Error, quality, and applied geography: An editorial on process

Introduction

At first glance, it may seem counter intuitive to write a theoretical reflection for an applied journal, but this seeming oddity will begin to make sense once entrenched and pernicious habits of thought are overcome and theory is understood not as the opposite of practice, but as a guide to good practice, as "a practical means of going on" (Thrift, 1999: 304). Indeed, in this editorial I want to show how the insights of error theory might help produce better applied geography by raising awareness of the deeper significance of the peer-review process in applied scholarship. The intent of this editorial is to provide readers with a novel lens through which to make sense of the practices that collectively create the subdiscipline of applied geography (Gatrell & Jensen, 2009; Martin, 2001; Murphy, 2006; Ward, 2006, 2007; see also Johnston & Sidaway, 2004; Olds, 2001; Simandan, 2002, 2005, 2011).

Peer-review in *Applied Geography* is particularly problematic as the venue must balance "contributions to the literature" (that are often methodological and/or conceptual) with the "APPLIED" world. In this sense, error theory, a philosophy of scientific induction spurred by the work of American philosopher Deborah Mayo (Mayo, 1996, 2008; Mayo & Spanos, 2006) is potentially useful. Before outlining the key tenets of error theory, it is worth pausing for a short discussion of a) the scholarly territory that *Applied Geography* covers and b) the information from the editorial statement about what it takes to get a manuscript published in this journal. The specifics of decision making for publication in *Applied Geography* are presented on the journal's webpage by first delineating the kind of topics that are covered by the subdiscipline of applied geography:

Papers are invited on any theme involving the application of geographical theory and methodology in the resolution of human problems. This may include papers on the techniques, problems and results of environmental and/or social research, as well as those concerned with the principles, policies and consequences of resource management and allocation. Articles are refereed before publication.

Following this description, the publisher provides a schematic portrayal of the specifics of the review process. Thus, authors are told that:

The practice of peer-review is to ensure that only good science is published…Our referees play a vital role in maintaining the high standards of Applied Geography and all manuscripts are peer-reviewed following the procedure outlined below. The Editor first evaluates all manuscripts. It is rare, but it is possible for an exceptional manuscript to be accepted at this stage. Manuscripts rejected at this stage are insufficiently original, have serious scientific flaws, have poor grammar or English language, or are outside the aims and scope of the journal. Those that meet the minimum criteria are normally passed on to at least 2 experts for review. Applied Geography employs double-blind reviewing, where both the referee and author remain anonymous throughout the process. Whenever possible, referees are matched to the paper according to their expertise. Referees are asked to evaluate whether the manuscript: – Is original – Is methodologically sound – Follows appropriate ethical guidelines – Has results which are clearly presented and support the conclusions – Correctly references previous relevant work.

Whereas the foregoing portrayal of the review process provides some guidance to the authors by telling them WHAT will happen with their submission (see Fig. 1), it remains frustrating because no deeper RATIONALE for that particular policy is being offered.

Given that more self-aware and self-reflective scholars make better scholars (Johnston & Sidaway, 2004; Rose, 1993), this editorial aims to fill that unaddressed problem and to show WHY that particular editorial policy makes sense epistemologically. As I am about to show, error theory is especially suited for this explanatory task. In what follows, I first present its key tenets, and then, armed with its original conceptual toolbox, I return to the problem of the peer-review process to expose its deeper rationale.

