive research, negative results, and critical tests.

Conducting high-quality research is a noble vocation that can be intensely satisfying. Researchers should collectively work towards setting clearer guidelines and promoting good practices to remove any self-doubt about fudging that makes papers appear more convincing. Most important, such guidelines can free researchers from the yoke of confirmatory bias in order to unleash their creative potential in theorizing.

NOTE

I thank Vivian Lun for her assistance in data analysis.

REFERENCES

Academy of Management. 2010. Information for contributors. [Last accessed 23 March 2011.] Available from URL: http://aom.pace.edu/amj/contributor_information.html

American Psychological Association. 2010. Journal of applied psychology. [Last accessed 24 March 2011.] Available from URL: http://www.apa.org/pubs/journals/apl/

Ayala, F. J. 2009. Darwin and the scientific method. ***Proceedings of National Academy of Sciences of the United States of America,*** 106(Suppl): 10033–10039.

Bem, D. J. 1987. Writing the empirical journal article. In M. Zanna & J. Darley (Eds.), ***The complete academic: A practical guide for the beginning social scientist:*** 171–201. New York: Random House.

Brockner, J., Chen, Y. R., Mannix, E. A., Leung, K., & Skarlicki, D. P. 2000. Culture and procedural fairness: When the effects of what you do depend on how you do it. ***Administrative Science Quarterly,*** 45(1): 138–159.

Canter, M., Bennett, B., Jones, S., & Nagy, T. 1994. ***Ethics for psychologists: Commentary on the APA ethics code.*** Washington, DC: American Psychological Association.

Chen, X. P. 2011. Author ethical dilemmas in the research publication process. ***Management and Organization Review,*** 7(3): 423–432.

Committee on Science, Engineering, and Public Policy. 1992. ***Responsible science: Ensuring the integrity of the research process.*** Washington, DC: National Academy Press.

Gerdtham, U. G., & Ruhm, C. J. 2006. Deaths rise in good economic times: Evidence from the OECD. ***Economics and Human Biology,*** 4(3): 298–316.

© 2011 The International Association for Chinese Management Research

https://doi.org/10.1111/j.1740-8784.2011.00222.x Published online by Cambridge University Press

Goerzen, A. 2007. Alliance networks and firm performance: The impact of repeated partnerships. ***Strategic Management Journal,*** 28(5): 487–509.

Hambrick, D. C. 2007. The field of management's devotion of theory: Too much of a good thing? ***Academy of Management Journal,*** 50(6): 1346–1352.

Jaccard, J., & Jacoby, J. 2010. ***Theory construction and model-building skills: A practical guide for social scientists.*** New York: The Guildford Press.

Kerr, N. L. 1998. HARKing: Hypothesizing after the results are known. ***Personality and Social Psychology Review,*** 2(3): 196–217.

Kirkman, B. L., & Chen, G. 2011. Maximizing your data or data slicing? Recommendations for managing multiple submissions from the same dataset. ***Management and Organization Review,*** 7(3): 433–446.

Latham, G. P., Erez, M., & Locke, E. A. 1988. Resolving scientific disputes by the joint design of crucial experiments by the antagonists: Application to the Erez–Latham dispute regarding participation in goal setting. ***Journal of Applied Psychology,*** 73(4): 753–772.

Locke, E. A. 2007. The case for inductive theory building. ***Journal of Management,*** 33(6): 867–890.

Popper, K. 1959. ***The logic of scientific discovery.*** New York: Basic Books.

Ruhm, C. J. 2000. Are recessions good for your health? ***Quarterly Journal of Economics,*** 115(2): 617–650.

Tsang, E. W. K., & Kwan, K. M. 1999. Replication and theory development in organizational science: A critical realist perspective. ***Academy of Management Review,*** 24(4): 759–780.

Turner, E. H., Matthews, A. M., Linardatos, E., Tell, R. A., & Rosenthal, R. 2008. Selective publication of antidepressant trials and its influence on apparent efficacy. ***New England Journal of Medicine,*** 358(3): 252–260.

Uchino, B. N., Thoman, D., & Byerly, S. 2010. Inference patterns in theoretical social psychology: Looking back as we move forward. ***Social and Personality Psychology Compass,*** 4(6): 417–427.

Kwok Leung (mkkleung@cityu.edu.hk) has a Ph.D. in psychology and is currently chair professor of management at City University of Hong Kong. His research areas include justice, conflict, and creativity, cross-cultural research methods, international business, and social axioms. He is Deputy-Editor-in-Chief of *Management and Organization Review* and the President of the International Association for Cross-Cultural Psychology. He is a fellow of the *Academy of International Business*, *Academy of Intercultural Research*, and *Association for Psychological Science.*

Manuscript received: October 21, 2010
Final version accepted: April 11, 2011
Accepted by: Maureen L. Ambrose and Marshall Schminke

© 2011 The International Association for Chinese Management Research

https://doi.org/10.1111/j.1740-8784.2011.00222.x Published online by Cambridge University Press