Outline of error theory

To understand error theory, it is important to call attention to the counter-intuitive fact that the accumulation of information about an event is best described as a process of reducing informational entropy, that is, of eliminating all hypothetical explanations of said event, except the one that is actually true (Lipton, 2004). It immediately follows that the most efficient way to gain valid information ensues from abiding in one's reasoning by those procedures of enquiry that are particularly likely to minimise informational entropy, i.e. to shrink the space of contending hypotheses and settle on the correct explanation of the event of interest. For the sake of brevity, convenience, and concreteness, all types of procedures of enquiry in applied geography will be referred to by the shorthand "tests" or "error probes" (Mayo & Spanos, 2006). Some tests perform very poorly the task of ruling out incorrect explanatory hypotheses and they might well be called inconclusive, or lenient, or insevere tests. By using such unreliable error probes one is still beset with too many alternative explanations of the phenomenon of interest, and therefore, with a too high degree of uncertainty or ignorance as to where the truth lies. Some other
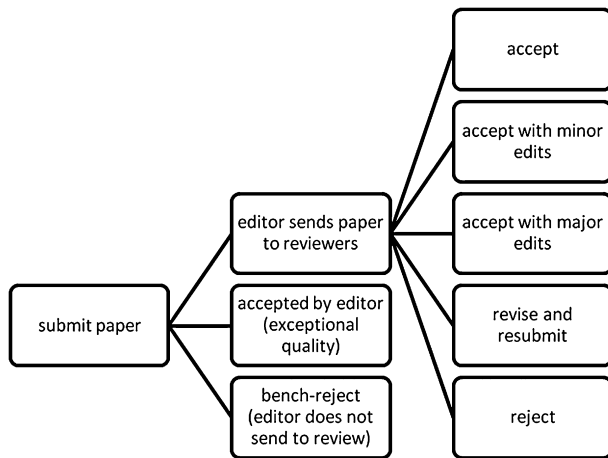
Fig. 1. The publication decision-process for *Applied Geography*.

tests turn out uncannily good at settling a dispute over competing hypotheses, and one would refer to them as conclusive, or stringent, or severe tests. By deft use of these highly reliable error probes, we can drastically contain our amount of uncertainty and, thereby, expand our practical knowledge commensurately.

Having introduced the distinction between lenient and severe tests, two related error-theoretical considerations will prove useful for the subsequent discussion. First, the very same error probe can be severe for some hypotheses, but lenient, or even minimally informative, for others. To give an example, a high score on a properly invigilated multiple-choice exam conclusively tests the hypothesis that the respective student knows the material, but cannot tell whether the knowledge was acquired by deep processing or merely by rote memorization (Mayo, 1996). The challenge is to discern exactly which hypothesis one wants to probe and then to design a test that is stringent for that hypothesis, specifically. Second, an effective and widespread tactic for building a severe test consists in amalgamating a collection of tests which, individually, have only low to moderate severity, but, taken together, perform just as well as a single stringent test. The warrant on which this tactic rests originates in Reichenbach's principle of the common cause (Reichenbach, 1956), itself grounded in the mathematical fact that the probability of a conjunction of moderately improbable individual events is itself extremely improbable (Tversky & Kahneman, 1983).

Applied geography through the lens of error theory

To begin to grasp how this novel theoretical vocabulary can deepen and refine our professional self-understanding as applied geographers, I will now return to analyse the practice discussed at the beginning of this essay, namely the submission of a paper for consideration of publication in *Applied Geography*. It is outside the scope of this editorial to review the vast interdisciplinary literature on the merits and demerits of the peer-review process for applied research. The aim here is to illustrate how a particular way of thinking – error theory – would frame and interpret the practice of peer-review. Translating this practice into a language amenable to the insights of error[1] theory is a rather straightforward

process. To be published in a leading applied geography journal, a paper must have successfully passed a battery of moderately stringent tests that, collectively, have severely tested the hypothesis "paper *x* is a world-class contribution to scholarship". More specifically, the testing consists in thoroughly probing the many ways in which the hypothesis may err from the truth, that is, the many ways in which the paper can fail to be a world-class contribution. Editors of major journals act as chief gatekeepers and coordinators of the testing process. Their own reputations depend on correctly accepting for publication all the papers that indeed are world-class contributions, but especially on correctly rejecting all the papers which fail this exacting standard of quality. To put this differently, as they go about doing their job of accepting and rejecting papers, the editors themselves become testable hypotheses. They all hope that, at the end of their mandate, fellow applied geographers will not bemoan the poor quality of the papers accepted for publication and that the citation impact of their journal will rise dramatically. These appear to be legitimate hopes[2] because, if materialised, they would count as highly stringent tests of the hypothesis "editor *y* is one of the best this journal has ever had". The authors of the submitted papers harbour their own kind of hopes, namely that their contributions will be accepted for publication; if true, this would amount to a severe test of the hypothesis "scholar *z* is a world-class applied geographer". In the short-term, publication in leading journals would directly increase scholars' self-esteem, while in the long-term, it would enable them to improve their standing in their home department (by applying for promotion, etc.) as well as in the whole discipline, more broadly (e.g. by moving to a better ranked department, etc.; see also Brunn, 1997).

If one proceeds analytically along the chain of publication, one notices that the severity of the error probes to which a paper is subjected tends to increase as one moves away from the author and further along the peer-reviewed publication process (see Fig. 2).

Indeed, the procedure of re-reading one's own drafts tends to be a rather lenient test of quality, for reasons that span both cognitive and emotional dimensions (narcissism, lack of critical distance, excessive enthusiasm and excitement, lack of sufficient knowledge, lack of perspective; the double curse of incompetence, etc.; see Ehrlinger, Johnson, Banner, Dunning, & Kruger, 2008; Clore & Huntsinger, 2007; Pronin, 2007). This widely acknowledged state of affairs explains not only why self-publication with vanity presses (i.e. a minimally severe test of quality) is unacceptable in academia, but also why authors frequently hear the advice to present their papers at conferences and to circulate their early drafts among their close colleagues (Brunn, 1997). Whereas it is generally correct to say that close colleagues and peers in attendance at conferences constitute better error probes of one's contribution, they can only seldom be relied on as truly severe tests. The easiest way to understand why this is so, is by comparing and contrasting them with the anonymous referees chosen by the editors themselves. The fact that the refereeing process is double-blind dramatically increases the stringency of the testing process because this procedural safeguard usually rules out the alternative explanations that the supposedly objective reports provided have been biased by personal conflicts or sympathies between referees and authors. In addition, the provision by the editor of a list with a taxonomy of errors to probe for

---

[1] Throughout her work, Deborah Mayo uses "error" as a broad term that encompasses the outcomes of human judgment as well (Mayo, 1996, 2008). This usage, while widely accepted in the philosophy of scientific induction (see e.g. Rescher, 2007), may be unfamiliar for some of the readers of this journal, who might prefer to think in terms of "reducing uncertainty" rather than "reducing error".

[2] It is worth keeping in mind, however, that: a) the inexorable tendency of impact factors to rise over time can, if not taken into account, compromise inferences about an editor's quality, and b) the motivations of authors and editors are not reducible to enhancing their own status or raising citation impacts. There are situations when authors and editors may have to disregard or temporarily suspend these particular motives. Along these lines, one reviewer noted that "there are other motivations, such as wanting to share interesting ideas or just wanting to do well, that drive some who publish".
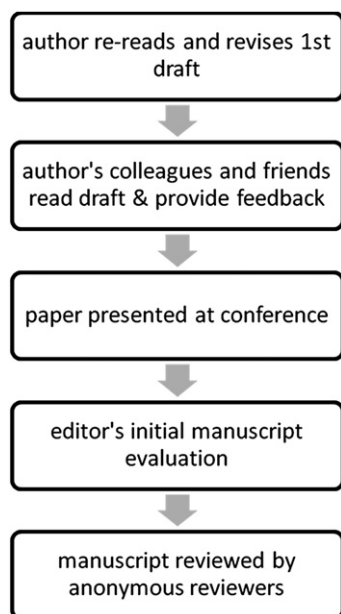
Fig. 2. The degree of severity with which a paper is tested increases as the manuscript moves away from the author and toward the anonymous reviewers.

(e.g. is the paper a significant original contribution?; is the exposition clear and does the conclusion logically follow from the argument?; is the methodology sound?; are the claims supported by appropriate evidence?; does the paper have broader social and policy applications?, etc.), further empowers the referees to operate as moderately-to-severe tests of the quality of a submitted paper. Since each referee is only boundedly rational (Simon, 1955, 2000), none of them has the omniscience and sagacity to detect each and all of the errors that may plague a particular submission. To address this all-too-human limitation, the editor compensates for the imperfect severity of each of the referees by selecting two or three of them. The higher their number, the higher the overall severity with which a paper is tested. Of course, since not all referees are created equal, a complementary tactic open to the editor is to increase severity by choosing only referees of the highest caliber. To be sure, these too are only boundedly rational, but their bounds happen to be more generous than those of the average scholar. The larger and better organised store of knowledge they possess (and see Ericsson, Nandagopal, & Roring, 2009) enables them to scrutinize arguments more perceptively and hence to operate as highly reliable, and, therefore, often sought after, error probes. Last but not least, the fact that the editor knows who the referees are, provides them with the incentive to act as severe tests of the submitted paper, because if they fail to do so and write shoddy or rabid reports instead (and see Dear, 2001), then the hypothesis "this referee is a serious, first-rate scholar" fails a highly probative test.

Before proceeding further, and for the sake of balanced judgment, it is important to highlight some of the weaknesses of the peer-review process, which jointly make it a less than maximally severe test of quality. The first weakness stems from the fact that the editor does know the identities and institutional affiliations of the authors, and hence is prone to ecologically fallacious judgments of the type "this author is from less prestigious institution and therefore the paper is probably weak; this author is from Elite University, and therefore the paper is probably strong" (see also the famous empirical study of this process by Peters & Ceci, 1982). The second weakness ensues from the fact that numerous factors conspire to undermine the anonymity of the authors to the reviewers themselves. These include the inherent nature of specialization (there is a strong chance that the reviewers have heard the paper presented earlier at a conference in their field) as well as the pervasive availability of search engines, which can easily tempt the reviewers to try to locate the respective paper's author. Finally, the third weakness results from the growing scale and faster pace of the review process. The surge in the number of academic outlets, the increased depersonalization of the editor–reviewer interaction, as well as the felt pressure to gain a competitive edge for journals by reducing the time granted to the reviewers to submit their reports might have led to an increased difficulty in finding appropriate referees (because they receive too many requests and because it is easier now to turn down an editor's request) and to a higher likelihood of receiving reviewer reports that are rushed and less-than-thorough.

A key tenet of error theory is that we learn by probing for error, acknowledging error, and addressing error, and nowhere is this insight more valuable than in explaining the fact that many, if not most, papers in applied geography are not simply accepted or rejected, but instead, undergo an elaborate social process of revision and resubmission. On one hand, because most submitted papers have passed only some lenient to moderately stringent tests (the authors, their colleagues, their audiences at conferences), they are likely to fail the more severe tests of the editor and of the anonymous referees. On the other hand, precisely because many papers have, nonetheless, already passed some such tests, they are likely to be *redeemable* failures. Seen as an actor in a broader network of scholars, the author receives the list of errors the editor and the referees have detected in one's redeemably failed paper, and addresses them. In the process, one learns new things and upgrades one's skill at uncovering and circumventing these newly comprehended types of error in one's future scholarship. In other words, the lens of error theory enables one to understand not only the moral (Livingstone, 2006; Ward, 2006, 2007), political (Castree, 2006; Sidaway, 2000; Ward, 2006, 2007) and instrumental (Bassett, 1999; Gatrell & Jensen, 2009; Martin, 2001; Murphy, 2006; Simandan, 2002; see also Frickel & Gross, 2005) dimensions of the careers of applied geographers but also their learning dimension. Given that we learn from our errors as we rewrite our papers for new resubmissions, and given that this rewriting is nothing less than the production of *better* (less error-ridden) geographical knowledge, it follows that error theory affords an alternative, and remarkably elegant, lens through which to understand just how we manage, incrementally, to achieve true progress in applied geography.

Acknowledgments

References

Bassett, K. (1999). Is there progress in human geography? The problem of progress in the light of recent work in the philosophy and sociology of science. *Progress in Human Geography, 23*(1), 27–47.

Brunn, S. (1997). Support for, and impacts of publishing in the Annals of the AAG: the authors and an editor speak. *The Professional Geographer, 3*, 364–371.

Castree, N. (2006). Geography's new public intellectuals. *Antipode, 38*(2), 396–412.

Clore, G., & Huntsinger, J. (2007). How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences, 11*(9), 393–399.

Dear, M. (2001). The politics of geography: hate mail, rabid referees, and culture wars. *Political Geography, 20*, 13–16.

Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2008). Why the unskilled are unaware? Further explorations of (lack of) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes, 105*, 98–121.

Ericsson, K. A., Nandagopal, K., & Roring, R. (2009). Toward a science of exceptional achievement: attaining superior performance through deliberate practice. *Annals of the New York Academy of Sciences, 1172*(1), 199–217.

Frickel, S., & Gross, N. (2005). A general theory of scientific/intellectual movements. *American Sociological Review, 70*, 204–232.

Gatrell, J., & Jensen, R. (Eds.). (2009). *Geotechnologies and the environment: Socioeconomic and planning applications*. Heidelberg: Springer.

Johnston, R. J., & Sidaway, J. (2004). *Geography and geographers: Anglo-American human geography since 1945* (6th ed.). London: Arnold.

Lipton, P. (2004). *Inference to the best explanation* (2nd ed.). London: Routledge.

Livingstone, D. N. (2006). Putting progress in its place. *Progress in Human Geography, 30*, 559–587, (with a commentary by Trevor Barnes and a rejoinder from the author).

Martin, R. (2001). Geography and public policy: the case of the missing agenda. *Progress in Human Geography, 25*, 189–209.

Mayo, D. (1996). *Error and the growth of experimental knowledge*. Chicago, IL: The University of Chicago Press.

Mayo, D. (2008). How to discount double-counting when it counts: some clarifications. *British Journal for the Philosophy of Science, 59*(4), 857–879.

Mayo, D., & Spanos, A. (2006). Severe testing as a basic concept in a Neyman–Pearson philosophy of induction. *British Journal for the Philosophy of Science, 57*, 323–357.

Murphy, A. (2006). Enhancing geography's role in public debate. *Annals of the Association of American Geographers, 96*(1), 1–13.

Olds, K. (2001). Practices for 'process geographies': a view from within and outside the periphery. *Environment and Planning D: Society and Space, 19*, 127–136.

Peters, D., & Ceci, S. (1982). Peer-review practices of psychological journals: the fate of published articles, submitted again. *Behavioral and Brain Sciences, 5*, 187–195.

Pronin, E. (2007). Perception and misperception of bias in human judgment. *Trends in Cognitive Sciences, 11*(1), 37–43.

Reichenbach, H. (1956). *The direction of time*. Berkeley, CA: University of California Press.

Rescher, N. (2007). *Error*. Pittsburgh, PA: University of Pittsburgh Press.

Rose, G. (1993). *Feminism and geography: The limits of geographical knowledge*. Minneapolis, MN: University of Minnesota Press.

Sidaway, J. D. (2000). Recontextualising positionality: geographical research and academic fields of power. *Antipode, 32*(3), 260–270.

Simandan, D. (2002). On what it takes to be a good geographer. *Area, 34*(3), 284–293.

Simandan, D. (2005). *New ways in geography*. Timisoara: West University Press.

Simandan, D. Is engaged pluralism the best way ahead for economic geography? Commentary on Barnes & Sheppard (2010). *Progress in Human Geography 35* (x): xxx-xxx, in press.

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics, 69*, 99–118.

Simon, H. A. (2000). Bounded rationality in social science: today and tomorrow. *Mind & Society, 1*(1), 25–39.

Thrift, N. (1999). Steps to an ecology of place. In J. Allen, D. Massey, & P. Sarre (Eds.), *Human geography today* (pp. 295–322). Cambridge: Polity Press.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment. *Psychological Review, 90*, 293–315.

Ward, K. (2006). Geography and public policy: towards public geographies. *Progress in Human Geography, 30*, 495–503.

Ward, K. (2007). Geography and public policy: activist, participatory and policy geographies. *Progress in Human Geography, 31*, 695–705.

Dragos Simandan*
Geography Department, Brock University, 500 Glenridge Avenue, St. Catharines, Ontario, Canada L2S 3A1
* Tel.: +1 905 688 5550x5010; fax: +1 905 688 6369.
E-mail address: simandan@brocku.ca